The Annals of Applied Probability
1995, Vol. 5, No. 4, 1008–1024

THE GROWTH AND SPREAD OF THE GENERAL BRANCHING RANDOM WALK[1]

BY J. D. BIGGINS

University of Sheffield

A general (Crump–Mode–Jagers) spatial branching process is considered. The asymptotic behavior of the numbers present at time t in sets of the form $[ta, \infty)$ is obtained. As a consequence it is shown that if B_t is the position of the rightmost person at time t, B_t/t converges to a constant, which can be obtained from the individual reproduction law, almost surely on the survival set of the process. This generalizes the known discrete-time results.

1. Introduction. This is a companion paper to Biggins (1996), which should be read for background information, additional motivation and examples, for the application of the results to m-ary search trees [a data-storage algorithm, see Devroye (1990)], for the multitype and d-dimensional extensions and for some discussion of the connections with the corresponding deterministic theory, as represented by van den Bosch, Metz and Diekmann (1990).

A general spatial branching process is considered, and the principal aim is to establish that if B_t is the position of the rightmost person at time t, then

$$\frac{B_t}{t} \to \gamma, \tag{1.1}$$

when the process survives, almost surely. Furthermore, a simple formula for γ is given.

The branching process is built up in the usual way. First the life history of an individual is described; then the process is constructed by letting each individual have an independent life history.

Life-histories consist of a triple (Z, M, χ), whose components describe the individual's reproduction, movement and importance, respectively. Here Z is a point process on $\mathbb{R} \times \mathbb{R}^+$, with each point corresponding to a child; the first coordinate gives the child's displacement from her parent's birth position and the second gives the parent's (strictly positive) age at that child's birth. The movement of the parent is described by the real-valued stochastic process M; a person that is born at position z will, at age a, be at $z + M(a)$. Finally χ is

Received August 1994; revised May 1995.

[1]This work was carried out while visiting the Institute of Mathematics and its Applications, University of Minnesota, as part of the program on "Emerging Applications of Probability Theory." I am happy to thank the IMA for its excellent support, and the NSF, NSA and ARO for their financial assistance via the IMA.

AMS 1991 *subject classification*. Primary 60J80.

Key words and phrases. Spatial spread, asymptotic speed, propagation rate, CMJ process.

a nonnegative stochastic process (usually, in this context, called a random characteristic) giving the importance in "counting" the population of the individual as she grows older. Both M and χ are assumed to have paths in Skorokhod D-space.

In the Ulam–Harris sample space, individuals are labelled by their line of descent, so xy is the yth child of x, and an independent copy of the basic triple is attached to each individual. Denote x's copy of Z by Z_x, with points $\{(z_{xy}, \tau_{xy})\}$. Let (p_x, σ_x) be the position and birth time of the person labelled x; the basic recursion defining these quantities, and therefore describing the population development, is

$$p_{xy} = p_x + z_{xy}, \qquad \sigma_{xy} = \sigma_x + \tau_{xy}.$$

Thus x's offspring, relative to x's own position and birth time, have positions and birth times given by the appropriate independent copy of Z.

Let U be the set of individuals that are born. Ignoring the spatial component, a general branching process counted by the characteristic χ, denoted by ξ^{χ}, can be defined by

$$\xi^{\chi}(t) = \sum_{x \in U} \chi_x(t - \sigma_x).$$

This gives the total weight (as measured by χ) of the population at time t. (Individuals make no contribution before they are born; that is, χ is zero for negative arguments.) This process has been extensively studied; see, for example, Jagers (1975), Nerman (1981) and Cohn (1985).

In several of the results for the spatial branching process considered here, attention will be confined to characteristics that take only the values 0 and 1, corresponding to "dead" and "alive," respectively. This is quite a natural assumption for the motivating question on the behavior of the most extreme (living) individual. Many of the results do not require such an assumption, but it seems to be important at one point in the argument, specifically, in Lemma 3. The matter is discussed further in Section 11.

Let the random measure N_t be defined by

$$N_t = \sum_{x \in U} \delta(p_x + M_x(t - \sigma_x))\chi_x(t - \sigma_x),$$

where $\delta(x)$ is a unit mass at x. Thus the mass of N_t is concentrated at the positions occupied by people at time t, with the mass at a point being the value of the corresponding person's characteristic at that time. The total mass of N_t develops like the general branching process $\xi^{\chi}(t)$. If χ is 0–1, N_t becomes a point process, and the position of the rightmost person at time t is given by

$$B_t = \sup\{p_x + M_x(t - \sigma_x) \colon x \in U, \chi_x(t - \sigma_x) = 1\}.$$

The result on the growth of B_t with t will be a consequence of an analysis of the behavior, for different a, of $N_t[ta, \infty)$ as t goes to infinity. Hence the route here to the behavior of B_t is like that adopted in Biggins (1977) for the discrete-time problem.

In the next two sections sufficient further notation will be developed to state the main theorems. The proofs are in the following seven sections. The last two sections give a brief discussion of the lattice case and more general notions of a random characteristic.

2. The growth of a general branching process. For this section the spatial element of the problem is ignored, with the focus being on the general branching process ξ^{χ}. Let $\tilde{\mu}$ be the intensity measure of the point process formed by the first-generation birth times and let $\tilde{m}$ be its Laplace transform. Thus

$$\tilde{m}(\phi) = \int e^{-\phi\tau}\tilde{\mu}(d\tau) = E\int e^{-\phi\tau}\tilde{Z}(d\tau),$$

where $\tilde{Z}$ is the point process formed by projecting Z onto the time axis. The Malthusian parameter is defined by

$$\alpha = \inf\{\phi\colon \tilde{m}(\phi) \le 1\}.$$

For supercritical processes, those for which $\tilde{m}(0) > 1$ or, equivalently, $\alpha > 0$, there is a positive probability that the process survives. Those theorems that consider sample path behavior, as opposed to estimates of expectations, will be for supercritical processes.

Because the proofs involve renewal theory, the lattice case needs to be handled separately. Consequently, it will be assumed that (Z, M, χ) is nonlattice, in the sense that Z is nonlattice in time. (This assumption could just as well have been that the measure $\tilde{\mu}$ is nonlattice.) Note that the spatial structure is irrelevant in this assumption. The lattice case is discussed briefly in Section 10.

A mild integrability condition on the characteristic χ is also required, for which the following definitions are needed. A function h will be called *moderately varying* if it is strictly positive and, for some $\varepsilon > 0$, satisfies

$$\sup_{t\ge 0}\left\{\frac{\sup\{h(s)\colon |s-t|\le\varepsilon,\ s\ge 0\}}{h(t)}\right\} < \infty. \tag{2.1}$$

A sufficient condition for h to be moderately varying is that $\log h$ is uniformly continuous on $[0,\infty)$. If (2.1) holds for any $\varepsilon > 0$, it holds for all $\varepsilon > 0$. A function h will be called a *regulator* if it is strictly positive, moderately varying and has an integral that does not grow too rapidly, in that

$$\frac{1}{t}\log\left(\int_0^t h(\sigma)\,d\sigma\right) \to 0, \tag{2.2}$$

as $t \to \infty$. Note in particular that $h(\sigma) \equiv 1$ is a regulator. The results in Biggins (1996) were, for simplicity, stated for this special case only.

The first theorem provides a crude estimate of the expected numbers in the general branching process, that is, of $E\xi^{\chi}(t)$, as $t \to \infty$. It is a special case of Theorem 3 in the next section, and it will be proved as part of that.

THEOREM 1. *Let ξ^χ be a nonlattice general branching process with Malthusian parameter α. We have*

(i)
$$\liminf_{t\to\infty} \frac{\log E\xi^\chi(t)}{t} \geq \alpha.$$

(ii) *If, for some regulator h,*

$$E\left(\sup_t \left\{\frac{e^{-\alpha t}\chi(t)}{h(t)}\right\}\right) < \infty, \tag{2.3}$$

then

$$\limsup_{t\to\infty} \frac{\log E\xi^\chi(t)}{t} \leq \alpha.$$

These properties of expected numbers are reflected in the sample paths of the process, as the next theorem shows.

THEOREM 2. *Let ξ^χ be a supercritical nonlattice general branching process with Malthusian parameter α. Then*

(i)
$$\liminf_{t\to\infty} \frac{\log \xi^\chi(t)}{t} \geq \alpha$$

when the process survives, almost surely.

(ii) *If, for some regulator h,* (2.3) *holds, then*

$$\limsup_{t\to\infty} \frac{\log \xi^\chi(t)}{t} \leq \alpha \quad a.s.$$

The proof is based on the estimates of expectations in the previous theorem and known results from Nerman (1981). Condition (2.3) provides control from above on the sample paths; it is a weaker version of Condition 5.2 in Nerman (1981).

3. Main theorems. Let the intensity measure of the point process Z be denoted by μ, with Laplace transform $m(\theta, \phi)$, so that

$$m(\theta, \phi) = \int e^{-\theta z - \phi\tau}\mu(dz, d\tau) = E\int e^{-\theta z - \phi\tau}Z(dz, d\tau).$$

This is assumed to be finite somewhere. Note that, for any fixed θ, m is a decreasing function of ϕ and that $\tilde{m}(\phi) = m(0, \phi)$. For supercritical processes, $m(0,0) > 1$ (but it need not be finite). Let

$$\alpha(\theta) = \inf\{\phi\colon m(\theta, \phi) \leq 1\}.$$

[Note that $\alpha(\theta)$ may be infinite.] It is easy to check that, because m is convex, $\alpha(\theta)$ is a convex function of θ. This implies that α is continuous on the interior of its domain of finiteness. Using the definition of m and both monotone and dominated convergence, it can be shown that α is actually continuous on the closure of its domain of finiteness. This means that α is a closed convex function as defined in Rockafellar [(1970), Section 7].

It has been assumed that, for some (θ_0, ϕ_0), $m(\theta_0, \phi_0)$ is finite; the convolution powers of μ are then well defined. Furthermore, for a sufficiently large value of ϕ_1, $m(\theta_0, \phi_1) < 1$, which implies that the "renewal" measure ν, formed by summing the convolution powers of μ, is also well defined.

It is a straightforward matter to obtain the Laplace transform of the intensity measure of N_t. It is

$$n_t(\theta) := E\int e^{-\theta p} N_t(dp) = \int e^{-\theta p} E\left(e^{-\theta M(t-\sigma)} \chi(t-\sigma)\right) \nu(dp, d\sigma).$$

It will be convenient to let $g_\theta(t) = E(e^{-\theta M(t)} \chi(t))$, so that

$$n_t(\theta) = \int g_\theta(t-\sigma) e^{-\theta p} \nu(dp, d\sigma). \tag{3.1}$$

Multiplying through by $e^{-\alpha(\theta)t}$ and integrating out p turns this into a renewal equation. Hence, if $g_\theta(\sigma)e^{-\alpha(\theta)\sigma}$ were directly Riemann integrable with a finite integral, precise asymptotics for $n_t(\theta)$ would result. Because only rather crude asymptotics are sought, less than this can be asked of g_θ. The next theorem describes the asymptotic behavior of n_t; the condition (3.2) supplies the necessary control over g_θ. Notice that, because $n_t(0) = E\xi^{\chi}(t)$, the special case $\theta = 0$ in this theorem is Theorem 1.

THEOREM 3. (i) *We have*

$$\liminf_{t\to\infty} \frac{\log n_t(\theta)}{t} \ge \alpha(\theta).$$

(ii) *If, for some regulator* h,

$$E\left(\sup_t \left\{\frac{e^{-\alpha(\theta)t} e^{-\theta M(t)} \chi(t)}{h(t)}\right\}\right) < \infty, \tag{3.2}$$

then

$$\limsup_{t\to\infty} \frac{\log n_t(\theta)}{t} \le \alpha(\theta).$$

Because the main results need the convergence of $t^{-1} \log n_t(\theta)$ for all $\theta \le 0$, it is natural to introduce the notion of (3.2) holding whenever needed. Specifically, the triple (Z, M, χ) will be called *regulated* if, for each $\theta \le 0$ in $\{\theta\colon \alpha(\theta) < \infty\}$, there is a regulator h such that, when $\phi = \alpha(\theta)$,

$$U(\theta, \phi) := \sup_t \left\{\frac{e^{-\phi t} e^{-\theta M(t)} \chi(t)}{h(t)}\right\}$$

has finite expectation. Note that $U(\theta, \alpha(\theta))$ is exactly the variable occurring in (3.2). The next result is an immediate consequence of this definition and Theorem 3.

COROLLARY 1. *Suppose that (Z, M, χ) is nonlattice and regulated. Then*

$$\frac{\log n_t(\theta)}{t} \to \alpha(\theta)$$

as $t \to \infty$, for all $\theta \le 0$.

Let the intensity measure of N_t be denoted by η_t, so that n_t is the transform of η_t. Estimates of η_t can now be based on Corollary 1 by an application of a variant of the large deviation results of Gärtner (1977) and Ellis (1984). A little additional notation is needed to state the result.

The concave dual (rate function, large deviation function, etc.) of a convex function k that is finite for some $\theta < 0$ is given by

$$k^*(x) = \inf_{\theta<0}\{x\theta + k(\theta)\}.$$

Because all the results will be formulated for right tails, attention has been confined to $\theta < 0$ here.

Let the right endpoint of the domain of finiteness of α^* be $\tilde{a}$, so that

$$\tilde{a} = \sup\{a: \alpha^*(a) > -\infty\}.$$

Usually this will be infinite. An example is given in Biggins (1996) that shows that it is necessary to exclude $a = \tilde{a}$ in the next two theorems. (In these $\log 0$ is to be interpreted as $-\infty$.)

THEOREM 4. *Suppose that (Z, M, χ) is nonlattice and regulated, and for some $\theta < 0$, $\alpha(\theta) < \infty$. Then, for all $a \neq \tilde{a}$,*

$$\frac{\log(\eta_t[ta,\infty))}{t} \to \alpha^*(a),$$

as $t \to \infty$.

The estimate of expected numbers in the previous theorem has a counterpart in the sample paths, described in the following result.

THEOREM 5. *Suppose that (Z, M, χ) is supercritical, nonlattice and regulated, and χ is a 0–1 characteristic. Suppose also that, for some $\theta < 0$, $\alpha(\theta) < \infty$.*

(i) *If $\alpha^*(a) < 0$, then, for any $\kappa > \alpha^*(a)$,*

$$e^{-\kappa t}N_t[ta,\infty) \to 0 \quad a.s.,$$

as $t \to \infty$.

(ii) *If $a \neq \tilde{a}$ and $\alpha^*(a) > 0$, then*

$$\frac{\log(N_t[ta,\infty))}{t} \to \alpha^*(a),$$

as $t \to \infty$, when the process survives, almost surely.

It will be clear from the proof of this result that, when $a = \tilde{a}$, $\alpha^*(\tilde{a})$ continues to provide an upper bound for $t^{-1}\log(N_t[ta,\infty))$, but it need no longer be its limit.

By looking at the a in the previous theorem for which $N[ta,\infty)$ decays and those for which it grows, the following corollary is established.

COROLLARY 2. *Suppose that (Z, M, χ) is supercritical, nonlattice and regulated, that χ is a 0–1 characteristic and that, for some $\theta < 0$, $\alpha(\theta) < \infty$. Then*

$$\frac{B_t}{t} \to \gamma := \inf\{a \colon \alpha^*(a) < 0\},$$

as $t \to \infty$.

There is an alternative formula for γ which is often simpler to compute. It already occurs in Biggins (1980), where upper bounds for the d-dimensional analogue of B_t/t were discussed.

PROPOSITION 1. *We have*

$$\gamma = \inf\left\{a \colon \inf_{\theta<0}\{\log m(\theta, -a\theta)\} < 0\right\}.$$

Theorem 3, with Theorem 1 as a special case, is proved first, for it plays a part in the proof of Theorem 2. Theorems 2 and 4, which both depend on Theorem 3, but are otherwise independent of each other, are proved next. The main proofs finish with that of Theorem 5, which draws on Theorems 2(i) and 4.

4. Preparatory lemmas on integrability. Two technical matters arise in the proofs. One is the Riemann approximation of certain functions; the other is controlling the process for all t through its values on a fine lattice. The objective here is to establish lemmas that, under the conditions imposed, ensure neither of these issues causes a problem. Key ideas are taken from Lemma 5.3 of Nerman (1981). The first two lemmas are needed for the proof of Theorem 3; the third lemma is important in controlling sample path behavior in Theorems 2 and 5.

For $\varepsilon > 0$, let, for $t \geq 0$,

$$\chi^{\varepsilon}(t) = \sup\{\chi(s) \colon |s - t| \leq \varepsilon,\ s \geq 0\}.$$

This is a sample-path upper approximation to χ. In a similar way, provided χ is 0–1, let, for t such that $\chi^{\varepsilon}(t) = 1$,

$$M^{\varepsilon}(t) = \sup\{M(s) \colon |s - t| \leq \varepsilon,\ \chi(s) = 1,\ s \geq 0\}.$$

For completeness, take $M^{\varepsilon}(t) = M(t)$, for all other t.

Recall that

$$U(\theta,\phi) := \sup_t \left\{ \frac{e^{-\phi t} e^{-\theta M(t)} \chi(t)}{h(t)} \right\},$$

where h is a regulator.

LEMMA 1. *If*

$$Y := \sup_{a \le t} \{ e^{-\theta M(a)} \chi(a) \}$$

has finite expectation, then g_θ is continuous almost everywhere on $[0, t]$, and hence is Riemann integrable there. In particular, if, for any ϕ, $U(\theta, \phi)$ has finite expectation, the conclusion holds.

PROOF. The fact that the paths of M and χ are D-valued and dominated convergence (using Y) show that $g_\theta(a)$ $[= Ee^{-\theta M(a)}\chi(a)]$ is D-valued, and hence continuous almost everywhere, on $[0, t]$. For the last part simply note that, for some constant K, $Y \le KU(\theta, \phi)$. □

LEMMA 2. *If $EU(\theta, \phi)$ is finite,*

$$\frac{1}{n} \log\left(\sum_{i=0}^{n} \sup_{i \le \sigma \le i+1} \{ g_\theta(\sigma) e^{-\phi\sigma} \} \right) \to 0,$$

as $n \to \infty$.

PROOF. Observe that

$$\begin{aligned} g_\theta(\sigma) e^{-\phi\sigma} &= E\left(e^{-\theta M(\sigma)} \chi(\sigma) e^{-\phi\sigma} \right) \\ &\le EU(\theta, \phi) h(\sigma). \end{aligned}$$

Now

$$\sum_{i=0}^{n} \sup_{i \le \sigma \le i+1} h(\sigma) \le K \int_0^{n+1} h(\sigma)\, d\sigma$$

as h is moderately varying. Thus, as h is a regulator and, therefore, by definition, satisfies (2.2), the result follows. □

LEMMA 3. *If χ is 0–1 and $U(\theta, \phi)$ has finite expectation, so has*

$$U^\varepsilon(\theta, \phi) := \sup_t \left\{ \frac{\exp(-\phi t) \exp(-\theta M^\varepsilon(t)) \chi^\varepsilon(t)}{h(t)} \right\}.$$

When $\theta = 0$ the conclusion holds for a general (not just 0–1) nonnegative χ.

PROOF. Note first that, when χ is a 0–1 characteristic and $\chi^\varepsilon(t) = 1$,

$$\exp(-\theta M^\varepsilon(t)) = \sup\{ \exp(-\theta M(s)) \chi(s) : |s - t| \le \varepsilon,\ s \ge 0 \}.$$

Hence,

$$\sup_t \left\{ \frac{\exp(-\phi t)\exp(-\theta M^{\varepsilon}(t))\chi^{\varepsilon}(t)}{h(t)} \right\}$$
$$= \sup_t \left\{ \frac{\exp(-\phi t)}{h(t)} \sup\{\exp(-\theta M(s))\chi(s) : |s-t| \le \varepsilon,\ s \ge 0\} \right\}$$
$$\le U(\theta,\phi)\exp(|\phi|\varepsilon) \sup_t \left\{ \frac{\sup\{h(s) : |s-t| \le \varepsilon,\ s \ge 0\}}{h(t)} \right\},$$

which is finite, as h is moderately varying. When $\theta = 0$, but χ is general, this bound still holds because the first equality is then just the definition of χ^{ε}. □

5. Proof of Theorem 3.

PROOF OF THEOREM 3(i). Suppose first that $\alpha(\theta) < \infty$ with $m(\theta, \alpha(\theta)) = 1$. This ensures that

$$e^{-\theta z - \alpha(\theta)\tau}\mu(dz, d\tau) \quad \text{and} \quad e^{-\theta p - \alpha(\theta)\sigma}\nu(dp, d\sigma)$$

are a probability measure and its associated renewal measure. Denote by $\tilde{\nu}_\theta$ the renewal measure that results by integrating out p here. Multiplying through by $e^{-\alpha(\theta)t}$ in the definition of n_t [(3.1)] and integrating out p gives

$$e^{-\alpha(\theta)t} n_t(\theta) = \int g_\theta(t-\sigma) e^{-\alpha(\theta)(t-\sigma)} \tilde{\nu}_\theta(d\sigma). \tag{5.1}$$

The aim is to bound this integral from below.

Replace $(M(t), \chi(t))$ by

$$(M(t), \chi(t) I(\chi(t) < B_1) I(t < B_2) I(|M(t)| < B_3)) \tag{5.2}$$

with the B's chosen large enough to ensure that the corresponding g_θ is not identically zero. (The indicator involving B_1 only matters when general characteristics are considered.) Obviously, because it makes the characteristic smaller, this replacement decreases $n_t(\theta)$, so it is enough to prove the result for characteristics of this form. Now, in such cases, Lemma 1 shows that g_θ is Riemann integrable. Taking suitable lower approximants to g_θ and using the renewal theorem shows that, for any finite T and some $K > 0$,

$$\int g_\theta(t-\sigma) e^{-\alpha(\theta)(t-\sigma)} \tilde{\nu}_\theta(d\sigma) \ge K \int_0^T g_\theta(\sigma) e^{-\alpha(\theta)\sigma} d\sigma > 0,$$

provided t is sufficiently large. Because, from (5.1),

$$t^{-1}\log n_t(\theta) = \alpha(\theta) + t^{-1}\log\left(\int g_\theta(t-\sigma) e^{-\alpha(\theta)(t-\sigma)} \tilde{\nu}_\theta(d\sigma) \right)$$
$$\ge \alpha(\theta) + t^{-1}\log\left(K\int_0^T g_\theta(\sigma) e^{-\alpha(\theta)\sigma} d\sigma \right),$$

for large t, and the second term on the right tends to zero as $t \to \infty$, the result is proved in this case.

To deal with the cases where $m(\theta, \alpha(\theta)) < 1$ or $\alpha(\theta) = \infty$, truncate the point process Z (and hence μ) by discarding points at a distance of more than T (in space or time) from the origin and children beyond the Tth in any family. Then, in an obvious notation, $m_T(\theta, \phi)$ is always finite, so $\alpha_T(\theta)$ must satisfy $m_T(\theta, \alpha_T(\theta)) = 1$. Also, as $T \uparrow \infty$, $\alpha_T(\theta) \uparrow \alpha(\theta)$. Hence the result already proved applies to the truncated process and numbers in it are obviously dominated by those in the original process. Thus

$$\liminf_{t \to \infty} \frac{\log n_t(\theta)}{t} \geq \alpha_T(\theta) \uparrow \alpha(\theta),$$

completing the proof of (i). □

PROOF OF THEOREM 3(ii). Assume $\alpha(\theta) < \infty$, for otherwise there is nothing to prove. Just as in the proof of the first part,

$$t^{-1} \log n_t(\theta) = \alpha(\theta) + t^{-1} \log\left(\int g_\theta(t - \sigma) e^{-\alpha(\theta)(t-\sigma)} \tilde{\nu}_\theta(d\sigma)\right).$$

Taking an upper Riemann approximant and using the fact that the renewal measure is uniformly bounded on intervals of fixed length shows that, for some finite K, the integral on the right here is less than

$$K \sum_{i=0}^{[t]+1} \sup_{i \leq \sigma \leq i+1} g_\theta(\sigma) e^{-\alpha(\theta)\sigma}.$$

[This estimate also covers the case where, because $m(\theta, \alpha(\theta)) < 1$, the renewal measure is defective.] Thus

$$t^{-1} \log n_t(\theta) \leq \alpha(\theta) + t^{-1} \log\left(K \sum_{i=0}^{[t]+1} \sup_{i \leq \sigma \leq i+1} g_\theta(\sigma) e^{-\alpha(\theta)\sigma}\right) \to \alpha(\theta),$$

when (3.2) holds, using Lemma 2. □

6. Proof of Theorem 2.

PROOF OF THEOREM 2(i). Suppose that $\chi(t)$ is replaced by

$$\tilde{\chi}(t) := \chi(t) I(t < a) I(\chi(t) < K),$$

with a and K chosen large enough to ensure that $\int E\tilde{\chi}(\sigma)\, d\sigma$ is positive. As in Section 5, truncate the point process Z (and hence μ) by discarding births later than T and all births after the Tth in any family. Then $\tilde{m}_T(\phi)$ is always finite and α_T must satisfy $\tilde{m}_T(\alpha_T) = 1$. By construction, the truncated process has moments of all orders and $\tilde{\chi}$ is bounded. Hence Theorem 5.4 of Nerman (1981) applies to give

$$\frac{\log \xi_T^{\tilde{\chi}}(t)}{t} \to \alpha_T$$

on the survival set of the truncated process, almost surely. Since, for $T_1 < T_2$, $\xi_{T_1}^{\tilde{x}}(t) \le \xi_{T_2}^{\tilde{x}}(t) \le \xi^{\chi}(t)$ and $\alpha_T \uparrow \alpha$ as $T \uparrow \infty$, it follows that

$$\liminf_{t \to \infty} \frac{\log \xi^{\chi}(t)}{t} \ge \alpha \quad \text{a.s.}$$

whenever a truncated process survives. Finally, as T increases, the survival sets for the truncated processes increase to that of the original process. This completes the proof of (i). □

PROOF OF THEOREM 2(ii). By Lemma 3, that (2.3) holds implies that (2.3) also holds with χ^{ε} in place of χ. Thus, replacing χ by χ^{ε}, Theorem 1 implies that $t^{-1} \log E\xi^{\chi^{\varepsilon}}(t) \to \alpha$, so that, as $t \to \infty$ along any fine lattice,

$$\limsup \frac{\log \xi^{\chi^{\varepsilon}}(t)}{t} \le \alpha \quad \text{a.s.,}$$

by Borel–Cantelli. Take a lattice with span less than ε and say t_1 and t_2 are neighboring lattice points. It is a consequence of the definition of χ^{ε} that, for $t_1 < t \le t_2$, $\xi^{\chi}(t) \le \xi^{\chi^{\varepsilon}}(t_2)$; hence

$$\limsup_{t \to \infty} \frac{\log \xi^{\chi}(t)}{t} \le \alpha \quad \text{a.s.} \qquad \square$$

7. Proof of Theorem 4. The large deviation result needed to make the link between the information in Corollary 1 and the result required is discussed first.

Let $\{\zeta_n\}$ be measures with logarithmic transforms $\{k_n\}$, so

$$k_n(\theta) = \log \int e^{-\theta x} \zeta_n(dx).$$

Assume that for some fixed sequence of positive numbers tending to infinity, $\{a_n\}$,

$$\lim_{n \to \infty} \frac{k_n(\theta)}{a_n} = k(\theta), \tag{7.1}$$

with $k(\theta) < \infty$, for some $\theta < 0$. A simple Markov bound shows that when (7.1) holds,

$$\limsup_{n \to \infty} \frac{\log(\zeta_n[a_n x, \infty))}{a_n} \le k^*(x).$$

In the problem considered here this upper bound is complemented by a lower bound obtained by truncation. To describe this, say that $\{\zeta_{n,T}\}$ is a *sequence of good minorants* for $\{\zeta_n\}$ if, for all Borel sets A, $\zeta_{n,T}(A) \le \zeta_{n,T+1}(A) \uparrow \zeta_n(A)$, as $T \uparrow \infty$, and

$$\lim_{n \to \infty} \frac{k_{n,T}(\theta)}{a_n} = k_T(\theta),$$

with k_T differentiable. The following result is a fairly straightforward consequence of Theorem 2 of de Acosta, Ney and Nummelin (1991).

PROPOSITION 2. *Suppose that there is a sequence of good minorants with $k_T \uparrow k$. Then*

$$\lim_{n\to\infty} \frac{\log(\zeta_n[a_n x, \infty))}{a_n} = k^*(x)$$

except, possibly, for $x = \sup\{a\colon k^*(a) > -\infty\}$.

An example given in Biggins (1996) illustrates that the right endpoint of the domain of finiteness of k^* really can be an exceptional point here.

Theorem 4 will follow immediately from Corollary 1 and Proposition 2 once a sequence of good minorants is identified. The truncations used in Section 5 are the obvious tools to manufacture these. First truncate (M, χ) as described at (5.2), but with B_1, B_2 and B_3 replaced by T. Clearly the resulting η_t is a minorant for the original one. Now truncate the point process Z also, in the way described in the proof of Theorem 3(i). This produces $m_T(\theta, \phi)$, which is always finite, with $m_T(\theta, \alpha_T(\theta)) = 1$ and $\alpha_T(\theta) \uparrow \alpha(\theta)$. As m_T is analytic, the implicit function theorem guarantees that α_T is differentiable. Corollary 1 applies to the truncated process for each T, showing that their intensity measures do indeed form a sequence of good minorants.

8. Proof of Theorem 5. Consider first the case where $\alpha^*(a) < 0$ and $\kappa > \alpha^*(a)$. Replace (M, χ) by $(M^\varepsilon, \chi^\varepsilon)$. It follows from Lemma 3 that Theorem 4 applies to this ε upper approximant, giving the behavior of its expected numbers. Therefore, by Borel–Cantelli,

$$e^{-\kappa t} N_t^\varepsilon[ta, \infty) \to 0 \quad \text{a.s.}$$

along any fine lattice, and the upper approximant at lattice values dominates $N_t[ta, \infty)$ for all t, much as in the proof of Theorem 2, giving the result claimed. Similarly, when $\alpha^*(a) > 0$,

$$\limsup_{t\to\infty} \frac{\log N_t[ta, \infty)}{t} \le \alpha^*(a) \quad \text{a.s.}$$

To establish that $\alpha^*(a) > 0$ is also the lower bound is rather harder. The idea is to find a general branching process embedded in the original one that has the property that, for any included individual, $p_x \ge a\sigma_x$, and that has a Malthusian parameter near to $\alpha^*(a)$. Theorem 2, on the growth of a general branching process, can then be applied to this embedded process to get a lower bound on numbers in the original process to the right of ta at time t. However, the details are more complicated than this outline suggests.

Assume that a is in the interior of $\{a\colon \alpha^*(a) > -\infty\}$; otherwise, there is nothing to prove. Note that, as a concave function, $\alpha^*(a)$ is continuous on this set. Truncate Z (and all its copies) by discarding all births that occur in $(0, \Delta)$.

The truncated process will be denoted by Z^Δ. It is not too hard to check that as $\Delta \downarrow 0$, $\alpha_\Delta^*(b) \uparrow \alpha^*(b)$, for all b in the interior of $\{a: \alpha^*(a) > -\infty\}$. For $\delta > 0$ and small enough that $\alpha^*(a) - \delta > 0$, choose $b > a$ and Δ small such that

$$\alpha^*(a) - \delta < \alpha_\Delta^*(b) \le \alpha^*(a). \tag{8.1}$$

The continuity of $\alpha^*(a)$ ensures this is possible. Discard the original movement process and characteristic for a moment and use instead

$$M^\Delta(t) = bt, \qquad \chi^\Delta(t) = I(0 \le t < \Delta),$$

so, at time t, only those recently born (for they satisfy $t - \Delta < \sigma_x \le t$) are registered in the count. It is clear that $(Z^\Delta, M^\Delta, \chi^\Delta)$ is regulated, so Theorem 4 gives

$$\frac{\log(\eta_t^\Delta[tb, \infty))}{t} \to \alpha_\Delta^*(b).$$

Now, using this and (8.1), choose t sufficiently large that

$$e^{(\alpha^*(a)-\delta)t} \le \eta_t^\Delta[tb, \infty) \le e^{(\alpha^*(a)+\delta)t}. \tag{8.2}$$

For this t, identify the set of individuals, denoted by $\mathscr{S}_0$, counted by the new characteristic at time t and to the right of tb. More formally,

$$\begin{aligned}\mathscr{S}_0 &= \{x: p_x + M^\Delta(t - \sigma_x) \ge tb, \chi^\Delta(t - \sigma_x) = 1\} \\ &= \{x: p_x \ge b\sigma_x, t - \Delta < \sigma_x \le t\}.\end{aligned}$$

Note that because Z^Δ has no births in $(0, \Delta)$, all members of $\mathscr{S}_0$ are on distinct lines of descent. It is easy to see that $\mathscr{S}_0$ forms an optional line in the sense of Jagers (1989), and so, by Theorem 4.14 of that paper, independent copies of the original process emanate from each member of this line. Now define a general branching process by considering $\mathscr{S}_0$ to be the direct descendants of the initial ancestor. From the process emanating from each of these, construct independent copies of $\mathscr{S}_0$ to form the next generation and so on. Observe that in the original process, every person selected to be in the embedded process automatically satisfies $p_x \ge b\sigma_x$.

The next thing to do is to bound the Malthusian parameter α_e for the embedded process. Note that by arrangement, all birth times in its first generation are near t, so the simple bounds

$$e^{-\phi t}\eta_t^\Delta[tb, \infty) \le E\left[\sum_{x \in \mathscr{S}_0} e^{-\phi\sigma_x}\right] \le e^{-\phi(t-\Delta)}\eta_t^\Delta[tb, \infty)$$

combine with (8.2) to give

$$\alpha_e \ge \alpha^*(a) - \delta. \tag{8.3}$$

Now that a suitable embedded process has been identified, it is necessary to count it in a way that respects the movement and counting in the original process. To do this, let

$$\hat{\chi}(t) := I(M(t) \ge -B + bt)\chi(t),$$

where B is positive and large enough to ensure that $\int E\hat{\chi}(\sigma)\,d\sigma$ is positive. Because everyone born in the embedded process has $p_x \geq b\sigma_x$, the positions of those registered by this characteristic at time t satisfy

$$p_x + M_x(t - \sigma_x) \geq b\sigma_x - B + b(t - \sigma_x) = bt - B, \tag{8.4}$$

and in addition they have to count (as judged by χ) in the original process. Applying Theorem 2(i) to the resulting process, obtained by counting the embedded process by $\hat{\chi}$, gives

$$\liminf_{t \to \infty} \frac{\log \xi^{\hat{\chi}}(t)}{t} \geq \alpha_e \quad \text{a.s.} \tag{8.5}$$

on the survival set of the embedded process. Using (8.4), $\xi^{\hat{\chi}}(t) \leq N_t[tb - B, \infty)$ and, since $b > a$, $N_t[tb - B, \infty) \leq N_t[ta, \infty)$, for large t. Thus it follows from (8.3) and (8.5) that

$$\liminf_{t \to \infty} \frac{\log N_t[ta, \infty)}{t} \geq \alpha^*(a) - \delta \quad \text{a.s.} \tag{8.6}$$

when the embedded process survives.

In fact the estimate (8.6) holds whenever the original process survives. The idea for showing this is simple. Instead of starting the embedded process from the initial ancestor, start several of them from some later individuals. This will not disturb the estimates too much, but will increase the part of the sample space where they hold. More precisely, an embedded process may be started from any individual in the original process. Denote the one started from x by $\xi_x^{\hat{\chi}}(t)$. A little thought shows that

$$\xi_x^{\hat{\chi}}(t) \leq N_t[b(t - \sigma_x) + p_x - B, \infty),$$

and again, for large t,

$$N_t[b(t - \sigma_x) + p_x - B, \infty) \leq N_t[ta, \infty),$$

so that (8.6) holds whenever the embedded process emanating from x survives. On any optional line each individual produces an independent embedded process, and clearly (8.6) holds whenever at least one of the associated embedded processes survives. By taking a large enough optional line, the event that one of these embedded processes survives can be made as close to the survival set of the original process as desired. Hence (8.6) holds on the survival set of the original process. Because δ was arbitrary, this completes the proof. □

9. Proofs of Corollary 2 and Proposition 1.

PROOF OF COROLLARY 2. Because χ is a 0–1 characteristic, $N_t[ta, \infty)$ is an integer, so when $e^{-\kappa t} N_t[ta, \infty) \to 0$ almost surely with $\kappa < 0$, it follows that $N_t[ta, \infty) = 0$, for all sufficiently large t, almost surely. Thus, because $I(B_t \geq ta) \leq N_t[ta, \infty)$, Theorem 5(i) yields $\limsup B_t/t \leq \gamma$ almost surely.

Standard convexity theory [Rockafellar (1970), Theorems 12.2 and 27.1(a)] yields that the supremum of the concave function α^* is $\alpha(0)$, which is greater than zero because the process is supercritical. This implies that $\alpha^*(a) > 0$, for all $a < \gamma$. Thus, for $a < \gamma$, using Theorem 5(ii), $I(N_t[ta, \infty) > 0)$ is, for all large t, the survival set. Because $I(B_t \geq ta) = I(N_t[ta, \infty) > 0)$, this shows that $\liminf B_t/t \geq \gamma$ almost surely on the survival set. □

PROOF OF PROPOSITION 1. Suppose $\alpha^*(a) < 0$, so that, for some $\theta' < 0$, $\theta' a + \alpha(\theta') < 0$. Then, because $m(\theta', \phi)$ is monotone decreasing in ϕ, $1 \geq m(\theta', \alpha(\theta')) > m(\theta', -\theta' a)$, so $\inf\{m(\theta, -a\theta): \theta < 0\} < 1$. Similarly, if $\alpha^*(a) > 0$, then, for all $\theta < 0$, $\theta a + \alpha(\theta) > 0$ and then $m(\theta, -\theta a) > m(\theta, \alpha(\theta))$. This implies that $m(\theta, -\theta a) > 1$, for all $\theta < 0$, as required. [The case where $m(\theta, \alpha(\theta)) < 1$ gives $m(\theta, -\theta a) = \infty$.] □

10. The lattice case. If birth times are lattice and M and χ only change on the same lattice of time points, nearly everything becomes easier; in particular, ε-approximants can be dispensed with. However, one minor aspect becomes more complicated.

Without loss of generality the lattice can be taken to be the integers. When the embedded process is constructed, once $\Delta < 1$, all the members of $\mathscr{S}_0$ have a single birth time l. Then, when applying Theorem 2(i) to establish (8.5), convergence only holds on the sublattice $l\mathbb{N}$, rather than on the original lattice. In deducing that the set on which (8.6) holds can be expanded to the whole of the survival set, attention can be confined to optional lines drawn only from individuals with birth times in the sublattice $l\mathbb{N}$. In this way, (8.6) is shown to hold throughout the survival set as t goes to infinity through $l\mathbb{N}$. Using the same argument, but confining attention, for fixed k, to individuals with birth times in the sublattice $k + l\mathbb{N}$ establishes the required result as t goes to infinity through $k + l\mathbb{N}$. Putting these together, for $k = 0, 1, \dots, l-1$, gives convergence along the full lattice.

11. More general characteristics. The development of lower bounds in Theorem 5 applies without change if the characteristic χ is general rather than 0–1. However, Lemma 3, which is important in establishing upper bounds on the sample paths, uses the fact that χ is 0–1 in an essential way.

One extension is very easy. Suppose χ is general, but there is a 0–1 characteristic $\chi^\dagger$ such that, for some constant C, $\chi \leq C\chi^\dagger$ and $(Z, M, \chi^\dagger)$ is regulated. Then the upper bounds for $N_t^\dagger$ give similar bounds for N_t, and Theorem 5 still holds. Note, however, that the conditions no longer involve (Z, M, χ), but rather $(Z, M, \chi^\dagger)$.

From the mathematical point of view, it seems natural to combine M and χ into a single entity. One way to do this is to widen the definition of a random characteristic so that it encompasses both elements. Then χ will, for each t, code how an individual of age t contributes weight to various sets. Thus, in this formulation, $\chi(t, du)$ is a random measure and its integral

$\int \chi(t, du)$ is a random characteristic in the usual (i.e., temporal) sense. The random measure N_t is now defined by

$$N_t(du) = \sum_{x \in U} \chi_x(t - \sigma_x, du - p_x).$$

For the questions addressed here, χ should take values in the measures that have Laplace transforms that are finite for all $\theta \leq 0$, and should, as a function of t, have left and right limits, all of this form. Conditions like this will be needed to push through the analogue of the results in Section 4. The treatment of expected values and lower bounds on the sample paths should extend without too much trouble to this broader framework, with the key condition (3.2) being replaced by

$$E\left(\sup_t \left\{ \frac{e^{-\alpha(\theta)t} \int e^{-\theta z} \chi(t, dz)}{h(t)} \right\} \right) < \infty.$$

However, at this level of generality, I could not see how to avoid explicit conditions on ε-approximants to the sample paths to obtain upper bounds that hold for all time. The difficulty already arises at the start of this section, where conditions on $(Z, M, \chi^\dagger)$ would be needed to push through upper bounds. This seemed a heavy price to pay for the extra generality, so the less sophisticated framework of 0–1 characteristics and individual movement was used in the main results.

Acknowledgments. I would like to thank Peter Ney for bringing de Acosta, Ney and Nummelin (1991) to my attention and a referee for useful comments on presentation.

REFERENCES

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

BIGGINS, J. D. (1980). Spatial spread in branching processes. *Biological Growth and Spread. Lecture Notes in Biomath.* **38** 57–67. Springer, Berlin.

BIGGINS, J. D. (1996). How fast does a general branching random walk spread? In *Classical and Modern Branching Processes* (K. B. Athreya and P. Jagers, eds.). Springer, New York.

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

DE ACOSTA, A., NEY, P. and NUMMELIN, E. (1991). Large deviation lower bounds for general sequences of random variables. In *Random Walks, Brownian Motion and Interacting Particle Systems* (R. Durrett and H. Kesten, eds.) 215–221. Birkhäuser, Boston.

DEVROYE, L. (1990). On the height of random m-ary search trees. *Random Structures and Algorithms* **1** 191–203.

ELLIS, R. S. (1984). Large deviations for a general class of random vectors. *Ann. Probab.* **12** 1–12.

GÄRTNER, J. (1977). On large deviations from the invariant measure. *Theory Probab. Appl.* **22** 24–39.

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.

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

ROCKAFELLAR, R. T. (1970). *Convex Analysis*. Princeton Univ. Press.

VAN DEN BOSCH, F., METZ, J. A. J. and DIEKMANN, O. (1990). The velocity of spatial population expansion. *J. Math. Biol.* **28** 529–565.

PROBABILITY AND STATISTICS SECTION
SCHOOL OF MATHEMATICS AND STATISTICS
HICKS BUILDING
UNIVERSITY OF SHEFFIELD
SHEFFIELD, S3 7RH
UNITED KINGDOM