J. reine angew. Math. **454** (1994), 181—217

Journal für die reine und
angewandte Mathematik

On a variational problem for an infinite particle system in a random medium

By *J.-B. Baillon* at Villeurbanne, *Ph. Clément* at Delft, *A. Greven* at Berlin and *F. den Hollander* at Utrecht

Part I. The global growth rate

---

1. Introduction

1.1. Motivation. In this paper we analyse and solve a *variational problem* that has been found by Greven and den Hollander [10] in a study of population growth in a random medium. We briefly describe the model and formulate the main result so as to keep our exposition self-contained. For further details, as well as for interpretation, we refer the reader to the original paper.

Briefly, the model is an infinite system of particles living on the integer lattice $\mathbb{Z}$ and subject to two random mechanisms:

(1) Particles *branch* according to *site-dependent* offspring distributions constituting a *random medium.*

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

One of the main points in the original paper was to show that the long term behaviour of the system can be extracted from an underlying variational principle. In particular, two variational formulas were derived, whose maxima are the exponential growth rate of the global resp. local population density and whose maximisers provide information about the path of descent of a typical particle in the global resp. local population. Here global refers to the population on the whole of $\mathbb{Z}$ and local to the population on a single site. However, both these formulas have a rather complex structure, and in order to get a clear picture of what is going on in the particle system a closer analysis via functional analytic techniques is required. It is the purpose of the present paper to carry out this analysis for the global variational formula. The local variational formula is of a different type and will be analyzed in a separate paper (Greven and den Hollander [11]). It turns out that the maximum and the maximisers in the global formula exhibit interesting *phase transitions* as the drift varies. This is due to the competition between the branching and the migration, which changes with the drift.

Our paper is organised as follows. In section 1 we define the model, formulate the global formula (Theorem 1 below), and present its solution (Theorem 2 below). The latter embodies the main result of this paper. In sections 2–5 we give the proof of Theorem 2. In section 2 we show how the variational formula can be transformed into an eigenvalue problem for a 1-parameter family of $\mathbb{N} \times \mathbb{N}$ matrices. In section 3 we do the spectral analysis. In section 4 we connect the results. In section 5 we prove an important inequality implying a monotonicity property as a function of the drift.

Much of the analysis in the present paper grew out of a study of a simpler version of the particle system in Baillon et al. [2]. The latter paper also contains a detailed evaluation of the role of the maximum and of the maximisers in the description of the particle system.

1.2. 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 marginal distribution α. F plays the role of a *random medium*. For fixed F, define a discrete time Markov process (η_n) on state space $\mathbb{N}^{\mathbb{Z}}$, with the interpretation

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

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

the evolution of which is as follows. At time $n = 0$ place one particle at every site, i.e., $\eta_0(x) \equiv 1$. 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, i.e. it consists of k new particles with probability $F_x(k)$, $k \geqq 0$. Immediately after creation each new particle independently decides to jump to one of the nearest-neighbour sites, choosing right with probability $\frac{1}{2}(1+h)$ and left with probability $\frac{1}{2}(1-h)$. The parameter $h \in [0, 1]$ is the *drift* and is the same for all x. The resulting sequence of particle numbers make up the state η_{n+1} at time $n+1$, etc. F stays fixed during the evolution.

Let

$$b_x = \sum_{k=0}^{\infty} k F_x(k) \tag{1.1}$$

denote the *mean offspring* at site x and let β denote the distribution of b_x induced by α. It is assumed that

$$0 < \inf_x b_x < \sup_x b_x = M < \infty, \tag{1.2}$$

i.e., β has bounded support (with maximal value M) and has strictly positive variance.

1.3. Growth rate of global particle density. For given F let

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

denote the *global particle density* at time n. From the properties of the evolution mechanism together with the individual ergodic theorem one deduces that

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

where E denotes the double expectation over the Markov process (η_n) given F as well as over F. The a.s. in (1.4) refers to the joint distribution of (η_n) and F. Thus, $D(F,\eta_n)$ a.s. does not depend on the realisation of F and η_n, although it does of course depend on their distribution via the two parameters β and h.

In Greven and den Hollander [10] it is shown that in the long time limit the global particle density grows exponentially fast at rate

(1.5) $$\varrho(\beta,h) = \lim_{n\to\infty} \frac{1}{n} \log D(F,\eta_n) \quad \text{a.s.}$$

and that $\varrho(\beta,h)$ can be computed in the form of a *variational formula*. To formulate this expression in Theorem 1 below we need the following symbols. Let $\mathscr{P}(\mathbb{N}^2)$ denote the set of probability measures on $\mathbb{N}^2$, $\langle\cdot,\cdot\rangle$ inner product over $\mathbb{N}^2$, $a(i,j) = i+j-1$ and $\hat{v}(i) = \sum_j v(i,j)$. Define

(1.6) $$M_\theta = \{v \in \mathscr{P}(\mathbb{N}^2) : \langle a, v\rangle = \theta^{-1}, \sum_{j\in\mathbb{N}} v(i,j) = \sum_{j\in\mathbb{N}} v(j,i) \text{ for } i \in \mathbb{N}\} \qquad (\theta\in(0,1]),$$

(1.7) $$f(i) = \log \int b^i \beta(db) \qquad (i\in\mathbb{N}),$$

(1.8) $$I_\theta(v) = \sum_{i,j\geqq 1} v(i,j) \log\left(\frac{v(i,j)}{\hat{v}(i)\,P_\theta(i,j)}\right) \qquad (\theta\in[0,1],\ v\in\mathscr{P}(\mathbb{N}^2)),$$

(1.9) $$I_h(\theta) = \frac{1}{2}(1+\theta)\log\left(\frac{1+\theta}{1+h}\right) + \frac{1}{2}(1-\theta)\log\left(\frac{1-\theta}{1-h}\right) \qquad (\theta\in[0,1],\ h\in[0,1)),$$

(1.10) $$P_\theta(i,j) = \binom{i+j-2}{i-1}\left[\frac{1}{2}(1+\theta)\right]^i \left[\frac{1}{2}(1-\theta)\right]^{j-1} \qquad (\theta\in[0,1],\ i,j\in\mathbb{N}).$$

Theorem 1 (Greven and den Hollander [10]). *For* $h \in (0,1)$

$$\varrho(\beta, h) = \sup_{\theta \in (0,1)} \sup_{\nu \in M_\theta} \{\theta[\langle f \circ a, \nu\rangle - I_\theta(\nu)] - I_h(\theta)\}, \tag{1.11}$$

and for $h = 0$ *and* $h = 1$

$$\begin{aligned} \varrho(\beta, 0) &= \log M, \\ \varrho(\beta, 1) &= \log \textstyle\int b\beta(db). \end{aligned} \tag{1.12}$$

The main difficulty in the analysis of (1.11) comes from the second supremum. This supremum involves a non-linear functional on an ∞-dimensional space subject to a linear constraint.

1.4. Solution of variational formula. Before we can state our solution of (1.11) we need to introduce the following operator:

$$A(i,j) = e^{-g(i+j-1)} P(i,j) \quad (i,j \in \mathbb{N}). \tag{1.13}$$

Here

$$g(i) = i \log M - f(i), \tag{1.14}$$

$$P(i,j) = \binom{i+j-2}{i-1} \left(\frac{1}{2}\right)^{i+j-1}. \tag{1.15}$$

Recall that M is the maximal value in (1.2). Note that (1.15) is the Markov transition kernel of (1.10) at $\theta = 0$.

In section 2 we show that (1.11) can be reduced to finding the largest eigenvalue $\lambda(r)$ and corresponding eigenvector τ_r in $l^2(\mathbb{N})$ of the 1-parameter family of matrices

$$A_r(i,j) = e^{-r(i+j-1)} A(i,j) \quad (r \geqq 0). \tag{1.16}$$

In section 3 we derive various properties of $\lambda(r)$ needed for the analysis of $\varrho(\beta, h)$. E.g. $\lambda(r)$ is a simple eigenvalue in $(0,1)$, and $r \to \lambda(r)$ is analytic, strictly decreasing and strictly log convex on $[0, \infty)$.

The formulation of our main result, namely the solution of (1.11) in Theorem 2 below, uses four more quantities, $\theta_c(\beta)$, $h_c(\beta)$, $r(\beta, h)$ and $\theta(\beta, h)$, defined in terms of $\lambda(r)$:

$$\theta_c = \theta_c(\beta) = \left(-\frac{\lambda'(0)}{\lambda(0)}\right)^{-1}, \tag{1.17}$$

$$h_c = h_c(\beta) = \frac{1 - \lambda^2(0)}{1 + \lambda^2(0)}, \tag{1.18}$$

$$(1.19)\qquad \text{if } h \leqq h_c\colon \quad r(\beta,h)=0\,,$$
$$\theta(\beta,h)=0\,,$$
$$\text{if } h > h_c\colon \quad r=r(\beta,h) \text{ is the unique solution of } h=\frac{1-\lambda^2(r)}{1+\lambda^2(r)}\,,$$
$$\theta(\beta,h)=\left(-\frac{\lambda'(r)}{\lambda(r)}\right)^{-1} \quad \text{at} \quad r=r(\beta,h)\,.$$

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

$$(1.20)\qquad \varrho(\beta,h)=\log\left[M(1-h^2)^{\frac{1}{2}}\right] \qquad \textit{if } 0\leqq h\leqq h_c\,,$$
$$=\log\left[M(1-h^2)^{\frac{1}{2}}\right]+r(\beta,h) \quad \textit{if } h_c<h<1\,.$$

(ii) *The maximisers* $\bar\theta=\bar\theta(\beta,h)$ *and* $\bar v=\bar v(\beta,h)$ *are given by*

$$(1.21)\qquad \bar\theta=\theta(\beta,h) \quad \textit{if } h\neq h_c\,,$$

$$(1.22)\qquad \bar v(i,j)=\frac{1}{\lambda(0)}\tau_0(i)A_0(i,j)\tau_0(j) \quad \textit{if } 0\leqq h\leqq h_c\,,$$
$$=\frac{1}{\lambda(r)}\tau_r(i)A_r(i,j)\tau_r(j) \qquad \textit{if } h_c<h<1\,.$$

Theorem 2B. (iii) $0<\theta_c\leqq h_c<1$.

(iv) $h\to\varrho(\beta,h)$ *is continuous and strictly decreasing on* $[0,1]$, *and is analytic on* $(0,h_c)$ *and on* $(h_c,1)$.

(v) *At* $h=h_c$

$$(1.23)\qquad \frac{\partial\varrho}{\partial h}(\beta,h_c+)-\frac{\partial\varrho}{\partial h}(\beta,h_c-)=\frac{\partial r}{\partial h}(\beta,h_c+)=\frac{\theta_c}{1-h_c^2}\,.$$

(vi) *If* $\log M>0>\log\int b\beta(db)$ *then* $\varrho(\beta,h)$ *changes sign at* $h=h_c^*$ *the unique solution of* $\varrho(\beta,h)=0$ *computable from* (1.20).

(vii) *The maximiser* $\bar\theta(\beta,h)$ *satisfies*

$$(1.24)\qquad \theta_c<\bar\theta<h \quad \textit{if} \quad h_c<h<1$$

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

The proof of Theorem 2 is given in section 4 and is based on the functional analyti results of sections 2 and 3. The following two figures display qualitatively $\varrho(\beta,h)$ an $\bar\theta(\beta,h)$ as functions of h for fixed β satisfying $\log M>0>\log\int b\beta(db)$:

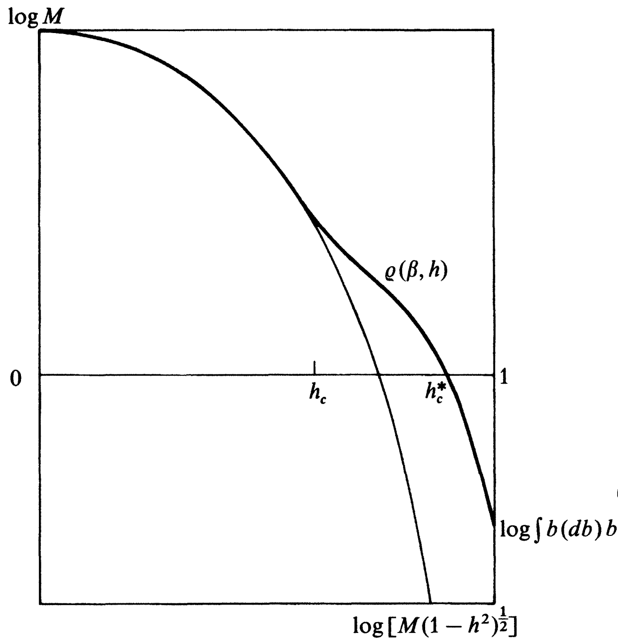

Figure 1

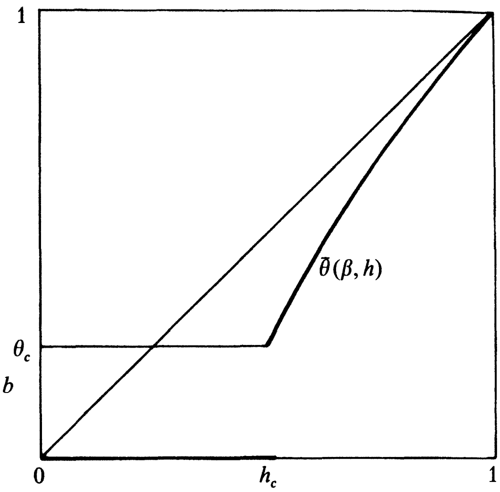

Figure 2

In section 5 we prove that $\lambda(r)$ satisfies the inequality $-\lambda'(r)/\lambda(r) > (1+\lambda^2(r))/(1-\lambda^2(r))$ for all $r > 0$. This is needed in section 4 to show that $h \to \varrho(\beta, h)$ is strictly decreasing and $\bar{\theta}(\beta, h) < h$ on $(h_c, 1)$. The proof uses tools like iterated maps, random continued fractions and Gibbs measures.

1.5. Reformulation of variational formula. We shall want to rewrite (1.11) in a slightly easier form in order to prepare for the variational analysis in section 2, namely

$$\varrho(\beta, h) = \sup_{\theta \in (0,1]} J_{\beta,h}(\theta) \tag{1.25}$$

with (recall (1.13–15))

$$J_{\beta,n}(\theta) = J_{\beta,h}(0) + \frac{1}{2}\theta \log\left(\frac{1+h}{1-h}\right) - \theta K(\theta)\,, \tag{1.26}$$

$$J_{\beta,h}(0) = \log[M(1-h^2)^{\frac{1}{2}}]\,, \tag{1.27}$$

$$K(\theta) = \inf_{\nu \in M_\theta} \sum_{i,j} \nu(i,j) \log\left(\frac{\nu(i,j)}{\hat{\nu}(i)\, A(i,j)}\right). \tag{1.28}$$

The above reformulation follows easily by substituting into (1.11) the identity

$$(1.29) \qquad I_0(\nu) - I_\theta(\nu) = \sum_{i,j} \nu(i,j) \log\left(P_\theta(i,j)/P(i,j)\right)$$

$$= \frac{1+\theta}{2\theta}\log(1+\theta) + \frac{1-\theta}{2\theta}\log(1-\theta)$$

$$= \frac{1}{2}\log\left(\frac{1+h}{1-h}\right) - \frac{1}{\theta}\left[I_h(0) - I_h(\theta)\right].$$

The second equality comes from the observation that $\nu \in M_\theta$ implies

$$\sum_{i,j} i\nu(i,j) = (1+\theta)/2\theta \quad \text{and} \quad \sum_{i,j} (j-1)\nu(i,j) = (1-\theta)/2\theta .$$

Finally, from (1.7) and (1.14) we have

$$(1.30) \qquad g(i) = -\log \int \left(\frac{b}{M}\right)^i \beta(db).$$

This implies the following properties for the function g (recall (1.2)):

(1.31)
(i) $g(0) = 0$,
(ii) g is strictly increasing,
(iii) $\lim_{i\to\infty} g(i)/i = 0$.

These properties will be essential and contain all about g that will be needed in sections 2–4. It is only in section 5 that we need the representation (1.30) in order to prove the differential inequality for $\lambda(r)$ mentioned below the figures.

2. Analysis of $K(\theta)$: variation over ν

2.1. A minimisation problem in ℓ^1. Throughout this section we assume that $\theta > 0$. In order to prepare for the analysis in sections 2.2 and 2.3 it will be convenient to reformulate (1.28) as a problem of minimisation on some appropriate compact convex set in $\ell^1(\mathbb{N}^2)$ (with ℓ^1 the space of absolutely summable sequences). This reformulation appears in Proposition 1 below.

Define first the symbols

$$(2.1) \qquad \lambda = \theta^{-1},$$

$$\mu(i,j) = (i+j-1)\nu(i,j)$$

and the sets

$$\Phi(\lambda) = \left\{ \mu \in \ell^1(\mathbb{N}^2) : \mu(i,j) \geqq 0, \sum_{i,j} \frac{\mu(i,j)}{i+j-1} = 1, \right. \tag{2.2}$$

$$\left. \sum_{i,j} \mu(i,j) \leqq \lambda, \sum_j \frac{\mu(i,j) - \mu(j,i)}{i+j-1} = 0 \right\},$$

$$\partial\Phi(\lambda) = \{\mu \in \Phi(\lambda) : \sum_{i,j} \mu(i,j) = \lambda\}. \tag{2.3}$$

Note that $\Phi(\lambda) \subset \bar{B}(\lambda)$, with $\bar{B}(\lambda)$ the closed ball in $\ell^1(\mathbb{N}^2)$ of radius λ, and

$$\partial\Phi(\lambda) = \Phi(\lambda) \cap \partial\bar{B}(\lambda),$$

with $\partial\bar{B}(\lambda)$ the surface of $\bar{B}(\lambda)$. Define

$$\phi(\mu) = \phi_1(\mu) + \phi_2(\mu) + \phi_3(\mu) + \phi_4(\mu), \tag{2.4}$$

$$\phi_1(\mu) = \sum_{i,j} \mu(i,j) \frac{g(i+j-1)}{i+j-1}, \tag{2.5}$$

$$\phi_2(\mu) = -\sum_{i,j} \mu(i,j) \frac{\log P(i,j)}{i+j-1}, \tag{2.6}$$

$$\phi_3(\mu) = \sum_{i,j} \frac{\mu(i,j)}{i+j-1} \log \frac{\mu(i,j)}{i+j-1}, \tag{2.7}$$

$$\phi_4(\mu) = -\sum_i \left(\sum_j \frac{\mu(i,j)}{i+j-1}\right) \log \left(\sum_j \frac{\mu(i,j)}{i+j-1}\right). \tag{2.8}$$

Then (1.28) reads, in the language of (2.1–8),

$$K(\lambda) = \inf_{\mu \in \partial\Phi(\lambda)} \phi(\mu). \tag{2.9}$$

Proposition 1 below says that the infimum may be extended from $\partial\Phi(\lambda)$ to $\Phi(\lambda)$ and that it is actually achieved on $\Phi(\lambda)$.

Proposition 1. *For* $\lambda \in [1, \infty)$

$$K(\lambda) = \min_{\mu \in \Phi(\lambda)} \phi(\mu). \tag{2.10}$$

The proof is done in several steps and is based on Lemma 1 and 2 below dealing with $\Phi(\lambda)$ resp. $\phi(\mu)$. We start by recalling some terminology. A sequence (μ_n) in $\ell^1(\mathbb{N}^2)$ is said to converge weakly * to some $\mu \in \ell^1(\mathbb{N}^2)$, written $\mu_n \xrightarrow{*} \mu$, if $\langle \mu_n, r \rangle \to \langle \mu, r \rangle$ for all $r \in c_0(\mathbb{N}^2)$ (with c_0 the space of sequences $(r(i,j))$ such that $r(i,j) \to 0$ as $i,j \to \infty$; ℓ^1 is the dual of c_0). Moreover, $\bar{B}(\lambda)$ is weak * compact, the weak * topology in $\bar{B}(\lambda)$ is metrisable, and on $\bar{B}(\lambda)$ weak * convergence and componentwise convergence are equivalent, i.e., $\mu_n \xrightarrow{*} \mu$ iff $\mu_n(i,j) \to \mu(i,j)$ for all i and j (see Rudin [17], 3.14–16).

Lemma 1. *For $\lambda \in [1, \infty)$: $\Phi(\lambda)$ is convex, is weak * compact and is the weak * closure of $\partial\Phi(\lambda)$.*

Proof. (1) Clearly $\Phi(\lambda)$ is convex.

(2) To see that $\Phi(\lambda)$ is weak * closed, let $\mu_n \in \Phi(\lambda)$ and $\mu \in \ell^1(\mathbb{N}^2)$ be such that $\mu_n \xrightarrow{*} \mu$, i.e., $\mu_n \to \mu$ componentwise. Then

$$\mu(i,j) \geqq 0 \quad \text{for all } i \text{ and } j,$$

$$\sum_{i,j} \mu(i,j) \leqq \liminf_{n\to\infty} \sum_{i,j} \mu_n(i,j) \leqq \lambda \quad \text{by Fatou's lemma},$$

$$\sum_{i,j} \frac{\mu(i,j)}{i+j-1} = \lim_{n\to\infty} \sum_{i,j} \frac{\mu_n(i,j)}{i+j-1} = 1 \quad \text{since} \left(\frac{1}{i+j-1}\right) \in c_0(\mathbb{N}^2).$$

Moreover, since $\left(\frac{\delta_{ki}}{i+j-1}\right) \in c_0(\mathbb{N}^2)$ for all k, it follows that

$$\sum_j \frac{\mu(k,j)-\mu(j,k)}{k+j-1} = \lim_{n\to\infty} \sum_{i,j} \frac{\delta_{ki}}{i+j-1}[\mu_n(i,j)-\mu_n(j,i)] = 0 \quad \text{for all } k.$$

Hence $\mu \in \Phi(\lambda)$. Since $\Phi(\lambda) \subset \bar{B}(\lambda)$, weak * closedness implies weak * compactness.

(3) To see that $\partial\Phi(\lambda)$ is weak * dense in $\Phi(\lambda)$ it suffices to find for every $\mu \in \Phi(\lambda)$ a sequence (μ_n) in $\partial\Phi(\lambda)$ such that $\mu_n \xrightarrow{*} \mu$. Indeed, pick any $\mu \in \Phi(\lambda)$ with

$$\sum_{i,j} \mu(i,j) = \tilde{\lambda} \leqq \lambda$$

and put

$$\mu_n(i,j) = t_n \mu(i,j) + s_n \delta_{ni} \delta_{nj}.$$

In order to have $\mu_n \in \partial\Phi(\lambda)$, t_n and s_n must satisfy

$$1 = t_n + s_n \frac{1}{2n-1},$$

$$\lambda = t_n \tilde{\lambda} + s_n.$$

This means

$$t_n = \frac{2n-1-\lambda}{2n-1-\tilde{\lambda}},$$

$$s_n = \frac{(2n-1)(\lambda-\tilde{\lambda})}{2n-1-\tilde{\lambda}}.$$

Since

$$t_n \to 1, \quad s_n \to \lambda - \tilde{\lambda}, \quad (\delta_{ni}\delta_{nj}) \xrightarrow{*} (0)$$

it follows that $\mu_n \xrightarrow{*} \mu$. □

Lemma 2. *For $\lambda \in [1, \infty)$: $\phi : \Phi(\lambda) \to \mathbb{R}$ is well-defined, nonnegative, weak * lower semicontinuous and convex.*

Proof. Recall (2.4–8).

(1) Since $g \geqq 0$ it is obvious that $\phi_1(\mu) \geqq 0$. In the notation with ν instead of μ (recall (2.1)) we have

$$\phi_2(\nu) + \phi_3(\nu) + \phi_4(\nu) = \sum_{i,j} \nu(i,j) \log\left(\frac{\nu(i,j)}{\hat{\nu}(i)\, P(i,j)}\right).$$

This is the relative entropy of $\nu(i,j)$ with respect to $\hat{\nu}(i)\, P(i,j)$, which are both probability measures on $\mathbb{N}^2$. Nonnegativity follows from Jensen's inequality via convexity of $x \to x \log x$.

(2) Property (1.31) (iii) implies that $\left(\frac{g(i+j-1)}{i+j-1}\right) \in c_0(\mathbb{N}^2)$, which makes ϕ_1 weak * continuous.

(3) Since $\left(\frac{-\log P(i,j)}{i+j-1}\right) \notin c_0(\mathbb{N}^2)$, ϕ_2 is not weak * continuous (recall (1.15)). However, if $\mu_n \to \mu$ componentwise, then $\phi_2(\mu) \leqq \liminf_{n \to \infty} \phi_2(\mu_n)$ by Fatou's lemma, which makes ϕ_2 weak * lower semicontinuous.

(4) Both ϕ_3 and ϕ_4 are entropy functions, and $0 \leqq \phi_4(\mu) \leqq -\phi_3(\mu)$ for all $\mu \in \Phi(\lambda)$ (the second inequality follows from Jensen's inequality via convexity of $x \to x \log x$). Split ϕ_3 into two parts

$$\phi_3(\mu) = \phi_3^1(\mu) + \phi_3^2(\mu)$$

with

$$\phi_3^1(\mu) = \sum_{i,j} \frac{1}{i+j-1}\, \mu(i,j) \log \mu(i,j),$$

$$\phi_3^2(\mu) = -\sum_{i,j} \frac{\log(i+j-1)}{i+j-1}\, \mu(i,j).$$

Next note that

$$\sum_{i,j} \left(\frac{1}{i+j-1}\right)^3 = \sum_k \frac{1}{k^2} = \frac{\pi^2}{6} < \infty,$$

$$\sum_{i,j} \mu^{\frac{3}{2}}(i,j) |\log \mu(i,j)|^{\frac{3}{2}} \leqq \lambda \max_{t \in [0,1]} t^{\frac{1}{2}} |\log t|^{\frac{3}{2}} = \lambda \left(\frac{3}{e}\right)^{\frac{3}{2}} < \infty.$$

Hence, by Hölder's inequality, ϕ_3^1 is bounded and therefore weak * continuous ($\ell^{\frac{3}{2}}$ is the dual of ℓ^3). Since $\left(\frac{\log(i+j-1)}{i+j-1}\right) \in c_0(\mathbb{N}^2)$, the same is true for ϕ_3^2. The weak * continuity of ϕ_4 follows by the same argument.

(5) Both $\phi_1(\mu)$ and $\phi_2(\mu)$ are linear in μ. We show that $\phi_3(\mu)+\phi_4(\mu)$ is convex in μ. In the notation with v and $\hat{v}$ (recall (2.1)) this sum equals

$$\phi_3(v)+\phi_4(v)=\sum_{i,j} v(i,j)\log\left(\frac{v(i,j)}{\hat{v}(i)}\right).$$

Convexity in v (and hence in μ) follows from Jensen's inequality via convexity of $x \to x\log x$. Note that there is no strict convexity because the sum is linear along lines where $v(i,j)/\hat{v}(i)$ is constant for all i and j: ϕ is positively homogeneous. □

Proof of *Proposition* 1. Since $\Phi(\lambda)$ is weak * compact (Lemma 1) and ϕ is weak * lower semicontinuous on $\Phi(\lambda)$ (Lemma 2), ϕ achieves its minimum on $\Phi(\lambda)$. It now suffices to find for every $\mu \in \Phi(\lambda)$ a sequence (μ_n) in $\partial\Phi(\lambda)$ such that $\mu_n \overset{*}{\to} \mu$ and $\phi(\mu_n) \to \phi(\mu)$. Indeed, in that case, if $\bar{\mu}$ is a minimiser of ϕ on $\Phi(\lambda)$ and $(\bar{\mu}_n)$ the corresponding approximating sequence in $\partial\Phi(\lambda)$, then

$$\begin{aligned}\inf_{\mu\in\partial\Phi(\lambda)} \phi(\mu) &\leqq \lim_{n\to\infty} \phi(\bar{\mu}_n)=\phi(\bar{\mu})\\ &= \min_{\mu\in\Phi(\lambda)} \phi(\mu) \leqq \inf_{\mu\in\partial\Phi(\lambda)} \Phi(\mu).\end{aligned}$$

To exhibit such a sequence (μ_n), pick the example in part (3) of the proof of Lemma 1. There we showed that $\mu_n \overset{*}{\to} \mu$. Since $\phi_k(\mu_n) \to \phi_k(\mu)$ for $k=1,3,4$ by weak * continuity, we need to worry about ϕ_2 only. But

$$\phi_2(\mu_n)=t_n\phi_2(\mu)-s_n\frac{\log P(n,n)}{2n-1}$$

and $P(n,n)\sim\left(\frac{1}{4\pi n}\right)^{\frac{1}{2}}$ by Stirling's formula. Hence also $\phi_2(\mu_n) \to \phi(\mu)$. □

2.2. Properties of $K(\lambda)$.

Proposition 2. *$\lambda \to K(\lambda)$ is non-increasing, convex and continuous on* $[1,\infty)$.

Proof. Let $1<\lambda_1<\lambda_2<\infty$. Then $\Phi(\lambda_1)\subseteqq\Phi(\lambda_2)$ and hence $K(\lambda_1)\geqq K(\lambda_2)$. Let $\mu_i\in\Phi(\lambda_i)$, $i=1,2$, and $0<t<1$. Then

$$\begin{gathered}(1-t)\mu_1+t\mu_2 \in \Phi\big((1-t)\lambda_1+t\lambda_2\big),\\ \phi\big((1-t)\mu_1+t\mu_2\big) \leqq (1-t)\phi(\mu_1)+t\phi(\mu_2)\end{gathered}$$

(ϕ is convex by Lemma 2). Pick $\mu_i = \bar{\mu}_i$ with $\bar{\mu}_i$ any minimiser, i.e., $K(\lambda_i) = \phi(\bar{\mu}_i)$ (see Proposition 1). It follows that

$$K\big((1-t)\lambda_1 + t\lambda_2\big) \leqq (1-t)K(\lambda_1) + tK(\lambda_2).$$

This says that $K(\lambda)$ is convex in λ and hence also continuous on $(1, \infty)$. One easily checks continuity at $\lambda = 1$, because this case is degenerate ($K(1) = g(1)$). □

Proposition 2 tells us that in principle there are two qualitatively different situations possible.

Case A. $K(\lambda)$ is strictly decreasing on $[1, \infty)$.

Case B. There exists $\lambda_c \in (1, \infty)$ such that $K(\lambda)$ is strictly decreasing on $[1, \lambda_c)$ and constant on $[\lambda_c, \infty)$.

We shall see in section 4 that actually only case B occurs. The transition at $\lambda = \lambda_c$ is connected with where the minimum is attained:

Proposition 3.

(2.11) *For* $\lambda \leqq \lambda_c$: $K(\lambda) = \min_{\mu \in \partial\Phi(\lambda)} \phi(\mu)$.

(2.12) *For* $\lambda > \lambda_c$: $K(\lambda) = \min_{\mu \in \partial\Phi(\lambda_c)} \phi(\mu)$.

Proof. To prove (2.11) let $\lambda \leqq \lambda_c$. Suppose that there is no $\bar{\mu} \in \partial\Phi(\lambda)$ such that $K(\lambda) = \phi(\bar{\mu})$. Since ϕ achieves its minimum on $\Phi(\lambda)$ (by Proposition 1), there must exist $\tilde{\lambda} < \lambda$ and $\tilde{\bar{\mu}} \in \partial\Phi(\tilde{\lambda})$ such that $K(\lambda) = \phi(\tilde{\bar{\mu}})$. However, $K(\tilde{\lambda}) = \inf_{\mu \in \partial\Phi(\tilde{\lambda})} \phi(\mu) \leqq \phi(\tilde{\bar{\mu}})$ and this contradicts $K(\tilde{\lambda}) > K(\lambda)$. To prove (2.12) let $\lambda > \lambda_c$. Then $K(\lambda) = K(\lambda_c)$ and $K(\lambda_c) = \min_{\mu \in \partial\Phi(\lambda_c)} \phi(\mu)$ by (2.11). □

Remark. Proposition 3 shows that for $\lambda \leqq \lambda_c$ the minimum is achieved on the boundary $\partial\Phi(\lambda)$, the set we started out with in our original variational formula (see (2.9)). If we would know that the minimiser of ϕ on $\Phi(\lambda)$ is *unique* for every $\lambda \in [1, \infty)$, then we could conclude from (2.12) that *for every* $\lambda > \lambda_c$ *this minimiser does not lie on the boundary* $\partial\Phi(\lambda)$. At this stage we cannot yet see uniqueness due to the fact that ϕ is convex but not strictly convex (see part (5) of the proof of Lemma 2). However, later we shall indeed establish uniqueness (Theorem 3 below), so that λ_c is indeed the value where the minimiser moves off the boundary into the interior.

2.3. Study of the minimiser (s). Next we study the minimiser(s). Lemmas 3–5 below list a few basic properties. Lemma 5 provides the connection with the eigenvalue problem as formulated later on in sections 2.4 and 2.5.

We return to the notation with θ, v instead of λ, μ (recall (2.1)), which was introduced only to make the link with ℓ^1 and c_0. Accordingly, we write $K(\theta)$, $\Phi(\theta)$, $\partial\Phi(\theta)$, $\phi(v)$ etc.

Lemma 3. *Any minimiser is symmetric, i.e., if $\bar{v} \in \Phi(\theta)$ and $K(\theta) = \phi(\bar{v})$ then $\bar{v}(i,j) = \bar{v}(j,i)$ for all i and j.*

Proof. Let $\bar{v}^s(i,j) = \bar{v}(j,i)$. Then $\bar{v}^s \in \Phi(\theta)$ and $\phi(\bar{v}^s) = \phi(\bar{v})$ by (1.15) and (2.1-8) (use that P is symmetric). Convexity of ϕ gives

$$\phi\left(\frac{1}{2}\bar{v} + \frac{1}{2}\bar{v}^s\right) \leqq \frac{1}{2}\phi(\bar{v}) + \frac{1}{2}\phi(\bar{v}^s) = \phi(\bar{v}) \leqq \phi\left(\frac{1}{2}\bar{v} + \frac{1}{2}\bar{v}^s\right)$$

hence

$$\phi\left(\frac{1}{2}\bar{v} + \frac{1}{2}\bar{v}^s\right) = \frac{1}{2}\phi(\bar{v}) + \frac{1}{2}\phi(\bar{v}^s).$$

As observed in part (5) of the proof of Lemma 2, this implies that

$$\frac{\bar{v}(i,j)}{\hat{\bar{v}}(i)} = \frac{\bar{v}^s(i,j)}{\hat{\bar{v}}^s(i)} \quad \text{for all } i \text{ and } j.$$

But $\hat{\bar{v}}(i) = \hat{\bar{v}}^s(i)$ because $\sum_j [v(i,j) - v(j,i)] = 0$ (recall (1.6)). □

In what follows we shall be able to get information about the minimiser(s) $\bar{v}$ by considering variations $\bar{v} + t\partial$, with $t > 0$ sufficiently small, of the form

$$\text{(2.13)} \qquad \partial(i,j) = \partial(j,i) \in \mathbb{R},$$

$$\partial(i,j) = 0 \quad \text{except at finitely many points},$$

$$\partial(i,j) \geqq 0 \quad \text{if } \bar{v}(i,j) = 0,$$

$$\text{(2.14)} \qquad \sum_{i,j} \partial(i,j) = 0,$$

$$\text{(2.15)} \qquad \sum_{i,j} (i+j)\,\partial(i,j) \leqq 0.$$

Indeed, (2.13–15) ensure that if $\bar{v} \in \Phi(\theta)$ then also $\bar{v} + t\partial \in \Phi(\theta)$. Variations ∂ with the latter property are called admissible. Note that if $\bar{v} \notin \partial\Phi(\theta)$ then we may even drop (2.15), because in that case still $\bar{v} + t\partial \in \Phi(\theta)$ for t sufficiently small under (2.14) alone. We shall need this observation later on.

The fact that $\bar{v}$ is a minimiser implies

$$\text{(2.16)} \qquad \phi(\bar{v}) \leqq \phi(\bar{v} + t\partial).$$

Together with the convexity of ϕ this implies

$$\text{(2.17)} \qquad \lim_{t \downarrow 0} \frac{1}{t}[\phi(\bar{v} + t\partial) - \phi(\bar{v})] \geqq 0.$$

Lemma 4. *For $\theta \in (0,1)$ any minimiser is strictly positive, i.e., if $\bar{\nu} \in \Phi(\theta)$ and $K(\theta) = \phi(\bar{\nu})$ then $\bar{\nu}(i,j) > 0$ for all i and j.*

Proof. Note that $\bar{\nu}(i,j) = \delta_{1i}\delta_{1j}$ at $\theta = 1$; in the lemma we excluded this degenerate case. Suppose that we have an admissible variation. Then from

$$\phi(\nu) = \sum_{i,j} \nu(i,j) \log\left(\frac{\nu(i,j)}{\hat{\nu}(i)\,A(i,j)}\right) \tag{2.18}$$

it is straightforward to deduce that for $t \to 0$

$$\begin{aligned}\frac{1}{t}[\phi(\bar{\nu} + t\partial) - \phi(\bar{\nu})] \sim &- \sum_{i,j} \partial(i,j) \log A(i,j) \\ &+ \sum_{i,j:\bar{\nu}(i,j)>0} \partial(i,j) \log\left(\frac{\bar{\nu}(i,j)}{\hat{\bar{\nu}}(i)}\right) \\ &+ \sum_{i,j:\hat{\bar{\nu}}(i)=0} \partial(i,j) \log\left(\frac{\partial(i,j)}{\hat{\partial}(i)}\right) \\ &+ \sum_{i,j:\bar{\nu}(i,j)=0,\hat{\bar{\nu}}(i)>0} \partial(i,j) \left[\log\left(\frac{t\partial(i,j)}{\hat{\bar{\nu}}(i)}\right) - 1\right].\end{aligned} \tag{2.19}$$

There is a contradiction with (2.17) if

$$\partial(i,j) > 0 \quad \text{for some } (i,j) \text{ with } \bar{\nu}(i,j) = 0, \hat{\bar{\nu}}(i) > 0\,. \tag{2.20}$$

We shall exclude $\bar{\nu}(i,j) = 0$ by finding admissible variations satisfying (2.20).

Since $\theta < 1$, there exists $(k,\ell) \neq (1,1)$, $k \leqq \ell$ such that $\bar{\nu}(k,\ell) = \bar{\nu}(\ell,k) > 0$. First we show that $\bar{\nu}(k,\ell-1) = \bar{\nu}(\ell-1,k) > 0$. Indeed, suppose the contrary. Then an admissible variation (for t sufficiently small) is

$$\begin{aligned}\partial(k,\ell) &= \partial(\ell,k) = -(1 + 1_{\{k=\ell\}})\,,\\ \partial(k,\ell-1) &= \partial(\ell-1,k) = 1 + 1_{\{k=\ell-1\}}\,,\\ \partial(i,j) &= 0 \quad \text{otherwise}\,.\end{aligned}$$

This satisfies (2.20) which provides the contradiction. By symmetry and backward induction it follows that $\bar{\nu}(i,j) = \bar{\nu}(j,i) > 0$ for all $i \leqq k$ and $j \leqq \ell$. Next we show that $\bar{\nu}(k,\ell+1) = \bar{\nu}(\ell+1,k) > 0$. Indeed, now the contrary is ruled out by the admissible variation

$$\begin{aligned}\partial(k,\ell+1) &= \partial(\ell+1,k) = 1 + 1_{\{k=\ell+1\}}\,,\\ \partial(k,\ell) &= \partial(\ell,k) = -2(1 + 1_{\{k=\ell\}})\,,\\ \partial(k,\ell-1) &= \partial(\ell-1,k) = 1 + 1_{\{k=\ell-1\}}\,,\\ \partial(i,j) &= 0 \quad \text{otherwise}\,.\end{aligned}$$

By symmetry and forward induction it follows that $\bar{v}(i,j) = \bar{v}(j,i) > 0$ for all $i \geqq k$ and $j \geqq \ell$. Combination of backward and forward induction proves the lemma. □

Lemma 5. *Fix $\theta \in (0,1)$. Let $\bar{v} \in \Phi(\theta)$ be a minimiser and define*

$$\bar{\xi}(i,j) = \log\left(\frac{\bar{v}(i,j)}{\hat{\bar{v}}(i)\, A(i,j)}\right), \tag{2.21}$$

$$\bar{\eta}(i,j) = \frac{1}{2}[\bar{\xi}(i,j) + \bar{\xi}(j,i)] .$$

Then there exists $r \geqq 0$ such that

$$\bar{\eta}(i,j) = \bar{\eta}(1,1) - r(i+j-2) \quad \textit{for all } i \textit{ and } j . \tag{2.22}$$

If $\bar{v} \notin \partial\Phi(\theta)$ then $r = 0$ and $\bar{\eta}$ is constant.

Proof. Consider variations of the form

$$\partial(i,j) = \partial(j,i),$$
$$\partial(i,j) \neq 0 \quad \text{if } (i,j) \text{ or } (j,i) \in \{(k,\ell), (m,n), (p,q)\},$$
$$\partial(i,j) = 0 \quad \text{otherwise}.$$

By Lemma 4 such variations are admissible (for t sufficiently small) provided (2.14) and (2.15) hold, i.e.,

$$\partial(k,\ell) + \partial(m,n) + \partial(p,q) = 0 , \tag{2.23}$$

$$(k+\ell)\,\partial(k,\ell) + (m+n)\,\partial(m,n) + (p+q)\,\partial(p,q) \leqq 0 . \tag{2.24}$$

Admissibility combined with (2.17) implies via (2.21) (recall (2.19))

$$\begin{aligned} 0 &\leqq \lim_{t \downarrow 0} \frac{1}{t}[\phi(\bar{v} + t\partial) - \phi(\bar{v})] \\ &= \sum_{i,j} \partial(i,j) \log\left(\frac{\bar{v}(i,j)}{\hat{\bar{v}}(i)\, A(i,j)}\right) \\ &= \sum_{i,j} \partial(i,j)\, \bar{\eta}(i,j) \\ &= \bar{\eta}(k,\ell)\,\partial(k,\ell) + \bar{\eta}(m,n)\,\partial(m,n) + \bar{\eta}(p,q)\,\partial(p,q) \end{aligned} \tag{2.25}$$

(for the second equality note that ∂ is symmetric). Now use (2.23) to eliminate $\partial(p,q)$ from (2.24) and (2.25) to obtain the following implication

$$\begin{aligned} &(k+\ell-p-q)\,\partial(k,\ell) + (m+n-p-q)\,\partial(m,n) \leqq 0 \\ \Rightarrow\ &(\bar{\eta}(k,\ell) - \bar{\eta}(p,q))\,\partial(k,\ell) + (\bar{\eta}(m,n) - \bar{\eta}(p,q))\,\partial(m,n) \geqq 0 . \end{aligned} \tag{2.26}$$

Here $\partial(k,\ell), \partial(m,n) \in \mathbb{R}$ are arbitrary. Via the elementary lemma

$$\left.\begin{array}{l} a,b,c,d \in \mathbb{R},\, a \neq 0,\, c \neq 0, \\ (ax+by \leqq 0 \Rightarrow cx+dy \geqq 0) \text{ for all } x,y \in \mathbb{R} \end{array}\right\} \Rightarrow \frac{c}{a} = \frac{d}{b} \leqq 0,$$

we obtain from (2.26)

$$\text{(2.27)} \qquad \frac{\bar{\eta}(k,\ell)-\bar{\eta}(p,q)}{k+\ell-p-q} = \frac{\bar{\eta}(m,n)-\bar{\eta}(p,q)}{m+n-p-q} = -r \quad \text{for some } r \geqq 0 .$$

Pick $(p,q)=(1,1)$ to read off (2.22). We can go further if $\bar{v} \notin \partial\Phi(\theta)$ since then the restriction (2.24) drops out (recall the remark below (2.15)). The remaining restriction (2.23) combined with (2.25) gives that for all $\partial(k,\ell), \partial(m,n) \in \mathbb{R}$

$$\big(\bar{\eta}(k,\ell)-\bar{\eta}(p,q)\big)\,\partial(k,\ell) + \big(\bar{\eta}(m,n)-\bar{\eta}(p,q)\big)\,\partial(m,n) \geqq 0 .$$

Hence $\bar{\eta}$ must be constant. □

2.4. The minimiser(s) v as solution of an eigenvalue problem. We are now in a position to exhibit $\bar{v}$ as solution of an eigenvalue problem.

Proposition 4. *Fix $\theta \in (0,1)$. Let $\bar{v} \in \Phi(\theta)$ be a minimiser and define*

$$\text{(2.28)} \qquad A_r(i,j) = e^{-r(i+j-1)} A(i,j) \quad (r \geqq 0) .$$

Then

$$\text{(2.29)} \qquad \bar{v}(i,j) = R_r [\hat{\bar{v}}(i)]^{\frac{1}{2}} A_r(i,j) [\hat{\bar{v}}(j)]^{\frac{1}{2}} \quad (i,j \geqq 1),$$

$$\text{(2.30)} \qquad [\hat{\bar{v}}(i)]^{\frac{1}{2}} = R_r \sum_j A_r(i,j) [\hat{\bar{v}}(j)]^{\frac{1}{2}} \quad (i \geqq 1)$$

where

$$\text{(2.31)} \qquad R_r = \frac{\bar{v}(1,1)}{\hat{\bar{v}}(1)\, A_r(1,1)} .$$

I. *If $\bar{v} \in \partial\Phi(\theta)$ then r is given by the constraint*

$$\text{(2.32)} \qquad \theta^{-1} = \sum_{i,j} (i+j-1)\, \bar{v}(i,j) .$$

II. *If $\bar{v} \notin \partial\Phi(\theta)$ then $r=0$.*

Proof. Combine (2.21) and (2.22). This gives via (1.16)

$$\bar{v}(i,j)\,\bar{v}(j,i) = R_r^2\, \hat{\bar{v}}(i)\, \hat{\bar{v}}(j)\, A_r(i,j)\, A_r(j,i)$$

with R_r given by (2.31). Next use that $\bar{v}$ and A_r are symmetric (recall Lemma 3 and (1.16)) to read off (2.29). Summing (2.29) on j we get (2.30) (recall Lemma 4). Equation (2.32) is obvious (recall (2.3)), while the last statement in the proposition is the same as that in Lemma 5. □

2.5. Identification with largest eigenvalue and corresponding eigenvector. Proposition 4 is an eigenvalue problem for the matrix A_r but it does not tell us what eigenvalue and eigenvector $(R_r^{-1}, \hat{\bar{v}}^{\frac{1}{2}})$ is. By doing a spectral analysis of the 1-parameter family of symmetric matrices $A_r (r \geqq 0)$ we shall be able to remove this obstacle and establish the following.

Proposition 5. *For every $r \geqq 0$ there exists a unique pair $(\lambda(r), \tau_r) \in \mathbb{R} \times \ell^2(\mathbb{N})$ solving*

$$A_r \tau_r = \lambda(r) \tau_r \tag{2.33}$$

under the restrictions $\tau_r(i) \geqq 0$ and $\sum_i \tau_r^2(i) = 1$; $\lambda(r)$ is the largest eigenvalue of A_r and is simple.

Moreover, for $r \geqq 0$ the following hold:

(i) $0 < \lambda(r) < 1$,

(ii) $r \to \lambda(r)$ *is continuous and strictly decreasing*,

(iii) $r \to \lambda(r)$ *is analytic*,

(iv) $r \to \log \lambda(r)$ *is strictly convex*,

(v) $0 < \tau_r < 1$ *for all* i,

(vi) $r \to \tau_r \in \ell^2(\mathbb{N})$ *is analytic*.

The proof of Proposition 5 is deferred to section 3. Let us first see what it does for our variational problem.

Theorem 3. *For every $\theta \in (0, 1)$ there is a unique minimiser $\bar{v} \in \Phi(\theta)$ given by*

$$\bar{v}(i,j) = \frac{1}{\lambda(r)} \tau_r(i) A_r(i,j) \tau_r(j) \tag{2.34}$$

where $\lambda(r)$ and τ_r are defined in Proposition 5.

Define θ_c as in (1.17). *Then*

$$\theta_c^{-1} = \lim_{r \downarrow 0} -\frac{d}{dr} \log \lambda(r) = -\frac{\lambda'(0)}{\lambda(0)}. \tag{2.35}$$

I. *If $\theta \geqq \theta_c$ then $\bar{v} \in \partial\Phi(\theta)$ and $r = r(\theta)$ is the unique solution of*

$$\theta^{-1} = -\frac{d}{dr} \log \lambda(r) = -\frac{\lambda'(r)}{\lambda(r)}. \tag{2.36}$$

Moreover, $\theta \to r(\theta)$ and $\theta \to K(\theta)$ are strictly increasing on $(\theta_c, 1)$, and

(2.37) $$K(\theta) = \phi(\bar{v}) = -\frac{r(\theta)}{\theta} - \log \lambda(r).$$

II. *If $\theta < \theta_c$ then $\bar{v} \notin \partial\Phi(\theta)$ and $r = r(\theta) = 0$. Moreover, $\bar{v}(\theta) = \bar{v}(\theta_c) \in \partial\Phi(\theta_c)$ and*

(2.38) $$K(\theta) = K(\theta_c) = -\log \lambda(0).$$

Finally, $\theta \to \theta K(\theta)$ is convex on $(0,1)$, with $\lim_{\theta \downarrow 0} \theta K(\theta) = 0$ and $\lim_{\theta \downarrow 1} \theta K(\theta) = K(1)$.

Proof. Return to Proposition 4. Equation (2.30) reads

$$A_r \hat{v}^{\frac{1}{2}} = R_r^{-1} \hat{v}^{\frac{1}{2}}.$$

From Proposition 5 (since $\hat{v}(i) \geqq 0$ and $\sum_i \hat{v}(i) = 1$) it follows that $R_r^{-1} = \lambda(r)$ and $\hat{v}^{\frac{1}{2}} = \tau_r$. Now (2.29) becomes (2.34).

The relation between r and θ is obtained as follows. Recall the distinction between cases A and B in section 2.2. These correspond to $\theta_c = 0$ and $\theta_c > 0$, respectively.

I. If $\theta \geqq \theta_c$ then $\bar{v} \in \partial\Phi(\theta)$. Hence by Proposition 4.I and (2.34)

$$\begin{aligned}
\theta^{-1} &= \sum_{i,j} (i+j-1)\, \bar{v}(i,j) \\
&= \frac{1}{\lambda(r)} \sum_{i,j} \tau_r(i)(i+j-1)\, A_r(i,j)\, \tau_r(j) \\
&= \frac{1}{\lambda(r)} \left\langle \tau_r, \left(-\frac{\partial A_r}{\partial r}\right) \tau_r \right\rangle.
\end{aligned}$$

Now, since $\lambda(r) = \langle \tau_r, A_r \tau_r \rangle$,

$$\begin{aligned}
\lambda'(r) &= \frac{d}{dr} \langle \tau_r, A_r \tau_r \rangle \\
&= \left\langle \tau_r, \frac{\partial A_r}{\partial r} \tau_r \right\rangle + \left\langle \frac{\partial \tau_r}{\partial \tau}, A_r \tau_r \right\rangle + \left\langle \tau_r, A_r \frac{\partial \tau_r}{\partial r} \right\rangle.
\end{aligned}$$

By the symmetry of A_r the last two terms are identical and equal to

$$\lambda(r) \left\langle \frac{\partial \tau_r}{\partial r}, \tau_r \right\rangle = \frac{1}{2} \lambda(r) \frac{d}{dr} \langle \tau_r, \tau_r \rangle = 0.$$

This proves (2.36).

The fact that (2.36) has a unique and strictly decreasing solution is a consequence of Propositions (5)(iii), (iv). Indeed, $-\frac{d}{dr}\log\lambda(r)$ is strictly decreasing, since $-\frac{d^2}{dr^2}\log\lambda(r) \leqq 0$ with equality at most in isolated points. Equation (2.37) follows by substitution of (2.34) into (2.18). Use Lemma 3, (1.16) and (2.32).

II. If $\theta < \theta_c$ then $\bar{v} \notin \partial\Phi(\theta)$ by the remark below (2.12). Hence $r = 0$ by Proposition 4.II, and $\bar{v} = \bar{v}(r = 0) = \bar{v}(\theta = \theta_c)$ with eigenvalue $\lambda(0)$. Equation (2.38) is just (2.37) at $r = 0$.

To get convexity of $\theta K(\theta)$ compute for $\theta \geqq \theta_c$ using (2.36) and (2.37)

$$\frac{d}{d\theta}(\theta K(\theta)) = -\log\lambda(r),$$

$$\frac{d^2}{d\theta^2}(\theta K(\theta)) = \frac{1}{\theta}\frac{dr(\theta)}{d\theta} > 0. \quad \square$$

3. Proof of Proposition 5: spectral analysis of A_r

Having thus reduced (1.20), via Theorem 3, to the eigenvalue problem of Proposition 5, we are now ready to give the proof of the latter. There are several lemmas on the way. We start by collecting properties of P, the matrix in (1.15) which appears in the definition of A_r in (1.13).

3.1. Properties of P. The matrix P is symmetric, strictly positive and stochastic, i.e.,

(3.1)
$$\text{(i)} \quad P(i,j) = P(j,i),$$
$$\text{(ii)} \quad P(i,j) > 0,$$
$$\text{(iii)} \quad \sum_j P(i,j) = 1.$$

The following identity will be important (in particular in section 5):

$$\sum_j P(i,j)\,a^{j-1} = \frac{1}{(2-a)^i} \quad (a \in [0,1]). \tag{3.2}$$

The next lemma summarizes the properties of P needed for the proof of Proposition 5.

Lemma 6. (i) *P is recurrent, i.e.* $(\sum_{n\geqq 0} P^n)(i,j) = \infty$ *for all* i, j.

(ii) *If* $u \in \ell^\infty$, $w \geqq 0$ *and* $(I - P)u = w$, *then* $w = 0$ *and* $u = c1\,(c \in \mathbb{R})$.

(I = identity matrix, 1 = vector with all components equal to 1.)

(iii) $P: \ell^2 \to \ell^2$ *is a bounded linear operator with*

$$\|P\|_2 = r(P) = r_{\text{ess}}(P) = 1\,, \tag{3.3}$$

where $\|\cdot\|_2$ *is the operator norm,* $r(\cdot)$ *is the spectral radius and* $r_{\text{ess}}(\cdot)$ *is the essential spectral radius.*

(iv) *P is not compact in* ℓ^2.

(For the definition of essential spectrum, see Kato [12], X. 1.11.)

Proof. (1) To prove part (i), define the sequence of vectors $(x_k)_{k \geqq 0}$ by

$$x_k(i) = \frac{1}{k+1}\left(\frac{k}{k+1}\right)^{i-1} \quad (i \geqq 1, k \geqq 0) \tag{3.4}$$

and use (3.2) to see that

$$x_{k+1} = Px_k \quad (k \geqq 0)\,. \tag{3.5}$$

Hence

$$\Big(\sum_{n \geqq 0} P^n x_0\Big)(i) = \sum_{n \geqq 0} \frac{1}{n+1}\left(\frac{n}{n+1}\right)^{i-1} = \infty \quad \text{for every } i = 1\,.$$

Since $x_0 = e_1 = (1, 0, 0, \ldots)$ this says that

$$\Big(\sum_{n \geqq 0} P^n\Big)(i,j) = \infty \quad \text{for every } i \geqq 1 \text{ when } j = 1\,.$$

But then the same holds for all i, j by (3.1)(ii).

Remark. P is in fact the transition matrix of a non-degenerate critical branching process with one immigrant (Greven and den Hollander [10]) and therefore is null-recurrent (Athreya and Ney [1], VI. 7).

(2) To prove part (ii) use the Poisson equation $(I-P)_u = w$ to write for $N \geqq 0$

$$\sum_{n=0}^{N} P^n w = u - P^{N+1} u\,.$$

Let $N \to \infty$ and use $\|u - P^{N+1}u\|_\infty \leqq 2\|u\|_\infty$ together with the recurrence of P, to get $w = 0$ and hence $Pu = u$. Since P is irreducible and recurrent it has no non-constant bounded harmonic functions (Neveu [15], 6.1) and hence $u = c1$ $(c \in \mathbb{R})$. Incidentally, the latter statement is non-trivial, but under the additional restriction that $u(i_0) = \|u\|_\infty$ for some i_0 it has an easy proof, namely

$$\|u\|_\infty = u(i_0) = (Pu)(i_0) \leqq \|u\|_\infty$$

by (3.1)(iii), with equality iff u is constant.

(3) To prove $\|P\|_2 \leqq 1$, note that by (3.1)(i), (iii)

$$P(\ell^1) \subseteqq \ell^1, \quad \|P\|_1 = 1 ,$$

$$P(\ell^\infty) \subseteqq \ell^\infty, \quad \|P\|_\infty = 1 .$$

Use the Riesz-Thorin interpolation theorem (Dunford-Schwartz [8], VI.10.11) to get

$$\|P\|_2 \leqq \sqrt{\|P\|_1 \|P\|_\infty} = 1 .$$

(4) To prove $r(P) = \|P\|_2 = 1$, note that because

$$r(P) = \sup_{x \in \ell^2} \frac{\langle Px, x\rangle}{\langle x, x\rangle} \leqq \|P\|_2 \leqq 1$$

it suffices to exhibit a sequence (x_k) in ℓ^2 such that

$$\frac{\langle Px_k, x_k\rangle}{\langle x_k, x_k\rangle} \to 1 \quad (k \to \infty) . \tag{3.6}$$

For this pick the sequence x_k defined in (3.4). The l.h.s. of (3.6) equals $\frac{2k+1}{2k+2}$.

(5) To prove part (iv) argue by contradiction. For a positive compact operator P, $r(P) \in \sigma(P)$ and $p_\sigma(P) \supseteqq \sigma(P) \setminus \{0\}$, with $\sigma(\cdot)$ the spectrum and $p_\sigma(\cdot)$ the point spectrum (see Zaanen [19], 12.4). Hence $r(P) = 1$ implies $1 \in p_\sigma(P)$, i.e., $Pu = u$ has a solution in ℓ^2. But this contradicts part (ii).

(6) To prove $r_{\text{ess}}(P) = 1$, combine $r_{\text{ess}}(P) \leqq r(P) = 1$ with $1 \notin p_\sigma(P)$. Together with (3) and (4) this gives part (iii). □

3.2. Distinction between compact and non-compact A_r. For the spectral analysis we shall need

Lemma 7. *$A_r : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ is a compact operator if and only if*

$$r > 0, \text{ or } r = 0 \text{ and } g(\infty) = \infty . \tag{C}$$

In the sequel we shall often refer to condition (C).

Proof. For $r > 0$ estimate the Hilbert-Schmidt norm of A_r (see Dunford and Schwartz [8], Part II, XI.6):

$$\begin{aligned}\|A_r\|_{HS}^2 &= \sum_{i,j} A_r^2(i,j) \\ &= \sum_{i,j} e^{-2[r(i+j-1)+g(i+j-1)]} P^2(i,j) \\ &\leqq \sum_{i,j} e^{-2r(i+j-1)} \\ &= e^{-2r}(1-e^{-2r})^{-2} < \infty .\end{aligned}$$

The inequality uses $g(i) \geqq 0$ and $P(i,j) \leqq 1$. Hence A_r is Hilbert-Schmidt and therefore compact.

Next consider the case $r = 0$. Use (1.13) to split (recall (2.28))

$$A_0 = A = e^{-g(\infty)} P + K \tag{3.7}$$

with

$$\begin{aligned} K(i,j) &= k(i+j-1)\, P(i,j)\,, \\ k(i) &= e^{-g(i)} - e^{-g(\infty)}\,. \end{aligned} \tag{3.8}$$

Observe that k is positive and strictly decreasing to zero (recall (1.31)(ii)). We shall prove that $K: \ell^2 \to \ell^2$ is compact. This will prove the lemma because P is not compact by Lemma 6 (iv).

First note that $K1 \in c_0(\mathbb{N})$, namely

$$(K1)(i) = \sum_j k(i+j-1)\, P(i,j) \leqq \sum_j k(i)\, P(i,j) = k(i)\,.$$

Let $\bar{B}_{c_0}$ be the closed unit ball in $c_0(\mathbb{N})$. Since

$$K(\bar{B}_{c_0}) \subseteqq [-K1, K1]\,,$$

K is compact in c_0. Since K is symmetric and ℓ^1 is the dual of c_0, K is also compact in ℓ^1 (Rudin [17], 4.19). From the interpolation theorem (Triebel [18], p. 117) it now follows that K is compact in ℓ^2 (since $\ell^1 \subseteqq \ell^2 \subseteqq c_0$). □

3.3. Spectral analysis of A_r (compact case). There are several lemmas on the way.

Lemma 8 (Standard). *Assume* (C). *Then*

$$\sup_{\|x\|=1} \langle x, A_r x\rangle = \max_{\|x\|=1, x \geqq 0} \langle x, A_r x\rangle = \max_{\lambda \in \sigma(A_r)} |\lambda| \in p_\sigma(A_r) \tag{3.9}$$

where $\|\cdot\|$ *is the* ℓ^2*-norm,* $\sigma(\cdot)$ *is the spectrum, and* $p_\sigma(\cdot)$ *is the point spectrum.*

Proof. By the strict positivity of A_r

$$\sup_{\|x\|=1} \langle x, A_r x\rangle = \sup_{\|x\|=1, x \geqq 0} \langle x, A_r x\rangle > 0\,.$$

Since A_r is compact, the supremum is a maximum and is an element of $p_\sigma(A_r)$. Indeed, the supremum is an element of $\sigma(A_r)$ and by compactness $\sigma(A_r) \setminus \{0\} \subseteqq p_\sigma(A_r)$ (see Zaanen [19], 12.4). Finally, $\mu \in p_\sigma(A_r)$ means that there exists some $y \in \ell^2$ with $\|y\| = 1$ such that $A_r y = \mu y$, and therefore

$$|\mu| = |\langle y, A_r y\rangle| \leqq \langle |y|, A_r |y|\rangle \leqq \max_{\|x\|=1, x \geqq 0} \langle x, A_r x\rangle\,,$$

so that the maximum is the spectral radius of A_r. □

From now on we denote by $\lambda(r)$ the largest eigenvalue (is spectral radius) of A_r.

Lemma 9 (Standard). *Assume* (C). *Then the algebraic multiplicity of the eigenvalue* $\lambda(r)$ *is* 1, *and there is a unique corresponding eigenvector* τ_r *under the restrictions that* $\tau_r \geqq 0$ *and* $\|\tau_r\| = 1$. *Moreover* $\tau_r > 0$.

Proof. From the symmetry of A_r we know that the algebraic and the geometric multiplicity of $\lambda(r)$ are equal (see Zaanen [19], 11.3). If $x \in \ell^2$ with $\|x\| = 1$ is an eigenvector associated with $\lambda(r)$, then so is $|x|$ because

$$\lambda(r) = \langle x, A_r x\rangle \leqq \langle |x|, A_r |x|\rangle \leqq \lambda(r).$$

Moreover, by the strict positivity of A_r,

$$A_r |x| = \lambda(r)|x| \Rightarrow |x| > 0,$$

so that there is at least one strictly positive eigenvector τ_r associated with $\lambda(r)$. Now let r be any eigenvector for $\lambda(r)$ such that $\|\tau\| = 1$ and $\tau \geqq 0$. We show that $\tau = \tau_r$. Indeed, pick i_0 such that $\tau(i_0) > 0$ and put $t = \tau_r(i_0)/\tau(i_0) > 0$. Clearly $\tau_r - t\tau$ is either zero or is an eigenvector for $\lambda(r)$. In the latter case also $|\tau_r - t\tau|$ is an eigenvector and hence $|\tau_r - t\tau| > 0$ by the above observation. But $\tau_r(i_0) - t\tau(i_0) = 0$, so that we have a contradiction. Thus $\tau_r = t\tau$. Hence $t = \|\tau_r\|/\|\tau\| = 1$ and therefore $\tau_r = \tau$. □

Lemma 10 (Standard). *Assume* (C). *Let* (μ, y) *be any pair of eigenvalue and eigenvector with* $\|y\| = 1$ *and* $y \geqq 0$. *Then* $\mu = \lambda(r)$ *and* $y = \tau_r$.

Proof. Simply note that

$$\mu\langle y, \tau_r\rangle = \langle A_r y, \tau_r\rangle = \langle y, A_r \tau_r\rangle = \lambda(r)\langle y, \tau_r\rangle.$$

Since $\langle y, \tau_r\rangle > 0$, we get $\mu = \lambda(r)$ and hence $y = \tau_r$ by Lemma 9. □

Lemma 11. *Assume* (C). *Then the following functions are analytic on* $[0, \infty)$:

$$\begin{aligned} &r \to A_r \in B(\ell^2(\mathbb{N})), \\ &r \to \lambda(r), \\ &r \to \tau_r \in \ell^2(\mathbb{N}). \end{aligned} \tag{3.10}$$

Here $B(\ell^2)$ *is the space of bounded linear operators from* ℓ^2 *into* ℓ^2.

Proof. (1) For $x, y \in \ell^2$ define x_n and y_n by

$$x_n(i) = x(i)1_{\{i \leqq n\}},$$
$$y_n(i) = y(i)1_{\{i \leqq n\}}.$$

Then $x_n \to x$ and $y_n \to y$ in ℓ^2 as $n \to \infty$. Since $r \to \langle y_n, A_r x_n \rangle$ is analytic on $\{r \in \mathbb{C}: \operatorname{Re} r > 0\}$ for every n and since

$$|\langle y_n, A_r x_n \rangle - \langle y, A_r x \rangle| \leqq \|y_n\| \cdot \|A_r\| \cdot \|x_n - x\| + \|y_n - y\| \cdot \|A_r\| \cdot \|x\|,$$

it follows from the Weierstrass theorem for normal families of analytic functions (Behnke and Sommer [4], II. 7.42) that $r \to \langle y, A_r x \rangle$ is analytic (note that $\|A_r\| \leqq e^{-r}\|P\| = e^{-r}$ by (3.3)), i.e. $r \to A_r x$ is weakly analytic. This implies that $r \to A_r x$ is strongly analytic (see Rudin [16], 10.28). Now use the Banach-Steinhaus theorem (see Rudin [16], 2.5) to get that $r \to A_r$ is analytic in $B(\ell^2)$.

(2) First we prove continuity of $r \to \lambda(r)$. Pick any r and r' and note that by Lemma 8

$$\langle \tau_{r'}, (A_r - A_{r'})\tau_{r'} \rangle \leqq \lambda(r) - \lambda(r') \leqq \langle \tau_r, (A_r - A_{r'})\tau_r \rangle .$$

Hence

$$|\lambda(r) - \lambda(r')| \leqq \|A_r - A_{r'}\| .$$

Now let $r' \to r$ and use part (1). From the continuity of $r \to \lambda(r)$ and the analyticity of $r \to A_r$ we obtain the analyticity of $r \to \lambda(r)$ and $r \to \tau_r$ by applying Lemma 1.3 of Crandall and Rabinowitz [5]. The latter is a perturbation theorem for algebraically simple eigenvalues. □

Lemma 12. *Assume* (C). *On* $[0, \infty)$: $r \to \lambda(r)$ *is strictly decreasing*, $r \to \log \lambda(r)$ *is strictly convex, and* $0 < \lambda(r) < 1$.

Proof. (1) For every $r > r'$

$$\lambda(r) - \lambda(r') \leqq \langle \tau_r, (A_r - A_{r'})\tau_r \rangle < 0$$

because $A_r(i,j) < A_{r'}(i,j)$ and $\tau_r(i,j) > 0$ for all i and j.

(2) For every $x \geqq 0$, $\langle x, A_r x \rangle$ is log convex because $A_r(i,j)$ is log linear for every i and j and because log convexity is preserved under taking positive combinations. It follows that (recall Lemma 8)

$$\lambda(r) = \sup_{\|x\| = 1, x \geqq 0} \langle x, A_r x \rangle$$

is log convex because log convexity is preserved under taking pointwise limits and suprema (see Kingman [14] and Kato [13]).

(3) Suppose that $\log \lambda(r)$ is not strictly log convex. Then by analyticity there exist α and β such that $\lambda(r) = \alpha e^{\beta r}$. Now estimate

$$\begin{aligned}\lambda(r) &= \sup_{\|x\|=1, x \geqq 0} \langle x, A_r x\rangle \\ &\leqq \sum_{i,j} A_r(i,j) \\ &\leqq A_r(1,1) + \sum_{i,j: i+j \geqq 3} e^{-r(i+j-1)} P(i,j) \\ &= e^{-r} A(1,1) + \sum_{k \geqq 1} e^{-r(k+1)} \left(\frac{1}{2}\right)^{k+1} \sum_{\ell=0}^{k} \binom{k}{\ell} \\ &= e^{-r} A(1,1) + \frac{1}{2} e^{-2r}(1-e^{-r})^{-1}.\end{aligned}$$

The second inequality uses $g \geqq 0$. On the other hand, put $x = e_1 = (1,0,0,\ldots)$ to get

$$\lambda(r) \geqq \langle e_1, A_r e_1\rangle = A_r(1,1) = e^{-r} A(1,1).$$

Combining upper and lower bound we get

$$0 < A(1,1) \leqq \alpha e^{(\beta+1)r} \leqq A(1,1) + \frac{1}{2} e^{-r}(1-e^{-r})^{-1}.$$

Pass to the limit $r \to \infty$ to see that $\alpha = A(1,1)$ and $\beta = -1$. Thus

$$\lambda(r) = e^{-r} A(1,1) = A_r(1,1).$$

But this contradicts $(A_r \tau_r)(1) = \lambda(r) \tau_r(1)$ because $\sum_{j \geqq 2} A_r(1,j) \tau_r(j) > 0$.

(4) Trivially, $\lambda(r) > 0$ for all $r \geqq 0$ by the strict positivity of A_r. From (1.31)(ii) follows that for all $r \geqq 0$

$$A_r(i,j) \leqq e^{-g(1)} P(i,j),$$

hence $\lambda(r) \leqq e^{-g(1)} < 1$ by (3.3) in Lemma 6 (iii). □

Lemma 13. *For every $r > 0$: $\tau_r(i)$ is non-increasing in i.*

Proof. Let $(e_i)_{i \geqq 1}$ be the canonical base of ℓ^2 and let $f_j = \sum_{i=1}^{j} e_i$. Define

$$K = \{x \in \ell^2 : x \geqq 0, x(i+1) \leqq x(i) \text{ for } i \geqq 1\},$$

$$K_0 = \{x = \sum_j c_j f_j : c_j \geqq 0, c_j \neq 0 \text{ finitely often}\}.$$

K is a closed convex cone and $K = \bar{K}_0$ (the weak closure of K_0). Since A_r is a continuous operator on ℓ^2 for all $r > 0$, it follows that

$$A_r K \subset K \quad \textit{iff} \quad A_r f_j \in K \quad \textit{for all } j \geqq 1. \tag{3.11}$$

Since A_r is symmetric and has a spectral gap we know that $\lim_{n\to\infty} \lambda^{-n}(r) A_r^n x = \tau_r$ for all $x \in \ell^2$, $x \geqq 0$, $x \neq 0$. Hence it will follow from (3.11) that $\tau_r \in K$ once we show that $A_r f_j \in K$ for all $j \geqq 1$.

We next show that the latter indeed is true as a consequence of the following convexity property of $A_r(i,j)$:

$$(3.12) \qquad A_r(i,j+1) + A_r(i+1,j) - 2A_r(i+1,j+1) \geqq 0 \quad \text{for all } i,j \geqq 0$$

(with the convention $A_r(i,0) = A_r(0,j) = 0$). The inequality in (3.12) is easily verified by recalling (1.13–16) and (1.31)(ii) and noting that $P(i,j)$ satisfies the same equation but with equality. Now write out

$$(A_r f_j)(i) - (A_r f_j)(i+1) = \sum_{k=1}^{j} A_r(i,k) - \sum_{k=1}^{j} A_r(i+1,k)$$

$$= \left\{ \sum_{k=1}^{j} [A_r(i,k) - A_r(i+1,k-1)] - 2A_r(i+1,j) \right\} + A_r(i+1,j).$$

Call the term between braces $a_i(j)$. Then

$$a_i(j+1) - a_i(j) = A_r(i,j+1) + A_r(i+1,j) - 2A_r(i+1,j+1) \geqq 0$$

by (3.12), and so $a_i(j)$ is non-decreasing in j for all i. Hence

$$(A_r f_j)(i) - (A_r f_j)(i+1) \geqq a_i(1)$$
$$= A_r(i,1) - A_r(i+1,0) - 2A_r(i+1,1).$$

The r.h.s. is $\geqq 0$ by (3.12) because $A_r(i+1,0) = 0$. This shows that $A_r f_j \in K$, as was needed to complete the proof. □

3.4. Spectral analysis of A_0 (non-compact case). Define (recall that $\lambda(r)$ is decreasing by Lemma 12)

$$(3.13) \qquad \lambda(0) = \lim_{r\downarrow 0} \lambda(r).$$

The purpose of this section is to show that even in the non-compact case $\lambda(0)$ is an eigenvalue of A_0 in ℓ^2.

Lemma 14. *$A_0 x = \lambda(0) x$ has a solution in ℓ^∞ with $x \geqq 0$, $x \neq 0$.*

Proof. To construct a solution, define for every $r > 0$

$$\mu_r(i) = \frac{\tau_r(i)}{\tau_r(1)}.$$

By Lemma 13, $\mu_r \in \bar{B}_{\ell^\infty}$ the closed unit ball in $\ell^\infty(\mathbb{N})$. Since $\bar{B}_{\ell^\infty}$ is weak * compact, there exists $\mu_0 \in \ell^\infty$ such that $\mu_n \xrightarrow{*} \mu_0$ as $r \to 0$ along a subsequence. Note that $\mu_0 \geqq 0$, $\mu_0(1) = 1$. We show that $A_0 \mu_0 = \lambda(0)\mu_0$.

Start from $A_r \mu_r = \lambda(r)\mu_r$ for $r > 0$. Consider

$$\langle A_r \mu_r, x\rangle = \lambda(r)\langle \mu_r, x\rangle \quad (x \in \ell^1) \tag{3.14}$$

(ℓ^∞ is the dual of ℓ^1). First note that

$$\langle \mu_r, x\rangle \to \langle \mu_0, x\rangle,$$

$$\langle A_0 \mu_r, x\rangle = \langle \mu_r, A_0 x\rangle \to \langle \mu_0, A_0 x\rangle = \langle A_0 \mu_0, x\rangle .$$

Then note that

$$\langle \mu_r, (A_0 - A_r)x\rangle \to 0$$

by Lebesgue's dominated convergence theorem, because $0 \leqq A_r(i,j) \uparrow A_0(i,j)$,

$$\sum_i A_0(i,j) \leqq 1$$

and $\mu_r(i) \leqq 1$. Thus

$$\langle (A_0 \mu_0 - \lambda(0)\mu_0), x\rangle = 0 \quad \text{for all } x \in \ell^1,$$

which completes the proof. □

Lemma 15. $\lambda(0) > e^{-g(\infty)}$.

Proof. First we prove $\lambda(0) \geqq e^{-g(\infty)}$. Indeed, since $\mu_0 \geqq x_0 = e_1 = (1, 0, 0, \dots)$ it follows from (3.4) and (3.5) together with the decomposition in (3.7) that

$$P^k \mu_0 \geqq P^k x_0 = x_k,$$

$$A_0 \geqq e^{-g(\infty)} P$$

hence

$$\lambda^k(0)\mu_0 = A_0^k \mu_0 \geqq e^{-kg(\infty)} x_k .$$

Now recall $\mu_0(1) = 1$ and $x_k(1) = \dfrac{1}{k+1}$, and let $k \to \infty$.

Suppose next that $\lambda(0) = e^{-g(\infty)}$. This would imply via $A_0 \mu_0 = \lambda(0)\mu_0$ and the decomposition (3.7) that

$$(I - P)\mu_0 = e^{g(\infty)} K \mu_0 .$$

Hence $K\mu_0 = 0$ by Lemma 6(ii). This contradicts the strict positivity of K. □

Lemma 16. *$A_0 x = \lambda(0)x$ has a unique solution in ℓ^2 with $\|x\| = 1$, $x > 0$.*

Proof. The key property is that the essential spectral radius of a bounded symmetric linear operator is invariant under compact perturbations (see Kato [12], X. 5.35). Hence

$$r_{\text{ess}}(A_0) = r_{\text{ess}}(e^{-g(\infty)}P + K) = r_{\text{ess}}(e^{-g(\infty)}P) = e^{-g(\infty)}$$

via $r_{\text{ess}}(P) = 1$ as in (3.3). Thus $\lambda(0) > r_{\text{ess}}(A_0)$ by Lemma 15. Therefore $\lambda(0)$ is an eigenvalue with finite multiplicity. Repeat the argument in the proof of Lemma 9 to see that $\lambda(0)$ is simple and that the corresponding eigenvector is strictly positive. □

Proof of Proposition 5. Combine Lemmas 9-12 and 16. In particular, note that Lemmas 10 and 11 extend to the non-compact case, as a result of Lemma 16, so that $\lambda(r)$ is analytic on the *closed* interval $[0, \infty)$. □

4. Analysis of $\varrho(\beta, h)$: variation over θ

With Theorem 3 we have all the information we need in order to solve (1.25). First we formulate what properties $J_{\beta,h}(\theta)$ has by substituting our results for $K(\theta)$ into (1.26). Propositions 5 (i), (iii) imply that θ_c and h_c defined in (1.17) and (1.18) are both in (0, 1).

Theorem 4. *$\theta \to J_{\beta,h}(\theta)$ is continuous and concave on $[0,1]$. Furthermore, it is strictly concave and analytic on $(\theta_c, 1)$ and is linear on $(0, \theta_c)$, where $\theta_c \in (0,1)$ is defined in* (1.17). *Define h_c as in* (1.18)

$$h_c = \frac{1-\lambda^2(0)}{1+\lambda^2(0)} \in (0,1)\,. \tag{4.1}$$

I. *If $h < h_c$ then $J_{\beta,h}(\theta)$ is strictly decreasing on $[0,1]$ and $\varrho(h) = J_{\beta,h}(0)$ with unique maximiser $\bar{\theta} = 0$.*

II. *If $h = h_c$ then $J_{\beta,h}(\theta)$ is flat on $[0, \theta_c]$ and is strictly decreasing on $(\theta_c, 1)$. The maximiser is not unique, but again $\varrho(h) = J_{\beta,h}(0)$.*

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

$$\bar{\theta} = -\frac{\lambda(r)}{\lambda'(r)} \tag{4.2}$$

with $r = r(\beta, h)$ the unique solution of

$$h = \frac{1-\lambda^2(r)}{1+\lambda^2(r)} \tag{4.3}$$

and the maximum is

$$\varrho(\beta, h) = J_{\beta,h}(0) + r\,. \tag{4.4}$$

Proof. All statements prior to (4.1) are obvious from (1.26) and the last statement in Theorem 3, except for the strict concavity on $(\theta_c, 1)$. To see the latter, compute from (1.26) for $\theta > \theta_c$ using (2.36) and (2.37)

$$\frac{\partial}{\partial\theta} J_{\beta,h}(\theta) = \frac{1}{2}\log\left(\frac{1+h}{1-h}\right) + \log\lambda(r)\,, \tag{4.5}$$

$$\frac{\partial^2}{\partial\theta^2} J_{\beta,h}(\theta) = -\frac{1}{\theta}\frac{dr}{d\theta} \tag{4.6}$$

(note cancellation of terms) and use that $\theta \to r(\theta)$ is strictly increasing by Theorem 3.I.

The slope of $J_{\beta,h}(\theta)$ at $\theta = 0$ equals (observe that $r(\theta) = 0$ for $\theta < \theta_c$ by Theorem 3)

$$\frac{1}{2}\log\left(\frac{1+h}{1-h}\right) + \log\lambda(0)\,,$$

which changes from negative to positive at $h = h_c$ defined in (1.18). Now parts I and II are obvious. To see part III, note that $J_{\beta,h}(\theta)$ by (4.5) reaches its maximum when (4.3) holds with $r = r(\theta)$ the unique solution of (2.36). This proves (4.2) and (4.3). Finally, (4.4) is found by substituting (2.37) into (1.26). □

Remark. Note the important qualitative change of $J_{\beta,h}(\theta)$ as h crosses the critical value h_c. Also note that the maximiser jumps from $\bar{\theta} = 0$ to $\bar{\theta}_c > 0$.

Proof of Theorems 2A *and* 2B. Combine Theorems 3 and 4 with Proposition 5. Part (i) is (4.4). Part (ii) is (4.2) and (2.34). Part (iv) follows from (1.19) and Proposition 5 (i)-(iv), except for the strictly decreasing property of $h \to \varrho(\beta, h)$ on $(h_c, 1)$. The latter, by part (i), is implied by the following computation:

$$\begin{aligned}\frac{\partial}{\partial h}\varrho(\beta,h) &= -\frac{h}{1-h^2} + \frac{\partial}{\partial h} r(\beta,h)\\ &= -\frac{(1-\lambda^2(r))(1+\lambda^2(r))}{4\lambda^2(r)} - \frac{(1+\lambda^2(r))^2}{4\lambda(r)\lambda'(r)}\\ &= -\left(\frac{1+\lambda^2(r)}{2\lambda(r)}\right)^2\left[\frac{1-\lambda^2(r)}{1+\lambda^2(r)} + \frac{\lambda(r)}{\lambda'(r)}\right] < 0\,,\end{aligned}$$

where $r = r(\beta, h)$ and we substitute (4.3). In the last step we have used the following proposition, which will be proved in section 5:

Proposition 6.

$$-\frac{\lambda'(r)}{\lambda(r)} > \frac{1+\lambda^2(r)}{1-\lambda^2(r)} \quad \textit{for all } r > 0\,. \tag{4.7}$$

Part (iii) follows from (4.7) via (1.17) and (1.18). Part (v) follows because via (1.17) and (1.18)

$$\frac{\partial}{\partial h} r(\beta, h_c +) = -\frac{(1+\lambda^2(0))^2}{4\lambda(0)\lambda'(0)} = \frac{\theta_c}{1-h_c^2}.$$

Part (vi) is obvious from part (iv). Part (vii) follows from (1.19) and Proposition 5 (i)-(iv), except for the upper bound. The latter follows from Proposition 6 via (4.2) and (4.3). □

5. Proof of Proposition 6

The proof requires a sequence of steps. First we show that A_r can be viewed as inducing a *random map* on $[0,1)$ and that therefore $A_r^n e_1$ $(e_1 = (1,0,0,\ldots))$ can be written as the expectation of some functional of a *continued fraction* with random coefficients. Next we show that $\log \lambda(r) = \lim_{n\to\infty} \frac{1}{n} \log (A_r^n e_1)_1$ can be identified as the pressure of a *Gibbs measure* for some potential that can be expressed in terms of the continued fraction. Finally we establish the FKG-*property* of this Gibbs measure and use it to prove inequality (4.7) via a *class argument.*

Step 1. Fix $r > 0$. Our starting point is the observation that A_r is a convex combination of simpler matrices. Namely, combine (1.13), (1.16) and (1.30) to write

$$A_r = \int_{[0,\infty)} P_{r+s}\gamma(ds) \tag{5.1}$$

where

$$P_t(i,j) = e^{-t(i+j-1)} P(i,j) \quad (t>0,\ i,j \in \mathbb{N}), \tag{5.2}$$

(5.3) γ is the distribution of $\log(M/b)$ induced by the distribution β of b.

Thus A_r can be viewed as the expectation of a random operator. The family $(P_t)_{t>0}$ preserves the cone of geometric vectors

$$\{c x_\xi : c \in \mathbb{R}^+;\ x_\xi(i) = \xi^{i-1},\ i \in \mathbb{N},\ \xi \in [0,1)\}. \tag{5.4}$$

Indeed, from (3.2)

$$P_t x_\xi = T(t,\xi)\, x_{T(t,\xi)}, \tag{5.5}$$

$$T(t,\xi) = \frac{1}{2e^t - \xi}. \tag{5.6}$$

For each $t > 0$, $\xi \to T(t,\xi)$ is a map from $[0,1)$ into itself, with a unique attracting fixed point $\phi(t) \in (0,1)$. Thus A_r restricted to the set (5.4) can be viewed as inducing a random map on $[0,1)$ via the equation

$$A_r x_\xi = \int_{[0,\infty)} T(r+s,\xi)\, x_{T(r+s,\xi)}\gamma(ds). \tag{5.7}$$

Therefore define the following objects:

(5.8) $Y = (Y_1, Y_2, \ldots)$ is a random process with values in $[0, \infty)$,

(5.9) $$f_k(r, y) = 1/2e^{r+y_k} - 1/2e^{r+y_{k-1}} - \cdots - 1/2e^{r+y_1} \quad (y \in [0, \infty)^{\mathbb{N}}, k \in \mathbb{N})$$

where $y = (y_1, y_2, \ldots)$. Nσte that in the truncated continued fraction (5.9) the y_i's appear in reversed order.

Lemma 17. *Let $E_{\gamma^{\mathbb{N}}}$ denote expectation over Y w.r.t. $\gamma^{\mathbb{N}}$. Then*

(5.10) $$A_r^n e_1 = E_{\gamma^{\mathbb{N}}}\left(\left[\prod_{k=1}^{n} f_k(r, Y)\right] x_{f_n(r,Y)}\right).$$

Proof. For $n = 1$

$$\begin{aligned} A_r e_1 &= \int_{[0,\infty)} T(r+s, 0)\, x_{T(r+s,0)}\, \gamma\,(ds) \\ &= E_{\gamma^{\mathbb{N}}}\left(T(r+Y_1, 0)\, x_{T(r+Y_1,0)}\right) \end{aligned}$$

and we have $T(r + Y_1, 0) = f_1(r, Y)$. The proof follows by induction using (5.7) and the observation that $T(r + Y_{k+1}, f_k(r, Y)) = f_{k+1}(r, Y)$. □

Step 2. The next step is to evaluate the growth rate of the r.h.s. of (5.10). Define

(5.11) $$f(r, y) = 1/2e^{r+y_1} - 1/2e^{r+y_2} - \cdots \quad (y \in [0, \infty)^{\mathbb{N}}).$$

Note that in the continued fraction (5.11) the y_i's appear in the original order as opposed to (5.9). Let σ denote the shift on $[0, \infty)^{\mathbb{N}}$ defined by $\sigma y = (y_2, y_3, \ldots)$. Use the same symbol σ for the induced shift on $\mathscr{P}([0, \infty)^{\mathbb{N}})$, the set of probability measures on $[0, \infty)^{\mathbb{N}}$. Define

(5.12) $$M = \{Q \in \mathscr{P}([0, \infty)^{\mathbb{N}}) : \sigma Q = Q\},$$

(5.13) $$h(Q \mid \gamma^{\mathbb{N}}) = \text{specific relative entropy of } Q \text{ w.r.t. } \gamma^{\mathbb{N}}$$

(see e.g. Georgii [9], 15.13).

Lemma 18. *Let E_Q denote expectation over Y w.r.t. Q. Then*

(5.14) $$\begin{aligned} \log \lambda(r) &= \lim_{n\to\infty} \frac{1}{n} \log (A_r^n e_1)_1 \\ &= \sup_{Q \in M} \left[E_Q(\log f(r, Y)) - h(Q \mid \gamma^{\mathbb{N}})\right]. \end{aligned}$$

Proof. The first equality in (5.14) is a standard property of self-adjoint positive compact operators on a Hilbert space obtained via the spectral representation (Kato [12], V. 2.3).

From (5.10) we obtain, by reversing the order of $(Y_1, \ldots, Y_n)$ and defining

$$Y^{(n)} = (Y_1, \ldots, Y_n, \infty^{\mathbb{N}}),$$

that

$$(A_r^n e_1)_1 = E_{\gamma^{\mathbb{N}}}\left(\prod_{k=0}^{n-1} f(r, \sigma^k Y^{(n)})\right). \tag{5.15}$$

Next we note that by the monotonicity of $y \to f(r, y)$ in each component

$$\begin{aligned} \sup_{\substack{y, y' \in [0,\infty)^{\mathbb{N}} \\ y_i = y'_i \text{ for } 1 \leqq i \leqq n}} |f(r,y) - f(r,y')| &= f(r, (0^n, 0^{\mathbb{N}})) - f(r, (0^n, \infty^{\mathbb{N}})) \\ &= T^n(r, \phi(r)) - T^n(r, 0) \end{aligned} \tag{5.16}$$

where T^n denotes the nth iterate of the map $\xi \to T(r, \xi)$ and $\phi(r) \in (0,1)$ its unique fixed point (recall (5.6)). Since $\phi(r)$ is attracting, it follows that the r.h.s. of (5.16) tends to zero as $n \to \infty$ and hence

$$\lim_{n\to\infty} \frac{1}{n}\left| \log E_{\gamma^{\mathbb{N}}}\left(\prod_{k=0}^{n-1} f(r, \sigma^k Y^{(n)})\right) - \log E_{\gamma^{\mathbb{N}}}\left(\prod_{k=0}^{n-1} f(r, \sigma^k Y)\right)\right| = 0 \tag{5.17}$$

(also note that $f(r, y)$ is bounded away from 0 because γ has bounded support; recall (1.2) and (5.3)). Therefore we must show that

$$\lim_{n\to\infty} \frac{1}{n} \log E_{\gamma^{\mathbb{N}}}\left(\exp\left[\sum_{k=0}^{n-1} \log f(r, \sigma^k Y)\right]\right) = \text{r.h.s. of (5.14)}. \tag{5.18}$$

But (5.18) is a standard application of Varadhan's theorem in *large deviation theory* at the level of the empirical process, associated with Y. We refer to Deuschel and Stroock [6], 2.1.10, 4.4.1 and 4.4.12. To apply this theorem we use that γ has bounded support and that $E_Q(\log f(r, Y))$ is bounded from above and is continuous on $\mathscr{P}([0,\infty)^{\mathbb{N}})$ in the weak topology. We also refer to Georgii [9], 15.16 for the identification of $h(Q|\gamma^{\mathbb{N}})$. □

Step 3. The r.h.s. of (5.14) has the shape of the *Gibbs Variational Formula* in the theory of Gibbs measures (Georgii [9], 15.39). This leads to the identification below in Lemma 19.

Lemma 19. *The supremum in* (5.14) *is attained at* $\bar{Q}$ *equal to the projection on* $[0,\infty)^{\mathbb{N}}$ *of the unique shift-invariant Gibbs measure on* $[0,\infty)^{\mathbb{Z}}$ *w.r.t. the reference measure* $\gamma^{\mathbb{N}}$ *and with interaction potential* $(\Phi_A)_{A \subseteq \mathbb{Z}, |A| < \infty}$ *given by*

$$\begin{aligned} \Phi_{\sigma A} &= \Phi_A \quad \textit{for all } A, \\ \Phi_{\{1\}}(y) &= -\log f(r, y^{(1)}), \\ \Phi_{\{1,\ldots,k\}}(y) &= -\log[f(r, y^{(k)})/f(r, y^{(k-1)})] \quad (k \geqq 2), \\ \Phi_A &= 0 \quad \textit{for all } A \neq \{1, \ldots, k\} \textit{ for some } k \end{aligned} \tag{5.19}$$

(*recall that* $y^{(k)} = (y_0, \ldots, y_k, \infty^{\mathbb{N}})$).

Proof. We refer to Georgii [9], 15.39, from which it follows that the set of maxima of the supremum in (5.14) coincides with the projection on $[0,\infty)^{\mathbb{N}}$ of the set of shift-invariant Gibbs measures w.r.t. the reference measure $\gamma^{\mathbb{N}}$ and with interaction potential solving

$$-\log f(r,y) = \sum_{\substack{A\subseteq \mathbb{N}\\ |A|<\infty\\ A\ni\{1\}}} \Phi_A(y)\,. \tag{5.20}$$

Clearly (5.19) satisfies (5.20). The maximum in (5.14) equals minus the pressure of the potential. The sum in (5.20) is the specific energy.

To get that the maximum in (5.14) exists and is unique we use the theory of Gibbs measures. Return to (5.16). Since $\xi \to T(r,\xi)$ is smooth, the r.h.s. of (5.16) decays to zero asymptotically as $[(\partial/\partial\xi)T(r,\xi)|_{\xi=\phi(r)}]^n = [\phi(r)]^{2n}$ (recall (5.6)). Hence from (5.19)

$$\|\Phi_{\{1,\dots,k\}}\|_\infty \leqq (\phi(r)+\varepsilon)^{2k} \quad \text{for any } \varepsilon>0 \text{ and } k \text{ large}\,.$$

Since $\phi(r)<1$ for all $r>0$, the latter says that the interaction decays exponentially. This implies, according to a classical theorem (see Georgii [9], 8.39), the existence and uniqueness of the Gibbs measure. □

Step 4. In order to be able to take advantage of the identification in Lemma 19 we shall need the following notion: Q is said to be an FKG-*measure* if for all functions $a,b:[0,\infty)^{\mathbb{N}} \to \mathbb{R}$ measurable, bounded and increasing (in each component)

$$E_Q(ab) \geqq E_Q(a)\,E_Q(b)\,. \tag{5.21}$$

Lemma 20. *$\bar{Q}$ is an* FKG-*measure.*

Proof. Since the reference measure $\gamma^{\mathbb{N}}$ is product measure and since the interaction potential $(\Phi_A)_{A\subseteq \mathbb{Z},|A|<\infty}$ is given by (5.20), a sufficient condition for $\bar{Q}$ to be an FKG-measure is the following property:

$$f(r,y\vee y')\,f(r,y\wedge y') \geqq f(r,y)\,f(r,y') \quad \text{for all } y,y',r \tag{5.22}$$

where $y\vee y'$ and $y\wedge y'$ are the coordinatewise maximum resp. minimum of y and y'. The above criterion can be found e.g. in Batty and Bollmann [3].

The proof of (5.22) proceeds by induction. Let $f^{(k)}(r,y)$ denote the continued fraction truncated after the kth term (i.e., (5.9) but with the y_i's in the original order). Since $\lim_{k\to\infty} f^{(k)}(r,y) = f(r,y)$ for all r,y (recall (5.16)) it suffices to prove that

$$f^{(k)}(r,y\vee y')\,f^{(k)}(r,y\wedge y') \geqq f^{(k)}(r,y)\,f^{(k)}(r,y') \quad \text{for all } y,y',r,k\,. \tag{5.23}$$

To get (5.23) we first note that from (5.11) we have the recursion relation

$$f^{(k)}(r,y) = \frac{1}{2e^{r+y_0} - f^{(k-1)}(r,\sigma y)} \quad (k\geqq 1) \tag{5.24}$$

with $f^{(0)}(r,y) \equiv 0$. Differentiating (5.24) w.r.t. y_i we get

$$\frac{\partial}{\partial y_0} f^{(k)}(r,y) = -[f^{(k)}(r,y)]^2, \tag{5.25}$$

$$\frac{\partial}{\partial y_i} f^{(k)}(r,y) = -[f^{(k)}(r,y)]^2 \frac{\partial}{\partial y_i} f^{(k-1)}(r,\sigma y) \quad (i \geqq 1).$$

The last derivative is zero when $i \geqq k$. It therefore follows from (5.25) by induction that

$$\frac{\partial}{\partial y_i} f^{(k)}(r,y) = -\prod_{j=0}^{i} [f^{(k-j)}(r,\sigma^j y)]^2 \quad (0 \leqq i < k). \tag{5.26}$$

This in turn implies

$$\frac{\partial}{\partial y_i} \log f^{(k)}(r,y) = -f^{(k)}(r,y) \prod_{j=1}^{i} [f^{(k-j)}(r,\sigma^j y)]^2 \quad (0 \leqq i < k). \tag{5.27}$$

Since $y \to f^{(k)}(r,y)$ is non-decreasing in each y_i for all $k \geqq 0$, we can now from (5.27) read off the following property:

(5.28) $y \to f^{(k)}(r, y \vee y')/f^{(k)}(r,y)$ is non-decreasing in each y_i

(and the same for $y \vee y'$ and y').

Property (5.28) is the key to (5.23). Namely, define

$$g^{(k)}(r,(y,y')) = \frac{f^{(k)}(r,y \vee y') f^{(k)}(r,y \wedge y')}{f^{(k)}(r,y) f^{(k)}(r,y')}. \tag{5.29}$$

Define the following operations T_j on (y,y'):

$$(y,y') \to T_j(y,y') = (z,z') \quad (j \geqq 1) \tag{5.30}$$

$$z_i = y_i,\ z'_i = y'_i \quad \text{for } i \neq j,$$

$$z_j = z'_j = y_j \wedge y'_j.$$

Then (5.28) yields

$$g^{(k)}(r,(y,y')) \geqq g^{(k)}(r, T_j(y,y')) \quad \text{for each } j \geqq 1. \tag{5.31}$$

By repeating T_j for $j = 1, \ldots, k$ we eventually get

$$g^{(k)}(r,(y,y')) \geqq g^{(k)}(r,(y \wedge y', y \wedge y')) = 1. \quad \square$$

Step 5. The next step is to combine Lemmas 18–20. Namely, with the abbreviations

$$\mu_Q(r) = \exp E_Q(\log f(r,Y)), \tag{5.32}$$

$$\hat{M} = \{Q \in M : Q \text{ is FKG}\}, \tag{5.33}$$

we can write

$$\lambda(r) = \sup_{Q \in \hat{M}} \left[\mu_Q(r)\, e^{-h(Q \mid \gamma^N)}\right]. \tag{5.34}$$

The restriction of the supremum to the set $\hat{M}$ is the key to finishing the proof of inequality (4.7) in Proposition 6, as we shall next see.

We continue with a *class argument*. Define the class of functions

$$\mathscr{C} = \left\{\mu : (0, \infty) \to (0,1) \text{ satisfying} : e^{-x}\left(\mu(x) + \frac{1}{\mu(x)}\right)\uparrow\right\} \tag{5.35}$$

($\uparrow$ means non-decreasing) and note that for any μ such that μ' exists the following holds:

$$\mu \in \mathscr{C} \Leftrightarrow -\frac{\mu'(r)}{\mu(r)} \geqq \frac{1+\mu^2(r)}{1-\mu^2(r)}. \tag{5.36}$$

The following lemma applied to (5.34) shows that the function λ is in $\mathscr{C}$ and hence, by (5.36), that (4.7) holds with $\geqq$ instead of $>$. Later we shall exclude $=$.

Lemma 21. (i) *$\mu_Q \in \mathscr{C}$ for every $Q \in \hat{M}$.*

(ii) *$\mathscr{C}$ is closed under multiplying by a constant in* $(0,1)$ *and under taking suprema.*

Proof. Part (ii) is trivial. Part (i) has two steps.

(1) From (5.32) follows

$$-\frac{\mu_Q'(r)}{\mu_Q(r)} = E_Q\left(-\frac{f'(r,Y)}{f(r,Y)}\right). \tag{5.37}$$

To compute the r.h.s. of (5.37) first note that from (5.11)

$$f(r,y) = 1/\left(2e^{r+y_0} - f(r,\sigma y)\right) \tag{5.38}$$

and hence

$$\begin{aligned} -\frac{f'(r,y)}{f(r,y)} &= f(r,y)\left(2e^{r+y_0} - f'(r,\sigma y)\right) \\ &= f(r,y)\left(\frac{1}{f(r,y)} + f(r,\sigma y) - f'(r,\sigma y)\right) \\ &= 1 + f(r,y)\, f(r,\sigma y)\left[1 - \frac{f'(r,\sigma y)}{f(r,\sigma y)}\right]. \end{aligned} \tag{5.39}$$

Iterate (5.39) to obtain the representation

$$(5.40) \qquad -\frac{f'(r,y)}{f(r,y)}$$

$$= \frac{1}{f(r,y)} \sum_{i=0}^{\infty} \left\{ \prod_{j=0}^{i-1} f^2(r,\sigma^j y) \right\} f^2(r,\sigma^i y)\left(1 + f(r,\sigma^i y) f(r,\sigma^{i+1} y)\right).$$

Since $f(r,y)$ is bounded away from 1, the r.h.s. of (5.40) converges and the r.h.s. of (5.37) is finite. Next use that the $f(r,\sigma^i y)$'s are decreasing functions of y. Therefore we can apply the FKG-property of Q together with $\sigma Q = Q$ to get the lower bound (recall (5.21))

$$(5.41) \quad E_Q\left(-\frac{f'(r,y)}{f(r,y)}\right) \geqq \sum_{i=0}^{\infty} c^{2i}(1+c^2) = \frac{1+c^2}{1-c^2}, \quad \text{with } c = E_Q(f(r,Y)).$$

(2) Jensen's inequality applied to (5.32) gives

$$(5.42) \qquad \mu_Q(r) \leqq E_Q(f(r,Y)).$$

Combine (5.37), (5.41) and (5.42) to read off via (5.36) that $\mu_Q \in \mathscr{C}$. □

Step 6. Finally we exclude = in (4.7). Consider the representation in (5.34). Define

$$(5.43) \qquad \tilde{\mu}_Q(r) = \mu_Q(r) e^{-h(Q|\gamma^{\mathbb{N}})}.$$

First we show that for every $\tilde{\mu}_Q \in \hat{M}$

(5.44) either $\tilde{\mu}_Q$ satisfies (4.7) with $>$,

or $h(Q|\gamma^{\mathbb{N}}) = 0$ and $Q = (\delta_{\{c\}})^{\mathbb{N}}$ for some c.

Indeed, since $\mu_Q \in \mathscr{C}$ by Lemma 21 (i), $\tilde{\mu}_Q(r)$ satisfies the inequality in the r.h.s. of (5.36) with $>$ as soon as

$$\exp\left(-h(Q|\gamma^{\mathbb{N}})\right) < 1$$

(recall (5.35)). Moreover, (5.37), (5.41) and (5.42) show that equality can hold in (5.36) if and only if equality holds in (5.42). The latter, by (5.32), requires $f(r,Y)$ to be Q-a.s. constant, which means that the marginal of Q is a point mass. So (5.44) holds.

Next we note that $h(Q|\gamma^{\mathbb{N}}) = 0$ requires $Q = \gamma^{\mathbb{N}}$, which is incompatible with $Q = (\delta_{\{c\}})^{\mathbb{N}}$ by (1.2) and (5.3). So the second option in (5.44) is not possible. Thus we have proved that for every $\tilde{\mu}_Q \in \hat{M}$

$$(5.45) \qquad -\frac{\tilde{\mu}'_Q(r)}{\tilde{\mu}_Q(r)} > \frac{1+\tilde{\mu}^2_Q(r)}{1-\tilde{\mu}^2_Q(r)} \quad \text{for all } r > 0.$$

Next pick any $r_0 > 0$ and consider $r \to \tilde{\mu}_{\bar{Q}_{r_0}}(r)$ where $\bar{Q}_{r_0} \in \hat{M}$ denotes the maximum in Lemma 19 at $r = r_0$. Clearly

$$(5.46) \qquad \tilde{\mu}_{\bar{Q}_{r_0}}(r_0) = \lambda(r_0).$$

Moreover, since $\tilde{\mu}_{\bar{Q}_{r_0}}$ and λ' exist and $\tilde{\mu}_{\bar{Q}_{r_0}}(r) \leqq \lambda(r)$ for all $r > 0$, we must also have

$$\tilde{\mu}_{\bar{Q}_{r_0}}(r_0) = \lambda'(r_0)\,. \tag{5.47}$$

Substitute (5.46) and (5.47) into (5.45) to get the assertion. □

Acknowledgement. Part of the work in this paper was carried out while J.-B. Baillon, Ph. Clément and F. den Hollander were guests at and A. Greven was a member of the Institut für Mathematische Stochastik of the Universität Göttingen in the spring of 1991. We thank Mike Keane for his input in the proof of Proposition 6 on this occasion.

References

[1] *K.B. Athreya* and *P.E. Ney*, Branching Processes, Springer, Berlin 1972.
[2] *J.-B. Baillon*, *Ph. Clément*, *A. Greven* and *F. den Hollander*, A variational approach to branching random walk in random environment, Ann. Probab. **21** (1993), 290–317.
[3] *C.J.K. Batty* and *H.W. Bollmann*, Generalised Holley-Preston inequalities on measure spaces and their products, Z. Wahrscheinlichkeitsth. verw. Geb. **53** (1980), 157–173.
[4] *H. Behnke* and *F. Sommer*, Theorie der Analytischen Funktionen einer komplexen Veränderlichen, Grundl. Math. Wiss. **77**, Springer, Berlin 1965.
[5] *M.G. Crandall* and *P.H. Rabinowitz*, Bifurcation, perturbation of simple eigenvalues, and linearized stability, Arch. Rat. Mech. Anal. **52** (1973), 161–180.
[6] *J.-D. Deuschel* and *D.W. Stroock*, Large Deviations, Academic Press, Boston 1989.
[7] *P.G. Dodds* and *D.H. Fremlin*, Compact operators in Banach lattices, Israel J. Math. **34** (1979), 287–320.
[8] *N. Dunford* and *J.T. Schwartz*, Linear Operators, Part I, II, Interscience Publishers, John Wiley & Sons, Inc., New York 1958, 1963.
[9] *H.-O. Georgii*, Gibbs Measures and Phase Transitions, De Gruyter Stud. Math. **9**, De Gruyter, Berlin 1988.
[10] *A. Greven* and *F. den Hollander*, Branching random walk in random environment: phase transitions for local and global growth rates, Probab. Th. Rel. Fields **91** (1992), 195–249.
[11] *A. Greven* and *F. den Hollander*, On a variational problem for an infinite particle system in a random medium, Part II: The local growth rate, Probab. Th. Rel. Fields **99** (1994).
[12] *T. Kato*, Perturbation Theory for Linear Operators, Grundl. Math. Wiss. **132**, Springer, Berlin 1966.
[13] *T. Kato*, Superconvexity of the spectral radius, and convexity of the spectral bound and the type, Math. Z. **180** (1982), 265–273.
[14] *J.F.C. Kingman*, A convexity property of positive matrices, Quart. J. Math. Oxford (2) **12** (1961), 283–284.
[15] *J. Neveu*, Potentiel Markovien récurrent des chaines de Harris, Ann. Inst. Fourier, Grenoble, (2) **22** (1972), 85–130.
[16] *W. Rudin*, Real and Complex Analysis, McGraw-Hill, New York 1966.
[17] *W. Rudin*, Functional Analysis, McGraw-Hill, New York 1973.
[18] *H. Triebel*, Interpolation Theory, Function Spaces, Differential Operators, North-Holland Mathematical Library, Amsterdam–New York–Oxford 1978.
[19] *A.C. Zaanen*, Linear Analysis, Bibliotheca Math. II, North-Holland Publ. Comp.-Noordhoff, Amsterdam–Groningen 1953.

Institut de Mathématiques et Informatique, Université de Lyon I, 43[Bd] du 11 novembre 89, Villeurbanne 69622 Cedex, France

Faculteit der Technische Wiskunde en Informatica, Technische Universiteit Delft, Mekelweg 4, 2600 GA Delft, The Netherlands

Institut für Stochastik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany

Mathematisch Instituut, Rijksuniversiteit Utrecht, P.O. Box 80.010, 3508 TA Utrecht, The Netherlands

Eingegangen 30. August 1991, in revidierter Fassung 11. November 1993