

# GENEALOGICAL CONSTRUCTIONS AND ASYMPTOTICS FOR CONTINUOUS-TIME MARKOV AND CONTINUOUS-STATE BRANCHING PROCESSES

BY THOMAS G. KURTZ

---

## Abstract

Genealogical constructions of population processes provide models which simultaneously record the forward-in-time evolution of the population size (and distribution of locations and types for models that include them) and the backward-in-time genealogies of the individuals in the population at each time  $t$ . A genealogical construction for continuous-time Markov branching processes from Kurtz and Rodrigues (2011) is described and exploited to give the normalized limit in the supercritical case. A Seneta–Heyde norming is identified as a solution of an ordinary differential equation. The analogous results are given for continuous-state branching processes, including proofs of the normalized limits of Grey (1974) in both the supercritical and critical/subcritical cases.

**Keywords:** Genealogical construction; Markov branching process; supercritical limit; Seneta–Heyde norming; continuous-state branching process

2010 Mathematics Subject Classification: Primary 60J80

Secondary 60F15; 60J25

---

## 1. A genealogical construction of branching processes

To characterize a Markov process as a solution of a martingale problem, we must specify a generator  $A$  by identifying the functions in the domain  $\mathcal{D}(A)$  and defining  $Af$  for each  $f \in \mathcal{D}(A)$ . A stochastic process  $X$  is then a solution of the corresponding martingale problem if there is a filtration  $\{\mathcal{F}_t\}$  such that  $X$  is  $\{\mathcal{F}_t\}$ -adapted and

$$M_f(t) = f(X(t)) - f(X(0)) - \int_0^t Af(X(s)) ds$$

is an  $\{\mathcal{F}_t\}$ -martingale for each  $f \in \mathcal{D}(A)$ .

We first consider a generator for a process whose state space is  $E = \bigcup_{n=0}^{\infty} [0, r]^n$  for some  $r > 0$ , that is,  $n$ -tuples of points which we call *particles* in  $[0, r]$ , with  $[0, r]^0$  denoting the state with no particles. The value of  $u \in [0, r]$  corresponding to a particle is called its *level*.

The domain of our generator is

$$\mathcal{D}(A_r) = \left\{ f(u, n) = \prod_{i=1}^n g(u_i) : 0 \leq g \leq 1, g \in C^1[0, r], g(r) = 1, g'(r) = 0 \right\},$$

where, for  $u \in [0, r]^n$ , we write  $f(u, n)$  rather than  $f(u)$  to emphasize the number of particles in the state. The boundary conditions  $g(r) = 1$  and  $g'(r) = 0$  essentially correspond to the assumption below that particles die when they hit the boundary.

For parameters  $a_k \geq 0$  and  $b \in \mathbb{R}$  satisfying

$$0 \leq r \sum_{k=1}^{\infty} ka_k - b < \infty, \tag{1.1}$$

define

$$\begin{aligned} A_r f(u, n) = & f(u, n) \sum_{i=1}^n \sum_{k=1}^{\infty} \frac{(k+1)a_k}{r^{k-1}} \int_{[u_i, r]^k} \left[ \left( \prod_{l=1}^k g(v_l) \right) - 1 \right] dv_1 \cdots dv_k \\ & + f(u, n) \sum_{i=1}^n \left( \sum_{k=1}^{\infty} r^2 a_k \left[ \left( 1 - \frac{u_i}{r} \right)^{k+1} - 1 + (k+1) \frac{u_i}{r} \right] - bu_i \right) \frac{g'(u_i)}{g(u_i)}. \end{aligned}$$

This generator looks more complicated than it is. The first term says that, for each  $k$ , the particle with level  $u_i$  gives ‘birth’ to  $k$  new particles at rate  $(k+1)a_k(r-u_i)^k/r^{k-1}$ , and the levels of the new particles are independently and uniformly distributed over the interval  $[u_i, r]$ . Note that the levels of the ‘offspring’ always lie above the level of the ‘parent’.

The second term says that the level  $U_i$  of each particle alive evolves as the solution of the differential equation

$$\dot{U}_i(t) = \sum_{k=1}^{\infty} r^2 a_k \left[ \left( 1 - \frac{U_i(t)}{r} \right)^{k+1} - 1 + (k+1) \frac{U_i(t)}{r} \right] - bU_i(t).$$

A particle ‘dies’ when its level hits  $r$ . Note that (1.1) ensures that particles cannot come into the interval  $[0, r)$  from above, and if equality holds in (1.1), that no particle can hit  $r$ , that is, no particle dies.

While the interpretation of the behavior of the process corresponding to  $A_r$  is straightforward, the point of defining the generator is not so clear: here is the reason. Let  $\alpha_r(n, du)$  be the joint distribution of  $n$  independent uniform  $[0, r]$  random variables, and, for  $f(u, n) = \prod_{i=1}^n g(u_i) \in \mathcal{D}(A)$ , define  $\lambda_g$  so that  $(1/r) \int_0^r g(v) dv = e^{-\lambda_g}$  and set

$$\hat{f}(n) = \int_{[0, r]^n} f(u, n) \alpha_r(n, du) = e^{-\lambda_g n}.$$

Note that  $\lambda_g \geq 0$ . A nice calculus exercise (see [14, Section 4.4]) gives

$$\begin{aligned} & \int_{[0, r]^n} A_r f(u, n) \alpha_r(n, du) \\ &= \sum_{k=1}^{\infty} n r a_k (e^{-\lambda_g(n+k)} - e^{-\lambda_g n}) + n \left( \sum_{k=1}^{\infty} r k a_k - b \right) (e^{-\lambda_g(n-1)} - e^{-\lambda_g n}) \\ &= \sum_{k=1}^{\infty} n r a_k (\hat{f}(n+k) - \hat{f}(n)) + n \left( \sum_{k=1}^{\infty} r k a_k - b \right) (\hat{f}(n-1) - \hat{f}(n)) \\ &\equiv C_r \hat{f}(n), \end{aligned} \tag{1.2}$$

where  $C_r$  is the generator for a continuous-time Markov branching process. The Markov mapping theorem of [12] (see [14, Section A.5]) implies that every solution  $X$  of the martingale

problem for  $C_r$ , that is, every continuous-time branching process, can be obtained from a solution  $U$  of the martingale problem for  $A_r$ . If the initial distribution of the branching process is  $\mu_0(dn)$  then the initial distribution of  $U$  is  $\int \alpha_r(n, du)\mu_0(dn)$ . In fact, if the distribution of  $X(t)$  is  $\mu_t(dn)$  then the distribution of  $U(t)$  is  $\int \alpha_r(n, du)\mu_t(dn)$ . More precisely, the conditional distribution of  $U(t)$  given  $\mathcal{F}_t^X$  is  $\alpha_r(X(t), du)$ .

Writing (1.2) as  $\alpha_r A_r f(n) = C_r \alpha f(n)$ , this identity can be interpreted as a generator version of the semigroup identity given in [17, Theorem 2]. That observation, however, probably does not provide any additional intuition regarding the meaning of these calculations. More insight may be provided by seeing the Markov mapping theorem as a corollary of a more general filtering result, as proved in [12].

Suppose that  $U$  with values in a state space  $E$  is a solution of the martingale problem for an operator  $A$  with corresponding filtration  $\{\mathcal{F}_t\}$ , and let  $\{\mathcal{G}_t\}$  be another filtration with  $\mathcal{G}_t \subset \mathcal{F}_t, t \geq 0$ . Let  $\pi_t$  be the conditional distribution of  $U(t)$  given  $\mathcal{G}_t$ . Then, for  $f \in \mathcal{D}(A)$ ,

$$\mathbb{E}[f(U(t)) \mid \mathcal{G}_t] - \mathbb{E}[f(U(0)) \mid \mathcal{G}_0] - \int_0^t \mathbb{E}[Af(U(s)) \mid \mathcal{G}_s] ds = \pi_t f - \pi_0 f - \int_0^t \pi_s Af ds$$

is a  $\{\mathcal{G}_t\}$ -martingale. Under some technical conditions on  $A$ , most importantly, that  $\mathcal{D}(A)$  is closed under multiplication, there is a converse to this observation. Specifically, if  $\{\tilde{\pi}_t, t \geq 0\}$  is a  $\mathcal{P}(E)$ -valued process adapted to a filtration  $\{\tilde{\mathcal{G}}_t\}$  and, for each  $f \in \mathcal{D}(A)$ ,

$$\tilde{\pi}_t f - \tilde{\pi}_0 f - \int_0^t \tilde{\pi}_s Af ds$$

is a  $\{\tilde{\mathcal{G}}_t\}$ -martingale, then there exists a solution  $U$  of the martingale problem for  $A$  and a filtration  $\{\mathcal{G}_t\}$  such that the  $\mathcal{P}(E)$ -valued process given by the conditional distributions  $\pi_t$  of  $U(t)$  given  $\mathcal{G}_t, t \geq 0$ , has the same finite-dimensional distributions as  $\{\tilde{\pi}_t, t \geq 0\}$  (see [13, Corollary 3.2]).

With reference to (1.2), if  $X$  is a solution of the martingale problem for  $C_r$  (that is,  $X$  is a corresponding branching process) then setting  $\tilde{\pi}_t(du) = \alpha_r(X(t), du)$ , for  $f \in \mathcal{D}(A_r)$ ,

$$\alpha_r f(X(t)) - \alpha_r f(X(0)) - \int_0^t C_r \alpha_r f(X(s)) ds = \tilde{\pi}_t f - \tilde{\pi}_0 f - \int_0^t \tilde{\pi}_s A_r f ds$$

is a martingale with respect to  $\{\tilde{\mathcal{G}}_t\} = \{\mathcal{F}_t^X\}$ , the filtration generated by  $X$ , and the general filtering result ensures the existence of  $U$ .

One should also note that (1.2) is not magic. Starting with  $A_r$  of the form

$$\begin{aligned} A_r f(u, n) = & f(u, n) \sum_{i=1}^n \sum_{k=1}^{\infty} \frac{(k+1)a_k}{r^{k-1}} \int_{[u_i, r]^k} \left[ \left( \prod_{l=1}^k g(v_l) \right) - 1 \right] dv_1 \cdots dv_k \\ & + f(u, n) \sum_{i=1}^n F(u_i) \frac{g'(u_i)}{g(u_i)}, \end{aligned}$$

to ensure that (1.2) holds, we compute

$$F(u_i) = \sum_{k=1}^{\infty} r^2 a_k \left[ \left( 1 - \frac{u_i}{r} \right)^{k+1} - 1 + (k+1) \frac{u_i}{r} \right] - bu_i.$$

The motivation for the constructions in [14] actually comes from branching Markov processes where the parameters  $a_k$  and  $b$  depend on the type or location of the particle (which evolves in time). These constructions are now known as *lookdown constructions* because each offspring has the same type as its parent, that is, each offspring, which always has its level above that of its parent, ‘looks down’ to the parent to determine its type.

Note that the path of a *level* does not correspond to a single particle, but rather a line of descent. The levels of the particles contain information about the future of the process. Roughly, the order of the levels is determined by the lengths of the lines of descent. In particular, if the line of descent is finite, the time at which the line terminates is a deterministic function of the level.

The process of levels records the genealogy of the branching process, at least if one marks a level each time the particle at that level gives birth. Now the levels at time  $t$  conditioned on  $\mathcal{F}_t^X$  are independent and identically distributed (i.i.d.) uniform, so, for  $k < X(t)$ , the particles with the  $k$  lowest levels give a random sample of size  $k$  from the population. Tracing the levels backward in time, we can identify the levels of the ancestors of these  $k$  particles at each time  $s < t$ , that is, we can identify the genealogical tree corresponding to the  $k$  ‘randomly’ selected particles. Note, however, that conditioned on  $\mathcal{F}_T^X$  for some  $T > t$ , the families of particles alive at time  $T$  that descend from each of the  $k$  particles with lowest levels at time  $t$  are not a random sample from the families at time  $T$  that have a common ancestor at time  $t$ . In particular, the particle at the lowest level at time  $t$  will have the stochastically largest family of descendants at time  $T$ . Consequently, it is more precise to say that the particles with the  $k$  lowest levels at a fixed time  $t$  have the statistical properties of a uniform random sample from the population at time  $t$  as far as the information in  $\mathcal{F}_t^X$  is concerned. See [5] for similar constructions for a large variety of stochastic population models.

Our goal in the next section is to show how to obtain the asymptotic behavior of a supercritical branching process from the asymptotic behavior of the particle levels. Kurtz and Rodrigues [14] treated the case  $a_k = 0$  for  $k > 1$  and promised to treat the general case ‘elsewhere’, which in fact is this paper. In particular, we treat the general case with  $\sum_k ka_k < \infty$ .

There is a long history of theorems of this type. For Galton–Watson processes (i.e. discrete-time Markov branching processes), convergence of the branching process scaled by its mean goes back to Harris [8]. For the discrete-time analogue of Theorem 3.1 considered in Section 3, the sufficiency of the  $X \log X$  condition was proved in [16] and the necessity in [11]. The continuous-time case discussed here is given in [1, Theorem III.7.2] and is a special case of results in [4]. The existence of a scaling giving a nontrivial limit for Galton–Watson processes with finite mean but for which the  $X \log X$  condition fails was shown in [19] (convergence in distribution) and [9] (almost-sure convergence). The existence of a corresponding scaling in the continuous-time setting follows from results in [3] on more general Bellman–Harris processes. Here, the scaling is obtained as a solution of an ordinary differential equation.

In Section 4, we let  $r \rightarrow \infty$  to obtain a genealogical construction for continuous-state branching processes. These processes were introduced in [10]. Results analogous to the Galton–Watson results described above were given by Grey [7]. Grey also proved scaling limits in the critical and subcritical cases under assumptions that the unscaled process converges to 0 but does not reach 0 in finite time. Here we rederive Grey’s results by analyzing the behavior of the levels in the genealogical construction.

To simplify notation, in the next section we assume that  $r = 1$ . The significance of  $r$  will be clear in the last section, but note that, under additional assumptions on the  $a_k$ ,  $C_r$  is scaled in such a way that, as  $r \rightarrow \infty$ , the corresponding branching process, normalized by  $r$ , should converge to a Feller diffusion process or a more general continuous-state branching process.

### 2. Asymptotic behavior of supercritical branching processes

The branching process is supercritical if

$$0 < b \leq \sum_{k=1}^{\infty} ka_k.$$

Since particles are continuously being born and dying, it is not convenient to index them by integers; however, with probability 1, no two particles have the same initial level, so we index the particles by their initial level  $v$ :  $\beta_v$  denotes the birth time of particle  $v$  ( $\beta_v = 0$  for particles in the initial population) and  $\tau_v$  is the time that  $U_v$  hits 1, with  $\tau_v = \infty$  if the particle never hits 1.

Each level evolves according to

$$\begin{aligned} \dot{U}_v(t) &= \sum_{k=1}^{\infty} a_k [(1 - U_v(t))^{k+1} - 1 + (k + 1)U_v(t)] - bU_v(t) \\ &\equiv F(U_v(t)) \\ &= \int_0^{U_v(t)} \Lambda(v) \, dv - bU_v(t) \end{aligned}$$

for  $\beta_v \leq t < \tau_v$ , and  $U_v(\beta_v) = v$ . Differentiating, we see that

$$\Lambda(v) = \sum_{k=1}^{\infty} (1 + k)a_k(1 - (1 - v)^k) > 0, \quad 0 < v \leq 1. \tag{2.1}$$

Note that  $\int_0^1 \Lambda(v) \, dv = \sum_{k=1}^{\infty} ka_k \geq b$ , and  $\Lambda(v) \downarrow 0$  as  $v \downarrow 0$ . Consequently, there exists  $0 < p \leq 1$  satisfying  $p^{-1} \int_0^p \Lambda(v) \, dv = b$ , that is,

$$F(p) = \sum_{k=1}^{\infty} a_k ((1 - p)^{k+1} - 1 + (k + 1)p) - bp = 0.$$

Let  $X$  be a branching Markov process with generator  $C$ , and assume that  $X(0) = 1$ . Let  $U_{v_0}$  be the level corresponding to the initial particle. Then  $v_0 = U_{v_0}(0)$  is uniformly distributed over  $[0, 1]$ . If  $U_{v_0}(0) > p$  then  $U_{v_0}(t)$  hits 1 in finite time, and since all descendants of  $U_{v_0}$  have higher levels, the population goes to extinction. In particular, the probability of extinction is  $q = 1 - p$ . If  $U_{v_0}(0) < p$  then  $U_{v_0}(t) \rightarrow 0$ . In particular,  $U_{v_0}$  survives for all time and has infinitely many offspring, or, more precisely, the particle at time 0 has infinitely many descendants.

**Theorem 2.1.** *Let  $m(t)$  satisfy  $m(0) = \frac{1}{2}p$  and  $\dot{m}(t) = F(m(t))$ . Then  $m(t) \rightarrow 0$ , and, for every level with  $U_v(\beta_v) < p$ ,*

$$\zeta_v = \lim_{t \rightarrow \infty} \frac{U_v(t)}{m(t)}$$

*exists and satisfies  $0 < \zeta_v < \infty$ . Furthermore,  $W \equiv \lim_{t \rightarrow \infty} m(t)X(t)$  exists almost surely, and  $W > 0$  if and only if  $U_{v_0}(0) < p$ , that is, if and only if  $X(t) \rightarrow \infty$ .*

*Let  $M_v$  be the solution of  $\dot{M}(t) = F(M(t))$  with  $M_v(0) = v$ ,  $0 < v < p$ , and define*

$$H(v) = \lim_{t \rightarrow \infty} \frac{M_v(t)}{m(t)}.$$

Then

$$\mathbb{E}[e^{-Wz} \mid W > 0] = 1 - p^{-1}H^{-1}(z),$$

giving the Laplace transform of the distribution of  $W$  conditioned on  $W > 0$ .

*Proof.* Note that

$$\frac{\dot{m}(t)}{m(t)} = \frac{1}{m(t)} \int_0^{m(t)} \Lambda(v) \, dv - b$$

and, for  $t \geq \beta_v$ ,

$$\frac{\dot{U}_v(t)}{U_v(t)} = \frac{1}{U_v(t)} \int_0^{U_v(t)} \Lambda(v) \, dv - b,$$

so, for  $t \geq 0$ ,

$$m(t) = m(0) \exp \left\{ \int_0^t \frac{1}{m(s)} \int_0^{m(s)} \Lambda(v) \, dv \, ds - bt \right\}$$

and, for  $t \geq \beta_v$ ,

$$U_v(t) = v \exp \left\{ \int_{\beta_v}^t \frac{1}{U_v(s)} \int_0^{U_v(s)} \Lambda(v) \, dv \, ds - b(t - \beta_v) \right\}.$$

Suppose that  $\frac{1}{2}p < v < p$ , and define  $\gamma_v^+ = \inf\{t : U_v(t) \leq \frac{1}{2}p\}$ . Then, for  $t > \gamma_v^+$ ,

$$\int_{\gamma_v^+}^t \frac{1}{U_v(s)} \int_0^{U_v(s)} \Lambda(v) \, dv \, ds = \int_0^{t-\gamma_v^+} \frac{1}{m(s)} \int_0^{m(s)} \Lambda(v) \, dv \, ds$$

and

$$\begin{aligned} \frac{U_v(t)}{m(t)} &= \frac{U_v(\beta_v)}{m(0)} \exp \left\{ \int_{\beta_v}^t \frac{1}{U_v(s)} \int_0^{U_v(s)} \Lambda(v) \, dv \, ds - \int_0^t \frac{1}{m(s)} \int_0^{m(s)} \Lambda(v) \, dv \, ds + b\beta_v \right\} \\ &= \frac{v}{m(0)} \exp \left\{ \int_{\beta_v}^{\gamma_v^+} \frac{1}{U_v(s)} \int_0^{U_v(s)} \Lambda(v) \, dv \, ds - \int_{t-\gamma_v^+}^t \frac{1}{m(s)} \int_0^{m(s)} \Lambda(v) \, dv \, ds + b\beta_v \right\} \\ &\rightarrow \frac{v}{m(0)} \exp \left\{ \int_{\beta_v}^{\gamma_v^+} \frac{1}{U_v(s)} \int_0^{U_v(s)} \Lambda(v) \, dv \, ds + b\beta_v \right\}. \end{aligned}$$

Similarly, suppose that  $0 < v < \frac{1}{2}p$ , and define  $\gamma_v^- = \inf\{t : m(t) \leq v\}$ . Then, as before, for  $t > \gamma_v^-$ ,

$$\begin{aligned} \frac{U_v(t)}{m(t)} &= \frac{U_v(\beta_v)}{m(0)} \exp \left\{ \int_{\beta_v}^t \frac{1}{U_v(s)} \int_0^{U_v(s)} \Lambda(v) \, dv \, ds - \int_0^t \frac{1}{m(s)} \int_0^{m(s)} \Lambda(v) \, dv \, ds + b\beta_v \right\} \\ &= \frac{v}{m(0)} \exp \left\{ \int_{t-\gamma_v^-}^t \frac{1}{U_v(s)} \int_0^{U_v(s)} \Lambda(v) \, dv \, ds - \int_0^{\gamma_v^-} \frac{1}{m(s)} \int_0^{m(s)} \Lambda(v) \, dv \, ds + b\beta_v \right\} \\ &\rightarrow \frac{v}{m(0)} \exp \left\{ - \int_0^{\gamma_v^-} \frac{1}{m(s)} \int_0^{m(s)} \Lambda(v) \, dv \, ds + b\beta_v \right\}. \end{aligned}$$

Conditioned on  $X(t)$ , the counting measure

$$\xi(t) = \sum_{\{v: \beta_v \leq t < \tau_v\}} \frac{\delta_{U_v(t)}}{m(t)}$$

has the distribution of a counting measure given by  $X(t)$  i.i.d. uniform  $[0, m(t)^{-1}]$  random variables. If  $U_{\nu_0} < p$  then  $X(t) \rightarrow \infty$ , and conditioned on  $m(t)X(t)$ ,  $\xi(t)$  is approximately a Poisson process with parameter  $m(t)X(t)$ , and  $U_{\nu_0}(t)/m(t)$  is approximately exponentially distributed with parameter  $m(t)X(t)$ . If we consider  $\xi(t)$  restricted to an interval  $[0, d]$ , the locations of the  $\xi(t, [0, d])$  points in the interval are distributed as  $\xi(t, [0, d])$  i.i.d. uniform  $[0, d]$  random variables. If  $\limsup_{t \rightarrow \infty} \xi(t, [0, d]) = \infty$  then

$$\liminf_{t \rightarrow \infty} \min\{x < d : \delta_x \text{ an atom of } \xi(t)\} = 0.$$

But  $U_{\nu_0}(t)/m(t) = \min\{x : \delta_x \text{ an atom of } \xi(t)\}$ , and  $\zeta_{\nu_0} = \lim_{t \rightarrow \infty} [U_{\nu_0}(t)/m(t)]$  exists and is positive, so  $\limsup_{t \rightarrow \infty} \xi(t, [0, d]) < \infty$ . Since, for all  $\nu < p$ ,  $\zeta_\nu = \lim_{t \rightarrow \infty} [U_\nu(t)/m(t)]$  exists, on the event  $\{U_{\nu_0} < p\}$ ,  $\xi(t)$  converges almost surely in the sense that, for each  $\varphi \in C_c[0, \infty)$ ,  $\varphi \geq 0$ ,

$$\lim_{t \rightarrow \infty} \int_0^{1/m(t)} \varphi(\nu)\xi(t, d\nu) = \sum_{\{\nu: \nu < p\}} \varphi\left(\lim_{t \rightarrow \infty} \frac{U_\nu(t)}{m(t)}\right) \equiv \int_0^\infty \varphi(\nu)\xi(\infty, d\nu).$$

(Observe that, for sufficiently large  $t$ , the only points  $U_\nu(t)/m(t)$  in the support of  $\varphi$  must have  $\nu < p$ .)

The conditional i.i.d. uniform property implies that

$$\begin{aligned} \mathbb{E}\left[\exp\left\{-\int_0^{1/m(t)} \varphi(\nu)\xi(t, d\nu)\right\} \middle| \mathcal{F}_t^{X(t)}\right] &= \left[m(t) \int_0^{1/m(t)} e^{-\varphi(\nu)} d\nu\right]^{X(t)} \\ &= \left[1 - m(t) \int_0^{1/m(t)} (1 - e^{-\varphi(\nu)}) d\nu\right]^{[1/m(t)]m(t)X(t)}. \end{aligned}$$

Letting  $t \rightarrow \infty$ , the left-hand side converges almost surely, so too does the right-hand side. Because

$$\left[1 - m(t) \int_0^{1/m(t)} (1 - e^{-\varphi(\nu)}) d\nu\right]^{1/m(t)} \rightarrow \exp\left\{-\int_0^\infty (1 - e^{-\varphi(\nu)}) d\nu\right\},$$

the limit  $W \equiv \lim_{t \rightarrow \infty} m(t)X(t)$  must exist, and we conclude that

$$\mathbb{E}\left[\exp\left\{-\int_0^\infty \varphi(\nu)\xi(\infty, d\nu)\right\} \middle| \vee_t \mathcal{F}_t^X\right] = \exp\left\{-W \int_0^\infty (1 - e^{-\varphi(\nu)}) d\nu\right\}.$$

Thus, conditioned on  $W$ ,  $\xi(\infty)$  is a Poisson point process on  $[0, \infty)$  with intensity  $W$ . If  $\nu_0 > p$ ,  $\xi(\infty, [0, \infty)) = 0$ , so  $W = 0$ . If  $\nu_0 < p$ , the left-hand side is almost surely less than 1 for any  $\varphi$  that is strictly positive on some interval, so  $W > 0$ .

Furthermore, conditioned on  $\{\nu_0 < p\} = \{W > 0\}$ ,  $H(\nu_0) = \lim_{t \rightarrow \infty} [U_{\nu_0}(t)/m(t)]$  must be exponentially distributed with parameter  $W$ . Then

$$\mathbb{P}\{H(\nu_0) > z \mid 0 < \nu_0 < p\} = \mathbb{E}[e^{-Wz} \mid W > 0] = 1 - p^{-1}H^{-1}(z). \quad \square$$

### 3. The $X \log X$ condition

Since  $\mathbb{E}[X(t)] = e^{bt}$ , the natural scaling of  $X$  should be  $e^{-bt} X(t)$ . Observing that

$$\lim_{t \rightarrow \infty} e^{-bt} X(t) = \lim_{t \rightarrow \infty} \frac{e^{-bt} W}{m(t)},$$

to understand the behavior of  $e^{-bt} X(t)$ , we must understand the behavior of  $e^{-bt} W/m(t)$ , or, equivalently,

$$e^{bt} m(t) = \frac{1}{2} p \exp \left\{ \int_0^t \frac{1}{m(s)} \int_0^{m(s)} \Lambda(v) \, dv \, ds \right\}. \tag{3.1}$$

Since the right-hand side of (3.1) is increasing, the problem reduces to whether or not the integral in the exponent is bounded.

Note that  $m(\cdot)$  is bounded above and below by functions of the form  $Ce^{-cs}$ , where we can always take  $0 < C < p$ . Then  $\int_0^t (1/m(s)) \int_0^{m(s)} \Lambda(v) \, dv \, ds$  is bounded above and below by functions of the form

$$\begin{aligned} \int_0^t \frac{e^{cs}}{C} \int_0^{Ce^{-cs}} \Lambda(v) \, dv \, ds &= \int_{Ce^{-ct}}^C \frac{1}{cz^2} \int_0^z \Lambda(v) \, dv \, dz \\ &= \frac{1}{cCe^{-ct}} \int_0^{Ce^{-ct}} \Lambda(v) \, dv - \frac{1}{cC} \int_0^C \Lambda(v) \, dv + \int_{Ce^{-ct}}^C \frac{\Lambda(z)}{cz} \, dz, \end{aligned}$$

where the first equality follows by substitution and the second by integration by parts. On the right-hand side, as  $t \rightarrow \infty$ , the first term tends to 0 (recall (2.1)), the second is constant, and the third converges to

$$\int_0^C \frac{\sum_{k=1}^{\infty} (k+1)a_k [1 - (1-z)^k]}{cz} \, dz = \sum_{k=1}^{\infty} (k+1)a_k \int_0^C \frac{1 - (1-z)^k}{cz} \, dz.$$

Applying l'Hôpital's rule,

$$\lim_{k \rightarrow \infty} \frac{\int_0^C z^{-1} (1 - (1-z)^k) \, dz}{\log k} = 1.$$

Thus,

$$\lim_{t \rightarrow \infty} e^{bt} m(t) = \frac{1}{2} p \exp \left\{ \int_0^{\infty} \frac{1}{m(s)} \int_0^{m(s)} \Lambda(v) \, dv \right\} < \infty$$

if and only if  $\sum_k (k \log k) a_k < \infty$ , and we have the following result.

**Theorem 3.1.** *The limit*

$$\hat{W} = \lim_{t \rightarrow \infty} e^{-bt} X(t)$$

exists almost surely and  $\mathbb{P}\{\hat{W} > 0\} = p$  if and only if  $\sum_{k=1}^{\infty} (k \log k) a_k < \infty$ . Otherwise,  $\mathbb{P}\{\hat{W} = 0\} = 1$ .



### 4. Asymptotic behavior of continuous-state branching processes

We obtain the genealogical construction for continuous-state branching processes by taking the limit of  $A_r$  as  $r \rightarrow \infty$ . To obtain the most general such process (satisfying a moment condition discussed below), we must let the  $a_k = a_k^{(r)}$  depend on  $r$ . Let

$$\Lambda^{(r)}(u) = \sum_{k=1}^{\infty} r(k+1)a_k^{(r)} \left[ 1 - \left( 1 - \frac{u}{r} \right)^k \right],$$

and assume that

$$\lim_{r \rightarrow \infty} \Lambda^{(r)}(u) = \lim_{r \rightarrow \infty} \sum_{k=1}^{\infty} r(k+1)a_k^{(r)} \left[ \sum_{l=1}^k \binom{k}{l} (-1)^{l+1} \left( \frac{u}{r} \right)^l \right] \equiv \Lambda(u) \tag{4.1}$$

exists uniformly for  $u$  in bounded intervals. This condition is essentially [6, Equation (9.4.36)].

Any limit of the form (4.1) has a completely monotone derivative, and by Bernstein’s theorem [18, Theorem 3.2],  $\Lambda$  can be written as

$$\Lambda(v) = 2a_0v + \int_0^{\infty} (1 - e^{-vy})\mu(dy) \tag{4.2}$$

for  $a_0 \geq 0$  and  $\mu$  a  $\sigma$ -finite measure satisfying  $\int_0^{\infty} (1 \wedge y)\mu(dy) < \infty$ . (Note that the corresponding statement in [14] does not characterize precisely the measure, denoted here by  $\mu$ , which is related to the measure  $\nu$  there by  $\mu(dy) = y^{-1}\nu(dy)$ .)

The limiting generator obtained in [14] has state space  $E = [0, \infty)^\infty \cup \bigcup_{k=0}^{\infty} [0, \infty)^k$ , and

$$\mathcal{D}(A) = \left\{ f(u) = \prod g(u_i) : g \in C^1[0, \infty), 0 \leq g \leq 1, \right. \\ \left. \text{there exists } v_g \text{ such that } g(v) = 1 (v \geq v_g) \right\}.$$

Then, if  $r > v_g$ ,

$$\begin{aligned} A_r f(u, n) &= f(u, n) \sum_{i=1}^n \sum_{k=1}^{\infty} \frac{(k+1)a_k}{r^{k-1}} \int_{[u_i, r]^k} \left[ \left( \prod_{l=1}^k g(v_l) \right) - 1 \right] dv_1 \cdots dv_k \\ &\quad + f(u, n) \sum_{i=1}^n \left[ \sum_{k=1}^{\infty} r^2 a_k \left( \left( 1 - \frac{u_i}{r} \right)^{k+1} - 1 + (k+1) \frac{u_i}{r} \right) - bu_i \right] \frac{g'(u_i)}{g(u_i)} \\ &= f(u, n) \sum_{i=1}^n \left[ \Lambda^{(r)}(u_i) - \Lambda^{(r)} \left( u_i + \int_{u_i}^{\infty} [1 - g(v)] dv \right) \right] \\ &\quad + f(u, n) \sum_{i=1}^n \left[ \int_0^{u_i} \Lambda^{(r)}(v) dv - bu_i \right] \frac{g'(u_i)}{g(u_i)}. \end{aligned}$$

Define  $\xi_n = \sum_{i=1}^n \delta_{u_i}$ , and assume that, as  $n \rightarrow \infty$ ,  $\xi_n \rightarrow \xi = \sum_{i=1}^{\infty} \delta_{u_i}$  such that  $\xi([0, d]) < \infty$  for all  $d > 0$ . Then, as  $r$  and  $n \rightarrow \infty$ ,  $A_r f(u, n)$  converges to

$$\begin{aligned} Af(u) &= f(u) \sum_{i=1}^{\infty} \left[ \Lambda(u_i) - \Lambda \left( u_i + \int_{u_i}^{\infty} [1 - g(v)] dv \right) \right] \\ &\quad + f(u) \sum_{i=1}^{\infty} \left[ \int_0^{u_i} \Lambda(v) dv - bu_i \right] \frac{g'(u_i)}{g(u_i)}. \end{aligned}$$

By (4.2),

$$\begin{aligned}
 Af(u) &= f(u) \sum_i 2a_0 \int_{u_i}^\infty [g(v) - 1] dv \\
 &\quad + f(u) \sum_i \int_0^\infty \left[ \exp \left\{ y \int_{u_i}^\infty [g(v) - 1] dv \right\} - 1 \right] e^{-yu_i} \mu(dy) \\
 &\quad + f(u) \sum_i \left( \int_0^{u_i} \Lambda(v) dv - bu_i \right) \frac{\partial_{u_i} g(u_i)}{g(u_i)}.
 \end{aligned}$$

For  $v_1 < v_2$ , the first term on the right-hand side states that a particle with level  $u_i$  produces particles with levels in  $[u_i + v_1, u_i + v_2]$  at rate  $2a_0(v_2 - v_1)$ . The second term states that a particle with level  $u_i$  produces bursts of particles where the levels of the new particles are given by a Poisson point process with intensity  $y$  on  $[u_i, \infty)$ . The rate at which such a burst is produced with  $y \in [y_1, y_2]$  is  $\int_{y_1}^{y_2} e^{-yu_i} \mu(dy)$ . As in the case of finite  $r$ , the third term states that each level evolves according to

$$\dot{U}_i(t) = \int_0^{U_i(t)} \Lambda(v) dv - bU_i(t) \equiv F(U_i(t)).$$

For  $x \in [0, \infty)$ , let  $\alpha(x, du)$  be the distribution of a Poisson point process on  $[0, \infty)$  with intensity  $x$ . For  $f(u) = \prod_i g(u_i)$ , define  $\lambda_g = \int_0^\infty [1 - g(v)] dv$ . Then  $\hat{f}(x) \equiv \alpha f(x) = e^{-\lambda_g x}$  and  $C \hat{f}(x)$  equals

$$\begin{aligned}
 \alpha Af(x) &= x \left( \int_0^{\lambda_g} \Lambda(v) dv - b\lambda_g \right) e^{-\lambda_g x} \\
 &= x \left( a_0 \lambda_g^2 - b\lambda_g + \int_0^\infty \left( \lambda_g - \frac{1}{y} [1 - e^{-\lambda_g y}] \right) \mu(dy) \right) e^{-\lambda_g x} \\
 &= x \left( a_0 \hat{f}''(x) + b \hat{f}'(x) + \int_0^\infty (\hat{f}(x + y) - \hat{f}(x) - \hat{f}'(x)y) \frac{1}{y} \mu(dy) \right),
 \end{aligned}$$

which is the generator of a Markov process obtained by solving the time-change equation

$$X(t) = X(0) + Y \left( \int_0^t X(s) ds \right),$$

where  $Y$  is a Lévy process with generator

$$Gf(x) = a_0 f''(x) + b f'(x) + \int_0^\infty (f(x + y) - f(x) - f'(x)y) \frac{1}{y} \mu(dy). \tag{4.3}$$

This time-change representation of a continuous-state branching process was first given in [15]. The Lévy measure for  $Y$  is  $\gamma(dy) = (1/y)\mu(dy)$ , which, by the conditions on  $\mu$ , clearly satisfies

$$\int_0^\infty (1 \wedge y^2) \gamma(dy) < \infty.$$

Note that  $G$  given by (4.3) is not the most general Lévy process with only positive jumps. The restriction given by the properties of  $\mu$  implies that  $\mathbb{E}[|Y(u)|] < \infty$  and  $\mathbb{E}[Y(u)] = bu$ ; hence,  $\mathbb{E}[X(t)] = \mathbb{E}[X(0)]e^{bt}$ . If we try to extend the genealogical construction to measures  $\mu$

which would give the most general Lévy process with only positive jumps,  $\Lambda$  given by (4.2) would blow up. We can obtain a genealogical construction for the general case by adding birth events of the form considered in [5, Section 3.3]. Using that construction, the levels would take random jumps. We have not attempted to study the asymptotic behavior of that construction.

As in Section 2, index the lines of descent by their levels at the times of their birth, and let  $v_0$  denote the lowest level at time 0. As before, all particles alive at time  $t$  satisfy  $U_v(t) \geq U_{v_0}(t)$ .

Assume that at least one of  $a_0$  and  $\mu$  is nonzero. Then, if  $b = 0$  (critical case) or  $b < 0$  (subcritical case),  $\lim_{t \rightarrow \infty} U_{v_0}(t) = \infty$ . A line of descent ends if  $U_v$  hits  $\infty$ , but, in general, that may not occur in finite time. In particular,

$$\int_{U_v(t_1)}^{U_v(t_2)} \frac{1}{F(v)} dv = t_2 - t_1,$$

so  $U_v$  hits  $\infty$  in finite time if and only if

$$\int_{U_v(\beta_v)}^{\infty} \frac{1}{F(v)} dv < \infty$$

(cf. [2, p. 167]), as holds, for example, when  $a_0 > 0$ .

In the supercritical case,  $b > 0$  as in the ordinary branching process case, and there exists a solution of

$$\int_0^\kappa \Lambda(v) dv - b\kappa = 0,$$

unless  $\lim_{v \rightarrow \infty} \Lambda(v) \leq b$ ; this last statement holds if and only if  $a_0 = 0$  and  $\mu([0, \infty)) \leq b$ , in which case, define  $\kappa = \infty$ . Then  $\lim_{t \rightarrow \infty} X(t) = \infty$  if and only if  $v_0 < \kappa$ . If  $X(0) = x > 0$  then

$$\mathbb{P}\{v_0 < \kappa\} = 1 - e^{-\kappa x} = \mathbb{P}\left\{\lim_{t \rightarrow \infty} X(t) = \infty\right\}.$$

The proof of the following result is essentially the same as that of Theorem 2.1. Only the definition of  $\Lambda$  has changed.

**Theorem 4.1.** *Assume that  $b > 0$ , and let  $\kappa$  be defined as above and  $X(0) = x > 0$ . Let  $m(t)$  satisfy  $m(0) = \frac{1}{2}\kappa$  for finite  $\kappa$ , else  $m(0) = 1$  if  $\kappa = \infty$ , and let*

$$\dot{m}(t) = F(m(t)).$$

*Then  $m(t) \rightarrow 0$  ( $t \rightarrow \infty$ ), and for every level  $v$  with  $U_v(\beta_v) < \kappa$ ,*

$$\zeta_v = \lim_{t \rightarrow \infty} \frac{U_v(t)}{m(t)}$$

*exists and satisfies  $0 < \zeta_v < \infty$ . Furthermore,*

$$W \equiv \lim_{t \rightarrow \infty} m(t)X(t)$$

*exists almost surely, and  $W > 0$  if and only if  $U_{v_0}(0) < \kappa$ , that is, if and only if  $X(t) \rightarrow \infty$ .*

*Let  $M_v$  be the solution of  $\dot{M}_v(t) = F(M_v(t))$ , where  $M_v(0) = v$ ,  $0 < v < \kappa$ , and define*

$$H(v) = \lim_{t \rightarrow \infty} \frac{M_v(t)}{m(t)}.$$

Then

$$\mathbb{E}[e^{-Wz} \mid W > 0] = \mathbb{P}\{v_0 > H^{-1}(z) \mid v_0 < \kappa\} = \frac{e^{-xH^{-1}(z)} - e^{-x\kappa}}{1 - e^{-x\kappa}}$$

is the Laplace transform of the distribution of  $W$  conditioned on  $W > 0$ .

The analog of the  $X \log X$  condition is given by the following result.

**Theorem 4.2.** *Under the assumptions of Theorem 4.1, the limit*

$$\hat{W} = \lim_{t \rightarrow \infty} e^{-bt} X(t)$$

exists almost surely and  $\mathbb{P}\{\hat{W} > 0\} = 1 - e^{-x\kappa}$  if and only if

$$\int_0^\infty \log(1+z)\mu(dz) < \infty, \tag{4.4}$$

or, equivalently,

$$\int_0^\infty z \log(1+z)\gamma(dz) < \infty.$$

*Proof.* As in the proof of Theorem 3.1, we must show that (4.4) is equivalent to

$$\int_0^C \frac{\Lambda(v)}{v} dv = \int_0^C \left[ 2a_0 + \int_0^\infty \frac{1 - e^{-vy}}{v} \mu(dy) \right] dv < \infty, \tag{4.5}$$

which holds if and only if

$$\int_0^\infty \int_0^C \frac{1 - e^{-vy}}{v} dv \mu(dy) < \infty.$$

Observing that

$$R(y) = \int_0^C \frac{1 - e^{-vy}}{v} dv = \int_0^{Cy} \frac{1 - e^{-w}}{w} dw,$$

we have

$$\lim_{y \rightarrow \infty} \frac{R(y)}{\log(1+y)} = \lim_{y \rightarrow \infty} (1+y)R'(y) = 1$$

and

$$\lim_{y \rightarrow 0} \frac{R(y)}{\log(1+y)} = C,$$

and it follows that (4.5) holds if and only if (4.4) holds. □

In addition, by similar arguments, we have the following result.

**Theorem 4.3.** *Suppose that  $b \leq 0$  or  $b > 0$  and  $0 < \kappa < \infty$ . Let  $m(0) = 1$  ( $b \leq 0$ ) or  $m(0) = 2\kappa$  ( $b > 0$ ), and let  $\dot{m}(t) = F(m(t))$ . Assume that  $\int_{m(0)}^\infty (1/F(v)) dv = \infty$ . Then  $m(t) < \infty$  for all  $t > 0$  and  $\lim_{t \rightarrow \infty} m(t) = \infty$ .*

*Assume that  $X(0) = x > 0$ , and if  $b > 0$ , assume that  $v_0 > \kappa$ . Then, for each line of descent,*

$$\zeta_v = \lim_{t \rightarrow \infty} \frac{U_v(t)}{m(t)}$$

exists and satisfies  $0 < \zeta_v < \infty$ . Furthermore,

$$W \equiv \lim_{t \rightarrow \infty} m(t)X(t)$$

exists almost surely and is positive. Let  $M_v$  satisfy  $\dot{M}_v(t) = F(M_v(t))$ , and define  $H(v) = \lim_{t \rightarrow \infty} M_v(t)/m(t)$ .

For  $b \leq 0$ ,  $\mathbb{P}\{W > 0\} = 1$ , and conditioned on  $W$ ,  $H(v_0)$  is exponentially distributed with parameter  $W$ . The distribution of  $W$  is determined by

$$\mathbb{E}[e^{-Wz}] = \mathbb{P}\{H(v_0) > z\} = \mathbb{P}\{v_0 > H^{-1}(z)\} = e^{-xH^{-1}(z)}, \quad z \geq 0.$$

For  $b > 0$ , conditioned on  $W$  and  $v_0 > \kappa$ ,  $H(v_0) - \kappa$  is exponentially distributed with parameter  $W$ . The distribution of  $W$  is determined by

$$\begin{aligned} \mathbb{E}[e^{-Wz} \mid W > 0] &= \mathbb{P}\{H(v_0) - \kappa > z \mid v_0 > \kappa\} \\ &= \mathbb{P}\{v_0 > H^{-1}(z + \kappa) \mid v_0 > \kappa\} \\ &= e^{-x[H^{-1}(z+\kappa) - \kappa]}. \end{aligned}$$

## References

- [1] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, New York.
- [2] BERTOIN, J. AND LE GALL, J.-F. (2006). Stochastic flows associated to coalescent processes. III. Limit theorems. *Illinois J. Math.* **50**, 147–181.
- [3] COHN, H. (1982). Norming constants for the finite supercritical Bellman–Harris process. *Z. Wahrscheinlichkeitsth.* **61**, 189–205.
- [4] DONEY, R. A. (1972). A limit theorem for a class of supercritical branching processes. *J. Appl. Prob.* **9**, 707–724.
- [5] ETHERIDGE, A. M. AND KURTZ, T. G. (2018). Genealogical constructions of population models. To appear in *Ann. Prob.*
- [6] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. John Wiley, New York.
- [7] GREY, D. R. (1974). Asymptotic behaviour of continuous time, continuous state-space branching processes. *J. Appl. Prob.* **11**, 669–677.
- [8] HARRIS, T. E. (1951). Some mathematical models for branching processes. In *Proc. 2nd Berkeley Symp. Math. Statist. Prob., 1950*, University of California Press, Berkeley and Los Angeles, pp. 305–328.
- [9] HEYDE, C. C. (1970). Extension of a result of Seneta for the super-critical Galton–Watson process. *Ann. Math. Statist.* **41**, 739–742.
- [10] JIŘINA, M. (1958). Stochastic branching processes with continuous state space. *Czechoslovak Math. J.* **8**, 292–313.
- [11] KESTEN, H. AND STIGUM, B. P. (1966). A limit theorem for multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37**, 1211–1223.
- [12] KURTZ, T. G. (1998). Martingale problems for conditional distributions of Markov processes. *Electron. J. Prob.* **3**, 29pp.
- [13] KURTZ, T. G. AND NAPPO, G. (2011). The filtered martingale problem. In *The Oxford Handbook of Nonlinear Filtering*, Oxford University Press, pp. 129–165.
- [14] KURTZ, T. G. AND RODRIGUES, E. R. (2011). Poisson representations of branching Markov and measure-valued branching processes. *Ann. Prob.* **39**, 939–984.
- [15] LAMPERTI, J. (1967). Continuous state branching processes. *Bull. Amer. Math. Soc.* **73**, 382–386.
- [16] LEVINSON, N. (1959). Limiting theorems for Galton–Watson branching process. *Illinois J. Math.* **3**, 554–565.
- [17] ROGERS, L. C. G. AND PITMAN, J. W. (1981). Markov functions. *Ann. Prob.* **9**, 573–582.
- [18] SCHILLING, R. L., SONG, R. AND VONDRAČEK, Z. (2012). *Bernstein Functions: Theory and Applications* (De Gruyter Stud. Math. **37**), 2nd edn. Walter de Gruyter, Berlin.
- [19] SENETA, E. (1968). On recent theorems concerning the supercritical Galton–Watson process. *Ann. Math. Statist.* **39**, 2098–2102.

THOMAS G. KURTZ, *University of Wisconsin–Madison*

Departments of Mathematics and Statistics, University of Wisconsin–Madison, 480 Lincoln Drive, Madison, WI 53706-1388, USA. Email address: kurtz@math.wisc.edu