The Annals of Probability
1997, Vol. 25, No. 1, 337–360

SENETA–HEYDE NORMING IN THE BRANCHING RANDOM WALK

BY J. D. BIGGINS[1] AND A. E. KYPRIANOU[2]

University of Sheffield

In the discrete-time supercritical branching random walk, there is a Kesten–Stigum type result for the martingales formed by the Laplace transform of the nth generation positions. Roughly, this says that for suitable values of the argument of the Laplace transform the martingales converge in mean provided an "$X \log X$" condition holds. Here it is established that when this moment condition fails, so that the martingale converges to zero, it is possible to find a (Seneta–Heyde) renormalization of the martingale that converges (in probability) to a finite nonzero limit when the process survives. As part of the proof, a Seneta–Heyde renormalization of the general (Crump–Mode–Jagers) branching process is obtained; in this case the convergence holds almost surely. The results rely heavily on a detailed study of the functional equation that the Laplace transform of the limit must satisfy.

1. Introduction. This paper considers the usual supercritical branching random walk. Thus, ignoring the spatial element, the population grows like a supercritical Galton–Watson process. The initial ancestor is at the origin of the real line, $\mathbb{R}$, and the positions of her children are given by a point process Z. Each of these children has children in the same way, in that the positions of each family relative to the parent are given by an independent copy of Z, and so on. Individuals are labelled by their line of descent, so if $u = i_1 \cdots i_n$ then u is the i_nth child of the i_{n-1}th child of $\dots$ the i_1th child of the initial ancestor. Now let $|u|$ be the generation in which u is born and write $v < u$ if v is a strict ancestor of u. Let T be the set of all people ever born, which can be viewed as a tree, with the population members as the nodes. Sums, products, sets and so on, defined with an index ranging over individual's labels will be restricted to those actually born, without this being made explicit. So $\{u: |u| = 1\}$ is the set of children born to the initial ancestor and is more accurately written as $\{u: |u| = 1, u \in T\}$ or $\{u: u = 1, 2, \dots, Z(\mathbb{R})\}$.

Denote the sigma-field generated by the process up to the nth generation by F^n. Let $Z^{(n)}$ be the point process formed by the nth generation, with

Received October 1995; revised June 1996.

[1] Some of this work was carried out while visiting the Institute of Mathematics and Its Applications, University of Minnesota, and the Mittag-Leffler Institute, Stockholm. I am grateful to both for their support.

[2] Supported by an EPSRC studentship.

AMS 1991 *subject classification*. Primary 60J80.

Key words and phrases. Martingales, functional equations, Seneta–Heyde norming, branching random walk.

points $\{z_u: |u| = n\}$; then, by definition, for any set A

$$Z^{(n+1)}(A) = \sum_{|u|=n} Z_u(A - z_u), \tag{1.1}$$

where, given F^n, Z_u are independent copies of Z. In general, a subscript u will be used to indicate quantities associated with u.

Suppose Z has intensity measure μ with Laplace–Stieltjes transform

$$m(\theta) = \int \exp(-\theta x)\mu(dx) = E\left[\int \exp(-\theta x) Z(dx)\right] = E\left[\sum_{|u|=1} \exp(-\theta z_u)\right].$$

The following assumptions will be in force except when explicitly discarded.

A1. $\mathrm{int}\{\phi: m(\phi) < \infty\}$ is non-empty.
A2. $m := m(0) > 1$.
A3. $P(Z(\mathbb{R}) = \infty) = 0$).
A4. $\theta \in \mathrm{int}\{\phi: m(\phi) < \infty\}$.
A5. $\theta > 0$.

The condition A1 (comfortably) ensures that convolutions of μ are well defined, and, in conjunction with A4, allows calculations involving the Laplace transform in a neighborhood of θ. The condition A2 is simply that the process is supercritical, so it survives for all generations with positive probability. A3 insists that family sizes are finite, which implies that every generation is finite (i.e., $P(Z^{(n)}(\mathbb{R}) = \infty) = 0$ for all n). Finally, A5 is for convenience; cases with $\theta < 0$ are transformed to satisfy A5 by reflection of peoples' positions through the origin.

It is well known and easily shown that

$$W^{(n)}(\theta) = m(\theta)^{-n} \int \exp(-\theta x) Z^{(n)}(dx) = \sum_{|u|=n} \frac{\exp(-\theta z_u)}{m^n(\theta)}$$

is a martingale with respect to the sigma-fields $\{F^n\}$. This martingale is positive and so has an almost sure limit $W(\theta)$ which, by Fatou's lemma, satisfies $E[W(\theta)] \le 1$. When $\theta = 0$, the study of this martingale goes back a long way, with the definitive result on the conditions needed for its L_1 convergence being given by the Kesten–Stigum theorem.

The L_1 convergence of the martingale $W^{(n)}(\theta)$, or variants of it, has also been considered by several authors; see in particular Kingman (1975), Kahane and Peyrière (1976), Biggins (1977a), Neveu (1988), Lyons (1996), Liu (1997), Waymire and Williams (1994, 1995, 1996). The next result is extracted from Biggins (1977a). To state it, a little further notation is useful. For a fixed θ, with $m(\theta)$ finite, let $X(\theta)$ be a random variable such that

$$\begin{gathered}(X(\theta) - \log m(\theta))/\theta \text{ has the distribution} \\ \exp(-z\theta - \log m(\theta))\mu(dz).\end{gathered} \tag{1.2}$$

The condition that $EX(\theta) > 0$ will appear quite often. It will be assumed throughout that $EX(\theta)$ is defined; this is implied by A4 when the assumptions are in force, and then

$$EX(\theta) = \log m(\theta) - \theta m'(\theta)/m(\theta).$$

THEOREM 1.1. *The martingale $W^{(n)}(\theta)$ converges in L_1, so that $EW(\theta) = 1$, if and only if*

$$EX(\theta) > 0 \tag{1.3}$$

and

$$E\left[W^{(1)}(\theta)\log^+ W^{(1)}(\theta)\right] < \infty, \tag{1.4}$$

and $EW(\theta) = 0$ otherwise.

In fact, under A1,

$$\text{int}\{\phi: m(\phi) < \infty\} \cap \{\phi: \log m(\phi) - \phi m'(\phi)/m(\phi) > 0\}$$

is an open interval, so that, (1.3) with A4 simply restricts θ to lie in this interval.

When $\theta = 0$ only condition (1.4) matters (the assumptions are also not needed; $1 < m < \infty$ suffices) and then Theorem 1.1 is the Kesten–Stigum theorem. For that case it was established by Seneta (1968) that it is always possible to find a sequence of constants $\{c_n\}$ such that $W^{(n)}(0)/c_n$ has a finite nonzero limit in distribution; Heyde (1970) strengthened this to almost sure convergence. Thus $\{c_n\}$ provides the Seneta–Heyde norming for the martingale $W^{(n)}(0)$. Of course when (1.4) holds, $\{c_n\}$ must converge to a finite nonzero constant. [It is more usual to rescale the generation size; that is, $Z^{(n)}(\mathbb{R})$ $(= m^n W^{(n)}(0))$, defining the Seneta–Heyde norming to be, in the notation just introduced, $m^n c_n$, but it will be more convenient here to think in terms of rescaling the martingale.] The main objective of this paper is to find a Seneta–Heyde norming for the martingale $W^{(n)}(\theta)$, that is, to prove the following theorem.

THEOREM 1.2. *When $EX(\theta) > 0$ there exists a sequence of constants $\{c_n\}$ such that*

$$\frac{W^{(n)}(\theta)}{c_n} \to \Delta \quad \textit{in probability},$$

where Δ is a finite random variable which is strictly positive when the process survives.

In general $\{c_n\}$ and Δ both depend on θ. Notice that the theorem only claims convergence in probability, rather than almost surely, for the renormalized martingale. The proof suggests that there may be cases where this is the best that can be done without further conditions, but more work on this aspect is required. Almost sure convergence of Seneta–Heyde renormalizations for related martingales will result from the method here. In particular, a new proof of the renormalization of the general (C-M-J) branching process, obtained by Cohn (1985), will be given. In fact the result obtained here (Corollary 7.2) makes weaker assumptions than were used by Cohn.

To simplify the notation, let

$$y_u(\theta) = \frac{\exp(-\theta z_u)}{m(\theta)^{|u|}}.$$

Furthermore, since for many of the arguments θ is fixed, it will be omitted whenever possible. Thus, following this convention, $W^{(n)} = \sum_{|u|=n} y_u$.

Let us note straightaway that, for fixed u,

$$\left\{\frac{y_{ui}}{y_u} : i = 1, 2, \dots, Z_u(\mathbb{R})\right\} = \left\{\frac{\exp(-\theta(z_{ui} - z_u))}{m(\theta)} : i = 1, 2, \dots, Z_u(\mathbb{R})\right\},$$

which by definition, given $F^{|u|}$, is a copy of $\{y_v : v = 1, 2, \dots, Z(\mathbb{R})\}$.

By looking at the branching processes stemming from each first generation person it is easily seen that

$$W^{(n+1)} = \sum_{|u|=1} \frac{\exp(-\theta z_u)}{m(\theta)} W_u^{(n)} = \sum_{|u|=1} y_u W_u^{(n)},$$

where, given F^1, $W_u^{(n)}$ are independent copies of $W^{(n)}$. Therefore, provided the constants $\{c_n\}$ satisfy $c_{n+1}/c_n \to 1$, the random variable Δ arising in Theorem 1.2 should satisfy the distributional equation

$$\Delta = \sum_{|u|=1} y_u \Delta_u,$$

where, given $\{y_u : |u| = 1\}$, Δ_u are independent copies of Δ. Expressing this distributional equation in terms of the Laplace transform of Δ it becomes

$$\Psi(x) = E\left[\prod_{|u|=1} \Psi(xy_u)\right]. \tag{1.5}$$

This and similar equations have been much studied; see for example Kahane and Peyrière (1976), Biggins (1977a), Durrett and Liggett (1983), Pakes (1992), Liu (1996).

The following results on the functional equation, which are central to the proof of the main result and of independent interest, will be established. In them, for simplicity, attention will be confined to solutions of (1.5) that lie in the set of Laplace transforms of nonnegative variables. There are interesting problems, which we hope to consider elsewhere, associated with the possibility of allowing the solution to lie in some larger set. A nontrivial solution to (1.5) is one that is the Laplace transform of a finite nonnegative variable that is not degenerate at zero.

THEOREM 1.3. *When $EX(\theta) > 0$, the functional equation* (1.5) *has a nontrivial solution.*

THEOREM 1.4. *When* $m(\theta) < \infty$ *and* $EX(\theta) > 0$, *any nontrivial solution* Ψ *to the functional equation* (1.5) *is such that* $L(x)[:= x^{-1}(1-\Psi(x))]$ *is slowly varying as* $x \downarrow 0$. (*A*1–*A*5 *are not needed*.)

THEOREM 1.5. *When* $EX(\theta) > 0$, *the nontrivial solution to the functional equation* (1.5) *is unique* (*up to a multiplicative constant in the argument*).

Existence follows from Theorem 1 of Liu's (1996) extension of the work of Durrett and Liggett (1983), on a rather more general functional equation, but will also be a by-product of results proved here; see Section 2. Notice that, in the framework adopted here,

$$E\sum_{|u|=1} y_u = E\sum_{|u|=1}\frac{\exp(-\theta z_u)}{m(\theta)} = 1,$$

so the function ρ in Liu (1996) is given by $\rho(x) = m(\theta x)/m(\theta)^x$ and the condition on ρ in Liu's Theorem 1 is automatically satisfied. Theorem 2 of Liu (1996) gives results on slow variation, but when specialized to this case it requires the extra condition that $m(0) < \infty$. [For the translation, note that $\log\rho$ there is convex (in x) with derivative at $x = 1$ given by $\theta m'(\theta)/m(\theta) - \log m(\theta)$.] The results in Liu (1996) on uniqueness within certain classes did not go far enough for our purposes, but analysis of certain multiplicative martingales eventually establishes Theorem 1.5.

In the Galton–Watson case, the functional equation (1.5) becomes the Poincaré functional equation, $\Psi(x) = f(\Psi(x/m))$ with f the probability generating function of the family size, the study of which goes back to the last century. For branching Brownian motion, the analogue of the functional equation is the Kolmogorov–Petrovski–Piscounov (or Fisher) equation. It was Neveu's (1988) use of solutions to the KPP equation to study branching Brownian motion and in particular the multiplicative martingales used in his study, that was the original inspiration for this study.

The proof of Theorem 1.2 is based on a method pioneered by Cohn in a series of papers; see Cohn (1985), for example. The idea is to find a normalization which prevents any limit being degenerate at zero or infinity, take a subsequence along which convergence in distribution holds, show that the Laplace transform of the limit of this subsequence satisfies the functional equation, use properties of the solution to show that the convergence along this subsequence can be strengthened to convergence in probability (Cohn usually phrases this in terms of a law of large numbers) and, finally, use uniqueness of the solution to show that convergence must hold along the whole sequence.

A continuous-time Markov version of the process, in which individuals move during their lifetime according to an independent increment process, is described in the final section of Biggins (1992). It is fairly straightforward to establish the analogue of Theorem 1.2 for such a process using the results given here, essentially by a discrete skeleton argument. The details of the argument can be found in Biggins and Kyprianou (1996), along with a discussion of the strategy of the proof given here.

2. A law of large numbers. The first ingredient in the proof of the existence of Seneta–Heyde norming constants for the martingales $W^{(n)}$ is a law of large numbers. An immediate consequence will be that the functional equation (1.5) does have nontrivial solutions. Before giving the law of large numbers, three lemmas are needed; the second of these will figure at several other places in the discussion too.

LEMMA 2.1. *Suppose* $\{c_i\}$ *is a sequence of nonnegative constants satisfying* $\sum_i c_i = 1$, *with* $a = \max c_i$. *Suppose* $\{Y_i\}$ *are independent identically distributed copies of a nonnegative random variable* Y *with* $E|Y| < \infty$ *and* $EY = 0$. *Then, for* $\varepsilon < 1/2$,

$$P\left(\left|\sum_i c_i Y_i\right| > \varepsilon\right) \le \frac{2}{\varepsilon^2}\left(\int_0^{1/a} atP(|Y| > t)\, dt + \int_{1/a}^{\infty} P(|Y| > t)\, dt\right).$$

This result is a special case of Lemma 2.2 in Kurtz (1972); see in particular the Remark at the end of Section 2 of that paper.

Let

$$l^{(n)} = \sup_{|u|=n} y_u$$

and let

$$L^{(n)} = \sup_{|u|=n} \frac{y_u}{\sum_{|v|=n} y_v} = \frac{l^{(n)}}{W^{(n)}}.$$

Both are taken to be zero if the process is extinct by the nth generation.

LEMMA 2.2. (i) *When* $EX \neq 0$, $(l^{(n)})^{1/n}$ *converges as* $n \to \infty$ *to a limit that is strictly less than* 1.

(ii) $l^{(n)} \to 0$.

PROOF. Let B_n be the position of the left-most nth generation person. By the Corollary to Theorem 2 in Biggins (1977b), B_n/n converges almost surely on the survival set to a constant γ, given by

$$\gamma = \inf_a \left\{a\colon \inf_\phi m(\phi)\exp(\phi a) > 1\right\};$$

consequently, $\exp(\theta\gamma)m(\theta) \ge 1$. Hence

$$\left(\sup_{|u|=n} y_u\right)^{1/n} = \left(\sup_{|u|=n} \frac{\exp(-\theta z_u)}{m(\theta)^n}\right)^{1/n}$$
$$= \left(\frac{\exp(-\theta B_n/n)}{m(\theta)}\right) \to \left(\frac{1}{\exp(\theta\gamma)\, m(\theta)}\right).$$

If $\exp(\theta\gamma)m(\theta) = 1$ then $\exp(\phi\gamma)m(\phi)$ attains its minimum when $\phi = \theta$ and calculus (justified by A4) shows that $EX = 0$, a contradiction; hence $\exp(\theta\gamma)m(\theta) > 1$, proving (i). Note that

$$l^{(n)} \le \sum_{|u|=n} y_u = W^{(n)} \to 0$$

when $EX = 0$ by Theorem 1.1. Combining this with (i) proves (ii). □

It is possible to prove that $l^{(n)} \to 0$ without A1–A5, but the result above will suffice for this study.

LEMMA 2.3. *When $EX > 0$, $(L^{(n)})^{1/n}$ converges as $n \to \infty$ when the process survives, almost surely, to a limit that is strictly less than* 1; *thus $L^{(n)} \to 0$.*

PROOF. The numerator in $L^{(n)}$ is dealt with by Lemma 2.2, so it is enough to discuss the denominator. As part of their Theorem 1, Chauvin and Rouault (1996) show that, $(W^{(n)})^{1/n} \to 1$ on the survival set, almost surely, when $\log m(\theta) - \theta m'(\theta)/m(\theta) > 0$ [which is equivalent to $EX > 0$ under A4, as has already been observed). Chauvin and Rouault assume throughout that $m(\phi) < \infty$ for all ϕ, but this condition is not needed in their proof of the result used here. □

THEOREM 2.4. *When $EX > 0$,*

$$\frac{W^{(n+1)}}{W^{(n)}} \to 1 \quad \textit{in probability}, \tag{2.1}$$

on the survival set of the process.

PROOF. Let S be the survival set and let $S_n = \{W^{(n)} > 0\}$ so that $S_n \downarrow S$. Then

$$E\left[I\{S\}I\left\{\left|\frac{W^{(n+1)}}{W^{(n)}} - 1\right| > \varepsilon\right\}\right] \le E\left[E\left[I\{S_n\}I\left\{\left|\frac{W^{(n+1)}}{W^{(n)}} - 1\right| > \varepsilon\right\}\middle| F^n\right]\right],$$

so it will suffice to show that the right-hand side here converges to zero.

It is easy to see that on S_n,

$$\frac{W^{(n+1)}}{W^{(n)}} - 1 = \sum_{|u|=n}\left(\frac{y_u}{\sum_{|v|=n} y_v}\right)\left(W_u^{(1)} - 1\right),$$

where $W_u^{(1)}$ are independent copies of $W^{(1)}$, given F^n.

Let $G(t) = P(W^{(1)} + 1 > t)$. Using Lemma 2.1 and $|W^{(1)} - 1| \le W^{(1)} + 1$ it follows that, on S_n,

$$P\left[\left|\frac{W^{(n+1)}}{W^{(n)}} - 1\right| > \varepsilon \middle| F^n\right] \le \frac{2}{\varepsilon^2}\left(\int_0^{1/L^{(n)}} L^{(n)} t G(t)\, dt + \int_{1/L^{(n)}}^{\infty} G(t)\, dt\right).$$

which, by Lemma 2.3 and dominated convergence, converges to zero almost surely as $n \to \infty$. The left-hand side here is also bounded by 1, so taking expectations and using dominated convergence again gives the required result. □

PROOF OF THEOREM 1.3. Let the Laplace transform of $W^{(n)}$ be $\Omega_n(x)$. Take c_n to be such that $\Omega_n(1/c_n) = \kappa$, where κ is fixed to be greater than the extinction probability but less than one. Choose a subsequence such that $W^{(n)}/c_n$ converges in distribution, with the transform of the limit being Ψ. Note that

$$W^{(n+1)} = \sum_{|u|=1} y_u W_u^{(n)} = \sum_{|u|=1} y_u W_u^{(n+1)} \frac{W_u^{(n)}}{W_u^{(n+1)}},$$

so, dividing through by c_{n+1}, taking Laplace transforms, letting n go to infinity through the selected subsequence, and using A3 and the law of large numbers proved in Theorem 2.4, it follows that Ψ satisfies the functional equation. Furthermore, $\Psi(0)$ must satisfy $\Psi(0) = f(\Psi(0))$ where f is the generating function of the family size and $\Psi(0) \geq \Psi(1) = \kappa$, which by arrangement exceed the extinction probability. Hence $\Psi(0) = 1$ so the limit along any subsequence must be proper, and it cannot be degenerate at zero because $\Psi(0) > \Psi(1)$. □

This shows that Theorem 1.3 holds whenever a result of the form given in Theorem 2.4 is available.

3. The functional equation and multiplicative martingales. The next theorem shows that solutions to the functional equation lead immediately to martingales. Neveu (1988) calls these martingales, which are formed by taking suitable products, "multiplicative martingales" and calls the $W^{(n)}$ "additive martingales" because they involve summing terms. (The empty product is 1.)

THEOREM 3.1. *If Ψ is a solution to the functional equation* (1.5) *then, for each $x > 0$,*

$$M^{(n)}(x) := \prod_{|u|=n} \Psi(xy_u) = \prod_{|u|=n} \Psi\left(x \frac{\exp(-\theta z_u)}{m(\theta)^n} \right)$$

is a martingale with respect to $\{F^n\}$.

PROOF. Splitting the $(n+1)$th generation into families yields

$$\begin{aligned} E\left[M^{(n+1)}(x) | F^n \right] &= E\left[\prod_{|u|=n} \prod_i \Psi(xy_{ui}) \middle| F^n \right] \\ &= \prod_{|u|=n} E\left[\prod_i \Psi(xy_u(y_{ui}/y_u)) \middle| F^n \right] \\ &= \prod_{|u|=n} \Psi(xy_u) = M^{(n)}(x), \end{aligned}$$

where the third equality uses the functional equation and the fact that, given F^n, for each nth generation u, $\{y_{ui}/y_u: i = 1, 2, \dots, Z_u(\mathbb{R})\}$ is an independent copy of $\{y_v: |v| = 1\}$. □

The martingale $\{M^{(n)}(x)\}$ is bounded and so has an almost sure (and L_1) limit $M(x)$; thus the following corollary is immediate.

COROLLARY 3.2. *For any* n,

$$\Psi(x) = E\left[\prod_{|u|=n} \Psi(xy_u)\right]$$

and $EM(x) = \Psi(x)$.

4. *L* is slowly varying. Both for the argument in this section and for later ones, we need to be able to estimate expressions like

$$E\left[\sum_{|u|=n} I\{y_u \le \beta\} y_u\right].$$

This is facilitated by expressing them in terms of the random variable X, defined at (1.2). Let S_n be the sum of n independent copies of X.

LEMMA 4.1 (A1–A5 are not required). *Suppose* $\theta > 0$ *and* $m(\theta) < \infty$.

(i) *When* $\theta \in \operatorname{int}\{\phi: m(\phi) < \infty\}$, *the Laplace transform of* X *is*

$$Ee^{-\phi X} = \frac{m((1+\phi)\theta)}{m(\theta)^{1+\phi}}.$$

(ii)
$$E\left[\sum_{|u|=n} I\{y_u \le \beta\} y_u\right] = P[S_n \ge -\log \beta].$$

(iii) *Given sets* $B_1, \dots, B_n$,

$$E\left[\sum_{|u|=n} I\left\{-\log y_v \in B_{|v|}, v \le u, |v| = 1, \dots, n\right\} y_u\right]$$
$$= P\left[S_j \in B_j, j = 1, \dots, n\right].$$

PROOF. The first two parts are straightforward calculations; similar results were used in Section 2 of Biggins (1977a). The final part is proved by induction on n. The sum is split according to the first generation, expectations are taken conditional on F^1 with the induction hypothesis and the branching property being used to compute the terms, then the overall expectation is computed and seen to be of the required form. □

PROOF OF THEOREM 1.4. Suppose, temporarily, that $\{|u| = n\}$ is ordered (arbitrarily) by $\prec$, then, from Corollary 3.2 and the definition of L,

$$L(x) = E\left[\frac{1 - \prod_{|u|=n}\Psi(xy_u)}{x}\right] = E\left[\sum_{|u|=n} \frac{1 - \Psi(xy_u)}{x} \prod_{v \prec u} \Psi(xy_v)\right],$$

where the second equality arises from a telescoping sum, giving the identity

$$1 = E\left[\sum_{|u|=n} y_u \frac{L(xy_u)}{L(x)} \prod_{v \prec u} \Psi(xy_v)\right]. \tag{4.1}$$

Because $L(x)$ is a Laplace transform of a positive measure [see Feller (1971), XIII.2] it is monotone decreasing as x increases. Suppose that $L(x)$ is not slowly varying, so there exists a constant $\beta < 1$ and a sequence $\{x_k\}$ with $x_k \downarrow 0$ such that $L(x_k \beta)/L(x_k) \to \eta > 1$; then, by monotonicity,

$$\liminf_{k\to\infty} \frac{L(x_k y)}{L(x_k)} \geq \eta \quad \text{for all } y \leq \beta.$$

Letting $x \downarrow 0$ through $\{x_k\}$ in (4.1) and using Fatou's lemma gives

$$1 \geq E\left[\sum_{|u|=n} \eta y_u I\{u\colon y_u \leq \beta\} + \sum_{|u|=n} y_u I\{u\colon \beta < y_u \leq 1\}\right],$$

which, because $E[\sum_{|u|=n} y_u] = 1$, implies that

$$(\eta - 1) E\left[\sum_{|u|=n} I\{u\colon y_u \leq \beta\} y_u\right] \leq E\left[\sum_{|u|=n} I\{u\colon y_u > 1\} y_u\right].$$

Provided $EX > 0$, the expectation on the left tends to one and that on the right tends to zero, by Lemma 4.1(ii) and the weak law of large numbers. This forces $\eta = 1$, which is a contradiction. □

5. The limit of the multiplicative martingales.

LEMMA 5.1.

$$-\log M(x) = \lim_n \sum_{|u|=n} xy_u L(xy_u). \tag{5.1}$$

PROOF. By Lemma 2.2, $\sup_{|u|=n} y_u = l^{(n)} \downarrow 0$. For any $\varepsilon > 0$, provided n is large enough to make $l^{(n)}$ sufficiently small,

$$\begin{aligned} -\log M^{(n)}(x) &= -\sum_{|u|=n} \log \Psi(xy_u) \\ &\geq -\sum_{|u|=n} \big(1 - \Psi(xy_u)\big) \\ &\geq -(1-\varepsilon) \sum_{|u|=n} \log \Psi(xy_u) = -(1-\varepsilon)\log M^{(n)}(x). \end{aligned}$$

Taking limits here and using the definition of L completes the proof. □

When $\Psi'(0) = -1$, $L(y) \uparrow 1$ as $y \downarrow 0$, so that (5.1) implies that $-\log M(x) = xW$, and hence, by Corollary 3.2, Ψ is the Laplace transform of W. In the general case, if $L(xy_u)$ in (5.1) could be approximated by $L(a_n)$ for some constants a_n, this would imply that $L(a_n)W^{(n)}$ converged (almost surely) to

$-\log M(1)$, solving the original problem. This idea seems not to work as just described, which, at least on our present understanding, prevents us from obtaining almost sure convergence in Theorem 1.2. However, the slow variation of L does allow the identification of Ψ as the Laplace transform of $-\log M(1)$ as the next lemma shows.

LEMMA 5.2. (i) $M(x) = M(1)^x$;
(ii) $-\log M(1)$ *has Laplace transform* Ψ;
(iii) $P(M(x) = 0) = 0$;
(iv) $\{M(x) < 1\}$ *is the survival set, almost surely.*

PROOF. Note first that

$$\left|\frac{\sum_{|u|=n} y_u L(xy_u)}{\sum_{|u|=n} y_u L(y_u)} - 1\right| = \left|\frac{\sum_{|u|=n} y_u L(y_u)\big((L(xy_u)/L(y_u)) - 1\big)}{\sum_{|u|=n} y_u L(y_u)}\right|$$
$$\leq \sup_{u:\,|u|=n}\left|\frac{L(xy_u)}{L(y_u)} - 1\right|,$$

which converges to zero as $n \to \infty$, because L is slowly varying and $\sup_{|u|=n} y_u = l^{(n)} \downarrow 0$ as $n \to \infty$, by Theorem 1.4 and Lemma 2.2, respectively. By Lemma 5.1, it follows that $M(x) = M(1)^x$, and taking expectations of this [using the fact noted in Corollary 3.2 that $EM(x) = \Psi(x)$] gives the second assertion. Since Ψ is the Laplace transform of a proper variable, (iii) holds. For the final part, note that $\{M(x) = 1\}$ must be at least the set of extinction, and that $P(M(x) = 1) = \Psi(\infty)$ which satisfies $f(\Psi(\infty)) = \Psi(\infty)$, hence $\Psi(\infty)$ actually is the extinction probability and (iv) must hold. □

To prove the uniqueness of solutions to the functional equation, other multiplicative martingales have to be introduced. Their form is similar to that of $M^{(n)}$, but the products, instead of being taken over $\{u: |u| = n\}$, will be taken over other sets of individuals.

6. General multiplicative martingales. This discussion draws on ideas and arguments in work on optional (stopping) lines by Chauvin (1988, 1991) and Jagers (1989).

A (stopping) line l is a set of individuals none of whom lies in the line of descent of any other; F^l contains full information on the life histories of all individuals that are neither in l nor a descendent of any member of l. The partial ordering of T by "is an ancestor of" ($<$) induces a partial order on lines, with $l_1 \leq l_2$ when every member of l_2 is a descendent (not necessarily strict) of some member of l_1. An optional line τ is a random line with the property that, for any fixed line l, $\{\tau \leq l\} \in F^l$, so, intuitively, the family trees descended from τ's members have no part in determining τ. It turns out that the martingale introduced in Theorem 3.1 is best viewed as arising as a particular case of products being taken over an increasing sequence of optional lines. The general case is described at the end of this section, after the necessary technical apparatus has been put in place.

The branching property, that different individuals in the same generation give rise to independent copies of the original tree, extends to individuals on an optional line, as is proved in Jagers [(1989), Theorem 4.14]. To be more precise, let T_u be the tree emanating from u viewed with u as the initial ancestor and let τ be an optional line. Then, conditioned on F^{τ}, the trees $\{T_u: u \in \tau\}$ are independent copies of the original tree T. Thus, for $u \in \tau$, $\{y_{uv}/y_u : v\}$ has, given F^{τ}, the same distribution as $\{y_v : v\}$. It will be relevant later that Jagers' result is actually for the multitype process with a general set of types, in which the distribution of a tree depends on the type of its initial ancestor.

Given any optional line τ, let

$$M^{(\tau)}(x) = \prod_{u \in \tau} \Psi(xy_u).$$

To simplify notation, the convention is adopted that products are over u when no variable is specified, so the "$u \in$" will often be dropped in $u \in \tau$, and so on. The argument "(x)" will also be suppressed when possible.

Notation for certain characteristics of lines is now introduced. Let $E_\tau(n)$ be the members of τ in the nth generation, and let $A_\tau(n)$ be the nth generation members who have no ancestors (including themselves) in τ. Let $\bar{g}(\tau)$ be $\sup\{n: A_\tau(n) \neq \emptyset\}$, so $\bar{g}(\tau)$ is the latest generation containing a member with no ancestor in τ, and let $\underline{g}(\tau)$ be $\inf\{|u|: u \in \tau\}$, so $\underline{g}(\tau)$ is the earliest generation containing a member of τ. If $\bar{g}(\tau)$ is finite then τ cuts right across the tree.

In the next lemma and the following theorem, M is the limit of the martingale $M^{(n)}$ introduced in Corollary 3.2.

LEMMA 6.1. *Assume that τ is optional and $\bar{g}(\tau)$ is finite almost surely. Then*

$$E[M|F^{\tau}] = M^{(\tau)}.$$

PROOF. Using the notation just introduced,

$$E[M^{(n)}|F^{\tau}] = E\left[\left(\prod_{j=1}^{n} \prod_{E_\tau(j)} \prod_{|v|=n-j} \Psi(xy_{uv})\right) \prod_{A_\tau(n)} \Psi(xy_u) \Big| F^{\tau}\right].$$

Thus, pulling the F^{τ}-measurable parts outside the expectation and applying the branching property together with Corollary 3.2,

$$\begin{aligned} E[M^{(n)}|F^{\tau}] &= \left(E\left[\prod_{j=1}^{n} \prod_{E_\tau(j)} \prod_{|v|=n-j} \Psi\left(xy_u\left(\frac{y_{uv}}{y_u}\right)\right) \Big| F^{\tau}\right]\right) \prod_{A_\tau(n)} \Psi(xy_u) \\ &= \left(\prod_{j=1}^{n} \prod_{E_\tau(j)} \Psi(xy_u)\right) \prod_{A_\tau(n)} \Psi(xy_u). \end{aligned}$$

In using the branching property [i.e., Jagers (1989), Theorem 4.14] two technical points arise. Firstly, $\bigcup_{j=1}^{n} E_{\tau}(j)$ is a stopping line by Jagers [(1989) Proposition 4.10], and its associated sigma-field contains F^{τ}. Secondly, the process must be considered as a multitype one in which individual u has type y_u, thereby allowing the function that is to be evaluated on u's daughter process to depend on y_u. When $n \to \infty$, the left-hand side converges in L_1 to $E[M|F^{\tau}]$ and, because $\bar{g}(\tau)$ is finite, the right-hand side converges almost surely to $M^{(\tau)}$. □

THEOREM 6.2. *Let $\tau(t)$ be an increasing sequence (indexed by $t \in [0,\infty)$) of optional lines, with*

$$\bar{g}(\tau(t)) < \infty \quad \text{for all } t, \text{ almost surely.} \tag{6.1}$$

Then

$$M^{(\tau(t))}(x) = \prod_{u \in \tau(t)} \Psi(xy_u) \tag{6.2}$$

is a (bounded) martingale with respect to $\{F^{\tau(t)}\}$. Furthermore, provided

$$\underline{g}(\tau(t)) \uparrow \infty \quad \text{as } t \uparrow \infty, \text{ almost surely,} \tag{6.3}$$

$M^{(\tau(t))}$ converges to M almost surely and in L_1 and

$$\lim_{t \to \infty} \sum_{u \in \tau(t)} y_u L(y_u) = -\log M(1). \tag{6.4}$$

PROOF. That (6.2) is a martingale follows immediately from the previous lemma. For the second part, note that, $M^{(\tau(t))} = E[M|F^{\tau(t)}] \to M$, almost surely, as $t \to \infty$, using (6.3). Finally, when (6.3) holds, Lemma 2.2 implies that

$$\sup_{u \in \tau(t)} y_u \downarrow 0 \quad \text{as } t \to \infty,$$

so the proof of the last part is just like that of Lemma 5.1. □

7. The general branching process. The general (C-M-J) branching process will play an important part in the study of the multiplicative martingales. The notation for this process, and the main results needed about it, are introduced in this section.

The process is constructed in the same way as a branching random walk. Associated with each individual is an independent copy of the reproduction point process ξ which gives that mother's age at the birth of each of her children. Individuals' birth times are computed by the obvious recursion, by adding the mother's age when that child is born to the mother's own birth time. It is also useful to have the notion of a characteristic, which is a mechanism for counting the population. Each individual has associated with it an independent copy of some function χ, and this function measures the contribution of the individual, as she grows older, to a count of the process.

These functions are zero for negative ages. Suppose the birth time of u is denoted by b_u. The χ-counted process is defined to be

$$\zeta_t^{\chi} = \sum_u \chi_u(t - b_u).$$

For example, if $\chi(a) = I\{a > 0\}$, then ζ_t^{χ} counts all those born before t. More extensive, and more careful, descriptions of the process can be found in Jagers (1975), Nerman (1981) and Asmussen and Hering (1983).

The intensity measure of the reproduction point process ξ is denoted by η, and there is assumed to be an $\alpha > 0$ for which $\int e^{-\alpha a}\eta(da) = 1$; so attention is fixed on supercritical processes with Malthusian parameter α. Note that by multiplying all birth times by α a supercritical process with Malthusian parameter α is transformed to one with Malthusian parameter equal to one.

For the treatment here the important theorem from the theory of general branching processes is the following, which is Theorem 6.3 of Nerman (1981) and is also given as Theorem X.5.1 in Asmussen and Hering (1983).

THEOREM 7.1. *Suppose there is a* $\beta < \alpha$ *such that* $\int e^{-\beta a}\eta(da) < \infty$, *that* ψ *and* χ *are two characteristics with* $E \sup e^{-\beta t}\psi(t)$ *and* $E \sup e^{-\beta t}\chi(t)$ *both finite* (*and with D-paths*). *Then, on the survival set of the process,*

$$\frac{\zeta_t^{\psi}}{\zeta_t^{\chi}} \to \frac{\int_0^\infty e^{-\alpha t}E\psi(t)\,dt}{\int_0^\infty e^{-\alpha t}E\chi(t)\,dt} \quad a.s.\ as\ t \to \infty.$$

The coming generation, which is, at time t, those individuals who are not yet born but whose mothers are, is of particular importance in the theory of the general branching process. Denote this set of individuals by $C(t)$. Nerman [(1981), Proposition 2.4] shows that

$$\sum_{u \in C(t)} \exp(-\alpha b_u)$$

is a martingale; this martingale can be written as $e^{-\alpha t}\zeta_t^{\chi}$ when χ is the characteristic

$$\chi(a) = I\{a > 0\}e^{\alpha a}\int_a^\infty e^{-\alpha t}\xi(dt).$$

Also, the martingale converges in L_1 under an "$X \log X$" condition [Nerman (1981), Corollary 3.3].

The following Seneta–Heyde result will be a by-product of the discussion in the next section.

THEOREM 7.2. *Consider a general* (C-M-J) *branching process with finite family sizes, reproduction intensity measure* η *and Malthusian parameter* $\alpha > 0$, *with birth times* $\{b_u\}$ *and coming generation* $C(t)$. *Suppose that there is* *a* $\beta < \alpha$ *such that* $\int \exp(-\beta\sigma)\eta(d\sigma) < \infty$, *then there is a slowly varying function* L *such that*

$$L(e^{-\alpha t}) \sum_{u \in \mathbf{C}(t)} \exp(-\alpha b_u)$$

has an (almost sure) limit that is finite and nonzero when the process survives.

Combining this with Theorem 7.1 gives the corresponding result for other ways of counting the process. This result extends Theorem 6.1 of Cohn (1985), where the same result is proved under the additional condition that η is a finite measure.

8. An embedded general branching process. To allow a good estimation of the terms on the left-hand side of (6.4) it will be useful if the y_u do not vary too much on $\tau(t)$. We consider a sequence of stopping lines picked to try to make sure this is the case: let $\mathcal{I}(t)$ be the set of individuals who are the first in their line of descent to have y_u less than e^{-t}, so

$$\begin{aligned}\mathcal{I}(t) &= \{u\colon y_u < e^{-t}, \text{ but } y_v \ge e^{-t} \text{ for } v < u\}\\ &= \{u\colon \theta z_u + |u|\log m(\theta) > t, \text{ but } \theta z_v + |v|\log m(\theta) \le t \text{ for } v < u\}.\end{aligned}$$

When the point process Z is concentrated on $(0,\infty)$ and θ is such that $m(\theta) = 1$ this is just the coming generation, defined in the previous section (but at time t/θ rather than at t). It will be shown, in Lemma 8.2, that $\mathcal{I}(t)$ is always the coming generation for a suitable general branching process.

LEMMA 8.1. *Theorem* 6.2 *applies to the optional lines* $\mathcal{I}(t)$.

PROOF. It is clear from the definition that $\mathcal{I}(t)$ increases with t. Using Lemma 2.2,

$$\bar{g}(\mathcal{I}(t)) \le \sup\Big\{n\colon \sup_{|u|=n} y_u > e^{-t}\Big\} = \sup\{n\colon l^{(n)} > e^{-t}\} < \infty,$$

so (6.1) holds. By A3, the nth generation is finite, so for sufficiently large t,

$$\{u\colon |u| \le n,\ y_u \ge e^{-t}\} = \{u\colon |u| \le n\}$$

and, for such t, $\underline{g}(\mathcal{I}(t)) \ge n$; thus (6.3) holds. □

The members of $\mathcal{I}(0)$ can now be considered to be the "children" of the initial ancestor, with intervening members on the line of descent being ignored. These will be called the i-children (i for indirect) of the initial individual, to distinguish them from the original children. Consider the i-child u (in the original labelling) to be born when her i-mother (the initial ancestor) has age

$$\sigma_u := -\log y_u = \theta z_u + |u|\log m(\theta),$$

which must be positive because $u \in I(0)$. Let ξ be the point process of the ages at i-child bearing, so ξ has the points $\{-\log y_u: u \in I(0)\}$, and let η be the intensity measure of ξ. Because $I(0)$ is an optional line, the trees emanating from its members are, given $F^{I(0)}$, independent copies of the original process. Hence each $u \in I(0)$ has associated with it the optional line $I_u(0)$ of its i-children, with associated births, and its reproduction point process, ξ_u, is given by $\{-\log(y_{uv}/y_u): v \in I_u(0)\}$. In this way a general (C-M-J) branching process with reproduction point process ξ, embedded in the original process, is constructed.

The birth time, b_u, of a person u occurring in the embedded process is, of course, obtained by adding the ages of the i-mothers in her ancestry when the appropriate child is born, so $b_u = -\log y_u$. However it is worth stressing that not all individuals in the original process occur in the embedded one. In fact, as the proof of the next lemma (or, better, a picture) shows, only those v giving a (strict ascending) ladder point of the sequence $\{-\log y_v: v \le u\}$ figure in the embedded process. (Note that this sequence is indexed by the individuals in the line of descent from the initial ancestor down to u.) Denote the set of individuals ever born in the embedded process by E and, for $u \in E$, let $m(u)$ be the i-mother of u. Then the coming generation at t for the embedded general branching process is

$$C(t) = \{u: u \in E,\, b_{m(u)} \le t < b_u\};$$

the connection between this and the optional lines already introduced is very simple.

LEMMA 8.2. $C(t) = I(t)$.

PROOF. Suppose that $u \in I(t)$ so that

$$-\log y_u > t, \quad \text{and} \quad -\log y_v \le t \quad \text{for } v < u; \tag{8.1}$$

thus u is a (strict ascending) ladder epoch for the sequence $\{-\log y_v: v \le u\}$. Let the ladder epochs of this sequence be $0 = v(0) < v(1) < \cdots < v(p) = u$. Then, by definition, for $v(i) < v < v(i+1)$,

$$-\log(y_v/y_{v(i)}) = -(\log y_v - \log y_{v(i)}) < 0$$

and

$$-\log(y_{v(i+1)}/y_{v(i)}) = -(\log y_{v(i+1)} - \log y_{v(i)}) > 0;$$

thus $v(i+1) \in I_{v(i)}(0)$ and so is an i-child of $v(i)$. Therefore the ladder epochs provide the ancestors for u in the embedded process, back to the initial ancestor, showing that $u \in E$. It remains to show that u is actually in $C(t)$, but this is immediate because $b_{m(u)} = b_{v(p-1)} = -\log y_{v(p-1)} \le t$ and $b_u = -\log y_u > t$.

Similarly, if $u \in C(t)$ its ancestors in the embedded process provide ladder epochs for $\{-\log y_v: v \le u\}$, which combine with $b_u > t \ge b_{m(u)}$ to show (8.1). □

THEOREM 8.3 (A1–A5 are not required). *Assume* $\theta > 0$, $m(\theta) < \infty$ *and* $EX \geq 0$. *The embedded general branching process has Malthusian parameter* 1, *that is to say*

$$\int \exp(-\sigma)\eta(d\sigma) = E\int \exp(-\sigma)\xi(d\sigma) = E\sum_{u\in I(0)} \exp(-b_u)$$
$$= E\sum_{u\in I(0)} y_u = 1.$$

PROOF. Simply compute, using Lemma 4.1(iii), that

$$E\sum_{I(t)} y_u = E\sum_{n\geq 1}\sum_{|u|=n} y_u I\{-\log y_u > t, -\log y_v \leq t \text{ for } v < u\}$$
$$= \sum_{n\geq 1} P\left[S_n > t, S_j \leq t \text{ for } j < n\right]$$
$$= P[S_n > t \text{ for some } n],$$

which is 1, provided $EX \geq 0$. □

COROLLARY 8.4. *Under the conditions of Theorem* 8.3, $\{\sum_{u\in I(t)} y_u\}$ *is a martingale.*

The correspondence between $C(t)$ and $I(t)$ means that, because the Malthusian parameter is 1 and $y_u = \exp(-b_u)$, this is Proposition 2.4 of Nerman (1981).

THEOREM 8.5. *When* $EX > 0$, *the Laplace transform of* η *converges for some value* $\beta < 1$, *that is, for some* $\beta < 1$,

$$\int \exp(-\beta\sigma)\eta(d\sigma) = E\int \exp(-\beta\sigma)\xi(d\sigma) = E\sum_{u\in I(0)} (y_u)^\beta < \infty.$$

Furthermore, η *is a finite measure if* $m(0) < \infty$.

PROOF. Temporarily, θ needs to figure more fully in the notation. Note first that

$$(y_u)^\beta = (y_u(\theta))^\beta = \left(\frac{\exp(-\theta z_u)}{m(\theta)^{|u|}}\right)^\beta = y_u(\beta\theta)\left(\frac{m(\theta\beta)}{m(\theta)^\beta}\right)^{|u|},$$

so, using Lemma 4.1(ii) but for $X(\theta\beta)$ rather than $X(\theta)$, and a Markov bound (for $\phi > 0$),

$$\begin{aligned}
E\sum_{I(0)}(y_u)^\beta &= E\sum_{I(0)} y_u(\beta\theta)\left(\frac{m(\theta\beta)}{m(\theta)^\beta}\right)^{|u|}\\
&= E\sum_u y_u(\beta\theta)\left(\frac{m(\theta\beta)}{m(\theta)^\beta}\right)^{|u|} I\{-\log y_u(\theta) > 0,\\
&\qquad\qquad -\log y_v(\theta) \le 0 \text{ for } v < u\}\\
&\le E\sum_u y_u(\beta\theta)\left(\frac{m(\theta\beta)}{m(\theta)^\beta}\right)^{|u|} I\{vi = u, -\log y_v(\theta) \le 0\}\\
&= E\sum_u y_u(\beta\theta)\left(\frac{m(\theta\beta)}{m(\theta)^\beta}\right)^{|u|} I\left\{vi = u,\ y_v(\beta\theta) \ge \left(\frac{m(\theta)^\beta}{m(\theta\beta)}\right)^{|v|}\right\}\\
&= \sum_{n\ge 1}\left(\frac{m(\theta\beta)}{m(\theta)^\beta}\right)^n P\left[S_{n-1}(\theta\beta) \le (n-1)\log\left(\frac{m(\theta\beta)}{m(\theta)^\beta}\right)\right]\\
&\le \sum_{n\ge 1}\left(\frac{m(\theta\beta)}{m(\theta)^\beta}\right)^n E[\exp(-\phi X(\beta\theta))]^{n-1}\left(\frac{m(\theta\beta)}{m(\theta)^\beta}\right)^{\phi(n-1)}\\
&= \frac{m(\theta\beta)}{m(\theta)^\beta}\sum_{n\ge 1}\left(\frac{m((1+\phi)\theta\beta)}{m(\theta)^{\beta(1+\phi)}}\right)^{n-1}.
\end{aligned}$$

Thus, to complete the proof of the first part, β must be chosen near enough to 1 to ensure, by A4, that $m(\theta\beta)$ is finite, and then ϕ must be chosen so that

$$\frac{m((1+\phi)\theta\beta)}{m(\theta)^{\beta(1+\phi)}} < 1.$$

Straightforward calculus establishes that, when $EX > 0$, $m(\theta\delta)/m(\theta)^\delta$ is strictly decreasing at $\delta = 1$. Hence, it is enough that $(1+\phi)\beta$ is slightly greater than 1.

For the last part, let $\beta \downarrow 0$ with ϕ chosen so that $(1+\phi)\beta$ is a constant slightly greater than 1, so that the sum above is convergent (and fixed). Now note that

$$\int \eta(d\sigma) = E\int \xi(d\sigma) = \lim_{\beta\downarrow 0} E\int e^{-\beta\sigma}\xi(d\sigma) = \lim_{\beta\downarrow 0} E\sum_{u\in I(0)}(y_u)^\beta < \infty. \quad \square$$

The next theorem provides the Seneta–Heyde renormalization for the coming generation martingale.

THEOREM 8.6. *When* $EX > 0$,

$$L(e^{-t}) \sum_{u \in I(t)} y_u \to -\log M(1)$$

as $t \to \infty$, *almost surely.*

PROOF. Lemma 8.1 shows that Theorem 6.2, applies to the optional lines $I(t)$, so

$$\lim_{t \to \infty} \sum_{u \in I(t)} y_u L(y_u) = -\log M(1).$$

Following Nerman (1981), decompose $I(t)$ into $I(t, c)$ given by,

$$I(t, c) = \{u \colon u \in E, -\log y_{m(u)} \le t, -\log y_u > t + c\}$$

(that is, those with overshoot at least c) and the remainder. Because L is monotone decreasing it clear that, on the survival set,

$$\begin{aligned}
1 &\le \frac{\sum_{I(t)} y_u L(y_u)}{L(e^{-t}) \sum_{I(t)} y_u} \\
&\le \frac{\sum_{I(t)} y_u L(y_u)}{L(e^{-t}) \sum_{I(t) \setminus I(t,c)} y_u} \\
&\le \frac{L(e^{-(t+c)})}{L(e^{-t})} + \frac{\sum_{I(t,c)} y_u L(y_u)}{L(e^{-t}) \sum_{I(t) \setminus I(t,c)} y_u} \\
&= \frac{L(e^{-(t+c)})}{L(e^{-t})} + \frac{\sum_{T(t,c)} y_u L(y_u)/L(e^{-t})}{\sum_{I(t) \setminus I(t,c)} y_u}
\end{aligned}$$

and, because L is slowly varying at zero, $L(e^{-(t+c)})/L(e^{-t})$ goes to 1 as t goes to infinity. Thus the result will be proved if

$$\lim_{c \to \infty} \lim_{t \to \infty} \frac{\sum_{I(t,c)} e^t y_u L(y_u)/L(e^{-t})}{\sum_{I(t) \setminus I(t,c)} e^t y_u} = 0 \tag{8.2}$$

on the survival set almost surely, where the outer limit need only be through the rationals.

By the integral representation of a slowly varying function [see, e.g., VIII.9 of Feller (1971)] for any ε, ε_1, both strictly positive, there is a δ such that for all $y < 1$,

$$\sup_{x < \delta} \frac{L(yx)}{L(x)} \le (1 + \varepsilon_1) y^{-\varepsilon}.$$

Thus

$$\sum_{I(t,c)} (e^t y_u) \frac{L(y_u)}{L(e^{-t})} = \sum_{I(t,c)} (e^t y_u) \frac{L((e^t y_u)e^{-t})}{L(e^{-t})}$$
$$\leq (1+\varepsilon_1) \sum_{I(t,c)} (e^t y_u)^{1-\varepsilon},$$

provided $e^{-t} < \delta$. Thus it suffices to show, in place of (8.2), that, for some $\varepsilon > 0$,

$$\lim_{c\to\infty} \lim_{t\to\infty} \frac{\sum_{I(t,c)} (e^t y_u)^{1-\varepsilon}}{\sum_{I(t)\setminus I(t,c)} (e^t y_u)} = 0. \tag{8.3}$$

The top line here can be written as $\sum_{I(t,c)} \exp((t-b_u)(1-\varepsilon))$, which is simply the embedded processes counted using the characteristic

$$\psi(a) = I\{a>0\} \int_{a+c}^{\infty} \exp(-(\sigma-a)(1-\varepsilon))\,\xi(d\sigma),$$

and the denominator is the embedded process counted using the characteristic

$$\chi(a) = I\{a>0\} \int_{a}^{a+c} \exp(-(\sigma-a))\,\xi(d\sigma).$$

It will now be shown that Theorem 7.1 applies to these two characteristics, provided ε is small enough. Note first that Theorem 8.5 holds with $\beta = 1-\varepsilon$, provided ε is sufficiently small. With this choice of β it remains to check the supremum condition on the characteristics. For ψ, note that, for $a > 0$,

$$\exp(-a\beta)\psi(a) = \int_{a+c}^{\infty} \exp(-\sigma(1-\varepsilon))\,\xi(d\sigma) \leq \int_0^{\infty} \exp(-\sigma(1-\varepsilon))\,\xi(d\sigma)$$
$$= \int_0^{\infty} \exp(-\sigma\beta)\,\xi(d\sigma),$$

which is independent of a and, using Theorem 8.5, has finite expectation. Similarly, for $a > 0$,

$$\exp(-a\beta)\chi(a) = \exp(-a\beta) \int_a^{a+c} \exp(-(\sigma-a))\,\xi(d\sigma)$$
$$\leq \int_a^{a+c} \exp(-\sigma\beta)\,\xi(d\sigma) \leq \int_0^{\infty} \exp(-\sigma\beta)\,\xi(d\sigma).$$

Thus

$$\lim_{t\to\infty} \frac{\sum_{I(t,c)} (e^t y_u)^{1-\varepsilon}}{\sum_{I(t)\setminus I(t,c)} e^t y_u} = \frac{\int_0^{\infty} \left(\int_{a+c}^{\infty} \exp(-(1-\varepsilon)\sigma)\,\eta(d\sigma)\right) \exp(-a\varepsilon)\, da}{\int_0^{\infty} \int_a^{a+c} \exp(-\sigma)\,\eta(d\sigma)\, da}$$

almost surely. The denominator here converges to $\int_0^{\infty} \sigma e^{-\sigma} \eta(d\sigma)$, which is finite, and the numerator goes to zero as c goes to infinity, provided ε is sufficiently small, using Theorem 8.5 in both cases. □

When the underlying point process Z is concentrated on $(0, \infty)$ and has Malthusian parameter equal to one, the embedded process and the original one are the same, so Theorem 7.2 is a simple consequence of this theorem.

PROOF OF THEOREM 1.5. Suppose we take two nontrivial solutions to (1.5), Ψ_1 and Ψ_2. Both of these can be used to construct multiplicative martingales; hence, applying Theorem 8.6,

$$\frac{-\log M_1(1)}{-\log M_2(1)} = \lim_{t\uparrow\infty} \frac{L_1(e^{-t})\Sigma_{u\in I(t)} y_u}{L_2(e^{-t})\Sigma_{u\in I(t)} y_u} = \lim_{t\uparrow\infty} \frac{L_1(e^{-t})}{L_2(e^{-t})} = c,$$

where c must be a constant. By Lemma 5.2(iv), both $-\log M_1(1)$ and $-\log M_2(1)$ are strictly positive and finite on the survival set, so $0 < c < \infty$. Thus $\log M_1(1) = c \log M_2(1)$, so Lemma 5.2(ii), which shows that $\log M(1)$ has Laplace transform Ψ, completes the proof. □

In particular this establishes that the functional equation satisfied by the Laplace transform of the limit variable obtained by Seneta–Heyde normalization of the general (C-M-J) branching process has a unique solution.

9. Seneta–Heyde norming constants in the BRW. As explained already, ideas of Cohn's are used to strengthen the convergence in distribution along a subsequence to convergence in probability along that subsequence. The following lemma provides the key; it is similar to Theorem 3.1 of Cohn (1985).

LEMMA 9.1. *Suppose that* $\{Y_n\}$ *is a sequence of nonnegative random variables adapted to the increasing sigma-fields* $\{G^n\}$. *Suppose also that along a fixed subsequence* $\{n(i)\colon i = 1, 2, \dots\}$ *for each* k *and* $x > 0$, *the conditional Laplace transform* $E[\exp(-xY_{n(i)})|G^k]$ *converges as* $i \to \infty$. *Denote the limit by* $\psi_k(x)$.

(i) *For each* $x > 0$, $\{\psi_k(x)\}$ *forms a bounded nonnegative martingale with respect to* $\{G^k\}$.

(ii) *Denote the limit of the martingale* $\{\psi_k(x)\}$ *by* $\psi(x)$. *If, for* $x > 0$, $\psi(x) = e^{-xX}$ *for a finite random variable* X (*that does not depend on* x) *then* $Y_{n(i)} \to X$ *in probability, as* $i \to \infty$.

PROOF. Using the definition of ψ_{k+1} and dominated convergence,

$$\begin{aligned} E\big[\psi_{k+1}(x)|G^k\big] &= E\Big[\lim_i E\big[\exp(-xY_{n(i)})|G^{k+1}\big]|G^k\Big] \\ &= \lim_i E\Big[E\big[\exp(-xY_{n(i)})|G^{k+1}\big]|G^k\Big] = \psi_k(x), \end{aligned}$$

proving (i).

It will be convenient to let $X_k = -\log \psi_k(1)$. Thus, because $\psi_k(x) \to \psi(x) = e^{-xX}$, $X_k \to X$ almost surely, and

$$\begin{aligned} 1 = \psi(x)e^{xX} &= \lim_k \left(\lim_i E\left[\exp(-xY_{n(i)})|G^k\right]\right)\exp(xX_k) \\ &= \lim_k \lim_i E\left[\exp\left(-x\left(Y_{n(i)} - X_k\right)\right)|G^k\right]; \end{aligned}$$

thus there exists a nonrandom sequence of integers $\{k(i)\}$ such that

$$\lim_i E\left[\exp\left(-x\left(Y_{n(i)} - X_{k(i)}\right)\right)|G^{k(i)}\right] = 1$$

in probability. If this equality were maintained on taking unconditional expectations then $Y_{n(i)} - X_{k(i)}$ would converge in distribution, and hence in probability, to zero, and the result would be proved. In fact a slightly modified argument is needed, where the convergence of $I\{X_{k(i)} \le u\}(Y_{n(i)} - X_{k(i)})$ for any $u > 0$ is considered.

Because $I\{X_k \le u\}$ is G^k measurable it follows that, for any finite u,

$$\begin{aligned} &E\left[\exp\left(-xI\left\{X_{k(i)} \le u\right\}\left(Y_{n(i)} - X_{k(i)}\right)\right)|G^{k(i)}\right] \\ &\quad = I\left\{X_{k(i)} \le u\right\}E\left[\exp\left(-x\left(Y_{n(i)} - X_{k(i)}\right)\right)|G^{k(i)}\right] + I\left\{X_{k(i)} > u\right\} \\ &\quad \to I\{X \le u\} + I\{X > u\} = 1 \end{aligned}$$

in probability, as i goes to infinity. For $x \ge 0$, the variables on the left are bounded by $\max\{e^{xu}, 1\}$, so taking unconditional expectations and applying dominated convergence shows that, for any u, $(Y_{n(i)} - X_{k(i)})I\{X_{k(i)} \le u\}$ converges in distribution, and hence in probability, to zero, as i goes to infinity. Now

$$\begin{aligned} \limsup_i P\left(|Y_{n(i)} - X| > \varepsilon\right) &\le \limsup_i \Big[P\left(|Y_{n(i)} - X_{k(i)}|I\left\{X_{k(i)} \le u\right\} > \varepsilon/2\right) \\ &\qquad + P\left(|X_{k(i)} - X| > \varepsilon/2\right) + P\left(X_{k(i)} > u\right)\Big] \\ &= 0 + 0 + P(X > u) \end{aligned}$$

which can be made arbitrarily small by taking u sufficiently large because, by assumption, X is finite almost surely. □

PROOF OF THEOREM 1.2. As in the proof of Theorem 1.3 in Section 2, let the Laplace transform of $W^{(n)}$ be $\Omega_n(x)$ and take c_n to be such that $\Omega_n(1/c_n) = \kappa$, where κ is fixed to be greater than the extinction probability but less than 1. The transform of the limit of any subsequence of $\{W^{(n)}/c_n\}$ that converges in distribution must satisfy the functional equation (1.5) with $\Psi(1) = \kappa$. By Theorem 1.5, the solution to the functional equation is unique, so $\{W^{(n)}/c_n\}$ converges in distribution along the full sequence.

Let $Y_n = W^{(n)}/c_n$ and $G^n = F^n$. To see that Lemma 9.1 applies, note that, for $k \le n$,

$$W^{(n)} = \sum_{|u|=k} y_u W_u^{(n-k)} = \sum_{|u|=k} y_u W_u^{(n)} \frac{W_u^{(n-k)}}{W_u^{(n)}},$$

so, dividing through by c_n, taking Laplace transforms, letting n go to infinity and using Theorem 2.4 gives that

$$\lim_n E\left[\exp\left(-xW^{(n)}/c_n\right)|F^k\right] = \prod_{|u|=k} \Psi(xy_u) = M^{(k)}(x).$$

[Thus Lemma 9.1(i) is, in this context, Theorem 3.1.] Now, by Lemma 5.2(i) and (iii), the martingales limits have the property required for Lemma 9.1(ii) to hold, and the result is proved. □

Note that the random variable Δ in Theorem 1.2 is $-\log M(1)$.

Acknowledgments. Conversations with A. Rouault, B. Chauvin, H. Cohn, P. Jagers, O. Nerman and J. Neveu clarified various points arising in the course of the work described here; we are grateful to all of them. We are also grateful to an anonymous referee for a careful reading of the manuscript.

REFERENCES

ASMUSSEN, S. and HERING, H. (1983). *Branching Processes*. Birkhäuser, Boston.

BIGGINS, J. D. (1977a). Martingale convergence in the branching random walk. *J. Appl. Probab.* **14** 25–37.

BIGGINS, J. D. (1977b). Chernoff's Theorem in the branching random walk. *J. Appl. Probab.* **14** 630–636.

BIGGINS, J. D. (1992). Uniform convergence of martingales in the branching random walk. *Ann. Probab.* **20** 137–151.

BIGGINS, J. D. and KYPRIANOU, A. E. (1996). Branching random walk: Seneta–Heyde norming. In *Trees*: *Proceedings of a Workshop, Versailles June 14–16, 1995* (B. Chauvin, S. Cohen and A. Rouault, eds.). Birkhäuser, Basel.

CHAUVIN, B. (1988). *Arbres et Processus de Branchement*. Ph.D. thesis, Univ. Paris 6.

CHAUVIN, B. (1991). Product martingales and stopping lines for branching Brownian motion. *Ann. Probab.* **19** 1195–1205.

CHAUVIN, B. and ROUAULT, A. (1996). Boltzmann–Gibbs weights in the branching random walk. In *Classical and Modern Branching Processes* (K. B. Athreya and P. Jagers, eds.) **84** 41–50. Springer, New York.

COHN, H. (1985). A martingale approach to supercritical (CMJ) branching processes. *Ann. Probab.* **13** 1179–1191.

DURRETT, R. and LIGGETT, M. (1983). Fixed points of the smoothing transform. *Z. Wahrsch. Verw. Gebiete* **64** 275–301.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, 2nd ed. Wiley, New York.

HEYDE, C. C. (1970). Extension of a result of Seneta for the supercritical branching process. *Ann. Math. Statist.* **41** 739–742.

JAGERS, P. (1975). *Branching Processes with Biological Applications*. Wiley, New York.

JAGERS, P. (1989). General branching processes as Markov fields. *Stochastic Process. Appl.* **32** 183–212.

KAHANE, J. P. and PEYRIÈRE, J. (1976). Sur certaines martingales de Benoit Mandelbrot. *Adv. Math.* **22** 131–145.

KINGMAN, J. F. C. (1975). The first birth problem for an age-dependent branching process. *Ann. Probab.* **3** 790–801.

KURTZ, T. G. (1972). Inequalities for the law of large numbers. *Ann. Math. Statist.* **43** 1874–1883.

LIU, Q. (1996). Fixed points of a generalized smoothing transformation and applications to branching processes. Unpublished manuscript.

LIU, Q. (1997). Sur une equation fonctionelle et ses applications: une extension du theoreme de Kesten–Stigum concernant des processus de branchement. *Adv. in Appl. Probab.* **29**.

LYONS, R. (1996). A simple path to Biggins' martingale convergence. In *Classical and Modern Branching Processes* (K. B. Athreya and P. Jagers, eds.) **84** 217–222. Springer, New York.

NERMAN, O. (1981). On the convergence of supercritical general (C-M-J) branching process. *Z. Wahrsch. Verw. Gebiete* **57** 365–395.

NEVEU, J. (1988). Multiplicative martingales for spatial branching processes. In *Seminar on Stochastic Processes, 1987* (E. Çinlar, K. L. Chung and R. K. Getoor, eds.) *Prog. Probab. Statist.* **15** 223–241. Birkhäuser, Boston.

PAKES, A. G. (1992). On characterizations via mixed sums. *Austral. J. Statist.* **34** 323–339.

SENETA, E. (1968). On recent theorems concerning the supercritical Galton–Watson process. *Ann. Math. Statist.* **39** 2098–2102.

WAYMIRE, E. C. and WILLIAMS, S. C. (1994). A general decomposition theory for random cascades. *Bull. Amer. Math. Soc.* (*N. S.*) **31** 216–222.

WAYMIRE, E. C. and WILLIAMS, S. C. (1995). Multiplicative cascades: dimension spectra and dependence. *J. Fourier Analysis and Applications. Special issue in honour of J.-P. Kahane* 589–609.

WAYMIRE, E. C. and WILLIAMS, S. C. (1996). A cascade decomposition theory with applications to Markov and exchangeable cascades. *Trans. Amer. Math. Soc.* **384** 585–632.

PROBABILITY AND STATISTICS SECTION
SCHOOL OF MATHEMATICS AND STATISTICS
HICKS BUILDING
UNIVERSITY OF SHEFFIELD
SHEFFIELD, S3 7RH
UNITED KINGDOM
E-MAIL: j.biggins@sheffield.ac.uk