Journal of Statistical Physics, Vol. 65, Nos. 5/6, 1991

Population Growth in Random Media. I. Variational Formula and Phase Diagram

A. Greven[1] and F. den Hollander[2]

We consider an infinite system of particles on the integer lattice $\mathbb{Z}$ that: (1) migrate to the right with a random delay, (2) branch along the way according to a random law depending on their position (random medium). In Part I, the first part of a two-part presentation, the initial configuration has one particle at each site. The long-time limit exponential growth rate of the expected number of particles at site 0 (local particle density) does not depend on the realization of the random medium, but only on the law. It is computed 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 earlier work the exponential growth rate of the Cesaro limit of the number of particles per site (global particle density) was studied and a different variational formula was found, but with similar structure, solution, and phases. Combination of the two results reveals an intermediate phase where the population globally survives but locally becomes extinct (i.e., dies out on any fixed finite set of sites).

KEY WORDS: Phase transition; variational formula; population growth; random medium.

0. INTRODUCTION AND MAIN RESULTS

0.1. Motivation

This and the following paper are part of a series in which we study population growth in random media via *variational techniques*. The models that we consider are of the following type. Infinitely many particles live on the integer lattice $\mathbb{Z}$ and are subject to two random mechanisms:

1. Particles *branch* according to a *site-dependent* offspring distribution chosen randomly for each site and kept fixed during the

[1] Institut für Mathematische Stochastik, Universität Göttingen, W-3400 Göttingen, Germany.
[2] Mathematisch Instituut, Rijksuniversiteit Utrecht, 3508 TA Utrecht, The Netherlands.

0022-4715/91/1200-1123$06.50/0

evolution. These offspring distributions play the role of the *random medium.*

2. Particles *migrate* by jumping to nearest-neighbor sites with site-independent probabilities. The migration has a *drift.*

We are interested in how the particle density behaves as a function of time and space. It turns out that such systems exhibit interesting *phase transitions* as the drift varies, due to the competition between the branching and the migration.[1-3]

The program of our series of papers is to analytically derive and explicitly solve variational formulas describing various aspects of the population growth. The value of the variational formula gives the exponential growth rate of the particle density, the maximizer provides information about the path of descent of a typical particle in the population. From these results we extract the *phase diagram.*

Part I is a continuation of Baillon *et al.*[2] and completes the description of the phase diagram. In particular, by combining results of both papers we establish the existence of an *intermediate phase* where the population globally survives but locally becomes extinct (extreme clustering).

Part II treats the phenomenon of wavefront propagation when the initial particle density has a block shape (like a Heaviside function). We calculate a *speed-dependent growth rate,* thereby obtaining a particle density profile after suitable scaling of time and space. This profile exhibits three phases separated by *two characteristic wavefront speeds.* In addition, its qualitative shape changes as the drift crosses a threshold.

All effects and phase transitions described are due to the randomness of the medium and disappear in the spatially homogeneous situation.

0.2. 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. Here F plays the role of the *random medium.* For fixed F, define a discrete-time Markov process $(\eta_n)_{n \geqslant 0}$ on $(\mathbb{N} \cup \{0\})^{\mathbb{Z}}$, with the interpretation

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

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

by specifying its one-step transition mechanism as follows. Given the state η_n at time n:

1. 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, i.e., it consists of k new particles with probability $F_x(k)$ $(k=0, 1, 2,...)$. Also, particles at the same site branch independently.
2. Immediately after creation each new particle at site x independently decides to either stay at x with probability $1-h$ or to jump to $x+1$ with probability h. The parameter $h \in [0, 1]$ is the *drift* and is the same for all x.

The resulting sequence of particle numbers after steps (1) and (2) make up the state η_{n+1} at time $n+1$. *F stays fixed during the evolution.*

Let

$$b_x = \sum_{k=0}^{\infty} kF_x(k) \tag{0.1}$$

denote the *mean offspring* at site x and let β denote the distribution of b_x induced by α (the distribution of F_x). It is assumed that β has bounded support and strictly positive variance, i.e.,

$$0 < \inf_x b_x < \sup_x b_x < \infty \tag{0.2}$$

We have thus specified the evolution mechanism of our particle system. It remains to fix the initial state η_0. We shall consider the following two starting configurations:

$$\text{Case I.} \qquad \eta_0(x) = 1 \qquad \text{for all} \quad x \in \mathbb{Z}$$

$$\text{Case II.} \qquad \eta_0(x) = \begin{cases} 1 & \text{for} \quad x \leqslant 0 \\ 0 & \text{for} \quad x > 0 \end{cases}$$

In case I we expect to find a globally uniform population, while in case II we expect to see a wavefront propagation phenomenon. The two cases are treated in Parts I and II, respectively of this two part presentation.

Our analysis will show that our i.i.d. assumption on the random medium F is not essential. In fact, the results to be discussed turn out to be the same for any stationary ergodic process $\{b_x\}_{x \in \mathbb{Z}}$ with marginal β, provided we take into account the entropy of this process. It is very crucial, though, that we assume (0.2): we shall see that the phenomena to be described would otherwise be absent.

In earlier papers we have also considered the situation where the migration in step 2 goes both ways, namely particles may jump to the left or to the right nearest-neighbor site with probabilities $\frac{1}{2}(1-h)$ and $\frac{1}{2}(1+h)$, respectively. This model is considerably more difficult and the analysis of the phase diagram is less complete.[1,3]

0.3. Local Particle Density and Speed-Dependent Growth Rate

Let E denote the double expectation over the Markov process (η_n) given F, as well as over F. Let

$$d_n(x, F) = E(\eta_n(x) \mid F) \tag{0.3}$$

denote the average particle density at site x at time n given F. We shall be interested in the following two quantities:

$$\text{Case I:} \qquad \lambda^{\mathrm{I}}(0, F) = \lim_{n \to \infty} \frac{1}{n} \log d_n^{\mathrm{I}}(0, F) \tag{0.4}$$

$$\text{Case II:} \qquad \lambda^{\mathrm{II}}(\tau, F) = \lim_{n \to \infty} \frac{1}{n} \log d_n^{\mathrm{II}}(\lfloor \tau n \rfloor, F) \tag{0.5}$$

i.e., the growth rate at $x=0$ for uniform starting configuration and the growth rate at $x=\lfloor \tau n \rfloor$ for block shape starting configuration. The latter is the *speed-τ growth rate*, observed at a site moving at speed $\tau \geqslant 0$. Theorem 1 below shows that the limits in (0.4) and (0.5) exist F-a.s., are constant F-a.s., and can be computed in terms of a *variational formula* depending on the two parameters β and h. To formulate the result, we need the following symbols. Let $\mathscr{P}(\mathbb{N} \times \operatorname{supp} \beta)$ denote the set of probability measures on the product of $\mathbb{N}$ and the support of β. Let $\theta \in (0, 1]$, $v \in \mathscr{P}(\mathbb{N} \times \operatorname{supp} \beta)$, and $i \in \mathbb{N}$, $j \in \operatorname{supp} \beta$. Define

$$M_{\theta,\beta} = \Big\{ v \in \mathscr{P}(\mathbb{N} \times \operatorname{supp} \beta)\colon \ \sum_{i,j} i v(i,j) = \theta^{-1}, \\ \sum_i v(i,j) = \beta(j) \text{ for all } j \Big\} \tag{0.6}$$

$$f(v) = \sum_{i,j} v(i,j)\, i \log j \tag{0.7}$$

$$I_{\theta,\beta}(v) = \sum_{i,j} v(i,j) \log\left(\frac{v(i,j)}{\pi_\theta(i)\, \beta(j)}\right) \tag{0.8}$$

$$I_h(\theta) = \theta \log\left(\frac{\theta}{h}\right) + (1-\theta)\log\left(\frac{1-\theta}{1-h}\right) \tag{0.9}$$

$$\pi_\theta(i) = \theta(1-\theta)^{i-1} \tag{0.10}$$

The notation is for the situation where supp β is countable. If supp β is continuous, then write integrals for the sums over j and in (0.8) use densities with the convention that $I_{\theta,\beta}(v) = \infty$ when v is not absolutely continuous w.r.t. $\pi_\theta \times \beta$.

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

$$\lambda^{\mathrm{I}}(0, F) = \lambda(\beta, h; 0) \quad F\text{-a.s.} \tag{0.11}$$

$$\lambda^{\mathrm{II}}(\tau, F) = \lambda(\beta, h; \tau) \quad F\text{-a.s.} \tag{0.12}$$

with

$$\lambda(\beta, h; \tau) = \sup_{\theta \in [\tau, 1] \cap (0, 1]} [J_\beta(\theta) - I_h(\theta)] \qquad (\tau \geqslant 0) \tag{0.13}$$

$$J_\beta(\theta) = \theta \sup_{v \in M_{\theta,\beta}} [f(v) - I_{\theta,\beta}(v)] \qquad (\theta \in (0, 1]) \tag{0.14}$$

The proof is given in Section 1. Incidentally, the proof will show that the same speed-τ growth rate is observed at site $x = \lfloor \tau_n n \rfloor$ for any $\tau_n \to \tau$ (i.e., the growth rate only varies on scale n).

0.4. Solution of Variational Formula

The variational expressions in Theorem 1 can be solved explicitly. This is carried out in Section 2. The following three quantities play a key role in the solution:

$$F(r) = \exp\left[\sum_j \beta(j) \log\left(\frac{(j/M)\, e^{-r}}{1 - (j/M)\, e^{-r}}\right)\right] \qquad (r > 0) \tag{0.15}$$

$$h_c = \lim_{r \downarrow 0} \frac{1}{1 + F(r)} \tag{0.16}$$

$$\theta_c = \lim_{r \downarrow 0} -\frac{F(r)}{F'(r)} \tag{0.17}$$

Here M denotes the supremum of supp β [recall (0.2)]. Before we formulate our main result we list a few technical properties of the function $F(r)$ and the constants h_c and θ_c.

Lemma 1. For $r>0$:

(i) $F(r)$ is analytic, strictly positive, and strictly decreasing.

(ii) $-F(r)/F'(r)$ is strictly positive, strictly increasing, invertible, and its inverse is analytic and strictly increasing.

(iii) $-F(r)/F'(r) < 1/[1+F(r)]$.

Lemma 2. For all β: $0 \leqslant \theta_c \leqslant h_c < 1$:

(a) $h_c = 0$ iff $\sum_j \beta(j) \log(1 - j/M) = -\infty$.

(b) $\theta_c = 0$ iff $\sum_j \beta(j)(1 - j/M)^{-1} = \infty$.

The proof is an elementary calculation and is left to the reader.

To express our results in a compact form, we introduce two more key quantities, $r^* = r^*(\beta, h)$ and $\theta^* = \theta^*(\beta, h)$, which are defined as follows:

$$\begin{aligned} h \leqslant h_c: &\quad r^* = 0 \\ h > h_c: &\quad r^* \text{ is the unique solution of } h = \frac{1}{1+F(r)} \end{aligned} \tag{0.18}$$

$$\begin{aligned} h \leqslant h_c: &\quad \theta^* = 0 \\ h > h_c: &\quad \theta^* = -\frac{F(r^*)}{F'(r^*)} \end{aligned} \tag{0.19}$$

In the rest of Part I we consider the uniform starting configuration I from Section 0.2. For this case the variational formula has the following solution:

Theorem 2 I. (i) The growth rate in (0.11) at the origin is

$$\lambda(\beta, h; 0) = \log[M(1-h)] + r^* \tag{0.20}$$

(ii) The maximizers $\bar{\theta} = \bar{\theta}(\beta, h)$ and $\bar{v} = \bar{v}(\beta, h)$ in (0.13) and (0.14) are

$$\bar{\theta} = \theta^* \tag{0.21}$$

$$\bar{v}(i, j) = \pi_{\xi^*(j)}(i)\, \beta(j) \tag{0.22}$$

with $\xi^*(j) = 1 - (j/M)\, e^{-r^*}$ and $\pi_\theta(i)$ as in (0.10).

Corollary 2 I. (iii) $h \to \lambda(\beta, h; 0)$ is continuous and strictly decreasing on $(0, 1)$, analytic on $(0, h_c)$ and $(h_c, 1)$, and at the boundary points $\lambda(\beta, 0; 0) = \log M$ and $\lambda(\beta, 1; 0) = \sum_j \beta(j) \log j$.

(iv) If $h_c > 0$, then

$$\frac{\partial}{\partial h} \lambda(\beta, h_c+; 0) - \frac{\partial}{\partial h} \lambda(\beta, h_c-; 0) = \frac{\theta_c}{h_c(1-h_c)} \tag{0.23}$$

(v) If $\log M > 0 > \sum_j \beta(j) \log j$, then $\lambda(\beta, h; 0)$ as function of h changes sign at $h = h_c^*$, the unique solution of $\lambda(\beta, h; 0) = 0$ computable from (0.20).

(vi) $h \to \bar{\theta}(\beta, h)$ is strictly increasing and analytic on $(h_c, 1)$. Moreover,

$$\theta_c < \bar{\theta} < h \qquad \text{for} \quad h \in (h_c, 1) \tag{0.24}$$

and $\bar{\theta}$ changes from 0 to $\theta_c \geqslant 0$ at $h = h_c$.

Thus we see that the growth rate and the maximizers display interesting behavior as a function of h for fixed β. The proof is given in Section 2.

0.5. Phase Diagram

In order to come to a description of the phase diagram, we first discuss some implications of Theorem 2I and Corollary 2I for the history of the particles. Our treatment here will be somewhat informal. For a rigorous proof one would need techniques developed in Baillon *et al.*,[2] Section 3.

With each particle at site 0 at time n we can associate its *path of descent* consisting of the positions of all its ancestors at times $n-1,..., 0$. We define the *typical path of descent* of the population at site 0 at time n as the path of descent of a particle drawn randomly from this population (conditioned on it not being empty), and we shall denote by

$$\hat{Z}^n = (\hat{Z}_i^n)_{i=0}^n \qquad (\hat{Z}_0^n = 0) \tag{0.25}$$

its backward displacements relative to site 0. Two important functionals of $\hat{Z}^n$ are

$$\hat{\theta}_n = \frac{1}{n} \hat{Z}_n^n \tag{0.26}$$

$$\hat{v}_n = \frac{1}{\hat{Z}_n^n + 1} \sum_{x=0}^{\hat{Z}_n^n} \delta_{(\hat{l}_n(x), b_{-x})} \tag{0.27}$$

with

$$\hat{l}_n(x) = |\{0 < i \leqslant n \colon \hat{Z}_i^n = x\}| \tag{0.28}$$

i.e., the *empirical drift* and the *empirical distribution of local times and medium.*

In the case of a homogeneous medium (i.e., $b_x \equiv \text{const}$), it is well known that as $n \to \infty$

$$\begin{aligned} \hat{\theta}_n &\to h && \text{a.s.} \\ \hat{v}_n &\to \pi_h \times \beta && \text{in law} \end{aligned} \tag{0.29}$$

expressing the fact that asymptotically $\hat{Z}^n$ looks like the underlying random walk with drift h. However, for an inhomogeneous medium the situation is quite different. Namely, it turns out that as $n \to \infty$

$$\begin{aligned} \hat{\theta}_n &\to \bar{\theta} && \text{a.s.} \quad \text{for } h \neq h_c \\ \hat{v}_n &\to \bar{v} && \text{in law} \quad \text{for } h > h_c \end{aligned} \tag{0.30}$$

where $\bar{\theta}$ and $\bar{v}$ are the maximizers of our variational formula (0.13) and (0.14). Indeed, the analysis in Section 1 shows that the collection of paths with $\hat{\theta}_n \to \theta$ and $\hat{v}_n \to v$ has offspring of size $\exp(n[\theta f(v)])$ and has probability $\exp(-n[\theta I_{\theta,\beta}(v) + I_h(\theta)])$ asymptotically. Therefore as $n \to \infty$ the main contribution to the average population at site 0 comes from paths with $\theta = \bar{\theta}$ and $v = \bar{v}$, and the growth rate is

$$\lambda(\beta, h; 0) = \bar{\theta}[f(\bar{v}) - I_{\bar{\theta},\beta}(\bar{v})] - I_h(\bar{\theta}) \tag{0.31}$$

Thus we see that the variational formula reflects a *selection mechanism*: the population predominantly consists of those particles whose path of descent happens to be best adapted to the given environment. The fact that $\bar{\theta}$ and $\bar{v}$ are unique implies that there is a notion of optimal path of descent as expressed in (0.30). The fact that $\bar{\theta} \neq h$ and $\bar{v} \neq \pi_h \times \beta$ shows that this optimal path of descent does not look like the underlying random walk.

Equations (0.30) and (0.31) are the key to our phase diagram. The most interesting behavior occurs when β is chosen such that

$$\begin{gathered} \sum_j \beta(j) \left(1 - \frac{j}{M}\right)^{-1} < \infty \\ \log M > 0 > \sum_j \beta(j) \log j \end{gathered} \tag{0.32}$$

implying that $0 < \theta_c < 1$ and $0 < h_c^* < 1$.

(I) Localization vs. Delocalization. For $h < h_c$ we have $\bar{\theta} = 0$, meaning that the typical path moves at sublinear speed (localization). On the other hand, for $h > h_c$ we have $\bar{\theta} > 0$, so that the typical path moves at linear speed (delocalization). At $h = h_c$ the speed makes a jump of size $\theta_c > 0$. Since $\bar{\theta} < h$, the effect of the random medium is to slow down the typical path compared to its behavior in the homogeneous medium (see Fig. 1).

The fact that $\hat{v}_n \to \bar{v}$ for $h > h_c$ provides us with finer information on the typical path, e.g., what fraction of time it spends in the level sets of the medium. There is no such result for $h < h_c$ because at the boundary $\theta = 0$ our variational formula degenerates.

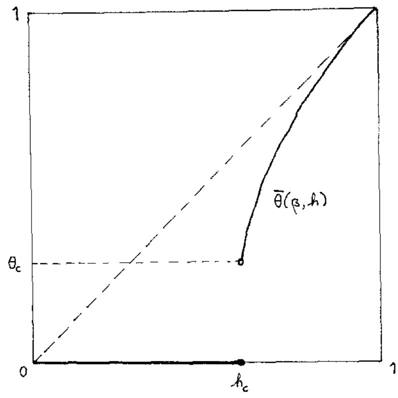

Fig. 1. The maximizer $\bar{\theta}(\beta, h)$ as a function of h for fixed β under the assumption (0.32).

(II) Survival vs. Extinction. For $h < h_c^*$ we have $\lambda(\beta, h; 0) > 0$ and therefore the average population at site $\hat{0}$ grows (survival). On the other hand, for $h > h_c^*$ we have $\lambda(\beta, h; 0) < 0$, so that the population dies out (extinction) (see Fig. 2, lower curve). To be able to speak of survival for $h < h_c^*$, we should really also establish that $\lambda(\beta, h; 0) > 0$ implies $\eta_n(0) \to \infty$ a.s. (and not only in expectation). We defer this point to a future paper.

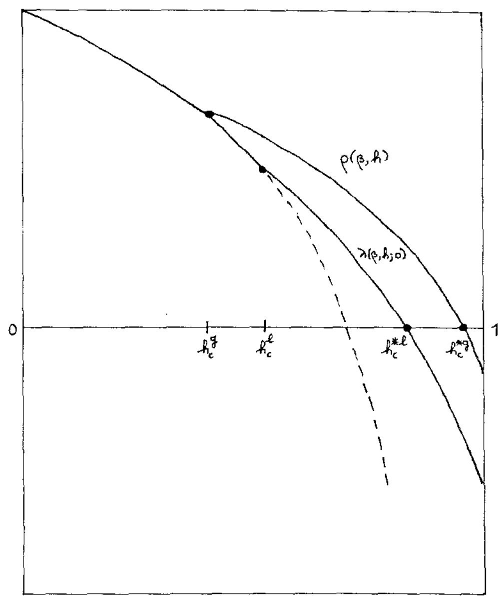

Fig. 2. The global growth rate $\rho(\beta, h)$ (upper curve) ad the local growth rate $\lambda(\beta, h; 0)$ (lower curve) as functions of h for fixed β under the assumption (0.41). The dashed curve is $\log[M(1-h)]$. The endpoints are $\rho(\beta, 1) = \log[\sum_j \beta(j) j]$ and $\lambda(\beta, 1; 0) = \sum_j \beta(j) \log j$.

0.6. Intermediate Phase

We shall now combine our results with earlier work. In Baillon *et al.*[2] we studied the *global particle density* at time n defined by

$$D(\eta_n, F) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{x=-N}^{N} \eta_n(x) \tag{0.33}$$

By applying the ergodic theorem, we showed that

$$D(\eta_n, F) = E(\eta_n(0)) \qquad \text{a.s.} \tag{0.34}$$

Thus, the *global* particle density is the expectation over the medium of the *local* particle density defined in (0.3), i.e.,

$$D(\eta_n, F) = E(d_n(0, F)) \qquad \text{a.s.} \tag{0.35}$$

We derived a variational formula for the global growth rate

$$\rho(\beta, h) = \lim_{n \to \infty} \frac{1}{n} \log D(\eta_n, F) \tag{0.36}$$

which has the same form as (0.13) and (0.14), but with (0.6)–(0.8) replaced by

$$M_\theta = \left\{ \nu \in \mathscr{P}(\mathbb{N}): \ \sum_i i\nu(i) = \theta^{-1} \right\} \tag{0.37}$$

$$f_\beta(\nu) = \sum_i \nu(i) \log \left[\sum_j \beta(j)\, j^i \right] \tag{0.38}$$

$$I_\theta(\nu) = \sum_i \nu(i) \log \left(\frac{\nu(i)}{\pi_\theta(i)} \right) \tag{0.39}$$

The fact that the two variational formulas turn out to be *different* is noteworthy, as it shows that the global and local structures of the population are controlled by different forces.

We obtained the solution of the global variational formula using functional analytic techniques. The striking analogy that we can now draw is that the solution is exactly the same as in Theorem 2I and Corollary 2I, but with $F(r)$ in (0.15) replaced by

$$G(r) = \sum_j \beta(j) \frac{(j/M)\, e^{-r}}{1 - (j/M)\, e^{-r}} \qquad (r > 0) \tag{0.40}$$

(see Fig. 2, upper curve). The function $G(r)$ has the same properties as formulated in Lemma 1 for $F(r)$.

Combination of the two explicit solutions with (0.15) and (0.40) [recall (0.15)–(0.20)] allows us now to obtain some interesting relations. We shall use upper indices g and l to distinguish between the global and the local cases. In addition, we strengthen (0.32) to the assumption

$$\begin{gathered}\sum_j \beta(j)\left(1-\frac{j}{M}\right)^{-2}<\infty \\ \log M>0>\log\left[\sum_j \beta(j)\, j\right]\end{gathered} \tag{0.41}$$

to ensure that $0<\theta_c^g<h_c^g<1$ and $0<h_c^{*g}<1$.

Lemma 3. For $r \geqslant 0$

$$-\frac{G'(r)}{G(r)}>1+G(r)=-\frac{F'(r)}{F(r)}>1+F(r) \tag{0.42}$$

The proof is elementary; the two inequalities are Lemma 1(iii) and its analogue for $G(r)$. Lemma 3 immediately leads to a comparison of the key quantities:

Theorem 3.

(i) $\theta_c^g<\theta_c^l=h_c^g<h_c^l$.

(ii) $r^{*g} \geqslant r^{*l}$ with strict inequality iff $h>h_c^g$.

(iii) $h_c^{*g} \geqslant h_c^{*l}$ with strict inequality iff $\rho(\beta, h_c^g)>0$.

Thus we find that $\rho(\beta, h) \geqslant \lambda(\beta, h; 0)$ with strict inequality iff $h>h_c^g$, and consequently an *intermediate phase* iff $\rho(\beta, h_c^g)>0$:

(III) Global Survival and Local Extinction. For $h_c^{*l}<h<h_c^{*g}$ we have $\rho(\beta, h)>0>\lambda(\beta, h; 0)$. The population experiences *extreme clustering*. Namely, the density of populated sites decays to zero, but the population on these sites grows so fast that even the overall particle density still grows (see Fig. 2).

Finally we offer the following remark. If the restriction $\sum_i \nu(i, j)=\beta(j)$ in the set $M_{\theta, \beta}$ in (0.6) is crossed out, then the local variational formula in Theorem 1 can be shown to reduce to the global variational formula in Baillon *et al.*[2] [recall (0.37)–(0.39)]. This can be interpreted as saying that the global particle density is carried by parts of the space where the random medium has a statistics that is optimal for the growth (and that is in general atypical under the law β^{Z}). Despite this clear link, we are not able to provide an intuitive explanation for the equality in (0.42).

1. PROOF OF THEOREM 1

1.1. Local Particle Density As Functional of Local Times of Random Walk

Let

$$\begin{gathered} S_0=0, \qquad S_n=X_1+\cdots+X_n \qquad (n\geqslant 1) \\ (X_i) \text{ i.i.d. with } P(X_i=1)=1-P(X_i=0)=h \end{gathered} \tag{1.1}$$

denote the random walk with drift h that serves as the underlying migration process of the particles. Let

$$l_n(x)=|\{0<i\leqslant n: S_i=x\}| \tag{1.2}$$

denote its *local time* at site x up to time n. The following proposition expresses $d_n(x, F)$ in (0.3) as the expectation of some exponential functional of the random sequence $\{l_n(x)\}_{x\geqslant 0}$. Let P_h and E_h denote probability and expectation w.r.t. (S_n).

Proposition 1.

$$d_n(x, F)=E_h\left(\prod_{y\geqslant 0}[b_{x-y}]^{l_n(y)}\,\eta_0(x-S_n)\right) \tag{1.3}$$

Proof. From the evolution mechanism of our process (recall steps (1) and (2) in Section 0.2) we obtain the following recursion relation:

$$E(\eta_n(x)\,|\,F)=(1-h)\,b_x E(\eta_{n-1}(x)\,|\,F)+hb_{x-1}E(\eta_{n-1}(x-1)\,|\,F) \tag{1.4}$$

[recall that b_x is the mean offspring of a particle at site x by (0.1)]. In the notation of (0.3) and with the help of (1.1) this may be rewritten as

$$d_n(x, F)=E_h(b_{x-X_1}d_{n-1}(x-X_1, F)) \tag{1.5}$$

Similarly,

$$d_{n-1}(x-X_1, F)=E_h(b_{x-X_1-X_2}d_{n-2}(x-X_1-X_2, F)) \tag{1.6}$$

and hence by iteration

$$d_n(x, F)=E_h\left(\prod_{i=1}^{n} b_{x-S_i}d_0(x-S_n, F)\right) \tag{1.7}$$

Now substitute $d_0(x, F)=\eta_0(x)$ and use (1.2) to write the product in terms of local times. ∎

From Proposition 1 we obtain the following expressions for the local particle densities figuring in (0.4) and (0.5):

$$d_n^{\mathrm{I}}(0, F) = E_h\left(\prod_{y \geq 0} [b_{-y}]^{l_n(y)}\right) \tag{1.8}$$

$$d_n^{\mathrm{II}}(\lfloor \tau n \rfloor, F) = E_h\left(\prod_{y \geq 0} [b_{\lfloor \tau n \rfloor - y}]^{l_n(y)} 1_{\{S_n \geq \lfloor \tau n \rfloor\}}\right) \tag{1.9}$$

By conditioning on the position of S_n and by using the one-sidedness of the random walk, we can turn the above expressions into a form that is more appropriate to apply techniques from the theory of large deviations. Let

$$l(x) = \lim_{n \to \infty} l_n(x) = |\{i > 0: S_i = x\}| \tag{1.10}$$

denote the *total local time* at site x.

Proposition 2.

$$d_n^{\mathrm{I}}(0, F) = \partial(0, n, \sigma_0 F) \tag{1.11}$$

$$d_n^{\mathrm{II}}(\lfloor \tau n \rfloor, F) = \partial(\lfloor \tau n \rfloor, n, \sigma_{\lfloor \tau n \rfloor} F) \tag{1.12}$$

with

$$\partial(\lfloor \tau n \rfloor, n, F) = \sum_{x \geq \lfloor \tau n \rfloor} P_h(S_n = x)\, E_h(x, n, F) \qquad (\tau \geq 0) \tag{1.13}$$

$$E_h(x, n, F) = E_h\left(\prod_{y=0}^{x} [b_y]^{l(y)} \,\middle|\, \sum_{y=0}^{x} l(y) = n\right) \tag{1.14}$$

and σ_y the map defined by $(\sigma_y F)_x = F_{y-x}$.

Proof. Equation (1.8) can be written as in (1.11) and (1.13) with

$$E_h(x, n, F) = E_h\left(\prod_{y \geq 0} [b_y]^{l_n(y)} \,\middle|\, S_n = x\right) \tag{1.15}$$

Then note that the r.h.s. equals

$$E_h\left(\prod_{y \geq 0} [b_y]^{l_n(y)} \,\middle|\, S_n = x, S_{n+1} > x\right)$$

$$= E_h\left(\prod_{y=0}^{x} [b_y]^{l(y)} \,\middle|\, \sum_{y=0}^{x} l(y) = n\right) \tag{1.16}$$

because $S_n = x$, $S_{n+1} > x$ implies $\sum_{y=0}^{x} l(y) = n$ and vice versa. Similarly for (1.9). ∎

1.2. Large Deviations

For a general reference to large-deviation theory we refer the reader to the books by Ellis[4] and by Deuschel and Stroock.[5]

The role of Proposition 2 is that it reduces the proof of Theorem 1 to a large-deviation problem for S_n and for $\{l(x)\}_{x \geqslant 0}$. Indeed, first note

$$P_h(S_n = x) = \binom{n}{x} h^x (1-h)^{n-x} \tag{1.17}$$

from which it follows that P_h induces a large-deviation family of probability measures for S_n/n on $[0, 1]$ with rate function

$$\lim_{n \to \infty} \frac{1}{n} \log P_h(S_n = \lceil \theta n \rceil) = -I_h(\theta) \tag{1.18}$$

with $I_h(\theta)$ given by (0.9). The same limit is obtained along any sequence $\theta_n \to \theta$, and $I_h(\theta)$ is bounded and continuous on $[0, 1]$. Next write the sum in (1.13) as an integral

$$\begin{aligned} \partial(\lfloor \tau n \rfloor, n, F) = &\int_{\theta \in (\tau_n, 1]} d(\theta n)\, P_h(S_n = \lceil \theta n \rceil)\, E_h(\lceil \theta n \rceil, n, F) \\ &+ 1_{\{\tau = 0\}} (1-h)^n\, b_0^n \end{aligned} \tag{1.19}$$

Here $\tau_n = \lfloor \tau n \rfloor / n$, and the last term corresponds to the event $\{S_n = 0\}$. The notation $\lfloor t \rfloor$ ($\lceil t \rceil$) denotes the largest (smallest) integer smaller (larger) than or equal to t. We shall prove in Section 1.3 the following result for the second factor $E_h(\lceil \theta n \rceil, n, F)$.

Proposition 3. For every $\theta \in (0, 1]$

$$\lim_{n \to \infty} \frac{1}{n} \log E_h(\lceil \theta n \rceil, n, F) = J_\beta(\theta) \qquad F\text{-a.s.} \tag{1.20}$$

with $J_\beta(\theta)$ given by (0.14). The same limit is obtained along any sequence $\theta_n \to \theta$, and $J_\beta(\theta)$ is bounded and continuous on $(0, 1]$.

Proposition 3 finishes the proof of Theorem 1 as follows. First of all, since the limit in (1.20) is F-a.s. constant, the same result holds when F is replaced by $\sigma_0 F$ or $\sigma_{\lfloor \tau n \rfloor} F$. Next, we apply Varadhan's theorem (see Deuschel and Stroock,[5] Theorem 2.1.10 and Exercise 2.1.20) to (1.19) and obtain

$$\begin{aligned} &\lim_{n \to \infty} \frac{1}{n} \log \partial(\lfloor \tau n \rfloor, n, \sigma_{\lfloor \tau n \rfloor} F) \\ &\quad = \sup_{\theta \in (\tau, 1]} [J_\beta(\theta) - I_h(\theta)] \qquad F\text{-a.s.} \end{aligned} \tag{1.21}$$

Here we use the continuity of $J_\beta(\theta)$ on $(0, 1]$ and of $I_h(\theta)$ on $[0, 1]$. Since $J_\beta(0) = \log b_0$ and $I_h(0) = -\log(1-h)$, we see that also the last term in (1.19) is taken care of. Finally, in Section 2 we shall see that $\lim_{\theta \downarrow 0} J_\beta(\theta) = \log M$. Since $b_0 \leqslant M$ and therefore $J_\beta(0) = \log b_0 \leqslant \log M$, it follows that we may replace $[\tau, 1]$ by $[\tau, 1] \cap (0, 1]$ under the supremum and thus circumvent the technical problem that (0.14) is not defined for $\theta = 0$. ∎

1.3. Proof of Proposition 3

Throughout this section we may assume that $\theta > 0$. The important fact about $E_h(x, n, F)$ in (1.14) is that this quantity is defined entirely in terms of the random sequence $\{l(x)\}_{x \geqslant 0}$, which has the nice property of being i.i.d. with a simple common distribution, namely $P_h(l(x) = i) = \pi_h(i) = h(1-h)^{i-1}$ given by (0.10). This makes it a suitable object for a large-deviation analysis.

First we note that $E_h(x, n, F)$ does not depend on the drift h. Indeed, (1.14) is a conditional expectation given $S_n = x$, $S_{n+1} > x$ [recall (1.16)] and all paths with this property have the same probability, namely $h^{x+1}(1-h)^{n-x}$. Thus we have

$$E_h(\lceil \theta n \rceil, n, F) = E_\theta(\lceil \theta n \rceil, n, F) \tag{1.22}$$

Changing the drift from h to θ [recall (1.1)] will be convenient because it allows us to write

$$E_\theta(\lceil \theta n \rceil, n, F) = \exp[o(n)]\, E_\theta\left(\prod_{y=0}^{\lceil \theta n \rceil} [b_y]^{l(y)}\, 1_{\{\sum_{y=0}^{\lceil \theta n \rceil} l(y) = n\}} \right) \tag{1.23}$$

The error term is just

$$P_\theta^{-1}\left(\sum_{y=0}^{\lceil \theta n \rceil} l(y) = n \right) = P_\theta^{-1}(S_n = \lceil \theta n \rceil, S_{n+1} > \lceil \theta n \rceil)$$

which is $\exp[o(n)]$ because of (1.18) and $I_\theta(\theta) = 0$.

Next, let

$$\nu_N = \frac{1}{N} \sum_{x=0}^{N-1} \delta_{(l(x), b_x)} \tag{1.24}$$

denote the *empirical distribution* over the interval $[0, N)$ of the total local time process and the random medium combined. Then we may write for the expectation in the r.h.s. of (1.23) the following integral:

$$S_\omega(n) = \int_{A_n} \exp[K_n f(\nu)]\, P_{\theta, \omega, K_n}(\nu_{K_n} \in d\nu) \tag{1.25}$$

Here $P_{\theta,\omega,N}$ denotes the probability distribution of ν_N living on $\mathscr{P}(\mathbb{N}\times\operatorname{supp}\beta)$ and induced by P_θ at fixed ω; f is given by (0.7); and

$$\begin{aligned} A_n &= \left\{ \nu \in \mathscr{P}(\mathbb{N}\times\operatorname{supp}\beta):\ \sum_{i,j} i\nu(i,j) = L_n \right\} \\ K_n &= \lceil \theta n \rceil + 1 \\ L_n &= n/K_n \end{aligned} \tag{1.26}$$

The index ω abbreviates the fixed medium $\omega = \{b_x\}_{x\in\mathbb{Z}}$. We want to apply Varadhan's theorem to the integral in (1.25), but there are three problems along the way:

I. ω is fixed. This means that in ν_N in (1.24) only the first of the indices $(l(x), b_x)$ is random.

II. A_n is a moving set and is not closed (in the weak topology).

III. $f(\nu)$ is not continuous in ν (in the weak topology).

We shall now explain how these problems may be circumvented. For more technical details we refer the reader to Greven and den Hollander.[1]

Problem I. The key observation is that the law of ν_N is *invariant under permutations* of $\{b_x\}_{x=0}^{N-1}$ due to the i.i.d. property of $\{l(x)\}_{x\geqslant 0}$. Hence (1.25) is the same for all realizations of the random medium with the same empirical distribution over the interval $[0, N)$ defined by

$$\tilde{\nu}_N(\omega) = \frac{1}{N}\sum_{x=0}^{N-1} \delta_{b_x} \tag{1.27}$$

Next we note that for every $\delta > 0$ by the strong law of large numbers

$$\begin{gathered} \tilde{\nu}_N(\omega) \in B^\delta \text{ for } N \text{ sufficiently large } \omega\text{-a.s.} \\ B^\delta = \{\tilde{\nu} \in \mathscr{P}(\operatorname{supp}\beta):\ \|\tilde{\nu} - \beta\| \leqslant \delta\} \end{gathered} \tag{1.28}$$

where $\|\cdot\|$ denotes total variation norm. Moreover, our assumption (0.2) on the boundedness of supp β is easily seen to yield the estimate

$$\begin{gathered} |\log S_\omega(n) - \log S_{\omega'}(n)| \leqslant \delta n \log\left(\frac{M}{m}\right) \\ \text{for all } \omega,\ \omega' \text{ such that } \tilde{\nu}_{K_n}(\omega),\ \tilde{\nu}_{K_n}(\omega') \in B^\delta \end{gathered} \tag{1.29}$$

Here we abbreviate

$$\begin{aligned} M &= \text{supremum of supp } \beta \\ m &= \text{infimum of supp } \beta \end{aligned} \tag{1.30}$$

and the estimate follows from (1.25) and (1.26) by noting that $f(v)$ cannot vary more than $L_n \delta \log(M/m)$ because of the restriction $\sum_{i,j} iv(i,j) = L_n$ [recall (0.7)]. By combining (1.28) and (1.29), we arrive at

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} |\log S_\omega(n) - \log S^\delta(n)| = 0 \qquad \omega\text{-a.s.} \tag{1.31}$$

where we introduce $S^\delta(n)$ as the integral of $S_\omega(n)$ over the ball B^δ

$$S^\delta(n) = \int_{B^\delta} S_\omega(n)\, P_{\beta, K_n}(\tilde{v}_{K_n}(\omega) \in d\tilde{v}) \tag{1.32}$$

Here $P_{\beta,N}$ denotes the probability measure for $\tilde{v}_N(\omega)$ on $\mathscr{P}(\text{supp}\,\beta)$ induced by the i.i.d. measure with marginal β for ω. The latter two equations are an important step: not only have we shown that the limit is a.s. *independent* of ω, by combining (1.25) and (1.32) we see that $S^\delta(n)$ is a double integral over walk and medium, i.e., an expectation over the *double process* $\{l(x), b_x\}_{x \geq 0}$, which is i.i.d. in both coordinates and therefore amenable to large-deviation arguments.

Problem II. For every $\varepsilon > 0$ since $L_n \to \theta^{-1}$ as $n \to \infty$

$$\begin{gathered} A_n \in A^\varepsilon \text{ for } n \text{ sufficiently large} \\ A^\varepsilon = \left\{ v \in \mathscr{P}(\mathbb{N} \times \text{supp}\,\beta) \colon \ \sum_{i,j} iv(i,j) \in [\theta^{-1} - \varepsilon, \theta^{-1} + \varepsilon] \right\} \end{gathered} \tag{1.33}$$

One easily sees that $f(v)$ cannot vary more than $\varepsilon \log(M/m)$ over the slab A^ε [recall (0.7)]. Therefore, uniformly in δ, since $K_n \sim \theta n$,

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} |\log S^\delta(n) - \log S^{\delta,\varepsilon}(n)| = 0 \tag{1.34}$$

where

$$S^{\delta,\varepsilon}(n) = \int_{M^{\delta,\varepsilon}_{\theta,\beta}} \exp[K_n f(v)]\, P_{\theta,\beta,K_n}(v_{K_n} \in dv) \tag{1.35}$$

Here $P_{\theta,\beta,N}$ denotes the probability measure for v_N on $\mathscr{P}(\mathbb{N} \times \text{supp}\,\beta)$, the empirical distribution of the double process, and we introduce the set

$$M^{\delta,\varepsilon}_{\theta,\beta} = \{ v \in A^\varepsilon \colon \tilde{v} \in B^\delta \} \tag{1.36}$$

We also once more use the remark below (1.23), which guarantees that $P_{\theta,\beta,K_n}(A_n)/P_{\theta,\beta,K_n}(A^\varepsilon)=\exp(c_\varepsilon n)$ with $c_\varepsilon\to 0$ as $\varepsilon\to 0$.

Problem III. The final step in the chain of approximations is to replace $\mathbb{N}$ by the finite set $\{1,..., R\}$. For simplicity, assume that $\operatorname{supp}\beta$ is finite. Then it may be shown that *uniformly* in δ and ε

$$\lim_{R\to\infty}\ \limsup_{n\to\infty}\frac{1}{n}\,|\log S^{\delta,\varepsilon}(n)-\log S^{\delta,\varepsilon,R}(n)|=0 \tag{1.37}$$

where

$$S^{\delta,\varepsilon,R}(n)=\int_{M_{\theta,\beta}^{\delta,\varepsilon,R}}\exp[K_n f(v)]\,P_{\theta,\beta,K_n}(v_{K_n}\in dv) \tag{1.38}$$

with

$$M_{\theta,\beta}^{\delta,\varepsilon,R}=M_{\theta,\beta}^{\delta,\varepsilon}\cap\mathscr{P}(\{1,..., R\}\times\operatorname{supp}\beta) \tag{1.39}$$

For the proof the reader is referred to Greven and den Hollander,[1] Lemma 15. To give just a rough idea: Any realization of $\{l(x)\}_{x\geqslant 0}$ inside $[0, K_n)$ with fixed $\sum_{x=0}^{K_n-1} l(x)$ has the same probability because $P_\theta(l(x)=i)=\pi_\theta(i)=\theta(1-\theta)^{i-1}$. Consider a level set $C_b=\{x\in\mathbb{Z}: b_x=b\}$ of the medium. For every site $x\in C_b\cap[0, K_n)$ where $l(x)>R$, remove the overshoot in units of size $\frac{1}{2}R$ and transport each unit to some $x'\in C_b\cap[0, K_n)$ where $l(x')\leqslant\frac{1}{2}R$. Since C_b has positive density (because $\operatorname{supp}\beta$ is finite), all the overshoot can be thus truncated and transported within the level sets, provided R is sufficiently large so that the number of units transported does not exceed $|C_b|$. The truncation has no effect on the exponential [leaves $f(v_{K_n})$ invariant]. No more than

$$\sum_{x=0}^{K_n-1} l(x)/\tfrac{1}{2}R\leqslant K_n(\theta^{-1}+\varepsilon)/\tfrac{1}{2}R\leqslant 4n/R$$

units are transported and so the truncation has entropy $\exp(c_R n)$ with $c_R\to 0$ as $R\to\infty$.

Having thus solved Problems I–III, we are now finally ready to apply Varadhan's theorem and complete the proof of Proposition 3. Take the integral in (1.38). Since $\{l(x), b_x\}_{x\geqslant 0}$ is i.i.d. with common distribution $\pi_\theta(i)\,\beta(j)$, it follows from Sanov's theorem that $P_{\theta,\beta,N}(v_N\in dv)$ is a large-deviation family on $\mathscr{P}(\mathbb{N}\times\operatorname{supp}\beta)$ with rate function $I_{\theta,\beta}(v)$ given by (0.8) (see Deuschel and Stroock,[5] Theorem 3.2.17). $I_{\theta,\beta}(v)$ is continuous in v and $M_{\theta,\beta}^{\delta,\varepsilon,R}$ is a *closed slab* in $\mathscr{P}(\mathbb{N}\times\operatorname{supp}\beta)$ equipped with the weak topology of measures. These two facts imply, via a standard argument, that the large-deviation principle holds on $M_{\theta,\beta}^{\delta,\varepsilon,R}$. Since f is bounded and continuous on $M_{\theta,\beta}^{\delta,\varepsilon,R}$, Varadhan's theorem applied to (1.38) gives

$$\lim_{n\to\infty} \frac{1}{n} \log S^{\delta,\varepsilon,R}(n) = \theta \sup_{v \in M_{\theta,\beta}^{\delta,\varepsilon,R}} [f(v) - I_{\theta,\beta}(v)] \tag{1.40}$$

It is easily checked that the r.h.s. of (1.40) converges to $J_\beta(\theta)$ given by (0.14) as $\delta \to 0$, $\varepsilon \to 0$, $R \to \infty$ (in this order). The reader is referred to Greven and den Hollander,[1] Lemma 20. Tracing back the various steps in the argument and recalling (1.31), (1.34), and (1.37), we see that we have proved (1.20) in Proposition 3.

The proof for $|\mathrm{supp}\, \beta|$ infinite follows trivially: approximate β by a distribution with finite support in total variation norm. Equation (1.20) carries over immediately, with $J_\beta(\theta)$ again given by (0.14).

The above proof clearly shows that the same limit is obtained in (1.20) along any sequence $\theta_n \to \theta$ (perturbation of K_n). The remaining claims, namely boundedness and continuity of $J_\beta(\theta)$ on $(0, 1]$, will follow from our analysis in Section 2. ▮

2. PROOF OF THEOREM 2I AND COROLLARY 2I

2.1. Reduction of Variational Formula

Our first step is to reduce the variational formula in Theorem 1 to a simpler form involving functions ξ: supp $\beta \to (0, 1]$. It will be clear later on that the reduced variational formula can be understood as a supremum over all *random walks in random environment*, pausing with probability $1-\xi(j)$ and stepping to the right with probability $\xi(j)$ on sites where the medium takes the value j, i.e., $\xi(j)$ plays the role of *local drift* of the walk.

Proposition 4. For $h \in (0, 1)$

$$\lambda(\beta, h; \tau) = \sup_{\theta \in [\tau,1] \cap (0,1]} J_{\beta,h}(\theta) \tag{2.1}$$

with

$$\begin{aligned} J_{\beta,h}(\theta) &= J_\beta(\theta) - I_h(\theta) \\ &= \log[M(1-h)] - \theta \log\left(\frac{1-h}{h}\right) - \theta K(\theta) \end{aligned} \tag{2.2}$$

$$\begin{aligned} K(\theta) = &\inf_{\xi \in N_{0,\beta}} \sum_j \beta(j)\, \xi^{-1}(j) \\ &\times \left\{ \log\left(\frac{M}{j}\right) + \xi(j) \log \xi(j) + [1-\xi(j)] \log[1-\xi(j)] \right\} \end{aligned} \tag{2.3}$$

where

$$N_{\theta,\beta}=\left\{\xi\colon \operatorname{supp}\beta\to(0,1]\text{ measurable: }\sum_j \beta(j)\,\xi^{-1}(j)=\theta^{-1}\right\} \tag{2.4}$$

Proof. We start with the observation that the set of ν-measures $M_{\theta,\beta}$ in (0.6) may be decomposed as

$$M_{\theta,\beta}=\bigcup_{\xi\in N_{\theta,\beta}} A_{\xi,\beta} \tag{2.5}$$

with

$$A_{\xi,\beta}=\{\nu(i,j)=\beta(j)\,\mu_{\xi(j)}(i)\colon \mu_{\xi(j)}\in B_{\xi(j)}\text{ for all }j\in\operatorname{supp}\beta\} \tag{2.6}$$

$$B_\zeta=\left\{\mu\in\mathscr{P}(\mathbb{N})\colon\ \sum_i i\mu(i)=\zeta^{-1}\right\}\qquad (\zeta\in(0,1]) \tag{2.7}$$

Define

$$I_\theta(\zeta)=\zeta\log\left(\frac{\zeta}{\theta}\right)+(1-\zeta)\log\left(\frac{1-\zeta}{1-\theta}\right) \tag{2.8}$$

$$\pi_\zeta(i)=\zeta(1-\zeta)^{i-1} \tag{2.9}$$

[compare with (0.9) and (0.10)] and

$$H(\mu\mid\pi)=\sum_i \mu(i)\log\left(\frac{\mu(i)}{\pi(i)}\right) \tag{2.10}$$

Note that for any $\mu\in B_\zeta$

$$\begin{aligned}\sum_i \mu(i)\log\left(\frac{\pi_\theta(i)}{\pi_\zeta(i)}\right)&=\left[\sum_i\mu(i)\right]\log\left(\frac{\theta}{\zeta}\right)+\left[\sum_i(i-1)\,\mu(i)\right]\log\left(\frac{1-\theta}{1-\zeta}\right)\\&=\log\left(\frac{\theta}{\zeta}\right)+[\zeta^{-1}-1]\log\left(\frac{1-\theta}{1-\zeta}\right)\\&=-\zeta^{-1}I_\theta(\zeta)\end{aligned} \tag{2.11}$$

This allows us now to rewrite (0.14). Indeed, substitute (0.7), (0.8), and (0.10) into (0.14) and use (2.11), to get

$$\begin{aligned}J_\beta(\theta)=\theta\sup_{\xi\in N_{\theta,\beta}}\ &\sup_{\substack{\mu_{\xi(j)}\in B_{\xi(j)}\\ j\in\operatorname{supp}\beta}}\sum_k\beta(k)\\&\times[\xi^{-1}(k)\{\log k-I_\theta(\xi(k))\}-H(\mu_{\xi(k)}\mid\pi_{\xi(k)})]\end{aligned} \tag{2.12}$$

The reason for introducing $\pi_{\xi(j)}$ in (2.12) is that we now can perform the second supremum, since

$$\inf_{\mu_{\xi(j)} \in B_{\xi(j)}} H(\mu_{\xi(j)} \mid \pi_{\xi(j)}) = 0 \qquad \text{for all } j \tag{2.13}$$

Indeed, $\pi_{\xi(j)} \in B_{\xi(j)}$, and the relative entropy $H(\mu \mid \pi)$ in (2.10) is non-negative and assumes the value zero iff $\mu = \pi$. Thus we obtain the reduced expression

$$J_\beta(\theta) = \theta \sup_{\xi \in N_{\theta,\beta}} \sum_j \beta(j)\, \xi^{-1}(j)\{\log j - I_\theta(\xi(j))\} \tag{2.14}$$

From this one easily gets (2.1)–(2.4) by substituting (0.9) and canceling a term $\theta \log \theta + (1-\theta)\log(1-\theta)$ via the property $\sum_j \beta(j)\,\xi^{-1}(j) = \theta^{-1}$ in the definition of $N_{\theta,\beta}$. ∎

Proposition 4 is an important step because we have performed the supremum over the first marginal of v. Now, if supp β is finite, then (2.3) is a finite variational problem which may be solved by the technique of Lagrange multipliers. This will be carried out in Proposition 5 below. If supp β is infinite, then this technique still provides us with the correct solution, but some mathematical justification is needed. We shall not insist on this point and refer the reader to Baillon *et al.*,[2] where the right functional analytic tools have been developed.

Incidentally, note that the above argument gives the maximizer $\bar{v}$ in (0.14) as a function of the minimizer $\bar{\xi}$ in (2.3), namely

$$\bar{v}(i,j) = \pi_{\bar{\xi}(j)}(i)\,\beta(j) \tag{2.15}$$

This proves (0.22) once we obtain $\bar{\xi}$.

2.2. Variation over ξ

A key role is played by the function $F(r)$ in (0.15) and the constant θ_c in (0.17). We recall that both $\theta_c = 0$ and $\theta_c > 0$ are possible.

Proposition 5. $\theta \to K(\theta)$ is continuous on $(0, 1]$, strictly increasing and analytic on $(\theta_c, 1)$, and constant on $(0, \theta_c]$. Furthermore, $\theta K(\theta)$ is convex on $(0, 1]$ and $\lim_{\theta \downarrow 0} \theta K(\theta) = 0$. There are two cases

I. If $\theta \in [\theta_c, 1] \cap (0, 1]$, then (2.3) achieves a minimum in $N_{\theta,\beta}$. The minimizer is unique and is given by

$$\bar{\xi}(j) = 1 - \frac{j}{M} e^{-r} \tag{2.16}$$

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

$$\theta = -\frac{F(r)}{F'(r)} \tag{2.17}$$

The minimum is

$$K(\theta) = -\frac{r}{\theta} + \log\frac{1}{F(r)} \tag{2.18}$$

II. If $\theta \in (0, \theta_c)$, then (2.3) does not achieve a minimum in $N_{\theta,\beta}$. The minimizer is independent of θ and is given by $\bar{\xi}(j) = 1 - j/M$, which is in $N_{\theta_c,\beta}$. The minimum is $K(\theta) = K(\theta_c) = \log[1/F(0)]$.

Proof. The Lagrangian of (2.3) equals

$$\begin{aligned}\mathscr{L}(\xi) = \sum_j \beta(j)\, \xi^{-1}(j) \Big\{ & \log\binom{M}{j} + \xi(j) \log \xi(j) \\ & + [1 - \xi(j)] \log[1 - \xi(j)] + r \Big\}\end{aligned} \tag{2.19}$$

with multiplier r. This yields

$$0 = \frac{\partial \mathscr{L}}{\partial \xi(j)} = -\beta(j)\, \xi^{-2}(j) \left\{ \log\binom{M}{j} + \log[1 - \xi(j)] + r \right\} \tag{2.20}$$

Hence

$$\xi(j) = 1 - \frac{j}{M} e^{-r} \tag{2.21}$$

The multiplier must be chosen such that

$$\theta^{-1} = \sum_j \beta(j)\, \xi^{-1}(j) = \sum_j \beta(j) \left(1 - \frac{j}{M} e^{-r}\right)^{-1} \tag{2.22}$$

One easily checks from (0.15) that the last sum equals $-F'(r)/F(r)$. This proves (2.16) and (2.17). Recall Lemma 1(ii) in Section 0.4 to see that $r = r(\theta)$ is unique and analytic. Substitution of (2.16) into (2.3) yields (2.18) via (2.22). This completes case I.

In case II for $\theta_c > 0$ a remarkable phenomenon occurs. Namely, (2.22) has no solution when $\theta \in (0, \theta_c)$ [recall (0.17) and Lemma 1(ii)]. This means that the minimum is not achieved in $N_{\theta,\beta}$. A closer inspection reveals that as θ crosses θ_c downward the minimizer sticks at its value for

$\theta=\theta_c$. That is, the minimizer is the one for $r=0$ in $N_{\theta_c,\beta}$ for all $\theta \leqslant \theta_c$. Again, for a detailed analysis one would have to use the functional analytic tools developed in Baillon *et al.*[2]

The remaining claims follow from (2.17) and (2.18). Namely,

$$K'(\theta)=\frac{r}{\theta^2}-\frac{dr}{d\theta}\left[\frac{1}{\theta}+\frac{F'(r)}{F(r)}\right]=\frac{r}{\theta^2} \tag{2.23}$$

which gives

$$(\theta K(\theta))'=\log\frac{1}{F(r)} \tag{2.24}$$

$$(\theta K(\theta))''=\frac{1}{\theta}\frac{dr}{d\theta}>0 \tag{2.25}$$

Recall Lemma 1(ii). One easily checks that $\lim_{\theta\downarrow 0}\theta K(\theta)=0$ using (2.18) together with (0.15) and (2.22). ∎

2.3. Variation over Θ

By combining (2.2) with what we know about $K(\theta)$ from Proposition 5, we get a picture of how $J_{\beta,h}(\theta)$ depends on θ.

Proposition 6. For every $h\in(0,1)$ fixed, $\theta\to J_{\beta,h}(\theta)$ is continuous and concave on $(0,1]$, strictly concave and analytic on $(\theta_c,1)$, and linear on $(0,\theta_c]$. Furthermore, $\lim_{\theta\downarrow 0}J_{\beta,h}(\theta)=\log[M(1-h)]$ and $J_{\beta,h}(1)=\log h+\sum_j \beta(j)\log j$.

I. If $h<h_c$, then $J_{\beta,h}(\theta)$ is strictly decreasing on $(0,1]$.

II. If $h=h_c$, then $J_{\beta,h}(\theta)$ is constant on $(0,\theta_c]$ and strictly decreasing on $(\theta_c,1]$.

III. If $h>h_c$, then $J_{\beta,h}(\theta)$ has strictly positive slope at $\theta=0$ and achieves a unique maximum at $\bar\theta\in(\theta_c,1)$ given by

$$\bar\theta=-\frac{F(r)}{F'(r)} \tag{2.26}$$

with r the unique solution of

$$h=\frac{1}{1+F(r)} \tag{2.27}$$

Proof. Most statements are immediate from Proposition 5. The strict concavity on $(\theta_c,1)$ follows from (2.25), because $dr/d\theta>0$ via (2.17) and

Lemma 1(ii). The distinction between cases I–III lies in the slope of $J_{\beta,h}(\theta)$ at $\theta = 0$. Indeed, from (2.2) we have

$$\frac{\partial}{\partial\theta} J_{\beta,h}(\theta) = \log\left(\frac{hF(r(\theta))}{1-h}\right) \tag{2.28}$$

with $r(\theta)$ obtained from (2.17). The limit $\theta \downarrow 0$ corresponds to $r \downarrow 0$ (irrespective of whether $\theta_c = 0$ or $\theta_c > 0$) ad hence the slope at $\theta = 0$ changes from positive to negative at $h = h_c$ gives by (0.16). From (2.28) we also read off (2.27) ad (2.26). ▮

With Proposition 6 we are now ready to perform the supremum over θ in (2.1) and compute $\lambda(\beta, h; 0)$.

Proof of Theorem 2I. The supremum runs over $\theta \in (0, 1]$. Equations (0.21) and (0.22) are immediate [recall (2.15) and (2.21)]. Equation (0.20) follows after substitution of (2.18) into (2.2). ▮

Proof of Corollary 2I. Continuity and analyticity are immediate via Lemma 1(i) and (ii). It is clear that $h \to \lambda(\beta, h; 0)$ is strictly decreasing on $(0, h_c]$. To see that it is also strictly decreasing on $(h_c, 1)$, differentiate (2.27) w.r.t. h and use (2.26) to get

$$\frac{dr}{dh} = -\frac{[1+F(r)]^2}{F'(r)} = \frac{\bar{\theta}}{h(1-h)} \tag{2.29}$$

Hence from (0.18) and (0.19)

$$\frac{\partial}{\partial h} \lambda(\beta, h; 0) = -\frac{1}{1-h} + \frac{dr}{dh} = -\frac{h-\bar{\theta}}{h(1-h)} \tag{2.30}$$

But $\bar{\theta} < h$ by combining (2.26) and (2.27) with Lemma 1(iii). The change of slope of $\lambda(\beta, h; 0)$ at $h = h_c$ equals $\lim_{h \downarrow h_c} dr/dh$. Use (2.29) and note that $\lim_{h \downarrow h_c} \bar{\theta} = \theta_c$ by (0.16), (0.17), (2.26), and (2.27). The remaining claims are obvious. ▮

REFERENCES

1. A. Greven and F. den Hollander, Branching random walk in random environment: phase transitions for local and global growth rates, *Prob. Th. Rel. Fields*, to appear.
2. J.-B. Baillon, Ph. Clément, A. Greven, and F. den Hollander, A variational approach to branching random walk in random environment, *Ann. Prob.*, to appear.
3. J.-B. Baillon, Ph. Clément, A. Greven, and F. den Hollander, On a variational problem for an infinite particle system in a random medium, preprint (1991).
4. R. S. Ellis, *Entropy, Large Deviations and Statistical Mechanics* (Springer, Berlin, 1985).
5. J.-D. Deuschel and D. W. Stroock, *Large Deviations* (Academic Press, Boston, 1989).