DYNAMICS OF THE TIME TO THE MOST RECENT COMMON ANCESTOR IN A LARGE BRANCHING POPULATION

STEVEN N. EVANS AND PETER L. RALPH

ABSTRACT. If we follow an asexually reproducing population through time, then the amount of time that has passed since the most recent common ancestor (MRCA) of all current individuals lived will change as time progresses. The resulting stochastic process has been studied previously when the population has a constant large size and evolves via the diffusion limit of standard Wright-Fisher dynamics. We investigate cases in which the population varies in size and evolves according to a class of models that includes suitably conditioned $(1 + \beta)$-stable continuous state branching processes (in particular, it includes the conditioned Feller continuous state branching process). We also consider the discrete time Markov chain that tracks the MRCA age just before and after its successive jumps. We find transition probabilities for both the continuous and discrete time processes, determine when these processes are transient and recurrent, and compute stationary distributions when they exist. We also introduce a new family of Markov processes that stand in a relation with respect to the $(1 + \beta)$-stable continuous state branching process that is similar to the one between the Bessel-squared diffusions and the Feller continuous state branching process.

1. INTRODUCTION

Any asexually reproducing population has a unique most recent common ancestor, from whom the entire population is descended. In sexually reproducing species, the same is true for each nonrecombining piece of DNA. For instance, our "mitochondrial Eve" from whom all modern-day humans inherited their mitochondrial DNA is estimated to have lived around 180,000 years ago [IKPG00], while our "Y-chromosomal Adam" is estimated to have lived around 50,000 years ago [TPS+00]. There have also been efforts to estimate the time since the MRCA lived (which we will also call the "age of the MRCA") in populations of other organisms, particularly pathogens [TCG+04, VBL+01]. These studies, using sophisticated models incorporating of demographic history, are focused on estimating the age of the MRCA at a single point in time (the present).

As time progresses into the future, eventually the mitochondrial lineages of all but one of the daughters of the current mitochondrial Eve will die out, at which point the new mitochondrial Eve will have lived somewhat later in time. The age of the MRCA is thus a dynamically evolving *process* that exhibits periods of upwards linear growth separated by downwards jumps.

Recently, [PW06] and [SD06] independently investigated the MRCA age process for the diffusion limit of the classical Wright-Fisher model. The Wright-Fisher model is perhaps the most commonly used model in population dynamics: each individual in a fixed size population independently gives birth to an identically distributed random number of individuals (with finite variance), and after the new offspring are produced some are chosen at random to survive so that the total population size remains constant. The diffusion limit arises by letting the population size go to infinity and taking the time between generations to be proportional to the reciprocal of the population size.

In this paper, we investigate the MRCA age process for a parametric family of population models in a setting that permits the population size to vary with time and, by suitable choice of parameters, allows control over the extent to which rare individuals can have large numbers of offspring that survive to maturity. The model for the dynamics of the population size is based on the *critical* $(1+\beta)$-*stable continuous-state branching process* for $0 < \beta \leq 1$. These processes arise as scaling limits of Galton-Watson branching processes as follows.

Write $Z_t^{(n)}$ for the number of individuals alive in a critical continuous-time Galton-Watson branching process with branching rate λ and offspring distribution γ. The distribution γ has mean 1 (and thus, the process is "critical"). Suppose that if W is a random variable with distribution γ, then the random walk with steps distributed as the random variable $(W-1)$ falls into the domain of attraction of a stable process of index $1 + \beta \in (1, 2]$. The case $\beta = 1$ corresponds to γ having finite variance and the random walk converging to Brownian motion after rescaling. Set $X_t^{(n)} = n^{-(1+\frac{1}{\beta})} Z_t^{(n)}$ and suppose that $X_0^{(n)} \to x$ as $n \to \infty$. Then, up to a time-rescaling depending on λ and the scaling of the stable process above, the processes $X^{(n)}$ converge to a Markov process X which is a critical $(1 + \beta)$-stable continuous-state branching process, and whose distribution is determined by the Laplace transform

$$(1.1) \qquad \mathbb{E}\left[e^{-\theta X_t} \,\middle|\, X_0 = x \right] = \exp\left(-\frac{\theta x}{(1 + \theta^\beta t)^{1/\beta}} \right).$$

If $\beta = 1$, this is Feller's critical continuous-state branching process [Fel51, Kni81]. (Note that time here is scaled by a factor of 2 relative to some other authors, so that the generator of our "Feller continuous-state branching process" is $x \frac{\partial^2}{\partial x^2}$.)

Let $\tau = \inf\{t > 0 : X_t = 0\}$ denote the *extinction time* of X (it is not hard to show that $X_t = 0$ for all $t \geq \tau$). Taking $\theta \to \infty$ in (1.1) gives

$$\mathbb{P}\{\tau > t \mid X_0 = x\} = 1 - \exp\left(-\frac{x}{t^{1/\beta}} \right),$$

so X dies out almost surely. However, it is possible to condition X to live forever in the following sense:

$$\lim_{T \to \infty} \mathbb{E}[f(X_t) \mid X_0 = x, \ \tau > T] = \frac{1}{x} \mathbb{E}[f(X_t) X_t \mid X_0 = x].$$

Thus, if $P_t(x', dx'')$ are the transition probabilities of X, then there is a Markov process Y with transition probabilities

$$Q_t(y', dy'') = \frac{1}{y'} P_t(y, y'') y''.$$

The process Y is the critical $(1+\beta)$-stable continuous-state branching process X *conditioned on non-extinction*. The distribution of Y_t is determined by its Laplace transform

$$(1.2) \qquad \mathbb{E}[\exp(-\theta Y_t) \mid Y_0 = y] = \exp\left(-\frac{y\theta}{(t\theta^\beta + 1)^{-1/\beta}}\right)\left(1 + t\theta^\beta\right)^{-\frac{\beta+1}{\beta}}$$

— see Section 6. Moreover, it is possible to start the process Y from the initial state $Y_0 = 0$, and the formula (1.2) continues to hold for $y = 0$. The super–process generalization of this construction was considered for $\beta = 1$ in [EP90, Eva93, Eva92] and for general β in [EW03].

For $\beta = 1$, the conditioned process Y can be described informally as a single "immortal particle" constantly throwing off infinitesimally small masses with each mass then evolving according to the dynamics of the unconditioned process. These infinitesimal masses can be interpreted as the single progenitors of families whose lineage splits from the immortal particle at the birth time of the progenitor and are eventually doomed to extinction. Most such families die immediately but a rare few live for a non-infinitesimal amount of time. More formally, there is a σ-finite measure ν on the space of continuous positive excursion paths

$$\mathcal{E}^0 := \{u \in C(\mathbb{R}_+, \mathbb{R}_+) : u_0 = 0 \ \& \ \exists \gamma > 0 \text{ s.t. } u_t > 0 \ \Leftrightarrow \ 0 < t < \gamma\}$$

such that if Π is a Poisson point process on $\mathbb{R}_+ \times \mathcal{E}^0$ with intensity $\lambda \otimes \nu$, where λ is Lebesgue measure, and $(\bar{X}_t)_{t \geq 0}$ is an independent copy of X begun at $\bar{X}_0 = y$, then the process

$$(1.3) \qquad \left(\bar{X}_t + \sum_{(s,u)\in\Pi} u_{(t-s)\vee 0}\right)_{t \geq 0}$$

has the same distribution as $(Y_t)_{t \geq 0}$ begun at $Y_0 = y$ — see [Eva93]. A point $(s, u) \in \Pi$ corresponds to a family that grows to non-negligible size: the time s records the moment the family splits off from the immortal particle, and the value u_r of the trajectory u gives the size of the family τ units of time after it split off. The family becomes extinct after the period of time $\gamma(u) := \inf\{r > 0 : u(t) = 0, \ \forall t > r\}$. The σ-finite measure ν is Markovian with transition probabilities the same as those of the unconditioned process X — in other words, ν arises from a family of entrance laws for the semigroup of X. The process $(\bar{X}_t)_{t \geq 0}$ records the population mass due to descendants of individuals other than the immortal particle who are present at time 0.

An analogous description of the conditioned process Y for the case $\beta \in (0, 1)$ is presented in [EW03]. There is again a single immortal lineage, but now families split off from that lineage with a non-infinitesimal initial size, reflecting the heavy-tailed offspring distributions underlying these models. More precisely, a decomposition similar to (1.3) holds, but the Poisson point process Π is now on $\mathbb{R}_+ \times \mathcal{E}$ where $\mathcal{E} = \{u \in D(\mathbb{R}_+, \mathbb{R}_+) : \exists \gamma > 0 \text{ s.t. } u_t > 0 \ \Leftrightarrow \ 0 \leq t < \gamma\}$, the set of càdlàg paths starting above zero that eventually hit zero. The non-decreasing process $(M_t)_{t \geq 0}$, where

$$M_t := \sum_{(s,u)\in\Pi} u_0.$$

is the total of the initial family sizes that split off from the immortal particle in the time interval $[0, t]$. It is a stable subordinator of index β.

Suppose now that $\beta \in (0,1]$ is arbitrary. Take $Y_0 = 0$, so that $\bar{X}_t \equiv 0$ in the decomposition (1.3) and all "individuals" belong to families that split off from the immortal particle at times $s \geq 0$. Extend the definition of $\gamma(u)$ given above for $u \in \mathcal{E}_0$ to $u \in \mathcal{E}$ in the obvious way. The individuals besides the immortal particle alive at time $t > 0$ belong to families that correspond to the subset $\mathcal{A}_t := \{(s,u) \in \Pi : 0 \leq t - s < \gamma(u)\}$ of the random set Π. At time t, the amount of time since the most recent common ancestor of the entire population lived is $A_t := \sup\{t - s : (s,u) \in \mathcal{A}_t\}$. As depicted in Figure 1, the MRCA age process $(A_t)_{t \geq 0}$ has saw-tooth sample-paths that drift up with slope 1 until the current oldest family is extinguished, at which time they jump downward to the age of the next-oldest family.

It is not necessary to know the Poisson point process Π in order to construct the MRCA age process $(A_t)_{t \geq 0}$. Clearly, it is enough to know the point process Λ on $\mathbb{R}_+ \times \mathbb{R}_{++}$ given by $\Lambda := \{(s, \gamma(u)) : (s,u) \in \Pi\}$. Indeed, if we define the *left-leaning wedge with base at (t,x)* to be the the set

$$(1.4) \qquad \triangle(t,x) := \{(u,v) \in \mathbb{R}^2 : u < t \ \& \ v - u > x - t\},$$

then

$$A_t = t - \inf\{s : \exists x > 0 \text{ s.t. } (s,x) \in \Lambda \cap \triangle(t,0)\}$$

— see Figure 1.

Note that Λ is a Poisson point process with intensity $\lambda \otimes \mu$, where μ is the push-forward of ν by γ; that is, $\mu((t,\infty)) = \nu(\{u : \gamma(u) > t\})$. We will show in Section 5 that $\mu((t,\infty)) = (1+\beta)/(\beta t)$.

With these observations in mind, we see that if Λ is now **any** Poisson point process on $\mathbb{R}_+ \times \mathbb{R}_{++}$ with intensity $\lambda \otimes \mu$, where μ is **any** σ-finite measure on $\mathbb{R}_{++}$ with $\mu(\mathbb{R}_{++}) = \infty$ and $0 < \mu((x,\infty)) < \infty$ for all $x > 0$, then the construction that built $(A_t)_{t \geq 0}$ from the particular point process Λ considered above will still apply, and produce an $\mathbb{R}_+$-valued process with saw-tooth sample paths. We are therefore led to the following general definition.

Definition 1.1. let Λ be a Poisson point process on $\mathbb{R}_+ \times \mathbb{R}_{++}$ with intensity measure $\lambda \otimes \mu$, where λ is Lebesgue measure and μ is a σ-finite measure on $\mathbb{R}_{++}$ with $\mu(\mathbb{R}_{++}) = \infty$ and $\mu((x,\infty)) < \infty$ for all $x > 0$. Define $(A_t)_{t \in \mathbb{R}_+}$ by

$$A_t := t - \inf\{s \geq 0 : \exists x > 0 \text{ s.t. } (s,x) \in \Lambda \cap \triangle(t,0)\},$$

where $\triangle(t,0)$ is defined by 1.4, and $A_t = 0$ if $\Lambda \cap \triangle(t,0)$ is empty.

We will suppose from now on that we are in this general situation unless we specify otherwise. We will continue to use terminology appropriate for the genealogical setting and refer to $(A_t)_{t \geq 0}$ as the *MRCA age process* and μ as the *lifetime measure*. We will assume for convenience that the measure μ is absolutely continuous, with a density m with respect to Lebesgue measure that is positive Lebesgue almost everywhere. It is straightforward to remove these assumptions.

The strong Markov property of the Poisson point processes Λ implies that $(A_t)_{t \geq 0}$ is a time-homogeneous strong Markov process. In particular, there is a family of probability distributions $(\mathbb{P}^x)_{x \in \mathbb{R}_+}$ on the space of $\mathbb{R}_+$-valued càdlàg paths, with $\mathbb{P}^x$ interpreted in the usual way as the "distribution of $(A_t)_{t \geq 0}$ started from $A_0 = x$". More concretely, the probability measure $\mathbb{P}^x$ is the distribution of the process $(A_t^x)_{t \geq 0}$ defined as follows. Let Λ^x be a point process on $[-x, \infty) \times \mathbb{R}_{++}$ that has the
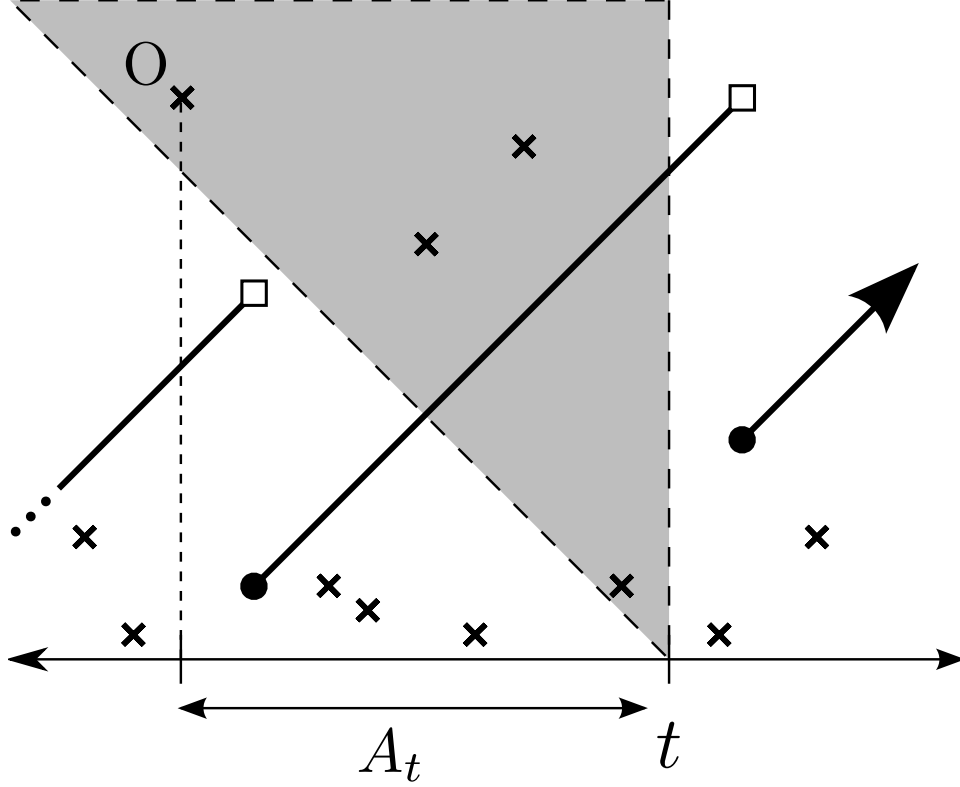
FIGURE 1. The points of Λ (marked "x"), the sample path of the process A (solid line), the oldest extant family at time t (labeled "O"), and the left-leaning wedge $\triangle(t, 0)$ (the shaded region).

distribution of the random point set $\{(t-x, y) : (t, y) \in \Lambda\} \cup \{(-x, Z)\}$, where Z is an independent random variable with values in the interval (x, ∞) and distribution

$$\mathbb{P}\{Z \leq z\} = \mu((x, z]) / \mu((x, \infty)).$$

Then we can define

$$A_t^x := t - \inf\{s \geq -x : \exists y > 0 \text{ s.t. } (s, y) \in \Lambda^x \cap \triangle(t, 0)\}, \quad t \geq 0$$

— see the proof of part (a) of Theorem 1.1 below for more detail. From now on, when we speak of the process $(A_t)_{t \geq 0}$ we will either be referring to the process constructed as in Definition 1.1 from the Poisson process Λ on a probability space equipped with the probability measure $\mathbb{P}$ or the canonical process on the space of càdlàg $\mathbb{R}_+$-valued paths equipped with the family of probability measures $(\mathbb{P}^x)_{x \geq 0}$. This should cause no confusion.

We prove the following properties of the process $(A_t)_{t \geq 0}$ in Section 2.

Theorem 1.1. (a) *The transition probabilities of the time-homogeneous Markov process $(A_t)_{t \geq 0}$ have an absolutely continuous part*

$$\mathbb{P}^x\{A_t \in dy\} = \frac{\mu((x, x+t])}{\mu((x, \infty))} \exp\left(-\int_y^{x+t} \mu((z, \infty)) \, dz\right) \mu((y, \infty)) \, dy,$$

for $y < x + t$ and a single atom

$$\mathbb{P}^x\{A_t = x + t\} = \frac{\mu((x+t, \infty))}{\mu((x, \infty))}.$$

(b) *The total rate at which the process $(A_t)_{t \geq 0}$ jumps from state $x > 0$ is*

$$\frac{m(x)}{\mu((x, \infty))},$$

and when the process jumps from state $x > 0$, the distribution of the state to which it jumps is absolutely continuous with density

$$y \mapsto \exp\left(-\int_y^x \mu((z, \infty))\, dz\right) \mu((y, \infty)), \quad 0 < y < x.$$

(c) *The probability $\mathbb{P}^0\{\exists t > 0 : A_t = 0\}$ that the process A returns to the state zero is positive if and only if*

$$\int_0^1 \exp\left(\int_x^1 \mu((y, \infty))\, dy\right)\, dx < \infty.$$

(d) *If*

$$\int_1^\infty \exp\left(-\int_1^x \mu((y, \infty))\, dy\right)\, dx = \infty,$$

then for each $x > 0$ the set $\{t \geq 0 : A_t = x\}$ is $\mathbb{P}^x$-almost surely unbounded. Otherwise, $\lim_{t \to \infty} A_t = \infty$, $\mathbb{P}^x$-almost surely, for all $x \geq 0$.

(e) *A stationary distribution π exists for the process $(A_t)_{t \geq 0}$ if and only if*

$$\int_1^\infty \mu((z, \infty))\, dz < \infty,$$

in which case it is unique, and

$$\pi(dx) = \mu((x, \infty)) \exp\left(-\int_x^\infty \mu((z, \infty))\, dz\right)\, dx.$$

(f) *If the process $(A_t)_{t \geq 0}$ has a stationary distribution π, then*

$$d_{TV}(\mathbb{P}^x\{A_t \in \cdot\}, \pi) \leq 1 - \exp\left(-\int_{t+x}^\infty \mu((y, \infty))\, dy\right) \times \frac{\mu([x, x+t))}{\mu([x, \infty))}.$$

where d_{TV} denotes the total variation distance. In particular, the distribution of A_t under $\mathbb{P}^x$ converges to π in total variation as $t \to \infty$.

Specializing Theorem 1.1 to the case when A is the MRCA age process of the conditioned critical $(1 + \beta)$-stable continuous state branching process gives parts (a) to (d) of the following result. Part (e) follows from an observation that a space-time rescaling of this MRCA age process is a time-homogeneous Markov process that arises from another Poisson process by the general MRCA age construction of Definition 1.1. The proof is in Section 5.

Corollary 1.1. *Suppose that A is the MRCA age process associated with the critical $(1 + \beta)$-stable continuous-state branching process.*

(a) *The transition probabilities of the process A have an absolutely continuous part*

$$\mathbb{P}^x\{A_t \in dy\} = \frac{(1 + \beta)t y^{1/\beta}}{\beta(x+t)^{2+1/\beta}}\, dy, \qquad 0 < y < x + t,$$

and a single atom

$$\mathbb{P}^x \{A_t = x + t\} = \frac{x}{x+t}.$$

(b) *The total rate at which the process A jumps from the state $x > 0$ is $1/x$, and when it jumps from state $x > 0$, the distribution of the state to which it jumps is absolutely continuous with density*

$$(1 + 1/\beta) \frac{y^{1/\beta}}{x^{1+1/\beta}}, \quad 0 < y < x.$$

(c) *The probability $\mathbb{P}^0\{\exists t > 0 : A_t = 0\}$ that the process A returns to the state zero is 0.*

(d) *For each $x \geq 0$, $\lim_{t \to \infty} A_t = \infty$, $\mathbb{P}^x$-almost surely.*

(e) *The process*

$$\left(e^{-t} A_{e^t}\right)_{t \in \mathbb{R}}$$

indexed by the whole real line is a time-homogeneous Markov process under $\mathbb{P}^x$ for any $x \geq 0$, and it is stationary when $x = 0$. Moreover, A_t/t converges in distribution to the $\mathrm{Beta}(1 + 1/\beta, 1)$ distribution as $t \to \infty$ under $\mathbb{P}^x$ for any $x \geq 0$, and A_t/t has the $\mathrm{Beta}(1 + 1/\beta, 1)$ distribution for all $t > 0$ when $x = 0$.

Note that the sample paths of $(A_t)_{t \geq 0}$ have local "peaks" immediately before jumps and local "troughs" immediately after. We investigate the discrete time Markov chain of successive pairs of peaks and troughs in Section 4. We also consider the jump heights and inter-jump intervals, and describe an interesting duality between these sequences in Section 3.

Finally, recall that the Bessel-squared process in dimension γ, where γ is an arbitrary non-negative real number, is the $\mathbb{R}_+$-valued diffusion process with infinitesimal generator $2x d^2/dx^2 + \gamma d/dx$. When γ is a positive integer, such a process has the same distribution as the square of the Euclidean norm of a Brownian motion in $\mathbb{R}^\gamma$. Feller's critical continuous-state branching process is thus, modulo a choice of scale in time or space, the zero-dimensional Bessel-squared process. It was shown in Example 3.5 of [PY82] that for $0 \leq \gamma < 2$ the Bessel-squared process with dimension γ conditioned on never hitting zero is the Bessel-squared process with dimension $4 - \gamma$. Thus, for $\beta = 1$, the conditioned process Y is the four-dimensional Bessel-squared process. We introduce a new family of processes in Section 6 that are also indexed by a non-negative real parameter and play the role of the Bessel-squared family for values of β other than 1. These processes will be studied further in a forthcoming paper.

We end this introduction by commenting on the connections with previous work. Firstly, we may think of each point $(s, x) \in \Lambda$ as a "job" that enters a queue with infinitely many servers at time s and takes x amount of time to complete. We thus have a classical $M/G/\infty$ queue [Tak62], except we are assuming that the arrival rate is infinite. With this interpretation, the quantity A_t is how long the oldest job at time t has been in the queue. Many properties of $M/G/\infty$ queues have been studied (see, for example, [Eli07]), but this work seems somewhat orthogonal to ours.

Secondly, note that the process $(A_t)_{t \geq 0}$ is an example of a *piecewise deterministic Markov process*: it consists of deterministic flows punctuated by random jumps. Such processes were introduced in [Dav84] and studied further in [Dav93] (see also

[JS96], where the nomenclature *jumping Markov processes* is used). The general properties of such processes have been studied further in, for example, [DC99, CD08, CDP01].

Lastly, piecewise deterministic Markov processes like $(A_t)_{t \geq 0}$ that have periods of linear increase interspersed with random jumps have been used to model many phenomena, such as stress in an earthquake zone [BVJ00], congestion in a data transmission network [DGR02], and growth-collapse [BPSZ06]. They also have appeared in the study of the additive coalescent [EP98] and $\mathbb{R}$-tree-valued Markov processes [EPW06].

2. Proof of Theorem 1.1

(a) Suppose that $(A_t)_{t \geq 0}$ is constructed from Λ as in Definition 1.1. For $s \geq x$ consider the conditional distribution of A_{s+t} given $A_s = x$. The condition $A_s = x$ is equivalent to the requirements that there is a point $(s-x, Z)$ in Λ for some $Z > x$ and that furthermore there are no points of Λ in the left-leaning wedge $\triangle(s-x, x)$. The conditional probability of the event $\{Z > z\}$ is $\mu((z, \infty))/\mu((x, \infty))$ for $z \geq x$. If $Z > x + t$, then $A_{s+t} = x + t$. Otherwise, $A_{s+t} < x + t$. The second claim of part (a) follows immediately.
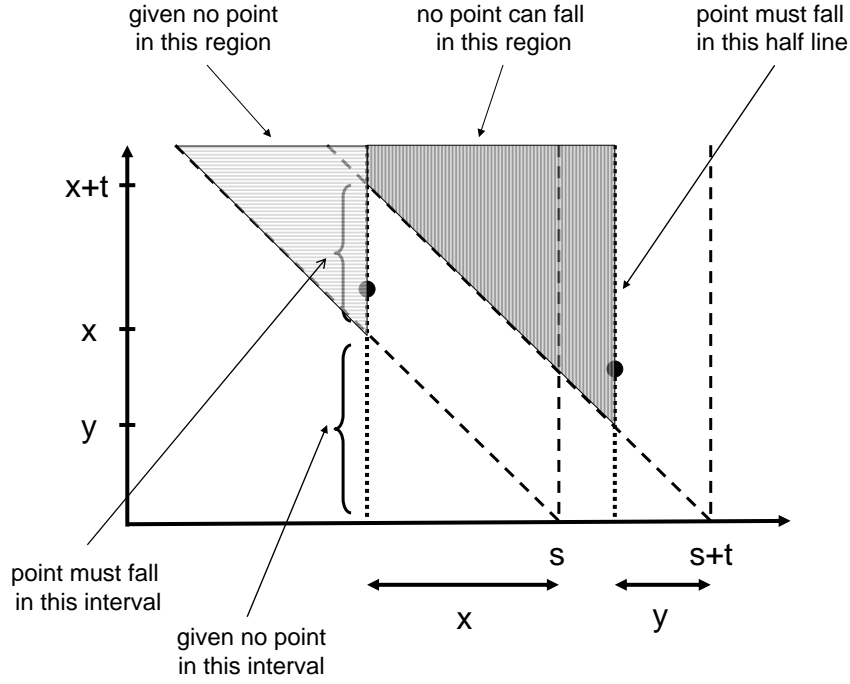

Figure 2. The computation of $\mathbb{P}\{A_{s+t} \in dy \mid A_s = x\}$.

Now consider $\mathbb{P}\{A_{s+t} \in dy \mid A_s = x\}$ for $y < x + t$. This case is depicted in Figure 2. By construction, $A_{s+t} = y$ if and only if there is a point $(s+t-y, W) \in \Lambda$ for some $W > y$, and there are no points of Λ in $\triangle(s + t - y, y)$. From above, the

condition $A_s = x$ requires there to be no points of Λ in the wedge $\triangle(s - x, x)$ (the lightly shaded region in Figure 2). Therefore, if $A_s = x$, then $A_{s+t} = y$ if and only if $Z \leq x + t$, there are no points of Λ in the darkly shaded region of Figure 2), and there is a point of Λ of the form $(s + t - y, w)$ with $w > y$.

Now, the conditional probability that $Z \leq x + t$ is

$$\frac{\mu((x, x + t])}{\mu((x, \infty))}.$$

The probability that no points of Λ are in the darkly shaded region is

$$\exp\left(-\int_y^{x+t} \mu((u, \infty))\, du\right).$$

The probability that Λ has a point in the infinitesimal region $[s + t - y, s + t - y + dy] \times (y, \infty)$ is $\mu((y, \infty))\, dy$. Multiplying these three probabilities together gives the first claim of part (a).

(b) Both claims follow readily by differentiating the formulae in (a) at $t = 0$. It is also possible to argue directly from the representation in terms of Λ.

(c) Suppose that $(A_t)_{t \geq 0}$ is constructed from Λ as in Definition 1.1. Note that

$$\{t > 0 : A_t = 0\} = \mathbb{R}_{++} \setminus \bigcup_{(t,x) \in \Lambda} (t, t + x).$$

The question of when such a "Poisson cut out" random set is almost surely empty was asked in [Man72], and the necessary and sufficient condition presented in part (c) is simply the one found in [She72] (see also [FFS85]).

(d) For $x > 0$, the random set $\{t \geq 0 : A_t = x\}$ is a discrete regenerative set under $\mathbb{P}^x$ that is not almost surely equal to $\{0\}$. Hence this set is a renewal process with an inter-arrival time distribution that possibly places some mass at infinity, in which case the number of arrivals is almost surely finite with a geometric distribution and the set is almost surely bounded.

Suppose that the set $\{t \geq 0 : A_t = x\}$ is $\mathbb{P}^x$-almost surely unbounded for some $x > 0$. Let $0 = T_0 < T_1 < \dots$ be the successive visits to x. It is clear that $\mathbb{P}^x\{\exists t \in [0, T_1] : A_t = y\} > 0$ for any choice of $y > 0$ and hence, by the strong Markov property, the set $\{t \geq 0 : A_t = y\}$ is also $\mathbb{P}^x$-almost surely unbounded. Another application of the strong Markov property establishes that the set $\{t \geq 0 : A_t = y\}$ is $\mathbb{P}^y$-almost surely unbounded. Thus, the set $\{t \geq 0 : A_t = x\}$ is either unbounded $\mathbb{P}^x$-almost surely for all $x > 0$ or bounded $\mathbb{P}^x$-almost surely for all $x > 0$.

A straightforward argument using the saw-tooth nature of the sample paths of A shows that the expected number of visits to x under $\mathbb{P}^x$ is a finite multiple of

$$\int_0^\infty \frac{\mathbb{P}^x\{A_t \in dx\}}{dx}\, dt.$$

Using the expression from part (a) and the argument above, this quantity is infinite, and hence the number of visits is P^x-almost surely infinite, if and only if

$$\int_1^\infty \exp\left(-\int_1^u \mu((y, \infty))\, dy\right) du = \infty.$$

If the set $\{t \geq 0 : A_t = x\}$ is $\mathbb{P}^x$-almost surely bounded for all $x > 0$, then, by an argument similar to the above, the set $\{t \geq 0 : A_t = y\}$ is $\mathbb{P}^x$-almost surely bounded for all $x, y > 0$. It follows that, for all $x > 0$, $\mathbb{P}^x$-almost surely all of the

sets $\{t \geq 0 : A_t = y\}$ are finite. This implies that $\lim_{t\to\infty} A_t$ exists $\mathbb{P}^x$-almost surely and takes values in the set $\{0, \infty\}$. However, it is clear from the Poisson process construction that 0 does not occur as a limit with positive probability.

(e) Suppose there exists a probability measure π on $\mathbb{R}_+$ such that

$$\int_{\mathbb{R}_+} \mathbb{P}\{A_t \in dy \mid A_0 = x\}\,\pi(dx) = \pi(dy), \quad y \in \mathbb{R}_+.$$

Taking $t \to \infty$ in part (a) gives

$$\begin{aligned}
\pi(dy) &= \lim_{t\to\infty} \int_{\mathbb{R}_+} \pi(dx) \frac{\mu((x, x+t])}{\mu((x, \infty))} \exp\left(-\int_y^{x+t} \mu((u, \infty))\,du\right) \mu((y, \infty))\,dy \\
&= \begin{cases} 0, & \text{if } \int_y^\infty \mu((u, \infty))\,du = \infty, \\ \exp\left(-\int_y^\infty \mu((u, \infty))\,du\right) \mu((y, \infty))\,dy, & \text{otherwise.} \end{cases}
\end{aligned}$$

Therefore, a stationary probability distribution exists if and only if $\int_y^\infty \mu((u, \infty))\,du < \infty$, and if a stationary distribution exists, then it is unique.

(f) It will be useful to begin with a concrete construction of a stationary version of the process A in terms of a Poisson point process. Suppose that $\int_x^\infty \mu((u, \infty))\,du < \infty$ for all $x > 0$, so that a stationary distribution exists. Let $\Lambda^{\leftrightarrow}$ be a Poisson point process on $\mathbb{R} \times \mathbb{R}_{++}$ with intensity measure $\lambda \otimes \mu$. Define $(A_t^{\leftrightarrow})_{t\in\mathbb{R}}$ by

$$A_t^{\leftrightarrow} := t - \inf\{s : \exists x > 0 \text{ s.t. } (s, x) \in \Lambda^{\leftrightarrow} \cap \triangle(t, 0)\}.$$

The condition on μ ensures that almost surely any wedge $\triangle(t, x)$ with $x > 0$ will contain only finitely many points of $\Lambda^{\leftrightarrow}$, and so $(A_t^{\leftrightarrow})_{t\in\mathbb{R}}$ is well-defined. The process $(A_t^{\leftrightarrow})_{t\in\mathbb{R}}$ is stationary and Markovian, with the same transition probabilities as $(A_t)_{t\geq 0}$.

Recall the construction of the process A^x started at x for $x > 0$ that was described preceding the statement of Theorem 1.1. Construct the point process Λ^x that appears there by setting $\Lambda^x := \{(t, y) \in \Lambda^{\leftrightarrow} : t > -x\} \cup \{(-x, Z)\}$, where Z is an independent random variable with values in the interval (x, ∞) and distribution $\mathbb{P}\{Z > z\} = \mu((z, \infty))/\mu((x, \infty))$.

By construction, $A_t^x = A_t^{\leftrightarrow}$ for all $t \geq T$, where T is the death time of all families alive at time $-x$ in either process:

$$T := \inf\{t > 0 : Z \leq t + x \text{ and } \Lambda^{\leftrightarrow} \cap \triangle(-x, t+x) = \emptyset\}.$$

Thus,

$$\begin{aligned}
d_{TV}(\mathbb{P}^x\{A_t \in \cdot\}, \pi) &\leq \mathbb{P}\{A_t^x \neq A_t^{\leftrightarrow}\} \\
&\leq \mathbb{P}\{T > t\} \\
&= 1 - \mathbb{P}\{Z \leq t + x\}\,\mathbb{P}\{\Lambda^{\leftrightarrow} \cap \triangle(-x, t+x) = \emptyset\},
\end{aligned}$$

and part (f) follows.

3. Duality and time-reversal

Suppose in this section that $\int_x^\infty \mu((y, \infty))\,dy < \infty$ for all $x > 0$, so that, by part (e) of Theorem 1.1, the process A has a stationary distribution. Let $(A_t^{\leftrightarrow})_{t\in\mathbb{R}}$ be the stationary Markov process with the transition probabilities of A that was constructed from the Poisson point process $\Lambda^{\leftrightarrow}$ in the proof of part (f) of Theorem 1.1.

Define the *dual process* $(\widehat{A}_t^{\leftrightarrow})_{t\in\mathbb{R}}$ by $\widehat{A}_t^{\leftrightarrow} := \inf\{s > 0 : \triangle(t,s) \cap \Lambda^{\leftrightarrow} = \emptyset\}$. See Figure 3. Thus, $\widehat{A}_t^{\leftrightarrow}$ is the amount of time that must elapse after t until the MRCA lived at a time after t. The càdlàg $\mathbb{R}_+$-valued process $(\widehat{A}_t^{\leftrightarrow})_{t\in\mathbb{R}}$ has saw-tooth sample-paths that drift down with slope -1 between upward jumps.
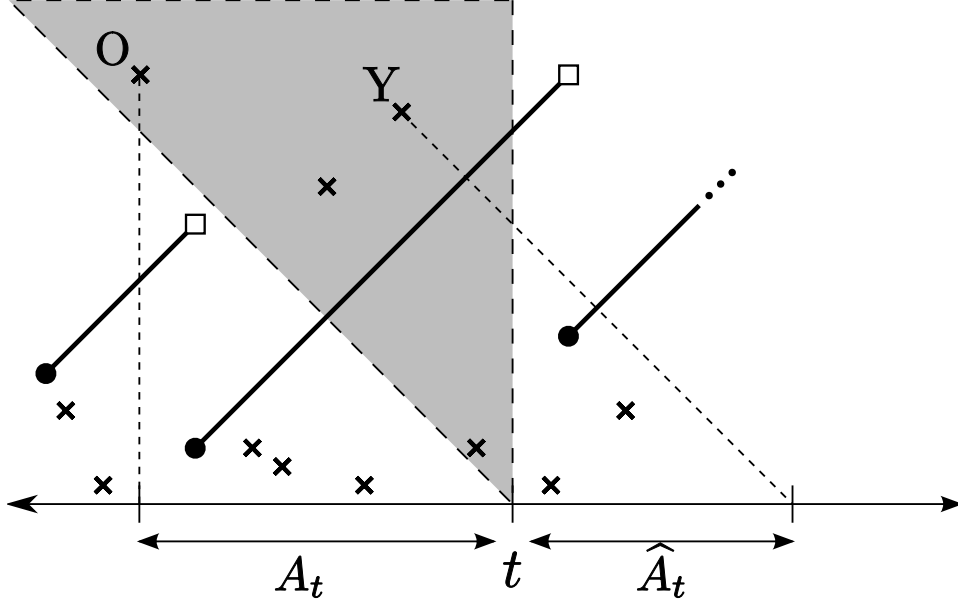

FIGURE 3. The process $A^{\leftrightarrow}$ and the dual process $\widehat{A}^{\leftrightarrow}$. The points in the shaded area represent the families alive at time t, and the solid line is the sample path of $A^{\leftrightarrow}$. The point marked "O" is the oldest living family at time t; the point marked "Y" is the family extant at time t that will live the longest into the future.

Proposition 3.1. *The dual process* $(\widehat{A}_t^{\leftrightarrow})_{t\in\mathbb{R}}$ *has the same distribution as the time-reversed process* $(\bar{A}_t^{\leftrightarrow})_{t\in\mathbb{R}}$, *where* $\bar{A}_t := \lim_{u\downarrow t} A_{-u}^{\leftrightarrow}$.

Proof. Define a bijection $\phi : \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R} \times \mathbb{R}_{++}$ by $\phi(t,x) := (t+x, x)$. Write $\Lambda^{\Leftrightarrow}$ for the image of $\Lambda^{\leftrightarrow}$ under ϕ. The map ϕ preserves the measure $\lambda \otimes \mu$, and hence $\Lambda^{\Leftrightarrow}$ has the same distribution as $\Lambda^{\leftrightarrow}$. Define $(A_t^{\Leftrightarrow})_{t\in\mathbb{R}}$ in terms of $\Lambda^{\Leftrightarrow}$ in the same way that $(A_t^{\leftrightarrow})_{t\in\mathbb{R}}$ was defined in terms of $\Lambda^{\leftrightarrow}$, so that $(A_t^{\Leftrightarrow})_{t\in\mathbb{R}}$ has the same distribution as $(A_t^{\leftrightarrow})_{t\in\mathbb{R}}$.

Let $\Gamma^{\leftrightarrow} := \{(s,t) \in \Lambda^{\leftrightarrow} : \triangle(s,t) \cap \Lambda^{\leftrightarrow} = \emptyset\}$, the set of points of $\Lambda^{\leftrightarrow}$ that, in genealogical terminology, correspond to families that at some time will be the oldest surviving family in the population. The linear segments of the paths of the dual process $(\widehat{A}_t^{\leftrightarrow})_{t\in\mathbb{R}}$ each begin at a point in $\Gamma^{\leftrightarrow}$ and descend with slope -1, whereas the of linear segments of the paths of the MRCA process $(A_t^{\leftrightarrow})_{t\in\mathbb{R}}$ ascend with slope $+1$ to points in $\Gamma^{\Leftrightarrow} := \{\phi(t,x) : (t,x) \in \Gamma^{\leftrightarrow}\}$ — see Figure 4.

Set $\bar{\Lambda}^{\leftrightarrow} := \{(-t,x) : (t,x) \in \Lambda^{\leftrightarrow}\}$, and define $\bar{\Gamma}^{\leftrightarrow}$ in terms of $\bar{\Lambda}^{\leftrightarrow}$ in the same manner that $\Gamma^{\leftrightarrow}$ was defined in terms of $\Lambda^{\leftrightarrow}$. Note that $\bar{\Lambda}^{\leftrightarrow}$ has the same distribution as $\Lambda^{\leftrightarrow}$ (because the map $(t,x) \mapsto (-t,x)$ preserves the measures $\lambda \otimes \mu$), and

hence $\bar{\Gamma}^{\leftrightarrow}$ has the same distribution as $\Gamma^{\leftrightarrow}$. To show that time–reversal exchanges the roles of A and $\widehat{A}$, it will thus suffice to check that $\Gamma^{\Leftrightarrow} = \bar{\Gamma}^{\leftrightarrow}$. However, this follows from the observations that the map ϕ converts left-leaning wedges to right-leaning wedges, so that a point in $\Gamma^{\Leftrightarrow}$ is a point of $\Lambda^{\Leftrightarrow}$ that has no other points of $\Lambda^{\Leftrightarrow}$ in the right-leaning wedge with it as the vertex, and the map $(t, x) \mapsto (-t, x)$ converts right-leaning wedges into left-leaning wedges. $\qquad\square$
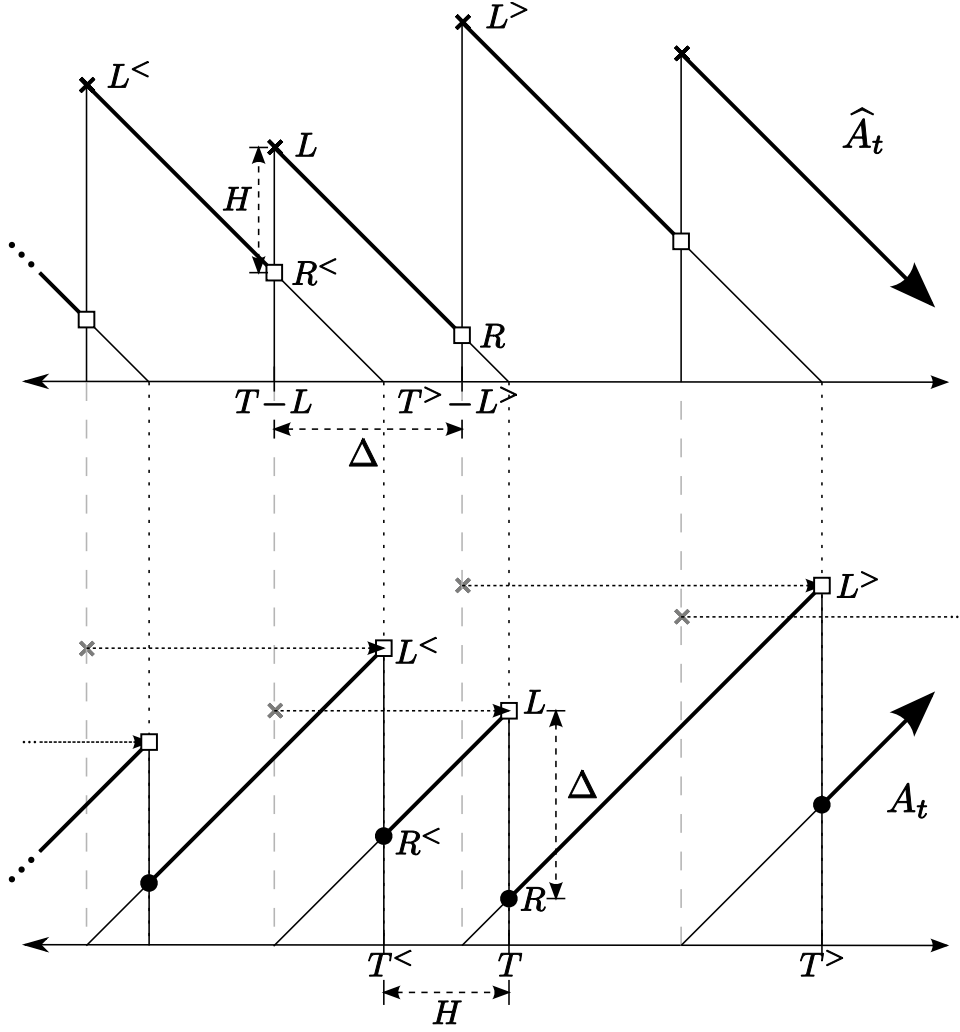

FIGURE 4. The (coupled) processes $A^{\leftrightarrow}$ and $\widehat{A}^{\leftrightarrow}$ for the same set of family lifetimes (the "x"s). The paths of $\widehat{A}^{\leftrightarrow}$ begin at points in $\Gamma^{\leftrightarrow}$ and descend, while the paths of $A^{\leftrightarrow}$ ascend to points in $\Gamma^{\Leftrightarrow}$. The mapping ϕ is shown by the horizontal dotted lines in the lower diagram.

Remark 3.1. There is an interesting connection between the jump sizes and the inter-jump intervals, stemming from the observation that the paths of $A^{\leftrightarrow}$ and $\widehat{A}^{\leftrightarrow}$ have the same sequences of "trough" and "peak" heights, while the roles of the jump sizes and inter-jump intervals for the two are exchanged — see Figure 4. To explain the connection, suppose that $T \in \{t \in \mathbb{R} : A_{t-}^{\leftrightarrow} \neq A_t^{\leftrightarrow}\}$ is a jump time for the process $A^{\leftrightarrow}$. Let $T^< := \sup\{t < T : A_{t-}^{\leftrightarrow} \neq A_t^{\leftrightarrow}\}$ and $T^> := \inf\{t > T : A_{t-}^{\leftrightarrow} \neq A_t^{\leftrightarrow}\}$ be the jump times on either side of T. Put $L := A_{T-}^{\leftrightarrow}$ and $R := A_T^{\leftrightarrow}$, and define $L^<$, $R^<$, $L^<$, and $R^>$ as the analogous values and left limits of $A^{\leftrightarrow}$ at the times $T^<$ and $T^>$. Write $\Delta := L - R$ for the size of the jump at time T and $H := T - T^<$ for the length of the time interval since the previous jump. Observe that $T - L$ is a jump time for the dual process $\widehat{A}^{\leftrightarrow}$, with $\widehat{A}_{(T-L)-}^{\leftrightarrow} = R^<$ and $\widehat{A}_{(T-L)}^{\leftrightarrow} = L$. Moreover,

$$H = L - R^< = \widehat{A}_{(T-L)}^{\leftrightarrow} - \widehat{A}_{(T-L)-}^{\leftrightarrow}$$

and

$$\Delta = (T^> - L^>) - (T - L) = \inf\{t > T - L : \widehat{A}_{t-}^{\leftrightarrow} \neq \widehat{A}_t^{\leftrightarrow}\} - (T - L).$$

Note that the map $T \mapsto T - L$ sets up a monotone bijection between the jump times of the process $A^{\leftrightarrow}$ and those of the process $\widehat{A}^{\leftrightarrow}$. It thus follows from Proposition 3.1 that the point processes

$$\left\{(T, \, T - T^<, \, A_{T-}^{\leftrightarrow} - A_T^{\leftrightarrow}) \, : \, A_{T-}^{\leftrightarrow} \neq A_T^{\leftrightarrow}\right\}$$

and

$$\left\{(T, \, A_{T-}^{\leftrightarrow} - A_T^{\leftrightarrow}, \, T - T^<) \, : \, A_{T-}^{\leftrightarrow} \neq A_T^{\leftrightarrow}\right\}$$

have the same distribution.

4. Jump chains

Suppose again that $\int_x^\infty \mu((y, \infty)) \, dy < \infty$ for all $x > 0$, so that the process A has a stationary distribution. Recall the stationary Markov process $(A_t^{\leftrightarrow})_{t \in \mathbb{R}}$ with the transition probabilities of A that was constructed from the Poisson point process $\Lambda^{\leftrightarrow}$ in the proof of part (f) of Theorem 1.1.

For $t \in \mathbb{R}$, denote by $J_t := \inf\{u > 0 : A_u^{\leftrightarrow} \neq A_{u-}^{\leftrightarrow}\}$ the next jump time of $A^{\leftrightarrow}$ after time t. Define an increasing sequence of random times $0 < T_0 < T_1 < \cdots$ by $T_0 := J_0$ and $T_{n+1} := J_{T_n}$ for $n \geq 0$. Put $L_n := A_{T_n-}^{\leftrightarrow}$ and $R_n := A_{T_n}^{\leftrightarrow}$. Thus, the sequences $(L_n)_{n=0}^\infty$ and $(R_n)_{n=0}^\infty$ record, respectively, the "peaks" and the "troughs" of the path of $A^{\leftrightarrow}$ that occur between the times 0 and $\sup_n T_n$.

The next result can be proved along the same lines as part (a) of Theorem 1.1, and we leave the proof to the reader.

Proposition 4.1. *The sequence* $(L_0, R_0, L_1, R_1, \ldots)$ *is Markovian with the following transition probabilities:*

$$\mathbb{P}\{R_n \in dy \mid L_n = x\} = \mu((y, \infty)) \exp\left(-\int_y^x \mu((u, \infty)) \, du\right) dy, \quad 0 < y \leq x,$$

and

$$\mathbb{P}\{L_{n+1} \in dz \mid R_n = y\} = \frac{m(z)}{\mu((y, \infty))} \, dz, \quad z > y.$$

In particular, the sequence of pairs $((L_n, R_n))_{n=0}^\infty$ *is a time-homogeneous Markov chain.*

Now we may compute the transition probabilities of the peak and trough chains. By Proposition 4.1,

$$
\begin{aligned}
&\mathbb{P}\{L_{n+1} \in dz \mid L_n = x\}/dz \\
&\quad = \int_0^x \exp\left(-\int_y^x \mu((u,\infty))\,du\right) \mu((y,\infty)) \frac{m(z)\mathbf{1}_{y \le z}}{\mu((y,\infty))}\,dy \\
&\quad = m(z) \int_0^{x \wedge z} \exp\left(-\int_y^x \mu((u,\infty))\,du\right)\,dy,
\end{aligned}
$$

(4.1)

and

$$
\begin{aligned}
&\mathbb{P}\{R_{n+1} \in dz \mid R_n = x\}/dz \\
&\quad = \int_x^\infty \frac{m(y)}{\mu((x,\infty))} \mu((z,\infty)) \exp\left(-\int_z^y \mu((u,\infty))\,du\right) \mathbf{1}_{y>z}\,dy \\
&\quad = \mu((z,\infty)) \int_{x \vee z}^\infty \frac{m(y)}{\mu((x,\infty))} \exp\left(-\int_z^y \mu((u,\infty))\,du\right)\,dy.
\end{aligned}
$$

(4.2)

It follows from (4.1) that the peak chain $(L_n)_{n=0}^\infty$ is λ-*irreducible*, where λ is Lebesgue measure on $\mathbb{R}_{++}$. That is, if A is a Borel subset of $\mathbb{R}_{++}$ with $\lambda(A) > 0$, then, for any $x \in \mathbb{R}_{++}$, there is positive probability that the the peak chain begun at x will hit A at some positive time — see Ch. 4 of [MT93] for more about this notion of irreducibility. It follows that the peak chain is either *recurrent*, in the sense that

$$
\sum_{n=0}^\infty \mathbb{P}\{L_n \in A \mid L_0 = x\} = \infty
$$

for all $x \in \mathbb{R}_{++}$ and all Borel subsets of $A \subseteq \mathbb{R}_{++}$ with $\lambda(A) > 0$, or it is *transient*, in the sense that there is a countable collection of Borel sets $(A_j)_{j=1}^\infty$ and finite constants $(M_j)_{j=1}^\infty$ such that $\bigcup_{j=1}^\infty A_j = \mathbb{R}_{++}$ and

$$
\sup_{x \in \mathbb{R}_{++}} \sum_{n=0}^\infty \mathbb{P}\{L_n \in A_j \mid L_0 = x\} \le M_j
$$

— see Theorem 8.0.1 of [MT93].

The peak chain is *strong Feller*; that is, the function

$$
x \mapsto \mathbb{E}[f(L_{n+1}) \mid L_n = x]
$$

is continuous for any bounded Borel function f. Also, because the support of λ is all of $\mathbb{R}_{++}$, if the peak chain is recurrent, then each point x of $\mathbb{R}_{++}$ is *topologically recurrent* in the sense that

$$
\sum_{n=0}^\infty \mathbb{P}\{L_n \in U \mid L_0 = x\} = \infty
$$

for every open neighborhood U of x. Hence, by Theorem 9.3.6 of [MT93], if the peak chain is recurrent, then it is *Harris recurrent*, which means that given any Borel set A with $\lambda(A) > 0$, the chain visits A infinitely often almost surely starting from any x. Moreover, the chain is recurrent (equivalently, Harris recurrent) if and only if it is *non-evanescent*; that is, started from any x there is zero probability that the chain converges to 0 or ∞ — see Theorem 9.2.2 of [MT93].

If the peak chain is recurrent (equivalently, Harris recurrent or non-evanescent), then it has an invariant measure that is unique up to constant multiples — see

Theorem 10.4.4 of [MT93]. If the invariant measure has finite mass, so that it can be normalized to be a probability measure, then the chain is said to be *positive*, otherwise the chain is said to be *null*.

Conversely, if the peak chain has an invariant probability measure, then it is recurrent (equivalently, Harris recurrent or non-evanescent) — see Proposition 10.1.1 of [MT93].

All of the remarks we have just made for the peak chain apply equally to the trough chain $(R_n)_{n=0}^{\infty}$. Recall that we are in the situation when A has a stationary version, so the transience or recurrence of L and R depends on their behavior near zero.

Proposition 4.2. *Consider the two Markov chains $(L_n)_{n=0}^{\infty}$ and $(R_n)_{n=0}^{\infty}$.*

(a) *Both chains are transient if and only if*

$$\int_0^1 \exp\left(\int_x^1 \mu((y,\infty))\,dy\right) dx < \infty.$$

(b) *Both chains are positive recurrent if and only if*

$$\int_0^1 m(x)\exp\left(-\int_x^1 \mu((y,\infty))\,dy\right) dx < \infty.$$

(c) *Both chains are null recurrent if and only if both*

$$\int_0^1 \exp\left(\int_x^1 \mu((y,\infty))\,dy\right) dx = \infty$$

and

$$\int_0^1 m(x)\exp\left(-\int_x^1 \mu((y,\infty))\,dy\right) dx = \infty.$$

Proof. Consider the set $\mathcal{Z} := \{t \in \mathbb{R} : A_t^{\leftrightarrow} = 0\}$. It follows from part (c) of Theorem 1.1 that $\mathbb{P}\{\mathcal{Z} \neq \emptyset\} > 0$ if and only if

(4.3) $$\int_0^1 \exp\left(\int_x^1 \mu(y,\infty)\,dy\right) dx < \infty.$$

By the stationarity of $(A_t^{\leftrightarrow})_{t\in\mathbb{R}}$ and the nature of its sample paths, it is clear that for $x > 0$ the set $\{t \in \mathbb{R} : A_t^{\leftrightarrow} = x\}$ is unbounded above and below almost surely (this also follows from part (d) of Theorem 1.1). It follows from a simple renewal argument that if (4.3) holds, then $\mathcal{Z}$ is unbounded above and below almost surely.

Because the paths of $(A_t^{\leftrightarrow})_{t\in\mathbb{R}}$ increase with slope 1 in the intervals $[T_n, T_{n+1})$, it follows that if (4.3) condition holds, then $\lim_{n\to\infty} T_n = \inf\{t > 0 : A_t^{\leftrightarrow} = 0\} < \infty$ almost surely and $\lim_{n\to\infty} L_n = \lim_{n\to\infty} R_n = 0$ almost surely. In this case, both chains are evanescent, and hence transient.

On the other hand, if (4.3) does not hold, then $\lim_{n\to\infty} T_n = \infty$. Moreover, the set $\{t \in \mathbb{R} : A_t^{\leftrightarrow} = x\}$ is almost surely unbounded above and below for any $x > 0$, as we observed above. If we split the path of $(A_t^{\leftrightarrow})_{t\in\mathbb{R}}$ into excursions away from x, then each excursion interval will contain only finitely many jumps almost surely and, because the excursions are independent and identically distributed, it cannot be the case that L_n or R_n converges to 0 or ∞ with positive probability. Thus, both chains are non-evanescent and hence recurrent.

It is clear from (4.1) that the kernel giving the transition densities of the peak chain $(L_n)_{n=0}^{\infty}$ is self-adjoint with respect to the measure having density

$$p(x) = m(x) \exp\left(-\int_x^{\infty} \mu((u,\infty))\,du\right),$$

with respect to Lebesgue measure, and so this measure is invariant for the peak chain. Clearly, $\int_0^{\infty} p(x)\,dx < \infty$ if and only if the condition in part (b) holds, in which case the peak chain is positive recurrent. Otherwise, the peak chain is either null recurrent or transient, and so part (a) shows that the peak chain is null recurrent if the two conditions in part (c) hold.

Similarly, it is clear from (4.2) that the kernel giving the transition densities of the trough chain $(R_n)_{n=0}^{\infty}$ is self-adjoint with respect to the measure having density

$$q(x) = \mu((x,\infty))^2 \exp\left(-\int_x^{\infty} \mu((u,\infty))\,du\right)$$

with respect to Lebesgue measure, and so this measure is invariant for the trough chain. An integration by parts shows that $\int_0^{\infty} q(x)\,dx < \infty$ if and only if the condition in part (b) holds, and so the trough chain is positive if and only if the peak chain is positive. Alternatively, we can simply observe from Proposition 4.1 that integrating the conditional probability kernel of R_n given L_n against an invariant probability measure for the peak chain gives an invariant measure for the trough chain, and integrating the conditional probability kernel of L_{n+1} given R_n against an invariant probability measure for the trough chain chain gives an invariant measure for the trough chain, so that one chain is positive recurrent if and only if the other is. □

Remark 4.1. If $m(x) = \alpha x^{-2}$ for $x \in (0,1]$, then both the peak and trough chains are

  (1) transient $\Leftrightarrow 0 < \alpha < 1$,
  (2) null recurrent $\Leftrightarrow \alpha = 1$,
  (3) positive recurrent $\Leftrightarrow \alpha > 1$.

Remark 4.2. It follows from parts (b) and (e) of Theorem 1.1 that the stationary point process $\{t \in \mathbb{R} : A_{t-}^{\leftrightarrow} \neq A_t^{\leftrightarrow}\}$ has intensity

$$\rho := \int m(x) \exp\left(-\int_x^{\infty} \mu((u,\infty))\,du\right)\,dx,$$

and so the peak and trough chains are positive recurrent if and only if ρ is finite. Suppose that ρ is finite and consider the point process

$$\Xi := \{(t, A_{t-}^{\leftrightarrow}, A_t^{\leftrightarrow}) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+ : A_{t-}^{\leftrightarrow} \neq A_t^{\leftrightarrow}\}.$$

The companion *Palm* point process Υ has its distribution defined by

$$\mathbb{P}\{\Upsilon \in \cdot\} = \rho^{-1}\mathbb{E}\left[\sum_{\{n\in\mathbb{Z}:0\leq T_n\leq 1\}} \mathbf{1}\{\theta_{T_n}\Xi \in \cdot\}\right],$$

where $\theta_s B = \{(t-s, \ell, r) : (t, \ell, r) \in B\}$ for $B \subset \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$. Enumerate the points of Υ as $((\tilde{T}_n, \tilde{L}_n, \tilde{R}_n))_{n\in\mathbb{Z}}$, where $\ldots < \tilde{T}_{-1} < \tilde{T}_0 = 0 < \tilde{T}_1 < \ldots$. A fundamental result of Palm theory for stationary point processes says that the random sequence $((\tilde{T}_n - \tilde{T}_{n-1}, \tilde{L}_n, \tilde{R}_n))_{n\in\mathbb{Z}}$ is stationary and that the distribution of the point process Ξ may be reconstructed from the distribution of this sequence

– see, for example, Theorem 12.3.II of [DVJ88] or [Kal00]. It is clear that the stationary random sequences $(\tilde{L}_n)_{n=0}^\infty$ and $(\tilde{R}_n)_{n=0}^\infty$ have the same distribution as the peak and trough chains started in their respective stationary distributions.

5. THE $(1+\beta)$-STABLE MRCA PROCESS

In this section we specialize to the motivating example of the MRCA process of a critical $(1+\beta)$-stable continuous-state branching process conditioned to live forever. Recall that the unconditioned continuous state branching process has Laplace transforms (1.1) and the conditioned process has Laplace transforms (1.2). For $\beta = 1$, the unconditioned process has generator $x\frac{\partial^2}{\partial x^2}$ with this choice of time scale.

As mentioned in the introduction, the set of points $(t, x) \in \mathbb{R}_+ \times \mathbb{R}_{++}$, where t is the time that a family splits from the immortal lineage and x is its total lifetime, is a Poisson point process with intensity measure $\lambda \otimes \mu$ for some σ-finite measure μ.

Lemma 5.1. *The lifetime measure μ associated with the critical $(1+\beta)$-branching process conditioned on non-extinction is given by*

$$\mu((x, \infty)) = \frac{1+\beta}{\beta x}, \quad x > 0.$$

Proof. As we remarked in the introduction

$$(5.1) \qquad \mathbb{P}\{X_t > 0 \mid X_0 = x\} = 1 - \exp\left(-\frac{x}{t^{1/\beta}}\right).$$

First consider the case of $\beta = 1$. Recall from the introduction that if Π is a Poisson point process on $\mathbb{R}_+ \times \mathcal{E}^0$ with intensity $\lambda \otimes \nu$, then

$$(5.2) \qquad \left(\sum_{(s,u)\in\Pi} u_{(t-s)\vee 0}\right)_{t \geq 0}$$

has the same distribution as the conditioned process $(Y_t)_{t\geq 0}$ with $Y_0 = 0$, and recall that μ is the push-forward of ν by the total lifetime function γ. Also,

$$(5.3) \qquad \left(\sum_{(s,u)\in\Pi:s\leq y/2} u_t\right)_{t \geq 0}$$

has the same distribution as the unconditioned process $(X_t)_{t\geq 0}$ with $X_0 = y$ — see [Eva93]. The factor of 2 differs from [Eva93] and arises from our choice of time scale. Therefore,

$$\begin{aligned} \mathbb{P}\{X_t > 0 \mid X_0 = y\} &= \mathbb{P}\{\exists (s, u) \in \Pi \;:\; s \leq y/2 \text{ and } \gamma(u) > t\} \\ &= 1 - \exp\left(-y\,\mu((t, \infty))/2\right), \end{aligned}$$

and comparing with (5.1) gives $\mu((t, \infty)) = 2/t$.

Now take $\beta \in (0, 1)$. It is shown in [EW03] that the mass thrown off the immortal lineage is determined by the jumps of a stable subordinator: if M_s is the amount of mass thrown off during the time interval $[0, s]$, then

$$\begin{aligned} \mathbb{E}\left[e^{-\theta M_s}\right] &= \exp\left(-s\frac{1+\beta}{\beta}\theta^\beta\right) \\ &= \exp\left(-s\int_0^\infty \left(1 - e^{-\theta x}\right)\nu(dx)\right), \end{aligned}$$

where $\nu(dx) = \frac{1+\beta}{\Gamma(1-\beta)} x^{-(1+\beta)} dx$ is the Lévy measure of the subordinator.

Since the jump size of the subordinator corresponds to the initial size of the new family, the lifetime measure μ is given by

$$\mu((t,\infty)) = \int_0^\infty \mathbb{P}\{X_t > 0 \mid X_0 = x\}\,\nu(dx).$$

and so, from the above and an integration by parts,

$$\begin{aligned}\mu((t,\infty)) &= \int_0^\infty \left(1 - e^{-x/t^{1/\beta}}\right)\nu(dx) \\ &= \frac{1+\beta}{\beta t}.\end{aligned}$$

$\square$

Proof of Corollary 1.1

Parts (a) to (d) follow immediately from Theorem 1.1. Part (e) will also follow from parts (e) and (f) of Theorem 1.1 after the following time and space change.

Define a new time parameter u by $t = e^u$. If the MRCA at time t lived at time $t - x$ on the original scale, then on the new time scale she lived at time $u - y$, where $t - x = e^{u-y}$. Solving for y, the MRCA age process in the new time scale is the process $(B_t)_{t\geq 0}$ given by $B_u = -\log(1 - e^{-u}A_{e^u})$. The process $(B_t)_{t\geq 0}$ is obtained by applying the construction (1.1) to the point process given by

$$\left\{\left(\log s, \log\left(1 + \frac{x}{s}\right)\right), (s,x) \in \Lambda\right\},$$

which is a Poisson point process on $\mathbb{R} \times \mathbb{R}_+$ with intensity measure $\lambda \otimes \rho$, where

$$\rho((y,\infty)) = \frac{1+\beta}{\beta(e^y - 1)}, \quad y \in \mathbb{R}_{++}.$$

Note that, in general, a time and space change of the Poisson process Λ gives a new Poisson point process, but the resulting intensity measure will not typically be of the form $\lambda \otimes \kappa$ for some measure κ: it is a special feature of μ and the transformation that the product measure structure is maintained in this case.

It is straightforward to check parts (e) and (f) of Theorem 1.1 that $(B_t)_{t\geq 0}$ has the stationary distribution

$$\pi(dx) = \frac{1+\beta}{\beta} e^{-x}(1 - e^{-x})^{\frac{1}{\beta}}\,dx,$$

and that the distribution of B_t converges to π in total variation as $t \to \infty$. Part (e) of the corollary then follows from the observation that $\frac{A_t}{t} = 1 - e^{-B_{\log(t)}}$ and an elementary change of variables. $\square$

6. An analogue of the Bessel-squared family

Recall that the transition probabilities of the unconditioned $(1+\beta)$-stable continuous state branching process $(X_t)_{t\geq 0}$ are characterized by the Laplace transforms

$$\mathbb{E}^x[\exp(-\theta X_t)] = \exp\left(-x\theta\left(t\theta^\beta + 1\right)^{-1/\beta}\right).$$

Hence, the transition probabilities of the conditioned process Y are characterized by the Laplace transforms

$$\mathbb{E}^y[\exp(-\theta Y_t)] = \frac{1}{y}\mathbb{E}^y[\exp(-\theta X_t)X_t]$$

$$= \frac{1}{y}\frac{\partial}{\partial\theta}\mathbb{E}^y[\exp(-\theta X_t)]$$

$$= \exp\left(-y\theta\left(t\theta^\beta + 1\right)^{-1/\beta}\right)\left(t\theta^\beta + 1\right)^{-\frac{\beta+1}{\beta}},$$

thus establishing (1.2).

Recall also that if $\beta = 1$, then $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are, up to a constant multiple, the Bessel-squared processes with dimensions 0 and 4 respectively. The Bessel-squared process $(Z_t)_{t\geq 0}$ with dimension d (not necessarily integral) is (up to constants) the Markov process characterized by the Laplace transforms

$$\mathbb{E}^z[\exp(-\theta Z_t)] = \exp\left(-z\theta\left(t\theta + 1\right)^{-1}\right)\left(t\theta + 1\right)^{-\frac{d}{2}}.$$

This suggests that for $0 < \beta < 1$ and $\delta \geq 0$ there might be a semigroup $(P_t)_{t\geq 0}$ such that

(6.1) $$P_t\exp(-\theta\cdot)(x) = \exp\left(-x\theta\left(t\theta^\beta + 1\right)^{-1/\beta}\right)\left(t\theta^\beta + 1\right)^{-\delta}.$$

We first verify that, for a fixed value of x, the right-hand side of (6.1) is the Laplace transform of a probability distribution (as a function of θ). We already know that

$$\exp\left(-x\theta\left(t\theta^\beta + 1\right)^{-1/\beta}\right)$$

is the Laplace transform of a probability measure, so it suffices to show that

$$\left(t\theta^\beta + 1\right)^{-\delta}$$

is also a Laplace transform of a probability distribution. Let $(S_t)_{t\geq 0}$ be the β-stable subordinator starting from $S_0 = 0$ normalized so that

$$\mathbb{E}[\exp(-\theta S_t)] = \exp(-\theta^\beta t)$$

and let $(T_t)_{t\geq 0}$ be the gamma subordinator starting from $T_0 = 0$ normalized so that for $t > 0$

$$\mathbb{P}\{T_t \in dy\} = \frac{y^{t-1}}{\Gamma(t)}\exp(-y)\,dy$$

and hence

$$\mathbb{E}[\exp(-\theta T_t)] = (\theta + 1)^{-t}.$$

Then if S and T are independent,

$$\mathbb{E}\left[\exp\left(-\theta S_{tT_\delta}\right)\right] = \mathbb{E}\left[\exp(-\theta^\beta t T_\delta)\right] = \left(t\theta^\beta + 1\right)^{-\delta}.$$

We next verify that $(P_t)_{t \geq 0}$ is a semigroup. Observe that

$$P_s P_t \exp(-\theta \cdot)(x)$$

$$= \left(t\theta^\beta + 1\right)^{-\delta} P_s \exp\left(-\theta \left(t\theta^\beta + 1\right)^{-1/\beta} \cdot\right)(x)$$

$$= \left(t\theta^\beta + 1\right)^{-\delta} \exp\left(-x\theta \left(t\theta^\beta + 1\right)^{-1/\beta} \left(s\theta^\beta \left(t\theta^\beta + 1\right)^{-1} + 1\right)^{-1/\beta}\right)$$

$$\times \left(s\theta^\beta \left(t\theta^\beta + 1\right)^{-1} + 1\right)^{-\delta}$$

$$= \exp\left(-x\theta \left((s+t)\theta^\beta + 1\right)^{-1/\beta}\right) \left((s+t)\theta^\beta + 1\right)^{-\delta}$$

$$= P_{s+t} \exp(-\theta \cdot)(x)$$

It is clear that $\lim_{t \downarrow 0} P_t \exp(-\theta \cdot)(x) = \exp(-\theta x)$ and so $\lim_{t \downarrow 0} P_t f(x) = f(x)$ for $f \in C_0(\mathbb{R}_+)$. Also, $\lim_{y \to x} P_t \exp(-\theta \cdot)(y) = P_t \exp(-\theta \cdot)(x)$, and so $\lim_{y \to x} P_t f(y) = f(x)$ for $f \in C_0(\mathbb{R}_+)$. The standard Feller construction gives that there is a strong Markov process $(Z_t)_{t \geq 0}$ with semigroup $(P_t)_{t \geq 0}$.

This family of Markov processes shares many features of the Bessel-squared family. For example, it follows for $a, b > 0$ that

$$\mathbb{E}^{az}[\exp(-\theta a^{-1} Z_{bt})] = \exp\left(-y\theta \left(bt \left(\frac{\theta}{a}\right)^\beta + 1\right)^{-1/\beta}\right) \left(bt \left(\frac{\theta}{a}\right)^\beta + 1\right)^{-\delta}.$$

Thus, the process $(b^{-1/\beta} Z_{bt})_{t \geq 0}$ is Markovian with the same transition probabilities as Z. Similarly, if $Z_0 = 0$, then the process $(e^{-t/\beta} Z_{e^t})_{t \in \mathbb{R}}$ is Markovian and stationary.

Furthermore, if $(Z'_t)_{t \geq 0}$ and $(Z'')_{t \geq 0}$ are two independent such processes with parameters δ' and δ'', then the process $(Z'_t + Z''_t)_{t \geq 0}$ also belongs to the family and has parameter $\delta' + \delta''$.

In a forthcoming paper, we will present a more thorough study of this family along the lines of [PY82, PY81].

REFERENCES

[BPSZ06] Onno Boxma, David Perry, Wolfgang Stadje, and Shelemyahu Zacks, *A Markovian growth-collapse model*, Adv. in Appl. Probab. **38** (2006), no. 1, 221–243. MR **MR2213972 (2007d:**60062)

[BVJ00]   K. Borovkov and D. Vere-Jones, *Explicit formulae for stationary distributions of stress release processes*, J. Appl. Probab. **37** (2000), no. 2, 315–321. MR **MR1780992 (2001e:**60146)

[CD08]   O. L. V. Costa and F. Dufour, *Stability and ergodicity of piecewise deterministic Markov processes*, SIAM J. Control Optim. **47** (2008), no. 2, 1053–1077. MR MR2385873

[CDP01]   G. Colombo and P. Dai Pra, *A class of piecewise deterministic Markov processes*, Markov Process. Related Fields **7** (2001), no. 2, 251–287. MR **MR1856497 (2002h:**60179)

[Dav84]   M. H. A. Davis, *Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models*, J. Roy. Statist. Soc. Ser. B **46** (1984), no. 3, 353–388, With discussion. MR **MR790622 (87g:**60062)

[Dav93]   ———, *Markov models and optimization*, Monographs on Statistics and Applied Probability, vol. 49, Chapman & Hall, London, 1993. MR **MR1283589 (96b:**90002)

[DC99]   François Dufour and Oswaldo L. V. Costa, *Stability of piecewise-deterministic Markov processes*, SIAM J. Control Optim. **37** (1999), no. 5, 1483–1502 (electronic). MR **MR1710229 (2000g:**60125)

[DGR02]   Vincent Dumas, Fabrice Guillemin, and Philippe Robert, *A Markovian analysis of additive-increase multiplicative-decrease algorithms*, Adv. in Appl. Probab. **34** (2002), no. 1, 85–111. MR **MR1895332 (2003f:**60168)

[DVJ88]   D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes*, Springer Series in Statistics, Springer-Verlag, New York, 1988. MR **MR950166 (90e:**60060)

[Eli07]   Iddo Eliazar, *The $M/G/\infty$ system revisited: finiteness, summability, long range dependence, and reverse engineering*, Queueing Syst. **55** (2007), no. 1, 71–82. MR **MR2293566 (2008b:**60199)

[EP90]   Steven N. Evans and Edwin Perkins, *Measure-valued Markov branching processes conditioned on nonextinction*, Israel J. Math. **71** (1990), no. 3, 329–337. MR **MR1088825 (92d:**60089)

[EP98]   Steven N. Evans and Jim Pitman, *Stationary Markov processes related to stable Ornstein-Uhlenbeck processes and the additive coalescent*, Stochastic Process. Appl. **77** (1998), no. 2, 175–185. MR **MR1649003 (99j:**60109)

[EPW06]   Steven N. Evans, Jim Pitman, and Anita Winter, *Rayleigh processes, real trees, and root growth with re-grafting*, Probab. Theory Related Fields **134** (2006), no. 1, 81–126. MR **MR2221786 (2007d:**60003)

[Eva92]   Steven N. Evans, *The entrance space of a measure-valued Markov branching process conditioned on nonextinction*, Canad. Math. Bull. **35** (1992), no. 1, 70–74.

[Eva93]   ———, *Two representations of a conditioned superprocess*, Proc. Roy. Soc. Edinburgh Sect. A **123** (1993), no. 5, 959–971. MR **MR1249698 (95e:**60082)

[EW03]   A. M. Etheridge and D. R. E. Williams, *A decomposition of the $(1 + \beta)$-superprocess conditioned on survival*, Proc. Roy. Soc. Edinburgh Sect. A **133** (2003), no. 4, 829–847. MR **MR2006204 (2004i:**60121)

[Fel51]   William Feller, *Diffusion processes in genetics*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950 (Berkeley and Los Angeles), University of California Press, 1951, pp. 227–246. MR MR0046022 (13,671c)

[FFS85]   P. J. Fitzsimmons, Bert Fristedt, and L. A. Shepp, *The set of real numbers left uncovered by random covering intervals*, Z. Wahrsch. Verw. Gebiete **70** (1985), no. 2, 175–189. MR **MR799145 (86k:**60017)

[IKPG00]   Max Ingman, Henrik Kaessmann, Svante Paabo, and Ulf Gyllensten, *Mitochondrial genome variation and the origin of modern humans*, Nature **408** (2000), no. 6813, 708–713.

[JS96]   J. Jacod and A. V. Skorokhod, *Jumping Markov processes*, Ann. Inst. H. Poincaré Probab. Statist. **32** (1996), no. 1, 11–67. MR **MR1373726 (97c:**60210)

[Kal00]   Olav Kallenberg, *An extension of the basic Palm measure correspondence*, Probab. Theory Related Fields **117** (2000), no. 1, 113–131. MR **MR1759510 (2001g:**60120)

[Kni81]   Frank B. Knight, *Essentials of Brownian motion and diffusion*, Mathematical Surveys, vol. 18, American Mathematical Society, Providence, R.I., 1981. MR **MR613983 (82m:**60098)

[Man72]   Benoit B. Mandelbrot, *Renewal sets and random cutouts*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **22** (1972), 145–157. MR MR0309162 (46 #8272)

[MT93]   S. P. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*, Communications and Control Engineering Series, Springer-Verlag London Ltd., London, 1993. MR **MR1287609 (95j:**60103)

[PW06]   P. Pfaffelhuber and A. Wakolbinger, *The process of most recent common ancestors in an evolving coalescent*, Stochastic Process. Appl. **116** (2006), no. 12, 1836–1859. MR MR2307061

[PY81]   Jim Pitman and Marc Yor, *Bessel processes and infinitely divisible laws*, Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980), Lecture Notes in Math., vol. 851, Springer, Berlin, 1981, pp. 285–370. MR **MR620995 (82j:**60149)

[PY82]   Jim Pitman and Marc Yor, *A decomposition of Bessel bridges*, Z. Wahrsch. Verw. Gebiete **59** (1982), no. 4, 425–457. MR **MR656509 (84a:**60091)

[SD06]   Damien Simon and Bernard Derrida, *Evolution of the most recent common ancestor of a population with no selection*, J. Stat. Mech. (2006), P05002 (electronic).

[She72]   L. A. Shepp, *Covering the line with random intervals*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **23** (1972), 163–170. MR MR0322923 (48 #1284)

[Tak62]   Lajos Takács, *Introduction to the theory of queues*, University Texts in the Mathematical Sciences, Oxford University Press, New York, 1962. MR MR0133880 (24 #A3704)

[TCG⁺04]   Simon A. A. Travers, Jonathan P. Clewley, Judith R. Glynn, Paul E. M. Fine, Amelia C. Crampin, Felix Sibande, Dominic Mulawa, James O. McInerney, and Grace P. McCormack, *Timing and reconstruction of the most recent common ancestor of the subtype C clade of Human Immunodeficiency Virus type 1*, J. Virol. **78** (2004), no. 19, 10501–10506.

[TPS⁺00]   Russell Thomson, Jonathan K. Pritchard, Peidong Shen, Peter J. Oefner, and Marcus W. Feldman, *Recent common ancestry of human Y chromosomes: Evidence from DNA sequence data*, Proceedings of the National Academy of Science **97** (2000), no. 13, 7360–7365.

[VBL⁺01]   Sarah K. Volkman, Alyssa E. Barry, Emily J. Lyons, Kaare M. Nielsen, Susan M. Thomas, Mehee Choi, Seema S. Thakore, Karen P. Day, Dyann F. Wirth, and Daniel L. Hartl, *Recent origin of Plasmodium falciparum from a single progenitor*, Science **293** (2001), no. 5529, 482 – 484.

STEVEN N. EVANS, DEPARTMENT OF STATISTICS #3860, UNIVERSITY OF CALIFORNIA AT BERKELEY, 367 EVANS HALL, BERKELEY, CA 94720-3860, U.S.A.
   E-mail address: `evans@stat.Berkeley.EDU`
   URL: `http://www.stat.berkeley.edu/users/evans/`

PETER L. RALPH, DEPARTMENT OF STATISTICS #3860, UNIVERSITY OF CALIFORNIA AT BERKELEY, 367 EVANS HALL, BERKELEY, CA 94720-3860, U.S.A.
   E-mail address: `plr@stat.Berkeley.EDU`
   URL: `http://www.stat.berkeley.edu/users/plr/`