MARKOVIAN PATHS TO EXTINCTION

PETER JAGERS,* ** *Chalmers University of Technology*
FIMA C. KLEBANER,*** *Monash University*
SERIK SAGITOV,* *Chalmers University of Technology*

Abstract

Subcritical Markov branching processes $\{Z_t\}$ die out sooner or later, say at time $T < \infty$. We give results for the *path to extinction* $\{Z_{uT}, \ 0 \le u \le 1\}$ that include its finite dimensional distributions and the asymptotic behaviour of $x^{u-1}Z_{uT}$, as $Z_0 = x \to \infty$. The limit reflects an interplay of branching and extreme value theory. Then we consider the population on the verge of extinction, as modelled by Z_{T-u}, $u > 0$, and show that as $Z_0 = x \to \infty$ this process converges to a Markov process $\{Y_u\}$, which we describe completely. Emphasis is on continuous time processes, those in discrete time displaying a more complex behaviour, related to Martin boundary theory.

1. Introduction

In an earlier paper [8] we considered big, general, nonlattice subcritical branching populations, starting from, say, x individuals. Here 'general' stands for single-type general branching processes, i.e. populations where no particular life span distribution is assumed and where individuals may give birth repeatedly and in random litters [7]. Writing r for the absolute value of the Malthusian parameter (which must be negative for subcritical populations), in [8] we showed that the time T to extinction must satisfy

$$T - \frac{\ln x}{r} \xrightarrow{\text{D}} \frac{\eta + \ln c}{r}, \qquad x \to \infty,$$

where the random variable η has a standard Gumbel distribution,

$$\text{P}(\eta \le y) = \exp(-\text{e}^{-y}), \qquad -\infty < y < \infty,$$

$0 < c < 1$ is the constant appearing in Yaglom's limit theorem [3] and '$\xrightarrow{\text{D}}$' denotes convergence in distribution. (For Markov branching processes and Galton–Watson processes these or corresponding asymptotics were obtained by Pakes [10, 11].)

The main results of [8], however, were those yielding population size 'half-way' to extinction, i.e. at times uT, for $0 < u < 1$, for example $u = \frac{1}{2}$. In the continuous time case, it was shown that if the population starts from x individuals at time zero and Z_t denotes its size at time

$t \geq 0$, then as $x \to \infty$ the properly scaled population size $x^{u-1} Z_{uT}$ converges weakly to a process of the form $Cc^{-u}\mathrm{e}^{-u\eta}$, for fixed u. In this, the constant C is defined by $C = \lim_{t\to\infty} \mathrm{e}^{rt} \mathrm{E}_1[Z_t]$, explicitly determined by life span and reproduction distributions. Integer suffices on expectations and probabilities refer to the starting number; in the formula for C the population thus starts from one (newborn) Eve. Here and in the sequel, E_x and P_x denote the expectation and probability, respectively, when the population is started from x individuals.

This was obtained through an L^2-argument presupposing, besides a bounded birth-and-death intensity, essentially a second moment of the number of offspring per individual. In the case of Markov branching processes, it turns out that the quadratic mean analysis can be replaced by conditioning at the nonstopping time uT and a total probability argument, which may have wider relevance for Markovian paths until hitting a designated state, zero or not. As a by-product, the second moment assumption can be relaxed to the (mathematically) natural condition of the offspring number X having a finite $x \log x$ moment (see (7), below). Furthermore, the time $uT < T$ can be replaced by others strictly preceding T, rendering it possible to study the path of the processes on the eve of extinction, i.e. at times $T - u$, $u > 0$, rather than long before.

This is the programme of the present paper, which will concentrate on subcritical continuous time Markov branching, but also deal with the Galton–Watson case.

2. Facts about Markov branching in continuous time

Continuous time Markov branching processes Z_t yield the number of particles at times $t \geq 0$ in a population of independently existing and reproducing particles. Each particle lives for an exponential time with parameter a and at the moment of death it splits into k offspring with probabilities p_k, $k = 0, 1, 2, \ldots$, and mean $m = \sum_{k=1}^{\infty} k p_k$. The word 'particle' is quite appropriate, reminding us that these are unbiological, nonageing, and splitting entities.

In this section we present some well-known results in the subcritical case $m < 1$, which are needed in the rest of the paper. The specific function defined in terms of the parameters of the process given by (4), below, is central for the theory and systematically used.

The offspring generating and process generating functions

$$f(s) = \sum_{k=0}^{\infty} p_k s^k,$$

$$F(s, t) = \mathrm{E}_1[s^{Z_t}]$$

are the classical tools of analysis. The latter satisfies the forward Kolmogorov equation

$$\frac{\partial F(s, t)}{\partial t} = \phi(s) \frac{\partial F(s, t)}{\partial s}, \tag{1}$$

where $\phi(s) = a(f(s) - s)$, and the backward Kolmogorov equation

$$\frac{\partial F(s, t)}{\partial t} = \phi(F(s, t)), \tag{2}$$

with the boundary condition $F(s, 0) = s$; see [3], [6], and [12] for this and a more leisurely presentation of Markov branching as a whole. Differentiating (2) with respect to s and then letting $s \uparrow 1$, we obtain a differential equation for the mean of Z_t,

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{E}_1[Z_t] = \phi'(1) \mathrm{E}_1[Z_t],$$

with the following solution:

$$E_1[Z_t] = e^{-rt}, \qquad r = a(1 - m). \tag{3}$$

In other words, the generally true asymptotic relation $E_1[Z_t] \sim Ce^{-rt}$ holds exactly for Markov branching, with $C = 1$. In the present subcritical case, $m < 1$, $\phi(0) = ap_0 > 0$, and the function

$$\pi(s) = \int_0^s \frac{dv}{\phi(v)}, \tag{4}$$

is well defined. It plays a crucial role in the rest of the paper.

Lemma 1. *In the subcritical case*

$$\pi(s) \sim -r^{-1} \ln(1 - s), \qquad s \uparrow 1. \tag{5}$$

Furthermore, there is a constant $1 < b < \infty$ such that

$$\pi(s) = -r^{-1}(\ln(1 - s) + \ln b) + o(1), \qquad s \uparrow 1, \tag{6}$$

if and only if the following $x \log x$-condition holds:

$$\sum_{k=2}^{\infty} (k \ln k) p_k < \infty. \tag{7}$$

Proof. Equation (5) holds since the positive function

$$D(s) = \frac{1}{(1 - m)(1 - s)} - \frac{1}{f(s) - s}$$

is $o(1/(1 - s))$ as $s \uparrow 1$. Relation (6) follows from the fact that the integral $\int_0^1 D(v) \, dv$ is finite if and only if (7) holds (see [3, Chapter I.10, Corollary 2]).

Proposition 1. *Suppose that $m < 1$ and (7) holds. Then the survival probability $Q(t) = P_1(Z_t > 0)$ has the asymptotics*

$$Q(t) \sim b^{-1} e^{-rt}, \qquad t \to \infty, \tag{8}$$

implying that

$$E_1[Z_t \mid Z_t > 0] \to b. \tag{9}$$

Moreover,

$$P_1(Z_t = k \mid Z_t > 0) \to b_k, \qquad k \geq 1, \, t \to \infty, \tag{10}$$

where $b = \sum_{k=1}^{\infty} k b_k$, and

$$B(s) := \sum_{k=1}^{\infty} b_k s^k = 1 - e^{-r\pi(s)}. \tag{11}$$

Thus, for Markovian branching processes, the Yaglom constant c is equal to b^{-1}.

Proof of Proposition 1. It follows from (2) that Harris's representation

$$\pi(F(s, t)) = \pi(s) + t \tag{12}$$

holds. Since $Q(t) = 1 - F(0, t)$, Lemma 1 and (12) imply that

$$-\ln Q(t) = rt + b + o(1), \qquad t \to \infty,$$

which in its turn yields (8).

On the other hand, for any fixed s, it follows from (12) that

$$F(s, t) = F(0, t + \pi(s)). \tag{13}$$

Together with (8), this implies that

$$1 - F(s, t) \sim b^{-1} e^{-r(t + \pi(s))}, \qquad t \to \infty.$$

Hence,

$$\frac{1 - F(s, t)}{Q(t)} \to e^{-r\pi(s)}, \qquad t \to \infty, \tag{14}$$

which leads further onto (10), since

$$\mathrm{E}_1[s^{Z_t} \mid Z_t > 0] = 1 - \frac{1 - F(s, t)}{Q(t)}.$$

The convergence in (9) follows from

$$\mathrm{E}_1[Z_t \mid Z_t > 0] = \frac{\mathrm{E}_1[Z_t]}{Q(t)},$$

(3), and (8). Finally, the equality $b = \sum_{k=1}^{\infty} k b_k$ is a consequence of (11), (6), and

$$\lim_{s \uparrow 1} B'(s) = \lim_{s \uparrow 1} r\pi'(s) e^{-r\pi(s)}$$
$$= b.$$

Proposition 1 is known as Yaglom's theorem. Assuming finite second moments, Yaglom proved the result for Galton–Watson processes, whereas the present case is due to Sevast´yanov. See [3, p. 114], for a brief proof and see [12, pp. 67–72] for a more complete exposition.

3. The path to extinction

As before, T denotes the extinction time of the Markov branching process Z_t, i.e. the time when it hits zero. If the initial population has only one member, the distribution function of the time to extinction is

$$G(t) = \mathrm{P}_1(T \le t) = \mathrm{P}_1(Z_t = 0) = F(0, t).$$

It will be convenient to write $g(t) = G'(t)$ and $F'(s, t) = (\partial/\partial s)F(s, t)$, so that

$$F'(s, t) = \frac{\phi(F(s, t))}{\phi(s)} \tag{15}$$

from (1) and (2). The backward equation (2) yields

$$g(t) = \phi(F(0, t)). \tag{16}$$

By (8) we obtain

$$g(t) \sim r(1 - F(0, t)) \sim rb^{-1}e^{-rt} \quad \text{as } t \to \infty. \tag{17}$$

The object studied in this section is the process $\{Z_{uT}, \ 0 \le u < 1\}$, depicting the *path to extinction*. In the following result, we see that its probability law permits a basic representation.

Theorem 1. *Let $x, y \in \{1, 2, \dots\}$, $u \in [0, 1)$, and $\kappa = u/(1 - u)$. The distribution of Z_{uT} is given by*

$$P_x(Z_{uT} = y) = \frac{y}{1 - u} \int_0^\infty G^{y-1}(t) \, P_x(Z_{\kappa t} = y) g(t) \, dt, \tag{18}$$

with the generating function

$$E_x[s^{Z_{uT}}] = \frac{sx}{1 - u} \int_0^\infty F^{x-1}(sG(t), \kappa t) F'(sG(t), \kappa t) g(t) \, dt. \tag{19}$$

Finite dimensional distributions are given by

$$P_x(Z_{u_1T} = y_1, Z_{u_2T} = y_2, \dots, Z_{u_nT} = y_n)$$
$$= \frac{y_n}{1 - u_n} \int_0^\infty G^{y_n-1}(t) \, P_x(Z_{\kappa_1 t} = y_1, \dots, Z_{\kappa_n t} = y_n) g(t) \, dt, \tag{20}$$

for $0 < u_1 < u_2 < \cdots < u_n < 1$ and $\kappa_i = u_i/(1 - u_n)$.

Proof. By the law of total probability we obtain

$$P_x(Z_{uT} = y) = \int_0^\infty P_x(Z_{uT} = y, T \in dt)$$
$$= \int_0^\infty P_x(Z_{ut} = y, T \in dt)$$
$$= \int_0^\infty P_x(Z_{ut} = y) \, P_x(T \in dt \mid Z_{ut} = y)$$
$$= \int_0^\infty P_x(Z_{ut} = y) \, P_y(T' + ut \in dt).$$

The last equality holds by the following probabilistic argument. Conditionally on $Z_{uT} = y$, the probability of the original population dying out at T is the same as the probability that a population of y individuals at uT dies out after time $T' = T(1 - u)$. Furthermore, by independence we obtain

$$P_y(T \le t) = G(t)^y.$$

Hence, this turns into

$$P_x(Z_{uT} = y) = \int_0^\infty P_x(Z_{ut} = y) y G^{y-1}(t(1 - u)) g(t(1 - u)) \, dt,$$

and (18) follows.

The generating function is obtained by direct calculations

$$\mathrm{E}_x[s^{Z_{uT}}] = \sum_y \mathrm{P}_x(Z_{uT} = y)s^y$$

from (18).

The very same total probability argument works for finite dimensional distributions. Indeed, with $0 < u_1 < u_2 < \cdots < u_n < 1$ and $\kappa_i = u_i/(1 - u_n)$, we obtain

$$\mathrm{P}_x(Z_{u_1T} = y_1, Z_{u_2T} = y_2, \ldots, Z_{u_nT} = y_n)$$
$$= \int_0^\infty \mathrm{P}_x(Z_{u_1t} = y_1, \ldots, Z_{u_nt} = y_n)y_n G^{y_n-1}((1 - u_n)t)g((1 - u_n)t)\,\mathrm{d}t$$
$$= \frac{y_n}{1 - u_n}\int_0^\infty G^{y_n-1}(t)\,\mathrm{P}_x(Z_{\kappa_1t} = y_1, \ldots, Z_{\kappa_nt} = y_n)g(t)\,\mathrm{d}t,$$

which is (20). The joint generating function can now be obtained similarly as in (19).

With the representation theorem in our hands we can state and prove the main result on the path between start and demise.

Theorem 2. *Suppose that $m < 1$ and that the $x \log x$ condition (7) holds. As $Z_0 = x$ tends to infinity, the finite dimensional distributions converge, i.e.*

$$\{x^{u-1}Z_{uT}, \ 0 \le u < 1\} \xrightarrow{\text{FD}} \{b^u \mathrm{e}^{-u\eta}, \ 0 \le u < 1\},$$

where η has a standard Gumbel distribution and '$\xrightarrow{\text{FD}}$' denotes convergence in finite dimensional distribution.

The question of tightness is left open here. We plan to address it in a forthcoming paper, in the context of general branching processes.

Proof of Theorem 2. The proof relies on dominated convergence. We show convergence of one-dimensional distributions by establishing the asymptotics of the function under the integral in (19) and exhibiting a dominating function. We then show how the Markov property can be used on the right-hand side of (20) to reduce the convergence of n-dimensional distributions to those of lesser dimensionality.

First, by (13) for $t \to \infty$ and $s_t \uparrow 1$, we obtain

$$1 - F(s_t, \kappa t) = Q(\kappa t + \pi(s_t)) \sim b^{-1}\exp(-r(\kappa t + \pi(s_t))) \sim (1 - s_t)\mathrm{e}^{-r\kappa t}. \qquad (21)$$

Taken together with (15), this gives

$$F'(s_t, t) \sim \frac{1 - F(s_t, t)}{1 - s_t} \sim \mathrm{e}^{-r\kappa t}. \qquad (22)$$

Combining (22), (21), and (17), we conclude that

$$xF^{x-1}(s_t, \kappa t)F'(s_t, \kappa t)g(t) \sim x(1 - (1 - s_t)\mathrm{e}^{-r\kappa t})^x \mathrm{e}^{-r\kappa t}rb^{-1}\mathrm{e}^{-rt}$$
$$\sim rb^{-1}x\exp(-x(1 - s_t)\mathrm{e}^{-r\kappa t})\exp\left(-\frac{r}{1 - u}t\right), \qquad (23)$$

as $s_t \uparrow 1$ and $t \to \infty$.

With $s_t = e^{-vx^{u-1}} G(t)$ and

$$rt = (1 - u)(z + \ln x - \ln b), \tag{24}$$

we obtain

$$e^{-rt} = e^{-z(1-u)} x^{u-1} b^{1-u}$$

and

$$1 - s_t \sim vx^{u-1} + e^{-z(1-u)} x^{u-1} b^{-u} \quad \text{as } x \to \infty,$$

and it follows from (23) that

$$x F^{x-1}(s_t, \kappa t) F'(s_t, \kappa t) g(t) \to r \exp(-vb^u e^{-uz}) e^{-z} \exp(-e^{-z}). \tag{25}$$

Using a change of variable from (24), (19) therefore yields

$$\begin{aligned}
\mathrm{E}_x[\exp(-vx^{u-1} Z_{uT})] &= \frac{\exp(-vx^{u-1})x}{1 - u} \int_0^\infty F^{x-1}(s_t, \kappa t) F'(s_t, \kappa t) g(t) \, \mathrm{d}t \\
&\to \int_{-\infty}^\infty \exp(-vb^u e^{-uz}) \exp(-e^{-z}) e^{-z} \, \mathrm{d}z \\
&= \mathrm{E}[\exp(-vb^u e^{-u\eta})],
\end{aligned}$$

provided that dominated convergence applies.

Dominated convergence does apply here. A function dominating the convergence (25) can be obtained in the following way from the monotonicity in s of the generating function and its derivative and the inequality $\ln x \le x - 1$ (where the c_i are suitable positive constants):

$$0 \le x F^{x-1}(s_t, \kappa t) F'(s_t, \kappa t) g(t) \le c_1 x \exp(-(x - 1)(1 - F(s_t, \kappa t))) F'(1, \kappa t) e^{-rt}.$$

Since

$$\begin{aligned}
x F'(1, \kappa t) e^{-rt} &= x e^{-rt} \mathrm{E}_1(Z_{\kappa t}) \\
&= x e^{-rt/(1-u)} \\
&= b e^{-z},
\end{aligned}$$

we obtain

$$x F^{x-1}(s_t, \kappa t) F'(s_t, \kappa t) g(t) \le c_2 \exp(-c_3 x(1 - F(s_t, \kappa t))) e^{-z}.$$

Now, by (21) it holds for some $c_3 > 0$ that

$$x(1 - F(s_t, \kappa t)) \ge cx(1 - s_t) e^{-r\kappa t} \ge cx Q(t) e^{-r\kappa t} \ge c_3 e^{-z}.$$

The reduction argument works in the same way from any dimension to the one before, so we content ourselves with the step of going from dimension two to dimension one. We let

$$\boldsymbol{F}(s_1, s_2; t_1, t_2) = \mathrm{E}_1[s_1^{Z_{t_1}} s_2^{Z_{t_2}}],$$

so that, for $0 < u_1 < u_2 < 1$ and κ_i as above,

$$\begin{aligned}
\mathrm{E}_x[s_1^{Z_{u_1 T}} &s_2^{Z_{u_2 T}}] \\
&= \sum_{y_1, y_2} s_1^{y_1} s_2^{y_2} \frac{y_2}{1 - u_2} \int_0^\infty G^{y_2 - 1}(t) \, \mathrm{P}_x(Z_{\kappa_1 t} = y_1, Z_{\kappa_2 t} = y) g(t) \, \mathrm{d}t \tag{26} \\
&= \frac{s_2 x}{1 - u_2} \int_0^\infty \boldsymbol{F}^{x-1}(s_1 s_2 G(t); \kappa_1 t, \kappa_2 t) \boldsymbol{F}_2(s_1 s_2 G(t); \kappa_1 t, \kappa_2 t) g(t) \, \mathrm{d}t,
\end{aligned}$$

where F_2 denotes the partial derivative with respect to the second argument of F. Here the Markov property yields

$$
\begin{aligned}
F(s_1, s_2; \kappa_1 t, \kappa_2 t) &= \mathrm{E}_1[s_1^{Z_{\kappa_1 t}} \, \mathrm{E}_{Z_{\kappa_1 t}}[s_2^{Z_{(\kappa_2 - \kappa_1)t}}]] \\
&= \mathrm{E}_1[s_1^{Z_{\kappa_1 t}} \, \mathrm{E}_1^{Z_{\kappa_1 t}}[s_2^{Z_{(\kappa_2 - \kappa_1)t}}]] \\
&= F(s_1 F(s_2, (\kappa_2 - \kappa_1)t), \kappa_1 t),
\end{aligned}
$$

which can be differentiated with the following result:

$$
F_2(s_1, s_2 G(t); \kappa_1 t, \kappa_2 t) = F'(s_1 F(s_2, (\kappa_2 - \kappa_1)t), \kappa_1 t) s_1 F'(s_2, (\kappa_2 - \kappa_1)t).
$$

To complete the proof, insert this into (26) and repeat the various estimates in the one-dimensional part.

4. On the eve of extinction

The limiting process of Theorem 2 does not tell what happens shortly before extinction. Indeed, this dramatic phase is abruptly summarized into the jump down to zero from the limit value when $u \uparrow 1$, i.e. $be^{-\eta}$.

To obtain a more detailed picture of the last stage of the population's life, we consider Z_{T-u}, for $u > 0$, and the limit of $P_x(Z_{T-u_1} = y_1, \ldots, Z_{T-u_n} = y_n)$ as $x \to \infty$ in the spirit of [2]. In Section 5 we exhibit a *time-reversed* process $\{Y_u, \ u \geq 0\}$, which is homogeneous and Markov and approximates Z_{T-u} for $Z_0 = x$ large.

Lemma 2. *The expected occupation time at the state $j \geq 1$ satisfies*

$$
\int_0^\infty P_x(Z_t = j) \, \mathrm{d}t \to \pi_j \quad \text{as } x \to \infty, \tag{27}
$$

where all $\pi_j > 0$ and the generating function

$$
\pi(s) = \sum_{i=1}^\infty \pi_i s^i \tag{28}
$$

coincides with (4). The sequence $\{\pi_j\}_{j \geq 1}$ defines a unique (up to a multiplying factor) stationary measure for Z_t in that

$$
\pi_j = \sum_{i=1}^\infty \pi_i p_{ij}(t), \qquad t \geq 0, \ j = 1, 2, \ldots, \tag{29}
$$

in terms of $p_{ij}(t) = P_x(Z_{t+u} = j \mid Z_u = i)$.
   Furthermore,

$$
p_{11}(t) = g(t)\pi_1 = \frac{g(t)}{a p_0}. \tag{30}
$$

Proof. The key asymptotic relation is that, for $rt = t_x - \ln b + z$,

$$
\begin{aligned}
x(1 - F(s, t)) &\sim x e^{-r\pi(s)} Q(t) \\
&\to (1 - B(s)) e^{-z} \quad \text{as } x \to \infty, \tag{31}
\end{aligned}
$$

which follows from (14) and (8). From (31) we obtain

$$
\begin{aligned}
\mathrm{E}_x[s^{Z_t}] - \mathrm{P}_x(Z_t = 0) &= F^x(s, t) - F^x(0, t) \\
&\to \exp(-(1 - B(s))\mathrm{e}^{-z}) - \exp(-\mathrm{e}^{-z}) \\
&= \exp(-\mathrm{e}^{-z})(\exp(B(s)\mathrm{e}^{-z}) - 1).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\sum_{y=1}^{\infty} s^y \int_0^{\infty} \mathrm{P}_x(Z_t = y)\,\mathrm{d}t &= \int_0^{\infty} (\mathrm{E}_x[s^{Z_t}] - \mathrm{P}_x(Z_t = 0))\,\mathrm{d}t \\
&\to r^{-1} \int_{-\infty}^{\infty} \exp(-\mathrm{e}^{-z})(\exp(B(s)\mathrm{e}^{-z}) - 1)\,\mathrm{d}z \\
&= r^{-1} \int_{-\infty}^{\infty} \exp(-\mathrm{e}^{-z}) \sum_{k=1}^{\infty} \frac{B^k(s)\mathrm{e}^{-kz}}{k!}\,\mathrm{d}z \\
&= r^{-1} \sum_{k=1}^{\infty} \frac{B^k(s)}{k!} \int_0^{\infty} \mathrm{e}^{-v} v^{k-1}\,\mathrm{d}v \\
&= r^{-1} \sum_{k=1}^{\infty} \frac{B^k(s)}{k} \\
&= r^{-1} \ln\left(\frac{1}{1 - B(s)}\right) \\
&= \pi(s).
\end{aligned}
$$

To prove (30), we write (15) in the form

$$
\sum_{j=1}^{\infty} j s^{j-1} p_{1j}(t) = \frac{\phi(F(s, t))}{\phi(s)}
$$

and note that $s = 0$ yields

$$
\begin{aligned}
p_{11}(t) &= \frac{\phi(F(0, t))}{a p_0} \\
&= g(t)\pi_1,
\end{aligned}
$$

by (16) and (4).

Proposition 2. *If $m < 1$ and the $x \log x$-condition (7) holds, then for $u > 0$ we obtain*

$$
\mathrm{P}_x(Z_{T-u} = y) \to \pi_y y G^{y-1}(u) g(u) \quad \text{as } x \to \infty.
$$

Similarly, for $u_1 > u_2 > \cdots > u_n > 0$, we obtain

$$
\begin{aligned}
\mathrm{P}_x(Z_{T-u_1} = y_1, \ldots, Z_{T-u_n} = y_n) & \\
\to \pi_{y_1} \mathrm{P}_{y_1}(Z_{u_1-u_2} = y_2, \ldots, &Z_{u_1-u_n} = y_n) y_n G^{y_n-1}(u_n) g(u_n),
\end{aligned}
$$

so that the measure $\{\pi_j\}_{j \geq 1}$ acts as an entrance law.
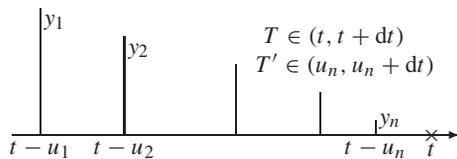
FIGURE 1.

Proof. The asserted convergence follows from Lemma 2 and the following equality (which is illustrated in Figure 1):

$$P_x(Z_{T-u_1} = y_1, \ldots, Z_{T-u_n} = y_n)$$

$$= \int_{u_1}^{\infty} P_x(Z_{t-u_1} = y_1, \ldots, Z_{t-u_n} = y_n) \, P_{y_n}(T' + t - u_n \in dt)$$

$$= y_n G^{y_n-1}(u_n) g(u_n) \int_{u_1}^{\infty} P_x(Z_{t-u_1} = y_1, \ldots, Z_{t-u_n} = y_n) \, dt$$

$$= y_n G^{y_n-1}(u_n) g(u_n) \, P_{y_1}(Z_{u_1-u_2} = y_2, \ldots, Z_{u_1-u_n} = y_n) \int_0^{\infty} P_x(Z_t = y_1) \, dt.$$

5. Looking back from extinction

The process $\{Z_t, \ t \geq 0\}$ has the following simple hold and jump description (which is illustrated in Figure 2):

- at state i it holds for an exponential time with parameter ai
- and then jumps to state $j \geq i - 1$ with probability $p_{ij} = p_{j-i+1}$.

As indicated by Figure 2 the population will consist of one single member during its last stage, before extinction. Turning our eyes back from that stage, we define a *time reversed* branching process, $\{Y_u, \ u \geq 0\}$, which will emerge as the large population limit of $\{Z_{T-u}, \ u \geq 0\}$.

Definition 1. The process $\{Y_u, \ u \geq 0\}$ starts from $Y_0 = 1$. At state i it holds for an exponential time with parameter ai, then it jumps to state j with probability

$$\tilde{p}_{ij} = \frac{j\pi_j}{i\pi_i} p_{i-j+1}, \qquad j = 1, \ldots, i + 1$$

(see Figure 3). To validate this, we need to check that the $\{\tilde{p}_{ij}\}_{j=1}^{i+1}$ are really probabilities.

Lemma 3. *For any $i \geq 1$, we obtain*

$$\sum_{j=1}^{i+1} \tilde{p}_{ij} = 1.$$

Proof. The transition probabilities of the original process $\{Z_t\}$ satisfy

$$p_{ji}(t) = ajp_{i-j+1}t + o(t) \quad \text{as } t \to 0,$$
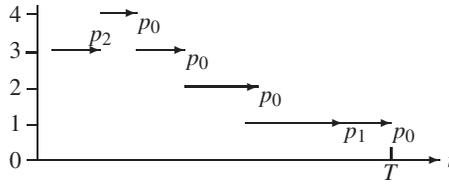
FIGURE 2.
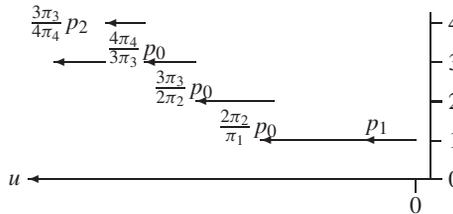
FIGURE 3.

for $i \neq j$, where $p_k = 0$ for negative k, and

$$p_{ii}(t) = 1 - ai(1 - p_1)t + o(t) \quad \text{as } t \to 0.$$

Now, we insert this into (29) to obtain

$$0 = \pi_i - \sum_{j=1}^{\infty} \pi_j p_{ji}(t)$$

$$= at\left(i\pi_i - \sum_{j=1}^{i+1} j\pi_j p_{i-j+1}\right) + o(t).$$

The result follows from the definition of $\tilde{p}_{ij}$.

Proposition 3. *The time-reversed branching process Y_u starting from $Y_0 = 1$ has transition probabilities*

$$\tilde{p}_{ij}(u) = \frac{\pi_j}{\pi_i} p_{ji}(u). \tag{32}$$

The marginal distribution of Y_u is given by

$$E[s^{Y_u}] = \frac{s\phi(G(u))}{\phi(sG(u))}. \tag{33}$$

Proof. The backward Kolmogorov equation for the original branching process

$$\frac{\mathrm{d}p_{ji}(u)}{\mathrm{d}u} = a\sum_{k=1}^{i+1} p_{jk}(u)kp_{i-k+1} - aip_{ji}(u)$$

together with (32) implies

$$\frac{\mathrm{d}\tilde{p}_{ij}(u)}{\mathrm{d}u} = ai\sum_{k=1}^{i+1} \tilde{p}_{ik}\tilde{p}_{kj}(u) - ai\tilde{p}_{ij}(u).$$

This is the forward Kolmogorov equation for the process Y_u and therefore (32) indeed gives the transition probabilities of the time-reversed process. Due to the independence between individuals, we have

$$p_{j1}(u) = jG^{j-1}(u)p_{11}(u)$$

and

$$\begin{aligned} P(Y_u = j) &= \tilde{p}_{1j}(u) \\ &= \frac{\pi_j}{\pi_1} p_{j1}(u) \\ &= \frac{\pi_j}{\pi_1} jG^{j-1}(u)p_{11}(u). \end{aligned}$$

Therefore, by (30) we obtain

$$P(Y_u = j) = jG^{j-1}(u)g(u)\pi_j \tag{34}$$

and

$$\begin{aligned} E[s^{Y_u}] &= \frac{d\pi(sG(u))}{du} \\ &= \frac{s\phi(G(u))}{\phi(sG(u))}. \end{aligned}$$

Theorem 3. *For subcritical Markov branching processes satisfying the $x \log x$ condition (7), the functional convergence $Z_{T-u} \xrightarrow{D} Y_u$, $u > 0$, as $Z_0 \to \infty$, holds. As $u \to \infty$, $e^{-ru}Y_u$ converges in distribution to an exponentially distributed random variable, with expectation b.*

Thus, looking backwards from the last survivor, population sizes constitute a time homogeneous Markov process, which in the long run grows exponentially like We^{ru}, with W exponential with mean b.

Proof of Theorem 3. For $0 < u_1 < u_2 < \cdots < u_n \le T$ (compare with the proof of Proposition 2), we have

$$P_x(Z_{T-u_1} = y_1, Z_{T-u_2} = y_2, \ldots, Z_{T-u_n} = y_n)$$
$$= y_1 G^{y_1-1}(u_1)g(u_1)p_{y_n,y_{n-1}}(u_n - u_{n-1}) \cdots p_{y_2,y_1}(u_2 - u_1) \int_0^\infty P_x(Z_t = y_n) \, dt.$$

Together with (32), (34), and (27) this yields the desired convergence, i.e.

$$\begin{aligned} P_x(Z_{T-u_1} = y_1, Z_{T-u_2} = y_2, \ldots, Z_{T-u_n} = y_n) \\ \to P(Y_{u_1} = y_1, Y_{u_2} = y_2, \ldots, Y_{u_n} = y_n). \end{aligned}$$

Now, it follows from (33) that

$$\begin{aligned} E[\exp(-\theta e^{-ru}Y_u)] &= \frac{\exp(-\theta e^{-ru})\phi(G(u))}{\phi(\exp(-\theta e^{-ru})G(u))} \\ &\sim \frac{(1-m)(1-G(u))}{(1-m)(1-\exp(-\theta e^{-ru})G(u))} \\ &\sim \frac{ce^{-ru}}{1 - \exp(-\theta e^{-ru}) + ce^{-ru}} \\ &\to \frac{c}{\theta + c} \quad \text{as } u \to \infty. \end{aligned}$$

It remains to prove tightness. We use [5, Theorem 2.1] concerning (not necessarily Markov) jump processes. This theorem and [5, Remark 4.2] suggest three conditions, (i$^\infty$), (ii$^\infty$), and (iii), that yield tightness. In our framework, the key condition (ii$^\infty$) requires that, for each $0 < a_1 < a_2 < \infty$,

$$\limsup_{x \to \infty} \sup_{a_1 \le s \le a_2} P_x(Z_{T-u} \text{ has at least two jumps in } u \in [s, s+\delta)) = o(\delta) \quad \text{as } \delta \to 0. \quad (35)$$

Condition (i$^\infty$) stipulates that for each $a, \eta > 0$ there exists a K such that

$$\sup_{x \ge 1} P_x\left(\sup_{0 \le u \le a} Z_{T-u} > K \right) \le \eta.$$

In accordance with [4, Corollary, p. 140], this condition follows from (35) and convergence of finite dimensional distributions. Finally, condition (iii) can be specified as follows. For each $\varepsilon > 0$, we have

$$\limsup_{x \to \infty} P_x\left(\sup_{0 < u \le a} |Z_{T-u} - 1| > \varepsilon \right) \to 0 \quad \text{as } a \to 0. \quad (36)$$

Now to the proof of (35). As in the deduction of Proposition 2,

$$P_x(Z_{T-u} \text{ has at least two jumps in } u \in [s, s+\delta])$$

$$= \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} \int_{s+\delta}^{\infty} P_x(Z_{t-s-\delta} = k_1, Z_{t-s} = k_2,$$

$$Z_{t-u} \text{ has at least two jumps in } u \in [s, s+\delta], T \in dt)$$

$$= \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} k_2 G^{k_2-1}(s) g(s) \int_{0}^{\infty} P_x(Z_t = k_1) \, dt$$

$$\times P_{k_1}(Z_\delta = k_2, Z_u \text{ has at least two jumps in } u \in [0, \delta]).$$

Since

$$P_{k_1}(Z_\delta = k_2, Z_u \text{ has at least two jumps in } u \in [0, \delta]) \le c_1 \sum_{k=k_1-1}^{\infty} k_1 k \delta^2 p_{k-k_1+1} p_{k_2-k+1},$$

for a suitable constant c_1, we have

$$\limsup_{x \to \infty} \sup_{a_1 \le s \le a_2} P_x(Z_{T-u} \text{ has at least two jumps in } u \in [s, s+\delta])$$

$$\le c_1 \delta^2 \sup_{a_1 \le s \le a_2} \sum_{k_1=1}^{\infty} \sum_{k=k_1-1}^{\infty} \sum_{k_2=k-1}^{\infty} \pi_{k_1} k_2 G^{k_2-1}(s) g(s) k_1 k p_{k-k_1+1} p_{k_2-k+1}.$$

If we recall that

$$\sum_{k_1=1}^{k+1} k_1 \pi_{k_1} p_{k-k_1+1} = k \pi_k,$$

this unwieldy expression simplifies in a drastic manner as follows:

$$\sum_{k_1=1}^{\infty} \sum_{k=k_1-1}^{\infty} \sum_{k_2=k-1}^{\infty} \pi_{k_1} k_2 G^{k_2-1}(s) g(s) k_1 k p_{k-k_1+1} p_{k_2-k+1}$$

$$= \sum_{k=1}^{\infty} \sum_{k_2=k-1}^{\infty} k \pi_k k_2 G^{k_2-1}(s) g(s) p_{k_2-k+1}$$

$$= \sum_{k_2=1}^{\infty} k_2 \pi_{k_2} G^{k_2-1}(s) g(s) = 1.$$

This takes us to the proof of (36). We then have to estimate

$$P_x(Z_{T-u} \text{ has at least one jump in } u \in (0, v], T \geq v)$$

$$= \sum_{k=1}^{\infty} \int_v^{\infty} P_x(Z_{t-v} = k, Z_{t-u} \text{ has at least one jump in } u \in (0, t], T \in dt)$$

$$\leq \sum_{k=2}^{\infty} \int_v^{\infty} P_x(Z_{t-v} = k, T \in dt)$$

$$+ \int_v^{\infty} P_x(Z_{t-v} = 1, Z_{t-u} \text{ has at least one jump in } u \in (0, t], T \in dt)$$

$$\leq c_1 \sum_{k=2}^{\infty} \pi_k k G^{k-1}(v) g(v)$$

$$+ c_2 P_1(Z_u \text{ has at least one jump in } u \in [0, v) \mid T = v).$$

Since $g(v) = \phi(G(t)) \to \phi(0) = ap_0 = 1/\pi_1$, as $t \downarrow 0$, we obtain

$$\sum_{k=2}^{\infty} \pi_k k G^{k-1}(v) g(v) = 1 - \pi_1 g(v) < \frac{\varepsilon}{2},$$

for small v, and it remains only to verify that

$$P_1(Z_u \text{ has at least one jump in } u \in [0, v) \mid T = v) \to 0 \quad \text{as } v \to 0.$$

Example 1. For birth and death processes we have

$$p_0 = \frac{1}{1+q}, \quad p_2 = \frac{q}{1+q}, \quad 0 < q < 1,$$

and all other reproduction probabilities vanish. Hence,

$$m = \frac{2q}{1+q}, \quad r = \frac{a(1-q)}{1+q}.$$

The function (4) can be found explicitly as follows:

$$\pi(s) = r^{-1} \ln\left(\frac{1-qs}{1-s}\right). \tag{37}$$

Due to (11) we have

$$B(s) = \frac{(1-q)s}{1-qs},$$

and the limit distribution is geometric, i.e. $b_k = q^{k-1}(1-q)$ and $b = 1/(1-q)$.

From (28) and (37) π_j can be explicitly computed as follows:

$$\pi_j = r^{-1} \frac{1-q^j}{j}.$$

Thus,

$$\tilde{p}_{i,i-1} = \frac{1-q^{i-1}}{1-q^i} p_2,$$

$$\tilde{p}_{i,i+1} = \frac{1-q^{i+1}}{1-q^i} p_0.$$

In particular $\tilde{p}_{10} = 0$ and $\tilde{p}_{12} = 1$. Notice the change upwards, which is however asymptotically negligible as $i \to \infty$,

$$\tilde{p}_{i,i-1} < p_{i-1,i},$$

$$\tilde{p}_{i,i+1} > p_{i+1,i}.$$

6. The lattice case

Somewhat strangely, in discrete time a limit distribution for the time to extinction exists only when the initial number of individuals tends to infinity along certain sequences of positive integers x_n (see [10]). The appropriate sequences have the form

$$x_n = m^{-n-z_n}, \quad z_n \to z, \quad z \in [0,1), n \to \infty. \tag{38}$$

Indeed, since $P(T_1 > n) \sim cm^n$

$$\lim_{n\to\infty} P(T_{x_n} \le n+k) = \lim_{n\to\infty} (1-cm^{n+k})^{x_n} = \exp(-cm^{k-z}),$$

for any integer k. leading to a discrete version of the extreme value distribution

$$\lim_{n\to\infty} P(T_{x_n} = n+k) = \exp(-cm^{k-z}) - \exp(-cm^{k-z-1}), \quad k = 0, \pm 1, \pm 2, \ldots,$$

which involves an extra parameter $z \in [0,1)$.

The Yaglom constant c in this is implicitly defined, in terms of the generating function $B(s) = \sum_{k=1}^{\infty} b_k s^k$ of the limit distribution (10) in the lattice case. Namely (see [3, p. 16]), with f and m as before, B is the unique solution to the equation

$$B(f(s)) = mB(s) + 1 - m$$

among probability generating functions vanishing at zero. Also, $c = 1/B'(1)$.

In the special case of a linear-fractional reproduction [3, p. 6]

$$f(s) = p_0 + \frac{(1-p_0)(1-p)s}{1-ps}, \tag{39}$$

both the limit distribution (10) and the constant c can be expressed explicitly in terms of the parameters $p_0 \in (0, 1)$ and $p \in (0, 1)$. This is the case of geometric reproduction modified at zero, with offspring probabilities p_k defined by

$$p_k = (1 - p_0)(1 - p)p^{k-1}, \qquad k = 1, 2, \ldots.$$

Since $m = (1 - p_0)/(1 - p)$, $p < p_0$ in the subcritical case. The superposition of two linear-fractional functions is again linear fractional, as is well known, implying that the solution $B(s)$ of (11) remains linear fractional. It is straightforward to verify that the solution is $B(s) = (p_0 - p)s/(p_0 - ps)$ corresponding to the geometric distribution

$$b_k = \left(1 - \frac{p}{p_0}\right)\left(\frac{p}{p_0}\right)^{k-1}$$

having $c = 1 - p/p_0$.

However, the normed process $\{x^{u-1}Z_{uT}\}$ does not converge as $Z_0 \to \infty$, not even along subsequences like (38), since

$$x_n \, \mathrm{E}_{x_n}[Z_{u(n+k)}] = m^{-u(n+z_n)}m^{[u(n+k)]} \sim m^{-uz}m^{[u(n+k)]-un} \quad \text{as } n \to \infty,$$

for any fixed u and k, the exponent $([u(n+k)] - un)$ is oscillating without convergence to any limit. So there is no lattice counterpart of Theorem 2.

7. A lattice version of Theorem 3

Alsmeyer and Rösler [1], [2] have a process Y_k approximating a subcritical Galton–Watson process Z_{T-k}, k seasons before extinction, $k \geq 0$, as a possibly random initial number $Z_0 = x$ tends to infinity. The time reversal $\{Y_k\}$ is a time-homogeneous Markov chain with $Y_0 = 0$ whose transition probabilities $Q_{ij} = \mathrm{P}(Y_{k+1} = j \mid Y_k = i)$ are related to the forward transition probabilities $P_{ij} = \mathrm{P}_j(Z_1 = i)$ through

$$Q_{ij} = \begin{cases} 0 & \text{if } i = j = 0, \\ \eta_j p_0^j & \text{if } i = 0, \, j \geq 1, \\ \dfrac{\eta_j}{\eta_i}P_{ji} & \text{if } i \geq 1, \, j \geq 1 \end{cases}$$

by means of a so-called quasi-invariant measure η_i of the branching process,

$$\sum_{i=1}^{\infty} \eta_i P_{ij} = \eta_j, \qquad j \geq 1.$$

It is straightforward to verify that the n-step backward and forward transition probabilities relate as

$$Q_{ij}^{(n)} = \frac{\eta_j}{\eta_i}P_{ji}^{(n)}, \qquad i \geq 1, \, j \geq 1.$$

An important family of quasi-invariant measures $\eta_i(z)$, $z \in [0, 1)$, is described by the generating function

$$\sum_{i=1}^{\infty} \eta_i(z)s^i = \sum_{k=-\infty}^{\infty} (\exp(-cm^{k-z}(1 - B(s))) - \exp(-cm^{k-z})).$$

It gives the time-reversal approximation under the initial condition (38). Every quasi-invariant measure η_i has a representation $\eta_i = \int_0^1 \eta_i(z)\lambda(\mathrm{d}z)$, in terms of this family and a finite Borel measure $\lambda(\mathrm{d}z)$ on $[0, 1)$, the latter reflecting randomness in the initial number of individuals. When nothing is known about the initial number, except that it is large, it may be natural to use the η_i with the uniform distribution $\lambda(\mathrm{d}z) = \mathrm{d}z$. In such a case we have

$$\sum_{i=1}^{\infty} \eta_i s^i = \sum_{k=-\infty}^{\infty} \int_0^1 (\exp(-cm^{k-z}(1 - B(s))) - \exp(-cm^{k-z}))\,\mathrm{d}z$$

$$= \frac{1}{\ln m^{-1}} \int_0^{\infty} (\exp(-c(1 - B(s))x) - \mathrm{e}^{-cx})x^{-1}\,\mathrm{d}x$$

$$= \frac{1}{\ln m^{-1}} \int_0^{\infty} \int_{c(1-B(s))}^c \mathrm{e}^{-ux}\,\mathrm{d}u\,\mathrm{d}x$$

$$= \frac{\ln(1 - B(s))}{\ln m}.$$

For linear fractional processes we can be much more specific about details. Recall that $B(s) = (p_0 - p)s/(p_0 - ps)$, $m = (1 - p_0)/(1 - p)$, and $c = 1 - p/p_0$, so that

$$\sum_{k=-\infty}^{\infty} (\exp(-cm^{k-z}(1 - B(s))) - \exp(-cm^{k-z}))$$

$$= \sum_{k=-\infty}^{\infty} \exp(-cm^{k-z}) \sum_{j=1}^{\infty} \frac{(cm^{k-z} B(s))^j}{j!}$$

$$= \sum_{k=-\infty}^{\infty} \exp(-cm^{k-z}) \sum_{j=1}^{\infty} \frac{c^{2j} m^{(k-z)j} s^j}{j!} \sum_{i=0}^{\infty} \frac{j(j+1)\cdots(j+i-1)}{i!}((1-c)s)^i$$

$$= \sum_{l=1}^{\infty} s^l \sum_{j=1}^{l} \binom{l-1}{j-1} \frac{(1-c)^{l-j}c^{2j}}{j!} \sum_{k=-\infty}^{\infty} \exp(-cm^{k-z})m^{(k-z)j}.$$

Hence,

$$\eta_i(z) = \sum_{k=-\infty}^{\infty} \exp(-cm^{k-z}) \sum_{j=1}^{i} \binom{i-1}{j-1} \frac{(1-c)^{i-j}(c^2 m^{k-z})^j}{j!}.$$

This generalizes the example in [3, pp. 71–72], with $p = 0$, which describes the moribund process where the number of offspring cannot exceed one.

In the linear fractional case, uniform prior reversal gives

$$\sum_{i=1}^{\infty} \eta_i s^i = \frac{\ln(1 - B(s))}{\ln m}$$

$$= \frac{\ln(1 - s)}{\ln m} - \frac{\ln(1 - (1 - c)s)}{\ln m}$$

$$= \frac{1}{\ln m^{-1}} \sum_{i=1}^{\infty} \frac{1 - \rho^{-i}}{i} s^i.$$

Here

$$\rho = \frac{1}{1-c} = \frac{p_0}{p}$$

is a fixed point of the generating function (39) in the subcritical case $r > 1$. The corresponding time-reversal has transition probabilities

$$Q_{ij} = \begin{cases} 0 & \text{if } i = j = 0, \\ \dfrac{(1 - \rho^{-j})p_0^j}{j \ln m^{-1}} & \text{if } i = 0, j \geq 1 \end{cases}$$

and if $i, j \geq 1$

$$Q_{ij} = \frac{i(1 - \rho^{-j})}{j(1 - \rho^{-i})} \sum_{l=1}^{i} \binom{i-1}{l-1}\binom{j}{l} p^{i-l}(1-p)^l p_0^{j-l}(1-p_0)^l.$$

In the last formula we used an expression for $P_{ij}^{(n)}$ derived in [9], which also entails a similar explicit expression for the n-step transition probabilities.

The generating function of the uniform prior reversed process has the simple expression

$$E[s^{Y_{k+n}} \mid Y_k = i] = \frac{f_n^i(\rho s) - f_n^i(s)}{r^i - 1}, \qquad i \geq 1, \tag{40}$$

implying (recall that $Y_0 = 0$)

$$E(s^{Y_n}) = \sum_{i=1}^{\infty} \frac{(f_{n-1}^i(\rho s) - f_{n-1}^i(s))p_0^i}{\rho^i i \ln m^{-1}}$$

$$= \frac{1}{\ln m} \ln \frac{1 - p f_{n-1}(\rho s)}{1 - p f_{n-1}(s)}.$$

Thus, the time-reversed chain grows exponentially

$$E[\exp(-\lambda m^n Y_n)] = \frac{1}{\ln m} \ln \frac{1 - p\rho \hat{f}_{n-1}(\exp(-\lambda m^n))}{1 - p f_{n-1}(\exp(-\lambda m^n))}$$

$$\to \frac{1}{\ln m} \ln \frac{1 - p_0 E(\exp(-\lambda W))}{1 - p} \quad \text{as } n \to \infty.$$

The conclusion builds upon the fact that the dual reproduction with generating function $\hat{f}(s) = f(\rho s)/\rho$ yields a supercritical Galton–Watson process $\{\hat{Z}_n\}$ with mean reproduction $\hat{m} = 1/m$ and $\hat{Z}_n/\hat{m}^n \to W$.

We conclude the paper by showing that (40) holds. From [9], we can conclude

$$\sum_{i=1}^{\infty} s^i P_{ij}^{(n)} = s\rho^{1-j} f_n'(\rho s) f_n^{j-1}(\rho s), \qquad j \geq 1.$$

For $j \geq 1$, we therefore obtain

$$\frac{\mathrm{d}}{\mathrm{d}s} \sum_{i=1}^{\infty} \frac{s^i}{i} P_{ij}^{(n)} = \rho^{1-j} f_n'(\rho s) f_n^{j-1}(\rho s)$$

$$= j^{-1}\rho^{-j} \frac{\mathrm{d}}{\mathrm{d}s}(f_n^j(\rho s)),$$

so that $\sum_{i=1}^{\infty} (s^i / i) P_{ij}^{(n)} = j^{-1} \rho^{-j} f_n^j(\rho s)$. Therefore, neglecting the constant $1/\ln m^{-1}$, we obtain

$$\sum_{i=1}^{\infty} s^i \eta_i P_{ij}^{(n)} = \sum_{i=1}^{\infty} \frac{s^i - (s/\rho)^i}{i} P_{ij}^{(n)}$$
$$= j^{-1} \rho^{-j} (f_n^j(\rho s) - f_n^j(s)).$$

Thus,

$$\sum_{j=1}^{\infty} \frac{s^j \eta_j}{\eta_i} P_{ji}^{(n)} = \frac{f_n^i(\rho s) - f_n^i(s)}{\rho^i - 1}, \qquad i \geq 1,$$

which is equivalent to (40).

Acknowledgements

References

[1] ALSMEYER, G. AND RÖSLER, U. (2002). Asexual versus promiscuous bisexual Galton–Watson processes: the extinction probability ratio. *Ann. Appl. Prob.* **12,** 125–142.

[2] ALSMEYER, G. AND RÖSLER, U. (2006). The Martin entrance boundary of the Galton–Watson process. *Ann. Inst. H. Poincaré Prob. Statist.* **42,** 591–606.

[3] ATHREYA, K. AND NEY, P. (1972). *Branching Processes*. Springer, Berlin.

[4] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.

[5] GUT, A. AND JANSON, S. (2001). Tightness and weak convergence for jump processes. *Statist. Prob. Lett.* **52,** 101–107.

[6] HARRIS, T. E. (1963). *The Theory of Branching Processes*. Springer, Berlin.

[7] JAGERS, P. (1975). *Branching Processes with Biological Applications*. John Wiley, Chichester.

[8] JAGERS, P., KLEBANER, F. C. AND SAGITOV, S. (2007). On the path to extinction. *Proc. Nat. Acad. Sci. USA* **104,** 6107–6111.

[9] KLEBANER, F. C., RÖSLER, U. AND SAGITOV, S. (2006). Transformations of Galton–Watson processes and linear fractional reproduction. Submitted.

[10] PAKES, A. G. (1989). Asymptotic results for the extinction time of Markov branching processes allowing emigration. I. Random walk decrements. *Adv. Appl. Prob.* **21,** 243–269.

[11] PAKES, A. G. (1989). On the asymptotic behaviour of the extinction time of the simple branching process. *Adv. Appl. Prob.* **21,** 470–472.

[12] SEVAST´YANOV, B. A. (1971). *Vetvyashchiesya Protsessy*. Nauka, Moscow.