The Annals of Probability
1993, Vol. 21, No. 1, 290–317

A VARIATIONAL APPROACH TO BRANCHING RANDOM WALK IN RANDOM ENVIRONMENT

BY J.-B. BAILLON, PH. CLÉMENT, A. GREVEN[1] AND F. DEN HOLLANDER[2]

Université de Lyon I, Technische Universiteit Delft, Universität Göttingen and Rijksuniversiteit Utrecht

This paper considers an infinite system of particles on the integers $\mathbb{Z}$ that: (1) step to the right with a random delay, and (2) split or die along the way according to a random law depending on their position. The exponential growth rate of the particle density is computed in the long time limit in the form of a variational formula that can be solved explicitly. The result reveals two phase transitions associated with localization vs. delocalization and survival vs. extinction. In addition, the system exhibits an intermittency effect. Greven and den Hollander considered the more difficult situation where the particles may step both to the left and right, but the analysis of the phase diagram was less complete.

0. Introduction. Interacting particle systems evolving in random media display new phenomena that are not encountered in their spatially homogeneous counterparts. The reason is that the random medium induces its own *selection of paths* by favoring certain patterns in the evolution over those that are typical in the homogeneous medium. An extreme example of this situation is mean-field branching on a large finite set. Here each site carries a branching rate that is an unbounded random variable. Particles split into two at the local rate of their current position and randomly change site at rate 1. It turns out that as time becomes large most of the population sits in the record points of the medium and most of the ancestors of the particles in the record points were created there and have not moved since. Thus the overall picture is that an event which is rare in the homogeneous medium, namely most particles standing still, now is the dominating event for the population growth, simply because it is favored by the random medium. The favorable spots in the random medium, namely the record points, are rare themselves and so it is not the average medium that governs the growth [Fleischmann and Molchanov (1990)]. For an extensive list of references on interacting particle systems in random media, we refer to Greven and den Hollander (1992).

The interplay between process and medium has a variational character: Among competing patterns in the evolution, one best adapted to the medium is selected. We see it as the main task in this area to uncover such *underlying variational structures*, to formulate them in a precise analytical form and to

Received August 1990; revised August 1991.
[1]Supported by the Deutsche Forschungsgemeinschaft through the SFB 123 at Heidelberg.
[2]Supported by the Royal Netherlands Academy of Arts and Sciences.
AMS 1991 *subject classifications*. Primary 60K35; secondary 82C22.
Key words and phrases. Branching random walk, random media, large deviations, variational problems, phase transitions.

attempt to analyze their properties. In Greven and den Hollander (1992), such a program was carried out for one-dimensional branching random walk with drift. It was shown how to extract the selection mechanism for population growth in an analytical form using large deviation techniques. Analysis of the resulting variational formulas led to various interesting phase transitions coming from nonanalyticities in the variational solution as a function of the drift. The only paper we are aware of where a similar program has been carried out to the full extent is Eisele and Lang (1987). They study the survival probability of a particle moving through a random trap field, which leads mathematically to asymptotics of exponential functionals of the Wiener sausage with drift.

In the present paper we study a simplified version of the model in Greven and den Hollander (1992). The reason is twofold. First, the variational formula is easier to extract. Second, the analysis of the variational formula can be carried out in complete detail, leading to a precise description of the phase diagram. We shall find phenomena like *localization vs. delocalization*, *intermittency* and *survival vs. extinction*, all coming out of the selection mechanism induced by the random medium. We shall see that localization vs. delocalization is a phase transition corresponding to a parameter value where the solution of the variational formula changes from being attained *at the boundary* of its domain to being attained *in the interior*. We shall be able to relate intermittency to the situation where the solution is attained *outside* the domain. This means that the system has a degenerate limiting behavior. We believe that such degeneracy is a guiding principle behind intermittency effects [for a mathematical approach to intermittency, see Gärtner and Molchanov (1990)]. Finally, we shall see that survival vs. extinction is a phase transition associated with a change of sign in the value of the solution.

The paper is organized as follows. There are three main theorems in Sections 1.4, 2.3 and 3.2. Section 1 contains the derivation of the variational formula, Section 2 its analysis and Section 3 its interpretation, that is, its implications for the branching random walk.

1. Derivation of variational formula.

1.1. *Model.* With each $x \in \mathbb{Z}$ is associated a random probability measure F_x on the nonnegative integers $\mathbb{N} \cup \{0\}$, called the *offspring distribution* at site x. The sequence

$$F = \{F_x\}_{x \in \mathbb{Z}}$$

is i.i.d. with common distribution α. F plays the role of a *random environment*. For *given F*, define a discrete-time Markov process (η_n) on $(\mathbb{N} \cup \{0\})^{\mathbb{Z}}$, with

$$\eta_n = \{\eta_n(x)\}_{x \in \mathbb{Z}},$$

$$\eta_n(x) = \text{number of particles at site } x \text{ at time } n,$$

the evolution of which is as follows. At time $n = 0$ place one particle at every site, that is, $\eta_0(x) \equiv 1$. Given the state η_n at time n, each particle is *independently* replaced by a new generation. The size of a new generation descending from a particle at site x has distribution F_x, that is, consists of k new particles with probability $F_x(k)$, $k = 0, 1, 2, \dots$. Immediately after creation each new particle *independently* decides to either stay at the site where it is or to jump to the right neighboring site. The jump probability is h and is the same for all x. The resulting sequence of particle numbers make up the state η_{n+1} at time $n+1$, and so forth.

Thus at each unit of time all particles independently *branch* and *walk*: $\eta_{n+1}(x)$ depends on $\eta_n(x-1), \eta_n(x)$ and on the *given* F_{x-1}, F_x, namely, it is the sum of the number of particles in the $\eta_n(x-1)$ new generations at $x-1$ that decide to jump to x and the number of particles in the $\eta_n(x)$ new generations at x that decide to stay at x.

The parameter h is the *drift* of the random walk. It will be assumed that

$$\sum_{k=0}^{\infty} kF_x(k) \le M < \infty \quad \alpha\text{-a.s.} \tag{1.1}$$

Write

$$b_x = \sum_{k=0}^{\infty} kF_x(k) \tag{1.2}$$

for the *mean offspring* at site x and let β denote the distribution of b_x induced by α. By (1.1), β has bounded support. Write

$$M = \sup_x b_x = \text{Supremum of supp } \beta \tag{1.3}$$

and assume b_x is not constant.

1.2. *Particle density at time n.* For given F, let

$$D(F, \eta_n) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{x=-N}^{N} \eta_n(x)$$

denote the *particle density* at time n. Our main result in Section 1, Theorem 1, gives a variational representation for the growth rate of this quantity. For each fixed n the state η_n depends on F. However, if η_n is also considered as a random variable in F, then its distribution is stationary and ergodic under translations in $\mathbb{Z}$. Hence, from the individual ergodic theorem,

$$D(F, \eta_n) = E(\eta_n(0)) \quad \text{a.s.,} \tag{1.4}$$

where E denotes the double expectation over the Markov process (η_n) given F as well as over F. The a.s. refers here and in the sequel to the joint distribution of (η_n) and F. The stationarity and ergodicity are an easy consequence of the i.i.d. property of F, the initial choice $\eta_0(x) \equiv 1$ and the evolution mechanism. In fact, because particles can move no more than one lattice spacing per unit of time, $\eta_n(x)$ and $\eta_n(y)$ are independent when $|x-y| > n$ so that η_n is even mixing. Note that (1.1) ensures that $E(\eta_n(x)) \le M^n < \infty$.

1.3. *Local times of random walk.* Let

$$S_0 = 0, \qquad S_n = X_1 + \cdots + X_n, \qquad n \geq 1,$$
$$X_i \text{ i.i.d. with } P(X_i = 1) = 1 - P(X_i = 0) = h$$

denote a single copy of the random walk with drift h starting at 0. Let

$$l_n(x) = |\{0 < i \leq n\colon S_i = x\}|, \qquad x \geq 0, n \geq 0, \tag{1.5}$$

denote its *local time* at site x up to time n. The following proposition expresses $D(F, \eta_n)$ as the expectation of some exponential functional of the random sequence $\{l_n(x)\}_{x=0}^{\infty}$. Let

$$f(k) = \log \int b^k \beta(db), \qquad k \geq 0, \tag{1.6}$$

and let P_h and E_h denote probability and expectation over the random walk (S_n).

PROPOSITION 1.

$$D(F, \eta_n) = E_h\left(\exp\left[\sum_{x=0}^{\infty} f(l_n(x))\right]\right) \quad \text{a.s.} \tag{1.7}$$

PROOF. First condition on F and prove by an elementary induction argument that

$$E(\eta_n(0)|F) = E_h\left(\sum_{y=-\infty}^{0} \prod_{i=0}^{n-1} b_{y+S_i} 1_{\{y+S_n=0\}}\right)$$

with b_x given by (1.2). Here think of the ancestors at time $n = 0$ at positions $y \leq 0$ and of how their progeny may reach $x = 0$ at time n. Next reverse time in the random walk to obtain

$$\begin{aligned} E(\eta_n(0)|F) &= E_h\left(\sum_{y=-\infty}^{0} \prod_{i=1}^{n} b_{-S_i} 1_{\{-S_n=y\}}\right) \\ &= E_h\left(\prod_{i=1}^{n} b_{-S_i}\right) \\ &= E_h\left(\prod_{x=0}^{\infty} [b_{-x}]^{l_n(x)}\right). \end{aligned}$$

The last step uses (1.5). Finally, take the expectation over the i.i.d. b_x, using (1.4) and (1.6). □

1.4. *Exponential growth rate.* Having thus reduced $D(F, \eta_n)$ to a problem involving local times of a random walk, we are now ready to formulate the variational formula that is the main subject in this paper and that will be analyzed and interpreted in detail in Sections 2 and 3. Let $\mathscr{P}(\mathbb{N})$ denote the set of probability measures on the positive integers $\mathbb{N}$. Let

$$(1.8)\qquad M_\theta = \Big\{\nu \in \mathscr{P}(\mathbb{N}) \colon \sum_{i \ge 1} i\nu(i) = \theta^{-1}\Big\}, \qquad \theta \in (0,1],$$

$$(1.9)\qquad I_\theta(\nu) = \sum_{i\ge 1} \nu(i) \log\left(\frac{\nu(i)}{\pi_\theta(i)}\right), \qquad \nu \in \mathscr{P}(\mathbb{N}),$$

$$(1.10)\qquad I_h(\theta) = \theta \log\left(\frac{\theta}{h}\right) + (1-\theta)\log\left(\frac{1-\theta}{1-h}\right), \qquad \theta \in [0,1],\ h \in (0,1),$$

$$(1.11)\qquad \pi_\theta(i) = \theta(1-\theta)^{i-1}, \qquad \theta \in (0,1],\ i \ge 1.$$

Write $\langle f, \nu\rangle = \sum_{i\ge 1} f(i)\nu(i)$ for the inner product.

THEOREM 1. *In the long time limit the particle density grows exponentially fast at rate*

$$(1.12)\qquad \lim_{n\to\infty} \frac{1}{n} \log D(F, \eta_n) = \rho(h) \quad a.s.$$

For $h \in (0,1)$,

$$(1.13)\qquad \rho(h) = \sup_{\theta\in(0,1]} \sup_{\nu \in M_\theta} \{\theta[\langle f,\nu\rangle - I_\theta(\nu)] - I_h(\theta)\}.$$

For $h = 0$ *and* $h = 1$,

$$(1.14)\qquad \begin{aligned} \rho(0) &= \log M,\\ \rho(1) &= f(1) = \log \textstyle\int b\beta(db). \end{aligned}$$

The growth rate depends on β and h. We write $\rho(h)$ because we shall be mainly interested in its dependence on h for fixed choice of β. The two boundary cases $\rho(0)$ and $\rho(1)$ are trivial because $l_n(x) = n 1_{\{0\}}(x)$ if $h = 0$ and $l_n(x) = 1_{(0,n]}(x)$ if $h = 1$, so that (1.14) follows by substitution into (1.7). Note that

$$\lim_{k\to\infty} f(k)/k = \log M$$

by (1.3) and (1.6).

1.5. *Proof of Theorem* 1: *Large deviations.* The proof consists of several steps and follows Greven and den Hollander (1992).

Our first step is to condition on the position of the random walk:

$$(1.15)\qquad \begin{aligned} &E_h\left(\exp\left[\sum_{x=0}^{\infty} f(l_n(x))\right]\right)\\ &\quad = \int_{\theta\in[0,1)} d(\theta n) P_h(S_n = [\theta n]) E_h\left(\exp\left[\sum_{x=0}^{\infty} f(l_n(x))\right] \Bigg| S_n = [\theta n]\right)\\ &\qquad + h^n \exp[nf(1)]. \end{aligned}$$

The last term corresponds to the event $\{S_n = n\}$. Next observe that the conditional expectation in the integrand is *independent* of h. The latter holds because all walks from 0 to $[\theta n]$ in n steps have the same probability $h^{[\theta n]}(1-h)^{n-[\theta n]}$ (so that the conditional expectation reduces to a combinatorial average). Replace E_h by E_θ and now assume that the following holds.

PROPOSITION 2. *There exists a bounded and continuous function $J\colon [0,1] \to \mathbb{R}$ such that*

$$(1.16)\qquad \lim_{n\to\infty} \frac{1}{n} \log E_\theta\left(\exp\left[\sum_{x=0}^{\infty} f(l_n(x))\right]\middle| S_n = [\theta n]\right) = J(\theta)$$

and the same limit is obtained along any sequence $\theta_n \to \theta$.

From Proposition 2 we proceed as follows.

PROPOSITION 3. *For every $h \in (0,1)$,*

$$(1.17)\qquad \lim_{n\to\infty} \frac{1}{n} \log E_h\left(\exp\left[\sum_{x=0}^{\infty} f(l_n(x))\right]\right) = \sup_{\theta\in[0,1]} (J(\theta) - I_h(\theta)).$$

PROOF. For every $h \in (0,1)$,

$$P_h(S_n = [\theta n]) = \binom{n}{[\theta n]} h^{[\theta n]}(1-h)^{n-[\theta n]}$$

is, by inspection, a large deviation family on $\mathbb{R}$ with rate function

$$(1.18)\qquad \lim_{n\to\infty} \frac{1}{n} \log P_h(S_n = [\theta n]) = -I_h(\theta)$$

given by (1.10). Since $I_h(\theta)$ is bounded and continuous in θ and since the same limit is obtained along any sequence $\theta_n \to \theta$, (1.17) follows by combining Proposition 2 and (1.18) and by applying Varadhan's theorem to (1.15) [see Deuschel and Stroock (1989), Theorem 2.1.10]. Note that (1.17) includes the boundary case $\theta = 1$ coming from the last term in (1.15) [$J(1) = f(1)$, $I_h(1) = -\log h$]. □

Propositions 1 and 3 prove Theorem 1, with the growth rate in (1.13) for $h \in (0,1)$ given by

$$(1.19)\qquad \rho(h) = \sup_{\theta\in[0,1]} (J(\theta) - I_h(\theta)).$$

Thus, to complete the proof of Theorem 1, our task is to prove Proposition 2 and to identify $J(\theta)$ in (1.16) and (1.19) as

$$(1.20)\qquad J(\theta) = \theta \sup_{\nu\in M_\theta} [\langle f, \nu\rangle - I_\theta(\nu)], \qquad \theta \in (0,1],$$

with M_θ and $I_\theta(\nu)$ given by (1.8), (1.9) and (1.11). A small technical point is that (1.20) is not defined for $\theta = 0$. However, we know that

$$J(0) = \log M \tag{1.21}$$

by the observation made below Theorem 1. The continuity of $J(\theta)$ at $\theta = 0$, claimed in Proposition 2, shows that the supremum in (1.19) may in fact be taken over $\theta \in (0,1]$, as was claimed in (1.13). Note also that $I_h(\theta)$ is continuous at $\theta = 0$. □

PROOF OF PROPOSITION 2. First assume $\theta \in (0,1]$. Observe that

$$\begin{aligned} &E_\theta\left(\exp\left[\sum_{x=0}^{\infty} f(l_n(x))\right]\middle| S_n = [\theta n]\right) \\ &\quad = E_\theta\left(\exp\left[\sum_{x=0}^{\infty} f(l_n(x))\right]\middle| S_n = [\theta n],\, S_{n+1} > [\theta n]\right) \\ &\quad = E_\theta\left(\exp\left[\sum_{x=0}^{[\theta n]} f(l(x))\right]\middle| \sum_{x=0}^{[\theta n]} l(x) = n\right). \end{aligned} \tag{1.22}$$

In the second equality we introduce the *total local time*

$$l(x) = \lim_{n\to\infty} l_n(x) \tag{1.23}$$

[see (1.5)] and we make essential use of the fact that the walk cannot step to the left.

Next, let

$$\nu_N = N^{-1} \sum_{x=0}^{N-1} \delta_{l(x)}$$

denote the *empirical distribution* of the total local time process $\{l(x)\}_{x=0}^{\infty}$ in the interval $[0, N)$. Then we can write

$$\begin{aligned} &E_\theta\left(\exp\left[\sum_{x=0}^{[\theta n]} f(l(x))\right]\middle| \sum_{x=0}^{[\theta n]} l(x) = n\right) \\ &\quad = E_\theta\left(\exp\left[K_n\langle f, \nu_{K_n}\rangle\right]\middle| \langle 1, \nu_{K_n}\rangle = L_n\right) \end{aligned} \tag{1.24}$$

with the abbreviations $K_n = [\theta n] + 1$, $L_n = n/K_n = n/([\theta n]+1)$ and 1 the identity function $1(i) = i$. Having thus, via (1.22) and (1.24), rewritten the l.h.s. of (1.16) as a conditional expectation of some exponential functional of ν_{K_n}, we are now ready to again apply a large deviation argument.

The key observation is that our total local time process is i.i.d. with common distribution π_θ given by (1.11) (recall that under P_θ the drift is θ). Hence by Sanov's theorem $P_\theta(\nu_N \in d\nu)$ is a large deviation family on $\mathscr{P}(\mathbb{N})$ with rate function $I_\theta(\nu)$ given by (1.9) [see Deuschel and Stroock (1989), Theorem 3.2.17]. Define the set

$$A_n = \{\nu \in \mathscr{P}(\mathbb{N}) : \langle 1, \nu\rangle = L_n\}$$

and write the r.h.s. of (1.24) as

$$\exp(o(n))\int_{A_n}\exp[K_n\langle f,\nu\rangle]P_\theta(\nu_{K_n}\in d\nu). \tag{1.25}$$

The error term is just $P_\theta^{-1}(\langle 1,\nu_{K_n}\rangle = L_n) = P_\theta^{-1}(S_n = [\theta n],\ S_{n+1} > [\theta n])$ and is subexponential because of (1.18) and $I_\theta(\theta) = 0$. [This is why we replaced E_h by E_θ in (1.15).] We want to apply Varadhan's theorem to the integral in (1.25) and conclude, with A_n "converging" to M_θ [note $L_n \to \theta^{-1}$ and recall (1.8)], that

$$\lim_{K_n\to\infty}\frac{1}{K_n}\log(1.25) = \sup_{\nu\in M_\theta}[\langle f,\nu\rangle - I_\theta(\nu)] \tag{1.26}$$

and thus achieve our goal (1.20) (note $K_n \sim \theta n$). However, there are two technical complications: (1) $\langle f,\nu\rangle$ is not continuous in ν, and (2) A_n is moving and is not closed (in the weak topology).

The first problem may be remedied by defining

$$g(k) = k\log M - f(k) \tag{1.27}$$

and replacing in (1.25)

$$\exp[K_n\langle f,\nu\rangle] = M^n\exp[-K_n\langle g,\nu\rangle], \qquad \nu\in A_n.$$

Indeed, since (1.3) and (1.6) give that g is sublinear, that is,

$$\lim_{k\to\infty} g(k)/k = 0, \tag{1.28}$$

we have $\langle g,\nu\rangle$ continuous in ν on A_n.

The second problem is harder to solve. Here the way to proceed is to truncate the local times and to thicken A_n to a thin slab. More precisely, one defines

$$M_\theta^{\varepsilon,R} = \{\nu\in\mathscr{P}(\{1,\dots,R\})\colon \langle 1,\nu\rangle\in[\theta^{-1}-\varepsilon,\theta^{-1}+\varepsilon]\}, \qquad \varepsilon>0,\ R<\infty,$$

and one shows that there exists $c^{\varepsilon,R}>0$, with $c^{\varepsilon,R}\to 0$ as $\varepsilon\to 0$ and $R\to\infty$, such that the integral in (1.25) when taken over $M_\theta^{\varepsilon,R}$ instead of A_n differs by at most a factor $\exp(c^{\varepsilon,R}n)$. To obtain the latter estimate, we refer the reader to Lemmas 6 and 7 in Greven and den Hollander (1992). The proof of these lemmas uses (1.28) and the following properties of g that follow from (1.3) and (1.6):

$$g(0) = 0, \quad g \text{ is strictly increasing and strictly concave.} \tag{1.29}$$

Since $M_\theta^{\varepsilon,R}$ is a closed set and since $I_\theta(\nu)$ is continuous on this set, we may apply Varadhan's theorem to the integral over $M_\theta^{\varepsilon,R}$ [see Deuschel and Stroock (1989), Theorem 2.1.10] and deduce

$$\begin{aligned}\lim_{K_n\to\infty}\frac{1}{K_n}\log(1.25) &\doteq O(\theta^{-1}c^{\varepsilon,R}) + \theta^{-1}\log M\\ &\quad + \sup_{\nu\in M_\theta^{\varepsilon,R}}[-\langle g,\nu\rangle - I_\theta(\nu)].\end{aligned} \tag{1.30}$$

Now, it is easily checked that the closure of $\bigcup_{R<\infty}\bigcap_{\varepsilon>0} M_\theta^{\varepsilon, R}$ is M_θ. This implies

$$\sup_{R<\infty} \inf_{\varepsilon>0} \sup_{\nu\in M_\theta^{\varepsilon, R}} \left[-\langle g,\nu\rangle - I_\theta(\nu)\right] = \sup_{\nu\in M_\theta} \left[-\langle g,\nu\rangle - I_\theta(\nu)\right]$$

by appealing to the continuity on M_θ of $\langle g,\nu\rangle$ and $I_\theta(\nu)$ (see Lemma 2). Finally, substitute

$$\langle g,\nu\rangle = \theta^{-1}\log M - \langle f,\nu\rangle, \qquad \nu\in M_\theta,$$

to obtain (1.26). This completes the proof of (1.16) and (1.20) for $\theta\in(0,1]$.

The last statement of Proposition 2 is an easy perturbation argument in K_n and L_n in (1.24). Namely, for $\theta\in(0,1]$ use the estimate mentioned below (1.28). For $\theta=0$ use (1.21), $f(k)\le k\log M$ and $\lim_{k\to\infty} f(k)/k = \log M$.

The remaining claims in Proposition 2, namely the boundedness and continuity of $J(\theta)$ on $[0,1]$, will follow from our analysis in Section 2. □

1.6. *Remarks.* In Greven and den Hollander (1992), the same model is considered except that the particles may jump both to left and right neighboring sites. A result like Theorem 1 is proved there, but the variational formula and its proof turn out to be considerably more involved. Part of the trouble comes from the fact that if the random walk can move both ways, then the total local time process $\{l(x)\}_{x=0}^{\infty}$ is no longer independent. In fact, it is not even Markov, but turns out to be a two-block functional of a Markov process. This causes trouble in establishing the appropriate large deviation principle and the resulting variational formula has a much more complex structure involving empirical distributions of pairs rather than singletons. There are several other complications. For instance, (1.22) no longer holds. How close are $l_n(x)$ and $l(x)$ for large n and x? Some work is needed to show that (1.22) holds asymptotically.

These problems motivated our study of the simpler model in the present paper. The proof in Section 1.5 has the benefit of being short and transparent, but more importantly, in Section 2 we shall be able to solve the variational formula (1.13) explicitly. This will lead us to a detailed description of *various phase transitions* associated with the particle behavior, as will be explained in Section 3. Similar results were also found for the two-sided model, but there the analysis of the phase diagram was inhibited by the lack of an explicit solution and therefore was less complete.

The two models differ in the way the phase diagram depends on the distribution β for the mean offspring. For example, in the one-sided model localization can only occur when β decays sufficiently fast close to the maximal value M, whereas in the two-sided model it always occurs.

2. Analysis of variational formula. The analysis of (1.19) and (1.20) proceeds in two steps: (i) supremum over ν; (ii) supremum over θ. The first is the essential step toward proving our main result in Section 2, Theorem 2.

2.1. *Supremum over ν.* We start by recalling (1.27)–(1.29), that is, $g(k) = k\log M - f(k)$ and

$$\lim_{k\to\infty} g(k)/k = 0, \tag{2.1}$$

$$g(0) = 0, \quad g \text{ is strictly increasing and strictly concave.} \tag{2.2}$$

Next use (1.8), (1.9) and (1.11) to rewrite (1.20) as

$$J(\theta) = \log M + \theta\log\theta + (1-\theta)\log(1-\theta) - \theta K(\theta), \tag{2.3}$$

where we introduce

$$K(\theta) = \inf_{\nu\in M_\theta}\left[\sum_{i\geq 1} g(i)\nu(i) + \sum_{i\geq 1}\nu(i)\log\nu(i)\right], \qquad \theta\in(0,1]. \tag{2.4}$$

The following proposition is the main part of our variational analysis and will be proved in Section 2.4. The proof uses only (2.1). Let

$$G(r) = \sum_{i\geq 1} e^{-ri-g(i)}, \qquad r\geq 0 \tag{2.5}$$

(see Lemma 10 in Section 3.3).

PROPOSITION 4. *$K(\theta)$ is well defined, continuous and nondecreasing on $(0,1]$. Furthermore, $\theta K(\theta)$ is convex on $(0,1]$, with $\lim_{\theta\downarrow 0}\theta K(\theta) = 0$ and $\lim_{\theta\uparrow 1}\theta K(\theta) = K(1) = g(1)$. Define*

$$\theta_c = -\frac{G(0)}{G'(0)} \in [0,1). \tag{2.6}$$

$K(\theta)$ is strictly increasing and analytic on $(\theta_c, 1)$ and is constant on $(0,\theta_c)$.

(i) *If $\theta\in[\theta_c,1]\cap(0,1]$, then the minimum in (2.4) is achieved in M_θ. The minimizer $\bar\nu = \bar\nu(\theta)$ is unique and is given by*

$$\bar\nu(i) = \frac{1}{G(r)}e^{-ri-g(i)}, \qquad i\geq 1, \tag{2.7}$$

with $r = r(\theta)$ the unique solution of

$$\theta = -\frac{G(r)}{G'(r)}, \tag{2.8}$$

and the minimum is

$$K(\theta) = -\frac{r}{\theta} + \log\frac{1}{G(r)}. \tag{2.9}$$

(ii) *If $\theta\in(0,\theta_c)$, then the minimum in (2.4) is not achieved in M_θ. The minimizer is given by*

$$\bar\nu(i) = \frac{1}{G(0)}e^{-g(i)}, \qquad i\geq 1,$$

which is in M_{θ_c}, *and the minimum is*

$$K(\theta) = K(\theta_c) = \log \frac{1}{G(0)}.$$

The function $r \to -G(r)/G'(r)$ is analytic, strictly positive and strictly increasing on $(0, \infty)$ and has range $(\theta_c, 1)$ (see Lemma 7 below). By θ_c in (2.6) we mean its limit as $r \downarrow 0$. A similar convention is used in (2.13) and (2.19). Note that $\theta_c = 0$ if and only if $G'(0) = -\infty$.

2.2. *Supremum over* θ. Fix $h \in (0, 1)$ and define

$$J_h(\theta) = J(\theta) - I_h(\theta), \tag{2.10}$$

so that, from (1.19),

$$\rho(h) = \sup_{\theta \in [0,1]} J_h(\theta). \tag{2.11}$$

Put (1.10) and (2.3) together to obtain

$$\begin{aligned} J_h(\theta) &= \log[M(1-h)] - \theta \log\left(\frac{1-h}{h}\right) - \theta K(\theta), \qquad \theta \in (0,1], \\ J_h(0) &= \log[M(1-h)]. \end{aligned} \tag{2.12}$$

The following proposition will be proved in Section 2.5.

PROPOSITION 5. *For* $h \in (0, 1)$, $J_h(\theta)$ *is continuous and concave in* θ *on* $[0, 1]$, *with* $J_h(0) = \log[M(1-h)]$ *and* $J_h(1) = f(1) + \log h$. *Furthermore,* $J_h(\theta)$ *is strictly concave and analytic on* $(\theta_c, 1)$ *and is linear on* $(0, \theta_c)$, *where* θ_c *is defined in* (2.6). *Define*

$$h_c = \frac{1}{1 + G(0)} \in [0, 1). \tag{2.13}$$

(i) *If* $h < h_c$, *then* $J_h(\theta)$ *is strictly decreasing on* $[0, 1]$ *and* $\rho(h) = J_h(0)$ *with unique maximizer* $\bar{\theta} = 0$.

(ii) *If* $h = h_c$, *then* $J_h(\theta)$ *is constant on* $(0, \theta_c)$ *and is strictly decreasing on* $(\theta_c, 1)$. *The maximizer is not unique, but again* $\rho(h) = J_h(0)$.

(iii) *If* $h > h_c$, *then* $J_h(\theta)$ *has strictly positive slope at* $\theta = 0$ *and achieves a unique maximum in* $(\theta_c, 1)$. *The maximizer is*

$$\bar{\theta} = -\frac{G(r)}{G'(r)} \tag{2.14}$$

with $r = r(h)$ *the unique solution of*

$$h = \frac{1}{1 + G(r)}, \tag{2.15}$$

and the maximum is

$$\rho(h) = J_h(0) + r. \tag{2.16}$$

Note that $h_c = 0$ if and only if $G(0) = \infty$.

2.3. *Implications for* $\rho(h)$. From Propositions 4 and 5 we now get a clear picture of how our growth rate $\rho(h)$ in Theorem 1 depends on the drift h.

THEOREM 2. (i) *The growth rate is given by*

$$\begin{aligned} \rho(h) &= \log[M(1-h)] && \text{if } 0 \le h \le h_c, \\ &= \log[M(1-h)] + r && \text{if } h_c < h < 1, \end{aligned} \tag{2.17}$$

with $r = r(h)$ *the unique solution of* (2.15). *Moreover*, $\rho(h)$ *is continuous and strictly decreasing on* $[0, 1]$, *with* $\rho(1) = f(1)$, *and is analytic on* $(0, h_c)$ *and on* $(h_c, 1)$.

(ii) *The maximizer* $\bar\theta = \bar\theta(h)$ *satisfies*

$$\begin{aligned} \bar\theta &= 0 && \text{if } 0 < h < h_c, \\ \theta_c < \bar\theta &< h && \text{if } h_c < h < 1. \end{aligned} \tag{2.18}$$

Moreover, $\bar\theta(h)$ *is strictly increasing and analytic on* $(h_c, 1)$.

(iii) *If* $h_c > 0$, *then*

$$\rho'(h_c+) - \rho'(h_c-) = -\frac{(1+G(0))^2}{G'(0)} = \frac{\theta_c}{h_c(1-h_c)}. \tag{2.19}$$

(iv) *If* $\log M > 0 > f(1)$, *then* $\rho(h)$ *changes sign at* $h = h_c^*$, *the unique solution of* $\rho(h) = 0$ *computable from* (2.17).

Theorem 2 is the key result about our particle system (see Figures 1 and 2). The proof will be given in Section 2.6. For further discussion and interpretation we refer to Section 3.

2.4. *Proof of Proposition* 4. This is the most involved part of Section 2. The argument will be broken up into several propositions and lemmas. It will be convenient to reformulate (2.4) as a problem of minimization on some appropriate compact convex set in $l^1(\mathbb{N})$ (l^1 is the set of absolutely summable sequences). Since we have information about $g(i)/i$ by (2.1), we define

$$\mu(i) = i\nu(i), \qquad i \ge 1.$$

For $\theta \in (0, 1]$ define the following subsets of $l^1(\mathbb{N})$:

$$\Phi(\theta) = \left\{\mu(i) \ge 0, \sum_{i\ge1} i^{-1}\mu(i) = 1, \sum_{i\ge1}\mu(i) \le \theta^{-1}\right\}, \tag{2.20}$$

$$\partial\Phi(\theta) = \left\{\mu(i) \ge 0, \sum_{i\ge1} i^{-1}\mu(i) = 1, \sum_{i\ge1}\mu(i) = \theta^{-1}\right\}. \tag{2.21}$$

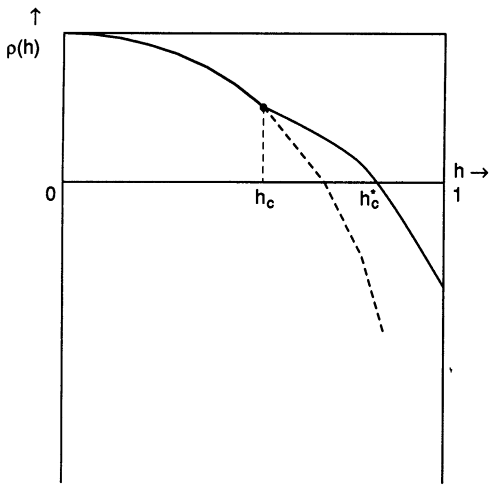

FIG. 1. *Qualitative picture of* $\rho(h)$ *as a function of* h. *The dashed curve is* $\log[M(1-h)]$. *The endpoints are* $\rho(0) = \log M$ *and* $\rho(1) = \log \int b\beta(db)$.

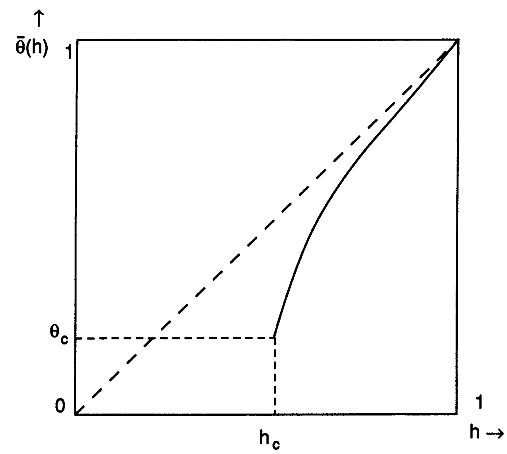

FIG. 2. *Qualitative picture of* $\bar{\theta}(h)$ *as a function of* h. *The dashed curve is the diagonal.*

Let

$$\varphi(\mu) = \sum_{i \geq 1} i^{-1}\mu(i)g(i) + \sum_{i \geq 1} i^{-1}\mu(i)\log\left[i^{-1}\mu(i)\right]. \tag{2.22}$$

Then (2.4) reads

$$K(\theta) = \inf_{\mu \in \partial\Phi(\theta)} \varphi(\mu). \tag{2.23}$$

PROPOSITION 6. *For* $\theta \in (0,1]$,

$$K(\theta) = \min_{\mu \in \Phi(\theta)} \varphi(\mu). \tag{2.24}$$

PROOF. The proof is done in several steps. Let $\overline{B}$ be the closed unit ball of l^1. We recall that a sequence (μ_n) in $\overline{B}$ converges weakly* to some $\mu \in \overline{B}$, written $\mu_n \xrightarrow{*} \mu$ and meaning

$$\langle \mu_n, \tau \rangle \to \langle \mu, \tau \rangle \quad \text{for all } \tau \in c_0(\mathbb{N})$$

(c_0 is the set of sequences tending to 0; l^1 is the dual of c_0), if and only if it converges componentwise, that is,

$$\mu_n(i) \to \mu(i) \quad \text{for all } i \in \mathbb{N}.$$

Moreover, we recall that $\overline{B}$ is weak* compact and that the weak* topology in $\overline{B}$ is metrizable [see Rudin (1973), 3.14–3.16].

LEMMA 1. *For every* $\theta \in (0,1]$, $\Phi(\theta)$ *in* (2.20) *is convex, weak* compact in* l^1 *and is the weak* closure of* $\partial\Phi(\theta)$ *in* (2.21):

$$\Phi(\theta) = \overline{\partial\Phi(\theta)}^{\,*}.$$

PROOF. Clearly $\Phi(\theta)$ is convex. To show that $\Phi(\theta)$ is weak* closed in l^1, let $\mu_n \in \Phi(\theta)$ and $\mu \in l^1$ be such that $\mu_n(i) \to \mu(i)$ for all i. Then $\mu(i) \geq 0$, $\sum_i \mu(i) \leq \liminf_{n\to\infty} \sum_i \mu_n(i) \leq \theta^{-1}$ by Fatou, and

$$\sum_i i^{-1}\mu(i) = \lim_{n\to\infty} \sum_i i^{-1}\mu_n(i) = 1 \quad \text{since } (i^{-1}) \in c_0.$$

Hence $\mu \in \Phi(\theta)$. To show that $\partial\Phi(\theta)$ is weak* dense in $\Phi(\theta)$, we want to find for any $\mu \in \Phi(\theta)$ a sequence $\mu_n \in \partial\Phi(\theta)$ such that $\mu_n \xrightarrow{*} \mu$. Choose $\mu_n = t_n\mu + s_n\varepsilon_n$, with $\varepsilon_n(i) = \delta_{ni}$ the usual basis in l^1,

$$t_n = \frac{n - \theta^{-1}}{n - \theta_0^{-1}}, \qquad s_n = \frac{n(\theta^{-1} - \theta_0^{-1})}{n - \theta_0^{-1}},$$

where $\theta_0^{-1} = \sum_i \mu(i)$ and $n > \theta^{-1} \geq \theta_0^{-1}$. We have

$$t_n \to 1, \qquad s_n \to \theta^{-1} - \theta_0^{-1}, \qquad \varepsilon_n \xrightarrow{*} (0),$$

and hence $\mu_n \xrightarrow{*} \mu$. □

LEMMA 2. *For every* $\theta \in (0,1]$, φ: $\Phi(\theta) \to \mathbb{R}$ *in* (2.22) *is well defined, strictly convex and weak* continuous in* l^1.

PROOF. Write φ as a sum of three functions $\varphi = \varphi_1 + \varphi_2 + \varphi_3$, where

$$\varphi_1(\mu) = \sum_i i^{-1}\mu(i)g(i),$$

$$\varphi_2(\mu) = \sum_i i^{-1}\mu(i)\log\mu(i),$$

$$\varphi_3(\mu) = -\sum_i i^{-1}\mu(i)\log i.$$

The functions φ_1 and φ_3 are linear in μ and weak* continuous since $(i^{-1}g(i))$ and $(i^{-1}\log i)$ belong to c_0. The function φ_2 is well defined since

$$\begin{aligned}\sum_i i^{-1}\mu(i)|\log\mu(i)| &\le \Big(\sum_i i^{-2}\Big)^{1/2}\Big(\sum_i \mu(i)\Big)^{1/2} \sup_i \mu^{1/2}(i)|\log\mu(i)|\\ &\le \frac{\pi}{\sqrt{6}}\frac{1}{\theta^{1/2}}\frac{2}{e} < \infty\end{aligned}$$

by Cauchy—Schwarz and the fact that $\sup_i \mu^{1/2}(i)|\log\mu(i)| \le \max_{t\in[0,1]} t^{1/2}|\log t| = 2/e$. Moreover, φ_2 is strictly convex since the function $t \to t\log t$ is strictly convex on $[0,1]$. To prove that φ_2 is weak* continuous, observe that $(i^{-1}) \in l^2(\mathbb{N})$. Hence it is enough to show that $\mu_n \xrightarrow{*} \mu \in \Phi(\theta)$ implies

$$\mu_n \log\mu_n \to \mu\log\mu \quad \text{weakly in } l^2.$$

For this to be true it is sufficient to have componentwise convergence and boundedness in l^2. But

$$\sum_i \mu_n^2(i)\log^2\mu_n(i) \le \Big(\sum_i \mu_n(i)\Big)\sup_i \mu_n(i)\log^2\mu_n(i) \le \frac{4}{\theta e^2} < \infty. \qquad \square$$

Now we can prove Proposition 6. Since $\Phi(\theta)$ is weak* compact and φ is weak* continuous, φ achieves its minimum on $\Phi(\theta)$. Moreover,

$$\begin{aligned}\varphi(\Phi(\theta)) &= \varphi\big(\overline{\partial\Phi(\theta)}^{*}\big) \subseteq \overline{\varphi(\partial\Phi(\theta))}\\ \varphi(\Phi(\theta)) &= \overline{\varphi(\Phi(\theta))}\\ \overline{\varphi(\partial\Phi(\theta))} &\subseteq \overline{\varphi(\Phi(\theta))}.\end{aligned} \tag{2.25}$$

Hence $\varphi(\Phi(\theta)) = \overline{\varphi(\partial\Phi(\theta))}$, which implies

$$\inf_{\mu\in\partial\Phi(\theta)} \varphi(\mu) = \min_{\mu\in\Phi(\theta)} \varphi(\mu). \qquad \square$$

From (2.24) we get the following properties of $K(\theta)$.

PROPOSITION 7. *K: $(0,1] \to \mathbb{R}$ is continuous and nondecreasing. Furthermore, $\theta K(\theta)$ is convex in θ. Moreover, there exists exactly one minimizer $\bar{\mu}(\theta) \in \Phi(\theta)$ in* (2.24).

PROOF. The uniqueness of the minimizer comes from the strict convexity of φ. Next let $0 < \theta_1 < \theta_2 < 1$. Then $\Phi(\theta_1) \supseteq \Phi(\theta_2)$, hence $K(\theta_1) \le K(\theta_2)$. Let $\mu_i \in \Phi(\theta_i)$, $i = 1, 2$, and $0 < t < 1$. Then

$$\begin{aligned}&(1-t)\mu_1 + t\mu_2 \in \Phi\Big(\big[(1-t)\theta_1^{-1} + t\theta_2^{-1}\big]^{-1}\Big),\\ &\varphi\big((1-t)\mu_1 + t\mu_2\big) \le (1-t)\varphi(\mu_1) + t\varphi(\mu_2).\end{aligned}$$

Hence with the choice $\mu_i = \bar{\mu}(\theta_i)$:

$$K\Big(\big[(1-t)\theta_1^{-1} + t\theta_2^{-1}\big]^{-1}\Big) \le (1-t)K(\theta_1) + tK(\theta_2),$$

which says that $K(\theta)$ is a convex function of θ^{-1}. This implies that $\theta K(\theta)$ is a convex function of θ and hence continuous on $(0,1)$. Continuity at $\theta = 1$ is checked trivially from (2.21)–(2.23) □

Next we study the minimizer $\bar{\mu}$. We identify $\bar{\mu}$ from the following variational argument. Denote by $t\partial$ the admissible variations of $\bar{\mu}$ with $t > 0$ [admissible means $\bar{\mu} + t\partial \in \Phi(\theta)$]. For such variations

$$\varphi(\bar{\mu}) \le \varphi(\bar{\mu} + t\partial). \tag{2.26}$$

Choose $\partial \in l^1$ of the form

(2.27) $\quad \partial(l) = 0$ except possibly for $l = i, j, k$ fixed and distinct.

We want to have $\bar{\mu} + t\partial \in \Phi(\theta)$ for t small enough, even if $\bar{\mu} \in \partial\Phi(\theta)$. This will be achieved by assuming [recall (2.20) and (2.21)]

$$\partial(i) + \partial(j) + \partial(k) \le 0, \tag{2.28}$$

$$\frac{\partial(i)}{i} + \frac{\partial(j)}{j} + \frac{\partial(k)}{k} = 0, \tag{2.29}$$

(2.30) $\quad$ if $\bar{\mu}(i) = 0$, then $\partial(i) \ge 0$, and similarly for j and k.

By using the convexity of φ and (2.26), we have

$$\lim_{t\downarrow 0} \frac{\varphi(\bar{\mu} + t\partial) - \varphi(\bar{\mu})}{t} \ge 0. \tag{2.31}$$

Define

$$\xi(l) = g(l) + \log\frac{\bar{\mu}(l)}{l} \quad \text{when } \bar{\mu}(l) > 0. \tag{2.32}$$

First we prove two technical lemmas.

Lemma 3. *If $\bar{\mu}(i)$ and $\bar{\mu}(j) > 0$, then $\bar{\mu}(k) > 0$.*

Proof. Suppose that $\bar{\mu}(k) = 0$. Consider the admissible variation with $\partial(i) + \partial(j) = -1$, $\partial(i)/i + \partial(j)/j = -1/k$, $\partial(k) = 1$. Then from (2.31) we get the contradiction

$$\begin{aligned}
0 &\le \lim_{t\downarrow 0} \frac{\varphi(\bar{\mu} + t\partial) - \varphi(\bar{\mu})}{t} \\
&= \frac{\partial(i)}{i}(\xi(i) + 1) + \frac{\partial(j)}{j}(\xi(j) + 1) + \frac{\partial(k)}{k}\left(g(k) + \lim_{t\downarrow 0}\log\frac{t\partial(k)}{k}\right) \\
&= -\infty.
\end{aligned}$$

□

LEMMA 4. *If $i < j$ and $\bar{\mu}(j) > 0$, then $\bar{\mu}(i) > 0$.*

PROOF. Suppose that $\bar{\mu}(i) = 0$. Consider the admissible variation $\partial(i) = 1$, $\partial(j) = -j/i$, $\partial(k) = 0$. Then from (2.31) and $\bar{\mu}(j) > 0$ we get the contradiction

$$0 \leq \frac{\partial(i)}{i}\left(g(i) + \lim_{t \downarrow 0} \log \frac{t\partial(i)}{i}\right) + \frac{\partial(j)}{j}(\xi(j) + 1) = -\infty. \qquad \square$$

We are now in a position to compute $\bar{\mu}$. Observe that Lemmas 3 and 4 imply that $\bar{\mu}(l) > 0$ for all $l \geq 1$ when $\theta < 1$. At $\theta = 1$ trivially $\bar{\mu}(l) = \partial_{1l}$; below we exclude this case.

LEMMA 5. *For all $i, j, k \geq 1$ distinct,*

$$\frac{\xi(i) - \xi(j)}{i - j} = \frac{\xi(j) - \xi(k)}{j - k} \leq 0. \tag{2.33}$$

PROOF. We use the following elementary fact. If $a, b, c, d \in \mathbb{R}$ with $a \neq 0$ and $b \neq 0$ satisfy the implication

$$ax + by \leq 0 \Rightarrow cx + dy \geq 0 \quad (\text{for all } x, y \in \mathbb{R}),$$

then

$$\frac{c}{a} = \frac{d}{b} \leq 0.$$

For our admissible variations we have from (2.31), (2.32) and $\bar{\mu}(i)$, $\bar{\mu}(j)$, $\bar{\mu}(k) > 0$,

$$\frac{\partial(i)}{i}(\xi(i) + 1) + \frac{\partial(j)}{j}(\xi(j) + 1) + \frac{\partial(k)}{k}(\xi(k) + 1) \geq 0.$$

If we use (2.29) to substitute for $\partial(k)$,

$$\partial(k) = -k\frac{\partial(i)}{i} - k\frac{\partial(j)}{j},$$

then we obtain via (2.28) the implication

$$(i - k)\frac{\partial(i)}{i} + (j - k)\frac{\partial(j)}{j} \leq 0$$

$$\Rightarrow (\xi(i) - \xi(k))\frac{\partial(i)}{i} + (\xi(j) - \xi(k))\frac{\partial(j)}{j} \geq 0$$

$$[\text{for all } \partial(i), \partial(j) \in \mathbb{R}].$$

Hence the claim follows from the remark at the beginning. $\square$

From Lemma 5 it now follows that there exists $r \geq 0$ such that

$$\frac{\xi(i) - \xi(j)}{i - j} = -r \leq 0 \quad \text{for all } i, j \geq 1, i \neq j.$$

Thus

$$\xi(i) = \xi(1) - r(i-1)$$

or, via (2.32),

$$\frac{\bar{\mu}(i)}{i} = \bar{\mu}(1) e^{-(i-1)r - [g(i) - g(1)]}.$$

From the normalization in (2.20),

$$\frac{\bar{\mu}(i)}{i} = \frac{1}{G(r)} e^{-ri - g(i)}, \tag{2.34}$$

where $G(r)$ is given by (2.5).

Now we distinguish two cases.

(i) $\bar{\mu} \in \partial\Phi(\theta)$. From (2.21) we have

$$\frac{1}{\theta} = \sum_{i \geq 1} \bar{\mu}(i) = \frac{1}{G(r)} \sum_{i \geq 1} i e^{-ri - g(i)}.$$

Hence

$$\theta = -\frac{G(r)}{G'(r)}.$$

So $\bar{\mu}$ is given by (2.34) with r the solution of the above equation, which is (2.8). (Lemma 7 below shows that the solution is unique.)

(ii) $\bar{\mu} \notin \partial\Phi(\theta)$. First we prove the following lemma.

LEMMA 6. *If $\bar{\mu} \notin \partial\Phi(\theta)$, then $\xi(j) = \xi(1)$ for all $j \geq 1$.*

PROOF. Take the variations $\partial(1) = \pm 1$, $\partial(j) = -j\partial(1)$, $\partial(l) = 0$ for all $l \neq 1, j$. This variation violates (2.28). However, $\bar{\mu} \notin \partial\Phi(\theta)$ guarantees that $\bar{\mu} + t\partial \in \Phi(\theta)$ for t sufficiently small, so that the restriction (2.28) drops out. Then, from (2.31),

$$\partial(1)(\xi(1) + 1) + \frac{\partial(j)}{j}(\xi(j) + 1) \geq 0.$$

Hence

$$\partial(1)(\xi(1) - \xi(j)) \geq 0. \qquad \square$$

From Lemma 6 it follows that if $\bar{\mu} \notin \partial\Phi(\theta)$, then $r = 0$ and $\bar{\mu}$ is given by

$$\frac{\bar{\mu}(i)}{i} = \frac{1}{G(0)} e^{-g(i)}. \tag{2.35}$$

Finally, we study the function

$$H(r) = -\frac{G(r)}{G'(r)}, \qquad r \geq 0.$$

LEMMA 7. *The function $H(r)$ is analytic, strictly positive and strictly increasing on $(0,\infty)$, with $\lim_{r\downarrow 0} H(r) = H(0) = \theta_c$ and $\lim_{r\uparrow\infty} H(r) = 1$. Moreover, $H(r)$ is invertible and its inverse H^{-1}: $[\theta_c, 1) \to [0,\infty)$ is analytic and strictly increasing on $(\theta_c, 1)$.*

PROOF. The proof is elementary and is left to the reader. Use the implicit function theorem plus the fact that $H'(r) > 0$ on $(0,\infty)$. □

We can now collect the above results and prove Proposition 4. From Lemma 7 combined with the above statements for cases (i) and (ii), we see that $\bar{\mu} \in \partial\Phi(\theta)$ if and only if $\theta \in [\theta_c, 1] \cap (0, 1]$. Combine $\mu(i) = i\nu(i)$ with (2.34) to get (2.7). Substitution of (2.7) into (2.4) immediately gives (2.9). Use (2.8), (2.9) and Lemma 7 to see that $K(\theta)$ is analytic on $(\theta_c, 1)$ because $r(\theta) = H^{-1}(\theta)$. For $\theta \in (0, \theta_c)$, on the other hand, $\bar{\nu}$ is given by (2.35) and $K(\theta) = K(\theta_c)$ is constant. One easily checks from (2.1), (2.2), (2.8) and (2.9) that $\theta K(\theta) \to 0$ as $\theta \to 0$, irrespective of whether $\theta_c = 0$ or $\theta_c > 0$.

2.5. *Proof of Proposition* 5. The statements of Proposition 5 prior to (2.13) are immediate from (2.12) and Proposition 4, except for the strict concavity on $(\theta_c, 1)$. To see the latter, compute from (2.12) for $\theta > \theta_c$, using (2.8) and (2.9),

$$\frac{\partial}{\partial\theta} J_h(\theta) = \log\left(\frac{hG(r)}{1-h}\right), \tag{2.36}$$

$$\frac{\partial^2}{\partial\theta^2} J_h(\theta) = -\frac{1}{\theta}\frac{dr}{d\theta} < 0. \tag{2.37}$$

The inequality follows from Lemma 7 with $r(\theta) = H^{-1}(\theta)$. Note cancellation of terms.

The slope of $J_h(\theta)$ at $\theta = 0$ is $\log(hG(0)/(1-h))$, irrespective of whether $\theta_c = 0$ or $\theta_c > 0$, and changes from plus to minus at $h = h_c$ defined in (2.13). Now parts (i) and (ii) of Proposition 5 are obvious. To see part (iii), note that $J_h(\theta)$ by (2.36) reaches its maximum when

$$h = \frac{1}{1 + G(r)},$$

with $r = r(\theta)$ the unique solution of (2.8):

$$\theta = -\frac{G(r)}{G'(r)}.$$

This proves (2.14) and (2.15). One finds (2.16) by substitution of $K(\bar{\theta})$ in (2.9) into (2.12) using (2.14) and (2.15).

Note the important qualitative change of $J_h(\theta)$ as h crosses the critical value h_c. Also note that the maximizer changes from $\bar{\theta} = 0$ to $\bar{\theta} = \theta_c \geq 0$, and therefore jumps when $\theta_c > 0$.

2.6. *Proof of Theorem* 2. From Proposition 5 we get (2.17), continuity of $\rho(h)$ on $[0, 1]$, as well as analyticity on $(0, h_c)$ and on $(h_c, 1)$. It is clear that $\rho(h)$ is strictly decreasing on $(0, h_c)$. To see that it is also strictly decreasing on $(h_c, 1)$, compute the following via (2.15) and (2.17):

$$\rho'(h) = -\frac{1}{1-h} + \frac{dr}{dh} = -\frac{1}{1-h} - \frac{(1+G(r))^2}{G'(r)}$$

$$= -\frac{1}{1-h}\left(1 + \frac{G(r)(1+G(r))}{G'(r)}\right).$$

Thus it suffices to prove the following lemma.

LEMMA 8. $-G'(r) > G(r)(1 + G(r))$ *for all* $r \geq 0$.

PROOF. Since $g(0) = 0$ and $g(i+j) < g(i) + g(j)$ for all $i, j \geq 1$ by (2.2), we argue

$$\begin{aligned}(1+G(r))^2 &= \sum_{i,j\geq 0} e^{-r(i+j)-g(i)-g(j)} \\ &< \sum_{i,j\geq 0} e^{-r(i+j)-g(i+j)} \\ &= \sum_{k\geq 0} (k+1)e^{-rk-g(k)} \\ &= -G'(r) + (1+G(r)).\end{aligned}$$

□

To get (2.19), note that

$$\lim_{h\downarrow h_c} \frac{dr}{dh} = -\frac{(1+G(0))^2}{G'(0)}.$$

and substitute (2.6) and (2.13). The remaining claims of Theorem 2 are obvious.

The upper bound $\bar{\theta}(h) < h$ follows from (2.14) and (2.15) with Lemma 8. Another way to see the latter inequality is by returning to (2.10).

LEMMA 9. *$J(\theta)$ is strictly decreasing on* $[0,1]$.

PROOF. Recall that $J(\theta)$ is continuous on $[0,1]$ [(2.3) and Proposition 4]. Differentiate (2.3) and use (2.8) and (2.9) to get

$$\begin{aligned} J'(\theta) &= \log\left(\frac{\theta G(r)}{1-\theta}\right) \quad \text{if } \theta > \theta_c, \\ &= \log\left(\frac{\theta G(0)}{1-\theta}\right) \quad \text{if } \theta < \theta_c. \end{aligned}$$

Both are strictly negative by Lemma 8 and by (2.8) and (2.6), respectively. □

Since (1.10) gives that $I_h(\theta)$ is decreasing in θ for $\theta < h$ and increasing in θ for $\theta > h$, we see from (2.10) and Lemma 9 that the maximizer $\bar{\theta}$ falls in $[0, h)$.

3. Interpretation of variational formula.

3.1. *Path of descent of a typical particle.* In this section we explore what Theorems 1 and 2 imply about the history of the particles in the population at a given time n. Our main result, Theorem 3 below, shows that in the limit as $n \to \infty$ almost all of the population has a very specific history, one that is singled out by the variational formula. This is the selection mechanism alluded to in the Introduction: *For large n the population predominantly consists of those particles whose history happened to be best adapted to the given environment.* In our situation "best adapted" means spending a lot of time on sites where b_x is large, little time on sites where b_x is small, and doing so in a way that is not too unlikely.

We start by introducing the notion of the *path of descent of a typical particle*. Let (η_n^z) be the process starting from one particle at site z and no particle elsewhere: a so-called single ancestor process. Then a version of our process is given by the independent sum

$$\eta_n(x) = \sum_{z \in \mathbb{Z}} \eta_n^z(x). \tag{3.1}$$

This allows us to construct the historical process for (η_n) in the standard way, as follows. For each single ancestor process we record the position of the ancestor at time 0, the positions of the particles in its offspring at time 1, and so on. That is, we have the whole family tree available on our probability space. For a formal construction, see, for instance, Harris (1963), Chapter 6. By (3.1), the historical process for (η_n) is simply the collection of family trees of all the single ancestor processes independently put together.

Now give each particle in the population at time n a label ω^n. Let $Z_i^n(\omega^n)$ denote the position of the ancestor at time i of particle ω^n and let

$$Z^n(\omega^n) = \left(Z_i^n(\omega^n) - Z_0^n(\omega^n)\right)_{i=0}^n, \tag{3.2}$$

which we call the path of descent of particle ω^n. Next let

$$\Omega_N^n = \{\omega^n \colon Z_n^n(\omega^n) \in [-N, N]\}$$

and define

$$\Gamma_N^n = \frac{1}{|\Omega_N^n|} \sum_{\omega^n \in \Omega_N^n} \delta_{Z^n(\omega^n)}. \tag{3.3}$$

Ω_N^n is the population at time n in the box $[-N, N]$. In Section 1.2 we saw that

$$\lim_{N\to\infty} |\Omega_N^n|/(2N+1) = E\big|\{\omega^n \colon Z_n^n(\omega^n) = 0\}\big| \quad \text{a.s.},$$

[see (1.4)], which is the expected population at the origin at time n. [We recall that the a.s. refers to the joint distribution of (η_n) and F.] By the same token we note that for each fixed n the vector-valued process Z^n is stationary and ergodic under translations in $\mathbb{Z}$ [by the same argument as in Section 1.2, this is a consequence of the i.i.d. property of F, the initial choice $\eta_0(x) \equiv 1$ and the evolution mechanism]. Hence from the individual ergodic theorem we have the existence of the limit

$$\lim_{N\to\infty} \Gamma_N^n = \Gamma^n \quad \text{a.s.} \tag{3.4}$$

Now let

$$\hat{Z}^n = \left(\hat{Z}_i^n\right)_{i=0}^n$$

be distributed according to Γ^n. This is what we define as the path of descent of a typical particle in the population at time n, in short *typical path*. The word "typical" is used here because (3.3) randomly selects a particle from the population in a large box $[-N, N]$.

We shall need the following functionals of $\hat{Z}^n$:

$$\hat{\theta}_n = n^{-1}\hat{Z}_n^n, \tag{3.5}$$

$$\hat{\nu}_n = \frac{1}{\hat{Z}_n^n + 1} \sum_{x=0}^{\hat{Z}_n^n} \delta_{\hat{l}_n(x)} \tag{3.6}$$

with

$$\hat{l}_n(x) = \left|\left\{0 < i \le n \colon \hat{Z}_i^n = x\right\}\right|.$$

That is, $\hat{\theta}_n$ is the *empirical drift* of the typical path and $\hat{\nu}_n$ is the *empirical distribution of local times* of the typical path over its range. Note that (3.6) is the a priori relevant definition only because paths cannot step to the left.

3.2. *Survival of the fittest.* Now we are ready to formulate our main result. As in Section 2, let $\bar{\theta} = \bar{\theta}(h)$ and $\bar{\nu} = \bar{\nu}(\theta)$ denote the maximizers of our variational formula (1.13) (see Theorems 1 and 2 and Propositions 4 and 5). We have seen that both are unique, except for $\theta_c > 0$ and $h = h_c$, when $\bar{\theta}$ is the whole interval $[0, \theta_c]$. Recall that if $\theta_c > 0$ and $\theta \in (0, \theta_c)$, then $\bar{\nu}(\theta)$ falls outside the domain M_θ of the variational formula. The following theorem will be proved in Section 3.4.

THEOREM 3. (i) *As* $n \to \infty$,

$$\hat{\theta}_n \to \bar{\theta} \quad a.s. \text{ for all } h \in (0,1) \setminus \{h_c\}, \tag{3.7}$$

$$\Gamma^n\left(\hat{\nu}_n | \hat{\theta}_n = \theta_n\right) \to \delta_{\bar{\nu}(\theta)} \quad \text{weakly along any sequence } \theta_n \to \theta > 0. \tag{3.8}$$

(ii) (*Localization vs. delocalization.*) *If* $h \in (0, h_c)$, *then* $\bar{\theta} = 0$, *so that the typical path has zero limiting drift. If* $h \in (h_c, 1)$, *then* $\theta_c < \bar{\theta} < h$, *so that the typical path has positive limiting drift but still slows down compared to the underlying random walk.*

(iii) (*Intermittency.*) *If* $\theta_c > 0$, *then for every* $\theta \in (0, \theta_c)$ *and every* $\theta_n \to \theta$ *there exist random sets* $A_n \subset [0, \hat{Z}_n^n]$ *on the events* $\{\hat{\theta}_n = \theta_n\}$ *such that*

$$\Gamma^n\left(|A_n| / \left|\left[0, \hat{Z}_n^n\right]\right| \,\middle|\, \hat{\theta}_n = \theta_n\right) \to \delta_0 \quad \text{weakly}, \tag{3.9}$$

but

$$\Gamma^n\left(n^{-1} \sum_{x \in A_n} \hat{l}_n(x) \geq (\theta_c - \theta)/\theta_c \,\middle|\, \hat{\theta}_n = \theta_n\right) \to 1. \tag{3.10}$$

No such sets exist if either $\theta_c > 0$ *and* $\theta \in (\theta_c, 1]$, *or* $\theta_c = 0$.

(iv) (*Survival vs. extinction.*) *For* $h \in [0, h_c^*)$ *the particle density tends to* ∞, *while for* $h \in (h_c^*, 1]$ *it tends to* 0 [*recall Theorem* 2(iv)].

REMARK. Note that if $h > h_c$, then $\bar{\theta} > 0$ and hence (3.7) and (3.8) imply

$$\Gamma^n(\hat{\nu}_n) \to \delta_{\bar{\nu}(\bar{\theta})} \quad \text{weakly}.$$

However, if $h < h_c$, then $\bar{\theta} = 0$ and so we get no information about the full limit law. The study of this object would require an analysis on a finer than exponential scale, which is beyond the scope of this paper. A similar refinement would be needed to settle the limit law of $\hat{\theta}_n$ at $h = h_c$.

3.3. *Description of phase transitions.* Before expanding on Theorem 3 we rewrite the function $G(r)$ in (2.5) in terms of the distribution β for the mean offspring (see Section 1.1).

LEMMA 10.

$$G(r) = \int \frac{b}{M} e^{-r} \left(1 - \frac{b}{M} e^{-r}\right)^{-1} \beta(db).$$

PROOF. Use (1.6) and (1.27) to write $\exp(-g(i)) = \exp(f(i) - i \log M) = \int (b/M)^i \beta(db)$. Substitute into (2.5). □

(I) *Localization vs. delocalization, and intermittency.* There are three cases to distinguish [see (2.6) and (2.13)]:

a. $\theta_c = h_c = 0 \Leftrightarrow \int (M-b)^{-1}\beta(db) = \infty$;
b. $0 = \theta_c < h_c \Leftrightarrow \int (M-b)^{-2}\beta(db) = \infty$, $\int (M-b)^{-1}\beta(db) < \infty$;
c. $0 < \theta_c < h_c \Leftrightarrow \int (M-b)^{-2}\beta(db) < \infty$.

Case a: For all $h > 0$ the typical path has a drift that tends to a strictly positive limit (i.e., its range is of order n): There is delocalization.

Case b: For $h < h_c$ the typical path localizes [i.e., its range is $o(n)$], while for $h > h_c$ it delocalizes. At $h = h_c$ the limiting drift is continuous.

Case c: The same as in case b, except that now the limiting drift makes a jump at $h = h_c$ of size θ_c.

If β has an atom at M, then we are in case a. The density of sites where the growth is maximal is strictly positive and therefore it does not pay for the particles to stand still: There are good growth spots in abundance. On the other hand, if β has a density w.r.t. Lebesgue measure tending to 0 sufficiently fast in the neighborhood of M, then we are in case b or c. The sites where the growth is maximal or close to maximal are too rare and the particles tend to stay close to them as much as possible, thereby losing their drift. In case c this effect is even so strong that for all paths with a limiting drift $\theta \in (0, \theta_c)$ a fraction at most θ/θ_c of the time is spent outside the thin sets A_n carrying these rare sites. This phenomenon is an example of intermittency. The jump in the limiting drift in case c can be explained as follows: Once the typical path moves it runs fast in order to find the rare good growth spots.

Note that for all $h \in (0, 1)$ the limiting drift is strictly smaller than h. That is, the typical path always has a tendency to *slow down* compared to its behavior in the spatially homogeneous medium. This is a somewhat subtle effect, because the typical path slows down on good growth sites but speeds up on bad growth sites. Apparently the net effect is to still slow down.

(II) *Survival vs. extinction.* At $h = h_c^*$ the growth rate of the particle density changes sign. This means that for $h > h_c^*$ the population dies out globally, so in particular at every given site there will be no particles from some random time onwards. For $h < h_c^*$, on the other hand, the population grows globally. This does not necessarily mean that at every given site we find a growing population: The assertion is about the *global* density only. In fact, in Greven and den Hollander (1991), it is shown that there is a range of h-values where the process *grows globally but dies out locally*. In this case the population experiences an extreme form of clustering: Huge peaks form on a thinning set. (This is in fact an intermittency-type phenomenon.) But the global picture is perhaps more relevant: The fact that elephants are extinct in northern Europe says little about elephants.

3.4. *Proof of Theorem* 3.

PROOF OF (3.7). Let

$$\theta_n(\omega^n) = n^{-1}(Z_n^n(\omega^n) - Z_0^n(\omega^n)).$$

Pick $\varepsilon > 0$ and define

$$\Omega_N^n(\varepsilon) = \{\omega^n \in \Omega_N^n: |\theta_n(\omega^n) - \bar{\theta}| > \varepsilon\}.$$

This is the collection of particles in the box $[-N, N]$ at time n having a path of descent with empirical drift away from $\bar{\theta}$. We have [recall the observations made below (3.3)]

$$\lim_{N\to\infty} \frac{1}{2N+1}|\Omega_N^n| = E|\{\omega^n: Z_n^n(\omega^n) = 0\}| \quad \text{a.s.,} \tag{3.11}$$

$$\lim_{N\to\infty} \frac{1}{2N+1}|\Omega_N^n(\varepsilon)| = E|\{\omega^n: Z_n^n(\omega^n) = 0, \quad |\theta_n(\omega^n) - \bar{\theta}| > \varepsilon\}| \quad \text{a.s.} \tag{3.12}$$

and

$$\lim_{N\to\infty} |\Omega_N^n(\varepsilon)|/|\Omega_N^n| = \Gamma^n(|\hat{\theta}_n - \bar{\theta}| > \varepsilon) \quad \text{a.s.}$$

by the definition of Γ^n, the law of the typical path at time n as defined in (3.4). The growth rate of the r.h.s. of (3.11) being ρ by Theorem 1, it suffices now to show that there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\frac{1}{n}\log[\text{r.h.s. } (3.12)] \le \rho - \delta \quad \text{for } n \text{ sufficiently large,}$$

that is, the density of such particles grows at a rate strictly smaller than that of the whole population.

As in Proposition 1 we can write

$$\text{r.h.s. } (3.12) = E_h\left(\exp\left[\sum_{x=0}^{\infty} f(l_n(x))\right] 1_{\{|n^{-1}S_n - \bar{\theta}| \ge \varepsilon\}}\right).$$

Then, retracing the steps of the proof in Section 1.5, we have

$$\lim_{n\to\infty} \frac{1}{n}\log[\text{r.h.s. } (3.12)] = \sup_{\substack{\theta\in[0,1] \\ |\theta-\bar{\theta}|\ge\varepsilon}} (J(\theta) - I_h(\theta)). \tag{3.13}$$

But this finishes the proof because we saw in Proposition 5 that $J(\theta) - I_h(\theta)$ achieves a unique maximum at $\bar{\theta}$ (for $h \ne h_c$). □

PROOF OF (3.8). Here the same remarks as in the previous proof apply. The quantity to consider this time is

$$\Gamma^n\Big(\|\hat{\nu}_n - \bar{\nu}\| \geq \varepsilon\Big),$$

where $\|\cdot\|$ is any metric that induces the weak* topology. This again is a ratio of two terms and it suffices to show that there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\frac{1}{n}\log E_h\left(\exp\left[\sum_{x=0}^{\infty} f(l_n(x))\right]1_{\{\|\nu_{n\theta_n+1} - \bar{\nu}\| \geq \varepsilon\}} | S_n = \theta_n n\right) \leq \rho - \delta$$

for n sufficiently large,

where we recall the definition of ν_N below (1.23). But the l.h.s. grows at rate

$$\sup_{\substack{\nu \in M_\theta \\ \|\nu - \bar{\nu}\| \geq \varepsilon}} \big\{\theta[\langle f, \nu\rangle - I_\theta(\nu)] - I_h(\theta)\big\}, \tag{3.14}$$

and so again this finishes the proof because we saw in Proposition 4 that, for every $\theta \in (0, 1]$, there is a unique maximizer $\bar{\nu} = \bar{\nu}(\theta)$ (irrespective of whether it is attained in M_θ or not). □

PROOF OF (3.9) AND (3.10). Every event in this proof is to be intersected with the events $\{\hat{\theta}_n = \theta_n\}$. From (3.7) and (3.8) it follows that

$$\sum_i i\hat{\nu}_n(i) = \frac{n}{\hat{Z}_n^n + 1} = \frac{n}{n\hat{\theta}_n + 1} \to \theta^{-1}, \tag{3.15}$$

$$\begin{aligned} &\hat{\nu}_n(i) \to \bar{\nu}(i) \quad \text{for all } i \\ &\text{with } \sum_i \bar{\nu}(i) = 1 \text{ and } \sum_i i\bar{\nu}(i) = \theta_c^{-1} < \theta^{-1} \end{aligned} \tag{3.16}$$

[recall Proposition 4(ii)]. Pick N and define the random set

$$B_n^N = \left\{x \in \left[0, \hat{Z}_n^n\right] : \hat{l}_n(x) \geq N\right\}.$$

Split $\hat{\nu}_n$ into the sum of two parts

$$\hat{\nu}_{n,1}^N = \frac{1}{\hat{Z}_n^n + 1} \sum_{x \in [0, \hat{Z}_n^n] \setminus B_n^N} \delta_{\hat{l}_n(x)},$$

$$\hat{\nu}_{n,2}^N = \frac{1}{\hat{Z}_n^n + 1} \sum_{x \in B_n^N} \delta_{\hat{l}_n(x)}.$$

As a consequence of (3.15) and (3.16), we have

$$\limsup_{n\to\infty} \sum_i i\hat{\nu}_{n,1}^N(i) \le \theta_c^{-1} \qquad \text{for all } N,$$

$$\liminf_{n\to\infty} \sum_i i\hat{\nu}_{n,2}^N(i) \ge \theta^{-1} - \theta_c^{-1} \quad \text{for all } N,$$

$$\lim_{N\to\infty} \limsup_{n\to\infty} \sum_i \hat{\nu}_{n,2}^N(i) = 0.$$

This in turn implies that there exists some random sequence $N(n) \to \infty$ such that

$$\sum_i \hat{\nu}_{n,2}^{N(n)}(i) \to 0, \tag{3.17}$$

$$\limsup_{n\to\infty} \left\{ \sum_i i\hat{\nu}_{n,1}^{N(n)}(i) \Big/ \sum_i i\hat{\nu}_{n,2}^{N(n)}(i) \right\} \le \frac{\theta_c^{-1}}{\theta^{-1} - \theta_c^{-1}}. \tag{3.18}$$

Now put

$$A_n = B_n^{N(n)}.$$

Then (3.17) is (3.9). Combine (3.18) with (3.16), substituting $\hat{\nu}_{n,1}^{N(n)}(i) = \hat{\nu}_n(i) - \hat{\nu}_{n,2}^{N(n)}(i)$, to get via (3.15):

$$\liminf_{n\to\infty} \frac{\hat{Z}_n^n + 1}{n} \sum_i i\hat{\nu}_{n,2}^{N(n)}(i) \ge \frac{\theta^{-1} - \theta_c^{-1}}{\theta^{-1}}.$$

This is (3.10).

In the complementary case, $\theta_c = 0$, or $\theta_c > 0$ and $\theta > \theta_c$, no such A_n can be found because, by Proposition 4(i) and (3.8),

$$\lim_{N\to\infty} \limsup_{n\to\infty} \sum_i i\hat{\nu}_{n,2}^N(i) = 0. \qquad \square$$

Acknowledgments. We thank E. Bolthausen for pointing out an error in an earlier version of Theorem 3. We thank the referee for detailed comments on the presentation.

REFERENCES

DEUSCHEL, J.-D. and STROOCK, D. W. (1989). *Large Deviations*. Academic, Boston.

EISELE, TH. and LANG, R. (1987). Asymptotics for the Wiener sausage with drift. *Probab. Theory Related Fields* **74** 125–140.

FLEISCHMANN, K. and MOLCHANOV, S. A. (1990). Exact asymptotics in a mean field model with random potential. *Probab. Theory Related Fields* **86** 239–251.

GÄRTNER, J. and MOLCHANOV, S. A. (1990). Parabolic problems for the Anderson model. I. Intermittency and related topics. *Comm. Math. Phys.* **132** 613–655.

GREVEN, A. and DEN HOLLANDER, F. (1992). Branching random walk in random environment: Phase transitions for local and global growth rates. *Probab. Theory Related Fields* **91** 195–249.

GREVEN, A. and DEN HOLLANDER, F. (1991). Population growth in random media. I. Variational formula and phase diagram. *J. Statist. Phys.* **65** 1123–1146.
HARRIS, T. E. (1963). *The Theory of Branching Processes*. Springer, Berlin.
RUDIN, W. (1973). *Functional Analysis*. McGraw-Hill, New York.

J.-B. BAILLON
INSTITUT DE MATHÉMATIQUES
ET INFORMATIQUE
UNIVERSITÉ DE LYON I
43 BD DU 11 NOVEMBRE 89
VILLEURBANNE, 69622 CEDEX
FRANCE

PH. CLÉMENT
FACULTEIT DER TECHNISCHE WISKUNDE
EN INFORMATICA
TECHNISCHE UNIVERSITEIT DELFT
MEKELWEG 4
2600 GA DELFT
THE NETHERLANDS

A. GREVEN
INSTITUT FÜR MATHEMATISCHE STOCHASTIK
UNIVERSITÄT GÖTTINGEN
LOTZESTRASSE 13
W-3400 GÖTTINGEN
GERMANY

F. DEN HOLLANDER
MATHEMATISCH INSTITUUT
RIJKSUNIVERSITEIT UTRECHT
P.O. BOX 80. 010
3508 TA UTRECHT
THE NETHERLANDS