Stochastic models of a parasitic infection, exhibiting three basic reproduction ratios

Running head: Stochastic models of a parasitic infection

C.J. Luchsinger

Abteilung Angewandte Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, SWITZERLAND, www.luchsinger-mathematics.ch, cl@luchsinger-mathematics.ch

Abstract. Two closely related stochastic models of parasitic infection are investigated: a non-linear model, where density dependent constraints are included, and a linear model appropriate to the initial behaviour of an epidemic. Host-mortality is included in both models. These models are appropriate to transmission between homogeneously mixing hosts, where the amount of infection which is transferred from one host to another at a single contact depends on the number of parasites in the infecting host. In both models, the basic reproduction ratio R_0 can be defined to be the lifetime expected number of offspring of an adult parasite under ideal conditions, but it does not necessarily contain the information needed to separate growth from extinction of infection. In fact we find three regions for a certain parameter where different combinations of parameters determine the behavior of the models. The proofs involve martingale and coupling methods.

Key words: Basic Reproduction Ratio - Host-parasite system - Group infection - Markov branching process - Coupling - infinitely many types

1 Introduction

In Barbour and Kafetzaki (1993) [BK], a model for the spread of a parasitic disease was introduced. The aim was to generate the highly over-dispersed distribution of numbers of parasites per host observed in schistosomiasis data. This goal was not to be achieved through heterogeneity in host susceptibility or through random (and over-dispersed) numbers of parasites acquired per infection, with the associated distribution being chosen so as to fit the data as well as possible. The drawback of such approaches is the danger that the resulting models may not adequately reflect what happens if parameter values change, since the distributions involved are chosen *ad hoc*, rather than being derived from an underlying biological mechanism. In [BK] the aim was to see how much variability in individual parasite loads can be achieved from chance interaction phenomena alone, without sacrificing the assumption of a population of homogeneous individuals.

The desired variability in parasite burden was achieved in the [BK]-model. However another important result emerged, namely that for some parameter values the natural candidate for the basic reproduction ratio R_0 does not necessarily contain the information needed to separate growth from extinction of infection. This result was further studied in Barbour (1994) [B2], where a *linear model*, approximating the initial behaviour of the original model in large populations, was investigated; the same phenomena were present there, too. Neither model included mortality of hosts, which however could be expected to have a substantial influence on the behaviour. In this paper, it is shown that in this case there are three separate ranges of parameters, in which different combinations of parameters are critical for separating growth from extinction.

2 The models and their behavior

We first consider a non-linear model with a fixed number M of individuals, each of which may carry parasites. Let $x^{(M)}$ be an infinite dimensional Markov process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$

$$x^{(M)}(\omega, t) : \Omega \times [0, \infty) \to \{[0,1] \cap M^{-1}\mathbb{Z}\}^{\infty},$$

in which $x_j^{(M)}(t)$, $j \geq 0$, denotes the *proportion* of individuals at time t, $t \geq 0$, that are infected with j parasites, so that $\sum_{j\geq 0} x_j^{(M)}(0) = 1$ and $x_j^{(M)}(0) \geq 0$, $j \geq 0$. We suppress the index M whenever possible. The parasites have independent lifetimes, exponentially distributed with mean $1/\mu$. Each infected individual makes contacts with other individuals with rate λ; but only those contacts that are with an uninfected individual lead to a new infection (concomitant immunity). If the infecting individual has j parasites, then the result of a contact is to establish a newly infected individual with a random number S_j of parasites, where $S_j := \sum_{i=1}^{j} Y_i$ and the Y_i are independent and identically distributed with mean θ and variance $\sigma^2 < \infty$. We define $p_{jk} := \mathbb{P}[S_j = k]$, and so

$$\sum_{k \geq 0} p_{jk} = 1 \quad \text{for each } j \text{ and } \sum_{k \geq 1} k p_{jk} = j\theta. \tag{2.1}$$

We assume that individuals have independent lifetimes, exponentially distributed with mean $1/\kappa$, no matter how high the parasite burden is. All parasites die if their host dies. We allow the possibility of having $\kappa = 0$, meaning that people live for an infinite length of time, as in [BK] and [B2].

In this non-linear model, we replace an individual that dies by an uninfected individual. The rates with which x changes are then as follows:

$$\begin{aligned} &x \to x + M^{-1}(e_{j-1} - e_j) \text{ at rate } jM\mu x_j \;;\; j \geq 1, \\ &x \to x + M^{-1}(e_k - e_0) \text{ at rate } \lambda M x_0 \sum_{l \geq 1} x_l p_{lk} \;;\; k \geq 1, \\ &x \to x + M^{-1}(e_0 - e_r) \text{ at rate } M x_r \kappa \;;\; r \geq 1, \end{aligned} \tag{2.2}$$

where e_i denotes the i-th co-ordinate vector in $\mathbb{R}^\infty$. We call this model N; this stands for **N**on-linear. We introduce a notation for the sigma-algebras too: $\mathcal{F}_s := \sigma\{x(u), 0 \leq u \leq s\}$. At this point it is convenient to explain why we have *these* rates: there are jMx_j parasites in individuals with j parasites and they all die at a rate of μ. If such a parasite dies, the proportion of individuals with j parasites decreases by $1/M$ and the proportion of individuals with $(j-1)$ parasites increases by $1/M$. This explains the first transition-rate. The second transition is an infection: there are $x_l M$ individuals with l parasites who make contacts according to a Poisson process of rate λ. But only those contacts that take place with uninfected individuals are infective. So the rate must be decreased by multiplying with the proportion of uninfected people x_0. Then we must include the probability that such an infection leads to an infection with k parasites, hence the probability p_{lk}. The last rate describes a death of a person: there are Mx_r individuals with r parasites and they die at a rate of κ.

The reason for modelling the lifetime of individuals and parasites with an exponential distribution is purely for mathematical simplicity. Careful study of schistosomiasis incidence data, led to the following conclusions for the infection process (Barbour (1977) [B1]): if individuals get infected, then they are not infected serveral times per year, but once every few years. On the other hand, children at the age of 12 may well have a large number of parasites in their body. So we presume that there is group infection. Then S_j is generated such that each of the j parasites in the body of the infecting individual independently produces a random number $Y_i, 1 \leq i \leq j$, of offspring in the newly infected host. The exact distribution of S_j would therefore have to be calculated as a convolution of the Y_i's, but this will never be necessary explicitly. Additionally we assume concomitant immunity in humans, meaning that people get infected and then are immune to further infections until they have fully recovered (see Bradley and McCullough (1973) [BM]).

The linear model, useful in modelling the initial phase of an epidemic outbreak, is defined as follows. Let X be an infinite dimensional Markov process

$$X(\omega, t) : \Omega \times [0, \infty) \to \{[0, \infty) \cap \mathbb{Z}\}^{\infty},$$

where $X_j(t),\ j \geq 1$, denotes the *number* of individuals at time t, $t \geq 0$, that are infected with j parasites. We assume that $0 < \sum_{j \geq 1} X_j(0) = M < \infty$ and $X_j(0) \geq 0$, $j \geq 1$. The rates at which X changes are as follows:

$$\begin{aligned}
&X \to X + (e_{j-1} - e_j) \text{ at rate } j\mu X_j \ ;\ j \geq 2, \\
&X \to X - e_1 \text{ at rate } \mu X_1 \ ;\ (j = 1), \\
&X \to X + e_k \text{ at rate } \lambda \sum_{l \geq 1} X_l p_{lk} \ ;\ k \geq 1, \\
&X \to X - e_r \text{ at rate } X_r \kappa \ ;\ r \geq 1.
\end{aligned} \tag{2.3}$$

We call this model L; this stands for **L**inear. We introduce a notation for the sigma-algebras too: $\mathcal{G}_s := \sigma\{X(u), 0 \leq u \leq s\}$. The difference between model N and L is the following: in model L the contact rate is λ and there is no limiting factor in the model. In model N the contact rate is altered from λ to λx_0, because only those infectious contacts that are with an uninfected individual lead to a new infection. In the remainder of this section we outline the behaviour of the two models; the proofs are deferred to section 3.

The basic reproduction ratios

Let us define $R_0 := \lambda\theta/(\mu + \kappa)$, $R_1 := (\lambda e \log \theta)/(\mu \theta^{\frac{\kappa}{\mu}})$ and $R_2 := \lambda/\kappa$. These are quantities which emerge as being critical in determining the behavior of the models as is seen in Theorems 2.1 and 2.3. R_0 is what would usually be called the basic reproduction ratio, because it denotes the average number of offspring of a single parasite during his whole lifetime in the absence of density dependent constraints. This can be seen in the following way. Since a parasite also dies if its host dies, a parasite has an exponentially distributed lifetime with parameter $\mu+\kappa$, and hence its expected lifetime is $(\mu+\kappa)^{-1}$. During its life, it makes contacts at rate λ per unit time, and on average these contacts result in infections with θ parasites. We do not have an obvious interpretation for R_1, but the reader is referred to [B2] for an interpretation if $\kappa = 0$. For $R_2 > 1$, R_2^{-1} denotes the probability that a pure birth and death process with contact rate λ and death rate κ dies out, beginning with one initial infected. As is seen later on, R_2 becomes critical when θ is 'large'. It seems that then the bulk of infected hosts die before they recover because they are infected with very large numbers of parasites. Therefore, in that case, if we are only interested whether the infection dies out or not, we almost have the same behaviour as in a pure birth and death process.

The (linear) system is essentially more complicated than the multitype branching process with finite number of types and it does not seem possible to simply characterise the growth rates in terms of Perron-Frobenius eigenvalues.

By the expression 'threshold behavior' we denote general statements of the following kind: if $R_0 > 1$ the epidemic has a positive probability to develop and if $R_0 \leq 1$ the epidemic dies out almost surely. As we see in what follows, the situation in our models is far more complex than that stated above.

Theorem 2.1 *Suppose $X(0) = y^{(0)}$ in model L such that $0 < \sum_{j\geq 1} y_j^{(0)} < \infty$ where $y^{(0)}$ is fixed. Then the following results hold:*

(1): $\log\theta \leq (1+\kappa/\mu)^{-1}$. Then $\mathbb{P}[\lim_{t\to\infty}\sum_{j\geq 1} X_j(t) = 0] = 1$ if and only if $R_0 \leq 1$.

(2): $(1+\kappa/\mu)^{-1} < \log\theta \leq \mu/\kappa$. Then $\mathbb{P}[\lim_{t\to\infty}\sum_{j\geq 1} X_j(t) = 0] = 1$ if and only if $R_1 \leq 1$.

(3): $\mu/\kappa < \log\theta$. Then $\mathbb{P}[\lim_{t\to\infty}\sum_{j\geq 1} X_j(t) = 0] = 1$ if and only if $R_2 \leq 1$.

In addition, the expected number of parasites in L grows with an exponential rate $(\lambda\theta -$

$\mu - \kappa$):

$$\mathbb{E}\Big[\sum_{j\geq 1} jX_j(t)\Big] = \Big(\sum_{j\geq 1} jy_j^{(0)}\Big)e^{(\lambda\theta-\mu-\kappa)t}. \tag{2.4}$$

Remarks 1) If $\kappa = 0$, these results stay true with the following adjustments: the third region for θ is shifted away to infinity. So we have only two regions for θ if $\kappa = 0$, namely: $\theta < e$ and $\theta \geq e$, and the basic reproduction ratios simplify to $R_0 = \lambda\theta/\mu$ and $R_1 = \lambda e \log\theta/\mu$. Then Theorem 2.1 is Theorem 2.1 in [B2].
2) The deterministic analogue is Remark 1) to Theorem 3.1 and equation (3.3) in [Ld].
3) Suppose $\mu/\kappa \geq \log\theta > (1+\kappa/\mu)^{-1}$, $R_0 > 1$ and $R_1 < 1$ (which is possible!). This implies that the epidemic dies out with probability one in model L; but it means too that the expected number of parasites tends to infinity.

If an epidemic outbreak is not in the initial phase anymore, the non-linear model is more appropriate. The following theorem shows that in the non-linear cases the epidemic eventually dies out with probability one, no matter what values the parameters take.

Theorem 2.2 *In the non-linear model N the infection dies out with probability one, that is, for all M, $1 \leq M < \infty$,*

$$\mathbb{P}[\lim_{t\to\infty} x(t)^{(M)} = e_0] = 1.$$

Remark There is no deterministic analogue of Theorem 2.2, however, the reader should observe Theorem 3.7 in [Ld] as a contrast.

Looking at Theorem 2.2, we see that in model N the epidemic *finally* dies out almost surely, no matter what values the parameters take. But the behavior of the non-linear model in finite time (and with M large) is quite different, depending on whether R_i, $i \in \{0,1,2\}$ is greater or smaller than one. This is made more precise in

Theorem 2.3 *Fix $y \in (\mathbb{N}\cup\{0\})^\infty$, such that $0 < Y := \sum_{j\geq 1} y_j < \infty$, and suppose that for each $M > Y$ we have $x_j^{(M)}(0) = y_j/M$ for all $j \geq 1$, $x_0^{(M)}(0) = 1 - Y/M$. Then in model N we have the following threshold behavior:*

Case 1): $\log\theta \le (1+\kappa/\mu)^{-1}$. *Then*

$$\lim_{t\to\infty}\lim_{M\to\infty}\mathbb{P}\big[\sum_{j\ge 1} x_j^{(M)}(t)=0\big]=1 \text{ if and only if } R_0\le 1.$$

Case 2): $(1+\kappa/\mu)^{-1} < \log\theta \le \mu/\kappa$. *Then*

$$\lim_{t\to\infty}\lim_{M\to\infty}\mathbb{P}\big[\sum_{j\ge 1} x_j^{(M)}(t)=0\big]=1 \text{ if and only if } R_1\le 1.$$

Case 3): $\log\theta > \mu/\kappa$. *Then*

$$\lim_{t\to\infty}\lim_{M\to\infty}\mathbb{P}\big[\sum_{j\ge 1} x_j^{(M)}(t)=0\big]=1 \text{ if and only if } R_2\le 1.$$

Remarks 1) We let M tend to ∞ first (with t fixed). In the linear models the contact rate λ stays the same no matter how many individuals are infected. But in the non-linear model this contact rate is altered by multiplying it with the proportion of uninfected $\lambda x_0^{(M)}$. As we increase M, we only increase the initial number of *uninfected* individuals: $Mx_0^{(M)}(0) = M - Y$, the initial *number* of infected individuals stays constant and equal to Y. Since the initial proportion $x_0^{(M)}(0)$ of uninfected tends to 1 as M tends to infinity, we almost have a linear model in the initial phase. So it is not too surprising that we have results analogous to those in Theorem 2.1. Note that it is vital to let M converge to infinity first and then let t converge to infinity, because of Theorem 2.2.

2) Again, as in Theorem 2.1, if $\kappa = 0$, these results stay true, with the interpretation that the third region for θ is shifted away to infinity.

3) The deterministic analogue is Theorem 3.7 in [Ld].

3 Proofs

We first have to be sure that the linear process is 'regular', in the sense that it makes only finitely many transitions in any finite time interval [0,T], almost surely. This is shown in the following

Lemma 3.1 *The process X that evolves according to L is regular.*

Proof of Lemma 3.1 If there are infinitely many transitions in a finite time interval [0,T], there must be infinitely many infections too in [0,T]. But this is impossible, as can be seen by comparison with a pure birth process of rate λ.

$\square$

Proof of Theorem 2.1 The case where $\kappa = 0$ was shown in [B2] as Theorem 2.1, except for equation (2.4). Therefore we may assume that $\kappa > 0$ except for the proof of (2.4). For the proof of Theorem 2.1 we first need four technical lemmas (Lemmas 3.2, 3.3, 3.4 and 3.5).

Lemma 3.2 *a) If* $\log\theta \leq (1+\kappa/\mu)^{-1}$ *and* $R_0 > 1$*, or if* $R_1 > 1$*, then* $R_2 > 1$*.*
b) If $\log\theta \leq (1+\kappa/\mu)^{-1}$ *and* $R_0 > 1$*; or if* $R_1 > 1$*; or if* $\mu/\kappa < \log\theta$ *and* $R_2 > 1$*, then* $\inf_{(0<\alpha\leq 1)} \lambda\theta^\alpha/(\mu\alpha+\kappa) > 1$.

Proof of Lemma 3.2 a) This follows from part b) because

$$R_2 = \frac{\lambda}{\kappa} = \left.\frac{\lambda\theta^\alpha}{\mu\alpha+\kappa}\right|_{\alpha=0} \geq \inf_{(0<\alpha\leq 1)} \frac{\lambda\theta^\alpha}{\mu\alpha+\kappa}.$$

We do not use part a) to prove part b).

b) In the first region we have $\log\theta \leq (1+\kappa/\mu)^{-1}$ and $\lambda\theta > \mu+\kappa$. We want to show that for $\alpha \in (0,1]$ we have $\lambda\theta^\alpha > \mu\alpha+\kappa$. We have

$$\lambda\theta^\alpha = \lambda\theta\theta^{\alpha-1} > (\mu+\kappa)\theta^{\alpha-1}$$

and therefore it is enough to show that $(\mu+\kappa)\theta^{\alpha-1} \geq \mu\alpha+\kappa$. We define $a := 1+\kappa/\mu$ and $b := 1-\alpha \geq 0$ and then all we have to show is that $a\theta^{-b} \geq a-b$ if $\theta \leq e^{\frac{1}{a}}$. We have finished this proof if we can show that $a \geq (a-b)e^{\frac{b}{a}}$. But this is obvious since dividing by a on both sides and choosing $x := b/a$ we need $(1-x) \leq e^{-x}$ which is true. In the second case we have $\lambda e\log\theta > \mu\theta^{\frac{\kappa}{\mu}}$. We want to show that for $\alpha \in (0,1]$ we have $\lambda\theta^\alpha > \mu\alpha+\kappa$. We have

$$\lambda\theta^\alpha > \frac{\mu\theta^{\frac{\kappa}{\mu}+\alpha}}{e\log\theta}$$

and therefore we only have to show that

$$\frac{\mu\theta^{\frac{\kappa}{\mu}+\alpha}}{e\log\theta} \geq \mu\alpha+\kappa.$$

We define $a := \alpha + \kappa/\mu$ and then all we have to show is that $\theta^a \geq ae\log\theta$. We define $b := a\log\theta$ and so we need to show that $e^b \geq eb$ which is true for all b. In the third region we have $\log\theta > \mu/\kappa$ and $\lambda > \kappa$. We want to show that for $\alpha \in (0,1]$ we have $\lambda\theta^\alpha > \mu\alpha + \kappa$. We have $\lambda\theta^\alpha > \kappa\theta^\alpha$ and therefore we only have to show that $\theta^\alpha > (\mu/\kappa)\alpha + 1$. If we define $a := (\mu/\kappa)\alpha$ and use $\log\theta > \mu/\kappa$ we only have to show that $e^a \geq a+1$ which is true.

□

For the following lemmas we define for $\delta > 0$

$$\begin{aligned} g_1(j) &:= \frac{1}{1+\delta j} \\ g_2(j) &:= \frac{1}{1+\delta j^{\alpha(j)}} \end{aligned} \tag{3.1}$$

and

$$\alpha(j) := \begin{cases} 1 & \text{if } j \leq K; \\ 1 - (1-\alpha_*)\left(1 - \frac{\log\log K}{\log\log j}\right)^2 & \text{if } j > K, \end{cases}$$

where $0 < \alpha_* < 1/6$ and α_* is made smaller if necessary later on; in what follows, δ is always smaller than 1 and $K \geq e^{e^3}$, even if we do not mention it every time.

Lemma 3.3 [[B2], page 108] *$\alpha(x)$ and g_2 have the following properties:*

a) $\alpha(x)\log(x)$ increases with x.

b) $\alpha(x)$ decreases with x.

c) $g_2(x)$ decreases with x.

d) For $x \geq K$,

$$0 \leq -\alpha'(x) \leq \frac{2}{x\log x\log\log K}.$$

e) For $c, x > 1$,

$$1 \geq x^{\alpha(cx)-\alpha(x)} \geq 1 - \frac{2(c-1)}{\log\log K}.$$

f) There exists a constant $k > 2$ such that

$$g_2''(x) \leq k\delta x^{\alpha(x)-2},$$

uniformly in $x > 0$, $\delta \leq 1$ and $K \geq e^{e^3}$.

Lemma 3.4 *a) For $j \geq 0$ the following inequality holds:*

$$1 - \mathbb{E}[g_1(S_j)] \geq \frac{\delta j\theta}{1+\delta j\theta}\left\{1 - \frac{\delta\sigma^2}{\theta(1+\delta j\theta)}\right\}.$$

b) For $j\theta \leq K$, k as in Lemma 3.3 f) and $\delta \leq k/(2K)$ we have

$$1 - \mathbb{E}[g_2(S_j)] \geq \frac{\delta j\theta}{1+\delta j\theta}\left\{1 - \frac{k^2\sigma^2}{\theta K}\right\}.$$

c) For $\delta(j\theta)^{\alpha(j\theta)} \leq 1$, k as in Lemma 3.3 f) and $s(k)$ a constant such that $s(k)k \geq 8$ and $\left(1 - \sqrt{2/s(k)k}\right)^2 \geq 3/s(k)$ we have

$$1 - \mathbb{E}[g_2(S_j)] \geq \frac{\delta(j\theta)^{\alpha(j\theta)}}{1+\delta(j\theta)^{\alpha(j\theta)}}\left\{1 - \frac{ks(k)\sigma^2}{\theta^2 j}\right\}.$$

d) Suppose δ is chosen so small that, if j satisfies $\delta(j\theta)^{\alpha(j\theta)} > 1$, then $\alpha(j) \leq 2\alpha_* < 1/3$ must be satisfied too (see the definition of $\alpha(j)$ for a definition of α_*). Then, for j such that $\delta(j\theta)^{\alpha(j\theta)} > 1$ is satisfied we have

$$1 - \mathbb{E}[g_2(S_j)] \geq \frac{\delta(j\theta)^{\alpha(j\theta)}}{1+\delta(j\theta)^{\alpha(j\theta)}}\left\{1 - O(j^{-2/3})\right\}.$$

Remark Lemma 3.4 allows us to replace $\mathbb{E}[g(S_j)]$ by $g(j\theta)$ with only small impact.

Proof of Lemma 3.4 See [B2], pages 107 ff.). There was a printing error in b) (k instead of k^2) and in d) it is possible to have $O(j^{-2/3})$ instead of $O(j^{-1/3})$ following exactly the same line of proof.

□

Define $M_\beta(t) := \sum_{j\geq 1} j^\beta X_j(t)$;

$$c_\beta(X) := \sum_{j\geq 1} j\mu X_j\{(j-1)^\beta - j^\beta\} + \lambda \sum_{k\geq 1}\sum_{j\geq 1} X_j p_{jk} k^\beta - \kappa \sum_{j\geq 1} j^\beta X_j$$

and

$$W_\beta(t) := M_\beta(t) - M_\beta(0) - \int_0^t c_\beta(X(u))du.$$

Lemma 3.5 *For $0 < \beta \leq 1$ and with the notation above, $W_\beta(t)$ is a $\mathcal{G}_t$-martingale where $\mathcal{G}_s := \sigma\{X(u), 0 \leq u \leq s\}$, whatever the value of $\kappa \geq 0$.*

Proof of Lemma 3.5 We can apply Theorem 2 in Hamza and Klebaner (1995) [HK], version for general state space. Choose $f(z) := \sum_{j \geq 1} j^\beta z_j$ and $c := (2\mu + \lambda\theta + \kappa)$, then $|L|f(z) \leq c(1 \vee |f(z)|)$ is satisfied, where L is the infinitesimal generator of the Markov process X.

□

Proof of Theorem 2.1 In part A) we prove extinction in all three cases 1) to 3) if the relevant $R_i \leq 1$. In part B) we prove that there is a positive probability that the epidemic develops in all three cases 1) to 3) if the relevant $R_i > 1$. In part C) we prove the fourth result.

A) In the first part of A) we assume that $R_0 \leq 1$. For $\beta = 1$ (we suppress the "1" in the next few steps) we can argue as follows ($W(0) = 0$):

$$M(t) = W(t) + M(0) + \int_0^t c(X(u))du. \tag{3.2}$$

Because W is a martingale, we therefore have for $0 < s < t$:

$$\mathbb{E}[M(t)|\mathcal{G}_s] = W(s) + M(0) + \mathbb{E}[\int_0^t c(X(u))du|\mathcal{G}_s],$$

and so finally by using again the definition of $W(s)$:

$$\mathbb{E}[M(t)|\mathcal{G}_s] = M(s) + \mathbb{E}[\int_s^t c(X(u))du|\mathcal{G}_s]. \tag{3.3}$$

But $c(X(u)) = (\lambda\theta - \mu - \kappa)M(u)$, and so we can derive

$$\mathbb{E}[M(t)|\mathcal{G}_s] = M(s) + \int_s^t (\lambda\theta - \mu - \kappa)\mathbb{E}[M(u)|\mathcal{G}_s]du.$$

So $\mathbb{E}[M(t)|\mathcal{G}_s] \leq M(s)$ for $0 < s < t$ if $R_0 \leq 1$ which means that M is a nonnegative supermartingale.

Now we observe that each $X \in \mathbb{N}^\infty \backslash \{0\}^\infty$ is transient. The communication structure of a Markov process divides the set of states into equivalence-classes. If a class is not closed, it is automatically transient. Here the set $\mathbb{N}^\infty \backslash \{0\}^\infty$ is an equivalence-class and is not closed (one can leave it by going to $\{0\}^\infty$, which is a separate absorbing class), and so each $X \in \mathbb{N}^\infty \backslash \{0\}^\infty$ is transient. But for each K the set $\{X \in \mathbb{N}^\infty \backslash \{0\}^\infty | \sum_{j\geq 1} jX_j \leq K\}$ is finite and transient, and hence is only visited finitely often a.s.. Hence it follows that $\lim_{t\to\infty} \sum_{j\geq 1} jX_j(t) = \lim_{t\to\infty} M(t)$ is almost surely either 0 or ∞.

Now, by the nonnegative (super)-martingale convergence theorem (see Revuz and Yor (1991) [RY], Corollary 2.11, § 2, Chapter II for example), we can conclude that M converges almost surely towards an a.s finite random variable which therefore must be 0, implying $\mathbb{P}[\lim_{t\to\infty} \sum_{j\geq 1} X_j(t) = 0] = 1$ if $R_0 \leq 1$ no matter what value θ has. This finishes the first direction ($R_0 \leq 1$) of the proof of 1) and those situations of 2) and 3) where $R_0 \leq 1$.

In the second part of A) we can therefore assume that $R_0 > 1$. We start with equation (3.3). Now β becomes vital for the proof and the reader can easily check that for any $\beta \in (0, 1]$ the calculations run through until equation (3.3). So we have

$$\mathbb{E}[M_\beta(t)|\mathcal{G}_s] = M_\beta(s) + \mathbb{E}[\int_s^t c_\beta(X(u))du|\mathcal{G}_s].$$

Now we prove that for each $\beta \in (0, 1]$ we have $c_\beta(X) \leq (\lambda\theta^\beta - \beta\mu - \kappa)M_\beta$. This goes as follows:

The function $f(y) := y^\beta$ is concave if $\beta \in [0, 1]$. So for y_1, y_2 we have

$$f(y_1) \leq f(y_2) + f'(y_2)(y_1 - y_2).$$

If we choose $y_1 = j - 1,\ y_2 = j$ we therefore get

$$\{(j-1)^\beta - j^\beta\} \leq -\beta j^{\beta-1},$$

and so we can derive

$$\sum_{j\geq 1} j\mu X_j\{(j-1)^\beta - j^\beta\} \leq \mu \sum_{j\geq 1} jX_j(-\beta j^{\beta-1}) \leq -\mu\beta \sum_{j\geq 1} j^\beta X_j.$$

Using Jensen's inequality for concave functions we have $\sum_{l\geq 0} p_{jl} l^\beta \leq (j\theta)^\beta$. So

$$\lambda \sum_{l\geq 1}\sum_{j\geq 1} X_j p_{jl} l^\beta = \lambda \sum_{j\geq 1} X_j \sum_{l\geq 1} p_{jl} l^\beta \leq \lambda\theta^\beta \sum_{j\leq 1} j^\beta X_j,$$

and so looking at the definition of c_β we can conclude

$$c_\beta(X) \le (\lambda\theta^\beta - \beta\mu - \kappa)\sum_{j\ge 1} j^\beta X_j.$$

We are free to choose $\beta \in (0,1)$. We want to argue just as we did in the first part of A) mutatis mutandis, for which it is enough to show that $(\lambda\theta^\beta - \mu\beta - \kappa) \le 0$ under the constraints of the theorem in cases 2) and 3) for suitably chosen β. Once accomplished, the proof of part A) is complete.

For case 2) we choose $\beta = \beta_0 := (1/\log\theta)\log\big(\mu/(\lambda\log\theta)\big)$. Elementary computations show that as $R_0 > 1$, $R_1 \le 1$ and $(1+\kappa/\mu)^{-1} < \log\theta \le \mu/\kappa$, we have $\beta_0 \in (0,1)$ and $\lambda\theta^{\beta_0} - \beta_0\mu - \kappa < 0$. So this ends the proof of the first direction ($R_1 \le 1$) of 2).

Case 3) is even simpler: $\mu/\kappa < \log\theta$ and therefore $\theta > 1$. Besides that we have $\lambda < \kappa$. We have to find a $\beta \in (0,1)$ such that $\lambda\theta^\beta - \beta\mu - \kappa < 0$. But this is clear ($\beta \to 0$ finally makes it). This ends the proof of the first direction ($R_2 \le 1$) of 3).

B) This proof consists of three parts. In part one (B1)) we derive the general strategy; in B2) we treat the case where $\theta \le 1$, and in B3) we treat the remaining case ($\theta > 1$).

B1) We think in terms of a discrete generation branching process with types $j = 1, 2, \ldots$. At each generation, each individual dies, an individual of type j being replaced either by one of type $j-1$ (death of a parasite) with probability $j\mu/(\lambda + j\mu + \kappa)$, or by one of type j and another of type k (infection) with probability $\lambda p_{jk}/(\lambda + j\mu + \kappa)$, or not replaced at all (death of an individual) with probability $\kappa/(\lambda + j\mu + \kappa)$ and type 0 individuals are not counted.

Then, if

$$q^{(n)}(j) := \mathbb{P}[\text{ extinction by generation } n \,|X(0) = e_j],$$

consideration of the first generation shows that $q^{(n+1)} = Tq^{(n)}$, where we have $(Tf)(0) = 1$ and

$$\begin{aligned}(Tf)(j) =& [j\mu/(\lambda + j\mu + \kappa)]f(j-1) + [\lambda/(\lambda + j\mu + \kappa)]f(j)\mathbb{E}[f(S_j)] \\ &+ [\kappa/(\lambda + j\mu + \kappa)],\ j \ge 1,\end{aligned}$$

where S_j is defined just before equation (2.1). Clearly, $q^{(0)}(0) = 1$ and $q^{(0)}(j) = 0$ for $j \geq 1$, and

$$q^{(n)}(j) \uparrow q(j) := \mathbb{P}[\text{ eventual extinction } | X(0) = e_j].$$

We wish to show that $q(j) < 1$ for $j \geq 1$ under the conditions stated in the theorem.

First observe that, if $f \geq h$ in the sense that $f(j) \geq h(j)$ for all $j \geq 0$, then $T^n f \geq T^n h$ for all $n \geq 1$ also. Hence, if we can find any f such that $f \geq q^{(0)}$ and $Tf \leq f$, it follows that $f \geq q$ also. If, in addition, $f(j) < 1$ for all $j \geq 1$, the same must be true of q. The remainder of the proof consists of finding a suitable function f.

But rather that looking for such an f directly, we look for a transformation of f. The heuristic idea is as that, for j very large, the probability $q(j)$ must be approximately κ/λ. That is, if we start with only one infected individual having a huge parasite burden, all infected individuals in the initial stages have large parasite burdens, and the only way that they then lose infectiousness is through death, since it takes much too long for the parasites to all die. Then the initial stages are well described by a pure birth and death process with birth rate λ and death rate κ, for which the probability of extinction is κ/λ. Lemma 3.2 a) guarantees us that this ratio is always smaller than 1 (in those cases relevant to us in part B) of the proof). So we expect that

$$\lim_{j \to \infty} q(j) = \frac{\kappa}{\lambda}.$$

For smaller values of j we expect values for $q(j)$ which are larger, because there are initially fewer parasites in the process, and for $j = 0$ we must even have $q(0) = 1$. We look for an f which is almost 1 if j is small and then decreases to the final limit κ/λ as j tends to infinity. So define

$$f(j) := (1 - \frac{\kappa}{\lambda})g(j) + \frac{\kappa}{\lambda}$$

and look for a g such that $g(0) = 1$ and $g(j)$ for $j \geq 1$ decreases slowly to 0.

What constraints must g satisfy in order that f should satisfy the conditions we asked for above? Let T operate on f successively, and define $f^{(n)} := T^n f$; set $f^{(n)} =$

$(1-\kappa/\lambda)g^{(n)}+\kappa/\lambda$. Then $g^{(n)}=\widetilde{T}^n g$, where

$$\begin{aligned}\widetilde{T}g(j)=&\frac{j\mu}{\lambda+j\mu+\kappa}g(j-1)+\frac{\kappa}{\lambda+j\mu+\kappa}g(j)\\&+\frac{\kappa}{\lambda+j\mu+\kappa}\mathbb{E}[g(S_j)]+\frac{\lambda-\kappa}{\lambda+j\mu+\kappa}g(j)\mathbb{E}[g(S_j)],\ j\geq 1.\end{aligned}$$

We must be sure that if we find a g such that for all $j\geq 1$ the three conditions

$$g(0)=1;\quad g(j)<1;\ \text{ and }\ \widetilde{T}g\leq g$$

are satisfied, then the corresponding conditions are true for f. The first two conditions are clearly satisfied: $f(0)=1$ and $f(j)<1$ for $j\geq 1$. The third condition is satisfied because

$$Tf=(1-\kappa/\lambda)\widetilde{T}g+\kappa/\lambda\leq(1-\kappa/\lambda)g+\kappa/\lambda=f.$$

As a conclusion of part B1) of the proof we now see that we have to find a (nonnegative) g such that for all $j\geq 1$ the following conditions

$$g(0)=1;\quad g(j)<1;\ \text{ and }\ \widetilde{T}g\leq g$$

are satisfied. The third condition can be explicitly rewritten as follows:

$$j\mu\big(g(j-1)-g(j)\big)+\kappa\big(1-g(j)\big)\leq\big(1-\mathbb{E}[g(S_j)]\big)\big(\kappa-g(j)\kappa+\lambda g(j)\big),\tag{3.4}$$

and if we talk about a g satisfying condition (3.4), we mean that g satisfies $g(0)=1$ and $g(j)<1$ for $j\geq 1$ too.

The computations that follow in B2) and B3) are awkward because we want to replace the expression $\mathbb{E}[g(S_j)]$ in (3.4) by $g(\theta j)$. This is justified up to a small error, but we therefore have to keep the error under control.

B2) In this part of the proof we suppose that $\theta\leq 1$. We now have to find a (nonnegative) g such that condition (3.4) is satisfied. We try $g_1(j):=(1+\delta j)^{-1}$, as defined in (3.1), for $\delta>0$ to be chosen later. With this choice of g and using Lemma 3.4 a) we see that (3.4) is satisfied if

$$\frac{\mu}{1+\delta(j-1)}+\kappa\leq\frac{\theta}{1+\delta j\theta}\left\{1-\frac{\delta\sigma^2}{\theta(1+\delta j\theta)}\right\}(\kappa\delta j+\lambda)$$

is satisfied. This equation is equivalent to

$$\mu\frac{1+\delta j\theta}{1+\delta(j-1)}\left\{1-\frac{\delta\sigma^2}{\theta(1+\delta j\theta)}\right\}^{-1} + \kappa\left\{(1+\delta j\theta)\left\{1-\frac{\delta\sigma^2}{\theta(1+\delta j\theta)}\right\}^{-1} - \theta\delta j\right\} \leq \lambda\theta.$$

As $R_0 > 1$ we can define $c := \lambda\theta - \mu - \kappa \ > 0$. Then the above inequality is equivalent to

$$\mu\frac{1+\delta j\theta}{1+\delta(j-1)}\left\{1-\frac{\delta\sigma^2}{\theta(1+\delta j\theta)}\right\}^{-1} - \mu + \kappa\left\{(1+\delta j\theta)\left\{1-\frac{\delta\sigma^2}{\theta(1+\delta j\theta)}\right\}^{-1} - \theta\delta j\right\} - \kappa \leq c$$

which is in turn equivalent to

$$\begin{aligned}&\mu\delta\frac{(\theta+\sigma^2-\delta\sigma^2)+j(\theta^2-\theta+\sigma^2\delta+\delta\theta^2)+j^2(\delta\theta^3-\delta\theta^2)}{(\theta-\delta\sigma^2-\delta\theta+\delta^2\sigma^2)+j(\delta\theta^2+\delta\theta-\sigma^2\delta^2-\delta^2\theta^2)+j^2\theta^2\delta^2}\\&+\kappa\delta\frac{\sigma^2+j\delta\theta\sigma^2}{\theta+\delta j\theta^2-\delta\sigma^2} \leq c.\end{aligned} \tag{3.5}$$

We now examine the first term of the left side of (3.5). As $\theta \leq 1$ we have $(\delta\theta^3 - \delta\theta^2) \leq 0$ (third term in the numerator). Now we choose $\delta < \min\big((\theta-\theta^2)/(\theta^2+\sigma^2), \theta/(\sigma^2+\theta)\big)$. With this choice, $\theta^2 - \theta + \sigma^2\delta + \delta\theta^2$ (second term in the numerator) is smaller than or equal to 0 and each term in the denominator is positive for all $j \geq 1$. So the first term of the left side of (3.5) is smaller than or equal to

$$\mu\delta\frac{\theta+\sigma^2-\delta\sigma^2}{\theta-\delta\sigma^2-\delta\theta+\delta^2\sigma^2}.$$

This term does not depend on j and so it is easily seen that δ can be made so small that the following inequality is satisfied

$$\mu\delta\frac{\theta+\sigma^2-\delta\sigma^2}{\theta-\delta\sigma^2-\delta\theta+\delta^2\sigma^2} < \frac{c}{2}.$$

Proceeding to the second part, choosing $\delta \leq \theta/2\sigma^2$ we have

$$\kappa\delta\frac{\sigma^2+j\delta\theta\sigma^2}{\theta+\delta j\theta^2-\delta\sigma^2} \leq 2\kappa\delta\frac{\sigma^2(1+2j\delta\theta)}{\theta(1+2j\delta\theta)} \leq \frac{c}{2}$$

for all $j \geq 1$ if we choose $\delta \leq c\theta/4\kappa\sigma^2$.

Combined, (3.5) is satisfied for all $j \geq 1$ which ends the proof of part B2).

B3) In this part of the proof we suppose that $\theta > 1$. Again, we have to find a (nonnegative) g such that condition (3.4) is satisfied.

In this part we cannot choose the simple function g_1, as before, because (3.4) is not satisfied for all j no matter how we choose δ. Instead we choose g_2 (see (3.1) for a definition of g_1 and g_2).

The construction of g_2 with an $\alpha(j)$ as exponent in a term of the denominator leads to a g with the same decay as g_1 as long as $j \leq K$ and then the decay is smaller. Heuristically spoken g_2 is (in comparison to g_1) somehow "lifted" over a critical region until it finally decays to 0 at a much slower rate than g_1. But the reader should be aware of the fact that for all $j \geq 0$ we nevertheless have $g_2(j) < g_2(j-1)$, as shown in Lemma 3.3.

With this choice of g we see that (3.4) is satisfied if

$$j\mu\left[\frac{g_2(j-1)}{g_2(j)} - 1\right] + \kappa\delta j^{\alpha(j)} \leq (1 - \mathbb{E}[g(S_j)])(\kappa\delta j^{\alpha(j)} + \lambda) \tag{3.6}$$

is satisfied. Again, if we talk about a g satisfying condition (3.6), we mean that g satisfies $g(0) = 1$ and $g(j) < 1$ for $j \geq 1$ too.

We introduce three regions for j and so B3) consists of 3 parts itself:

B3.1) Here we presume that $1 \leq j \leq K/\theta$. Then as $\theta > 1$ we are in a region where g_2 and g_1 are identical ($\alpha(j) = 1$) and so we have

$$\frac{g_2(j-1)}{g_2(j)} - 1 \leq \delta;$$

using Lemma 3.4 b), it is enough to show that

$$\mu + \kappa \leq \frac{\theta}{1 + \delta j\theta}\left\{1 - \frac{k^2\sigma^2}{\theta K}\right\}(\kappa\delta j + \lambda) \tag{3.7}$$

for (3.6) to be satisfied. Until now, we need $\delta < \delta_1 := \min\big(1\ ,\ k/(2K)\big)$. In all three regions we have $R_0 > 1$ and so we can define $c := \lambda\theta - \mu - \kappa > 0$. (3.7) is then equivalent to

$$\mu\delta j\theta + \frac{\kappa\sigma^2 k^2\delta j}{K} + \frac{\lambda\sigma^2 k^2}{K} \leq c. \tag{3.8}$$

With the choices $\delta < \delta_2 := \min\big(\delta_1 \ , \ c/(3K\mu) \ , \ (c\theta)/(3\kappa k^2\sigma^2)\big)$ and $K \geq K_1 := \max\big((3k^2\sigma^2\lambda)/c \ , \ e^{e^3}\big)$ equation (3.8) is satisfied which ends the proof of B3.1).

B3.2) Here we presume that $K/\theta < j \leq J+1$, with $J := J(K)$ such that $\alpha(J) \leq 2\alpha_*$. Elementary calculations show that

$$\frac{g_2(j-1)}{g_2(j)} - 1 \leq \delta\big(j^{\alpha(j)} - (j-1)^{\alpha(j-1)}\big) \leq \delta\alpha(j-1)(j-1)^{\alpha(j-1)-1}. \tag{3.9}$$

We choose $\delta < \delta_3 := \min\big(\delta_2 \ , \ (KJ\theta)^{-1}\big)$. Then Lemma 3.4 c) can be applied. As $\delta < (KJ\theta)^{-1}$ we can incorporate the denominator $1+\delta(j\theta)^{\alpha(j\theta)}$ of the right side of Lemma 3.4 c) in the correction term $(1 - O(K^{-1}))$ which allows us to rewrite this lemma in the following way:

$$1 - \mathbb{E}[g_2(S_j)] \geq \delta(j\theta)^{\alpha(j\theta)}(1 - O(K^{-1})).$$

Together with (3.9) we see that (3.6) is satisfied if

$$j\mu\alpha(j-1)(j-1)^{\alpha(j-1)-1} + \kappa j^{\alpha(j)} \leq (j\theta)^{\alpha(j\theta)}(1 - O(K^{-1}))(\kappa\delta j^{\alpha(j)} + \lambda)$$

is satisfied. The term $\kappa\delta j^{\alpha(j)}$ on the right side is of order $O(K^{-1})$ and so we skip it, we do not need it. We therefore have to show that

$$j\mu\alpha(j-1)(j-1)^{\alpha(j-1)-1} + \kappa j^{\alpha(j)} \leq \lambda(j\theta)^{\alpha(j\theta)}(1 - O(K^{-1})) \tag{3.10}$$

is satisfied. If we can show that

$$\begin{aligned}\frac{\lambda(j\theta)^{\alpha(j\theta)}(1 - O(K^{-1}))}{j\mu\alpha(j-1)(j-1)^{\alpha(j-1)-1} + \kappa j^{\alpha(j)}} &\geq \frac{\lambda\theta^{\alpha(j)}}{\mu\alpha(j)+\kappa}(1 - O((\log\log K)^{-1})) \\ &\geq 1,\end{aligned} \tag{3.11}$$

then (3.10) is satisfied. The last inequality of (3.11) is surely true by Lemma 3.2 b) for all K large enough and so we can concentrate on the first inequality. The first inequality is true if we can show that the following two inequalities hold:

$$\alpha(j)(j\theta)^{\alpha(j\theta)}(1 - O(K^{-1})) \geq \theta^{\alpha(j)}(1 - O((\log\log K)^{-1}))j\alpha(j-1)(j-1)^{\alpha(j-1)-1} \tag{3.12}$$

and

$$(j\theta)^{\alpha(j\theta)}(1 - O(K^{-1})) \geq \theta^{\alpha(j)}(1 - O((\log\log K)^{-1}))j^{\alpha(j)}. \tag{3.13}$$

Equation (3.12) is satisfied because the following three relations (3.14), (3.15) and (3.16) hold. Because of Lemma 3.3 d), we have

$$\begin{aligned}\frac{\alpha(j)}{\alpha(j-1)} &= 1 - \big(\alpha(j-1) - \alpha(j)\big)\frac{1}{\alpha(j-1)} \\ &\geq 1 - \frac{2}{\alpha(j-1)(j-1)\log(j-1)\log\log K}.\end{aligned} \tag{3.14}$$

Then, again by Lemma 3.3 d), we have

$$\begin{aligned}\theta^{\alpha(j\theta)-\alpha(j)} &= \exp([\alpha(j\theta)-\alpha(j)]\log\theta) \\ &\geq 1 + [\alpha(j\theta)-\alpha(j)]\log\theta \geq 1 - \log\theta\frac{2j(\theta-1)}{j\log j\log\log K}.\end{aligned} \tag{3.15}$$

Finally, again by Lemma 3.3 d) we can derive

$$\begin{aligned}\frac{j^{\alpha(j\theta)-1}}{(j-1)^{\alpha(j-1)-1}} &\geq (j-1)^{\alpha(j\theta)-\alpha(j-1)} \\ &= \exp\big(\log(j-1)[\alpha(j\theta)-\alpha(j-1)]\big) \\ &\geq 1 + \log(j-1)[\alpha(j\theta)-\alpha(j-1)] \\ &\geq 1 - \log(j-1)\frac{2(1+j(\theta-1))}{(j-1)\log(j-1)\log\log K}.\end{aligned} \tag{3.16}$$

Therefore (3.12) is satisfied. Furthermore, (3.15) and Lemma 3.3 e) show immediately that (3.13) is satisfied, which finishes the proof of B3.2)

B3.3) Finally we presume that $j > J+1$. By looking at the derivative of $j^{\alpha(j)}$ and using Lemma 3.3 b) we immediately gain

$$\frac{g_2(j-1)}{g_2(j)} - 1 \leq \left(\frac{\delta(j-1)^{\alpha(j-1)}}{1+\delta(j-1)^{\alpha(j-1)}}\right)\frac{\alpha(j-1)}{(j-1)}.$$

For $j > J+1$, we first have $\delta(j\theta)^{\alpha(j\theta)} \leq 1$ and then we get into the area where $\delta(j\theta)^{\alpha(j\theta)} > 1$. But the inequality of Lemma 3.4 d) is weaker than the inequality of Lemma 3.4 c). So, after making δ even smaller if necessary, we may use

$$1 - \mathbb{E}[g_2(S_j)] \geq \frac{\delta(j\theta)^{\alpha(j\theta)}}{1+\delta(j\theta)^{\alpha(j\theta)}}\left\{1 - O(j^{-2/3})\right\}$$

during the whole part of B3.3). Again, for the last time we want inequality (3.6) to be satisfied. All we need to show is therefore that

$$\begin{aligned} &j\mu\left(\frac{(j-1)^{\alpha(j-1)}}{1+\delta(j-1)^{\alpha(j-1)}}\right)\frac{\alpha(j-1)}{(j-1)}+\kappa j^{\alpha(j)} \\ &\leq \frac{(j\theta)^{\alpha(j\theta)}}{1+\delta(j\theta)^{\alpha(j\theta)}}\left\{1-O(j^{-2/3})\right\}(\kappa\delta j^{\alpha(j)}+\lambda). \end{aligned} \tag{3.17}$$

We want to get rid of the denominators: Equation (3.17) is equivalent to the following long expression:

$$\begin{aligned} &j\mu(j-1)^{\alpha(j-1)}\alpha(j-1)+j\mu(j-1)^{\alpha(j-1)}\alpha(j-1)\delta j^{\alpha(j\theta)}\theta^{\alpha(j\theta)} \\ &+\kappa j^{\alpha(j)}(j-1)+\kappa j^{\alpha(j)}\delta(j-1)^{\alpha(j-1)+1}+\kappa j^{\alpha(j)+\alpha(j\theta)}\delta\theta^{\alpha(j\theta)}(j-1) \\ &+\kappa j^{\alpha(j)+\alpha(j\theta)}(j-1)^{\alpha(j-1)+1}\delta^2\theta^{\alpha(j\theta)} \\ &\leq \big(1-O(j^{-2/3})\big)\Bigg(\kappa j^{\alpha(j)+\alpha(j\theta)}\delta\theta^{\alpha(j\theta)}(j-1)+(j-1)j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}\lambda \\ &+\kappa j^{\alpha(j)+\alpha(j\theta)}(j-1)^{\alpha(j-1)+1}\delta^2\theta^{\alpha(j\theta)}+j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}\lambda(j-1)^{\alpha(j-1)+1}\delta\Bigg). \end{aligned}$$

This is equivalent to

$$\begin{aligned} &j\mu(j-1)^{\alpha(j-1)}\alpha(j-1)+j\mu(j-1)^{\alpha(j-1)}\alpha(j-1)\delta j^{\alpha(j\theta)}\theta^{\alpha(j\theta)} \\ &+\kappa j^{\alpha(j)}(j-1)+\kappa j^{\alpha(j)}\delta(j-1)^{\alpha(j-1)+1} \\ &\leq \big(1-O(j^{-2/3})\big)\Big((j-1)j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}\lambda+j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}\lambda(j-1)^{\alpha(j-1)+1}\delta\Big) \\ &-O(j^{-2/3})\Big(\kappa j^{\alpha(j)+\alpha(j\theta)}\delta\theta^{\alpha(j\theta)}(j-1)+\kappa j^{\alpha(j)+\alpha(j\theta)}(j-1)^{\alpha(j-1)+1}\delta^2\theta^{\alpha(j\theta)}\Big). \end{aligned} \tag{3.18}$$

This inequality is satisfied if the following two inequalities are satisfied:

$$\begin{aligned} &j\mu\alpha(j-1)j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}+\kappa j^{\alpha(j)}(j-1) \\ &\qquad\leq \big(1-O(j^{-2/3})\big)\big(j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}\lambda(j-1)\big) \\ &\qquad\quad -O(j^{-2/3})\kappa j^{\alpha(j)+\alpha(j\theta)}(j-1)\delta\theta^{\alpha(j\theta)}, \end{aligned} \tag{3.19}$$

(we have divided by $\delta(j-1)^{\alpha(j-1)}$) and

$$\begin{aligned} &j\mu(j-1)^{\alpha(j-1)}\alpha(j-1)+\kappa j^{\alpha(j)}(j-1) \\ &\qquad\leq \big(1-O(j^{-2/3})\big)\big((j-1)j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}\lambda\big) \\ &\qquad\quad -O(j^{-2/3})\kappa j^{\alpha(j)+\alpha(j\theta)}\delta\theta^{\alpha(j\theta)}(j-1). \end{aligned} \tag{3.20}$$

The separation of inequality (3.18) is such that in inequality (3.19) we have all terms with a j to the power of "1 plus two α's" except in the last term where we have "1 plus three α's"; in inequality (3.20) we have all terms with a j to the power of "1 plus one α" except in the last term where we have "1 plus two α's".

We first show that (3.19) is satisfied. We divide inequality (3.19) by $j^{1+\alpha(j)}$. Then it is enough to show that the following inequality is satisfied:

$$\begin{aligned}
&\mu\alpha(j-1)\theta^{\alpha(j\theta)}+\kappa \\
&\qquad\qquad \leq \big(1-O(j^{-2/3})\big)\big(j^{\alpha(j\theta)-\alpha(j)}\theta^{\alpha(j\theta)}\lambda(1-O(j^{-1}))\big) \\
&\qquad\qquad\quad -O(j^{-2/3})\kappa j^{\alpha(j\theta)}\delta\theta^{\alpha(j\theta)}.
\end{aligned}$$

We can apply Lemma 3.3 e) to the right hand side, showing that it is enough to have

$$\begin{aligned}
&\mu\alpha(j-1)\theta^{\alpha(j\theta)}+\kappa \\
&\qquad\qquad \leq \big(1-O(j^{-2/3})\big)\big((1-O(1/\log\log K))\theta^{\alpha(j\theta)}\lambda(1-O(j^{-1}))\big) \\
&\qquad\qquad\quad -O(j^{-2/3})\kappa j^{\alpha(j\theta)}\delta\theta^{\alpha(j\theta)}.
\end{aligned}$$

As $\alpha(J)\leq 2\alpha_*<1/3$, the last term tends to 0. On the other hand, we have $\lambda>\kappa$. So, up to asymptotics in j, we only need to ensure that

$$\mu\alpha(j-1)\theta^{\alpha(j\theta)}+\kappa<\lambda\theta^{\alpha(j\theta)}$$

for $j>J+1$. As $\theta>1$, we only have to make α_* small enough; then the inequality above is satisfied, and hence (3.19) is satisfied also.

We now have to show that (3.20) is satisfied too. But (3.20) is almost the same as (3.19); it is enough to show that, for large j, we have

$$(j-1)^{\alpha(j-1)}\leq j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}.$$

We have

$$\begin{aligned}
j^{\alpha(j-1)-\alpha(j\theta)} &= \exp(\log j[\alpha(j-1)-\alpha(j\theta)]) \\
&\leq \exp\Big(\frac{2\log j(1+j(\theta-1))}{(j-1)\log(j-1)\log\log K}\Big)
\end{aligned}$$

which is near 1 for K large and is therefore finally smaller than $\theta^{\alpha(j\theta)}$. This shows that (3.20) is satisfied too. This ends the proof of B3.3) and therefore the proof of part B).

C) Observe that the following part runs through with $\kappa = 0$ too. We can use equation (3.2) ($\beta = 1$) and take the expectation, giving

$$\mathbb{E}[M(t)] = M(0) + \int_0^t \mathbb{E}[c(x(u))]du.$$

As $c(X(u)) = (\lambda\theta - \mu - \kappa)M(u)$ we have the integral equation

$$y(t) = M(0) + \int_0^t (\lambda\theta - \mu - \kappa)y(u)du$$

where $y(t) = \mathbb{E}[M(t)]$. But this immediately leads to (2.4) completing the proof of Theorem 2.1.

□

Proof of Theorem 2.2 The proof where $\kappa = 0$ was made in [BK, Theorem 2.3]. We may therefore assume that $\kappa > 0$.

First we find a lower bound for the probability that the epidemic dies out in an arbitrary, single time-interval of length 1, given that it has not died out yet. The probability that a given person dies in the next time interval and that the new-born does not have any infectious contacts at all in this interval is at least $(1 - e^{-\kappa})e^{-\lambda} > 0$. The probability that this happens to all M individuals in the same time-interval is at least $[(1 - e^{-\kappa})e^{-\lambda}]^M$. So the probability that the infection dies out in the next time-interval (given that it has not died out before) is at least $p_M := [(1 - e^{-\kappa})e^{-\lambda}]^M$. There are other ways that it can die out too, but we already have enough.

Let B_n be the event that the epidemic dies out in the time-interval $[0, n+1)$ for $n \geq 0$. Let us define the set $A := \{\lim_{t\to\infty} x(t) = e_0\} = \cup_{i\geq 0} B_i$. We have $B_i \subseteq B_{i+1}$. Let us look at $\mathbb{P}[B_n^c]$. We have to prove that $\mathbb{P}[B_n^c]$ converges to 0 as $n \to \infty$ to show the first part of Theorem 2.2. We have

$$\begin{aligned}\mathbb{P}[B_n^c] &= \mathbb{P}[B_n^c | B_{n-1}]\mathbb{P}[B_{n-1}] + \mathbb{P}[B_n^c | B_{n-1}^c]\mathbb{P}[B_{n-1}^c] \\ &= \mathbb{P}[B_n^c | B_{n-1}^c]\mathbb{P}[B_{n-1}^c] \leq (1 - p_M)\mathbb{P}[B_{n-1}^c].\end{aligned}$$

As a consequence, $\mathbb{P}[B_n^c] \leq (1 - p_M)^n \to 0$ as $n \to \infty$, completing the proof.

□

Remark to Theorem 2.2 As a particular consequence of Theorem 2.2, the process N is 'regular'.

Proof of Theorem 2.3 The idea of the proof is to show that for fixed M there exists a linear process X/M which is *in all components* larger than our original $x^{(M)}$, and such that, the larger we choose M, the more $x^{(M)}$ behaves like X/M. Then we use Theorem 2.1. Note that the proof works with $\kappa = 0$ too.

1. First we have to find that linear process X. For this we define a trivariate Markov process $(X^{(nl)}(t), X^{(r)}(t), R'(t))$. "$nl$" stands for non-linear, "r" stands for residual and the meaning of R' is explained later. In fact, each of the components in $(X^{(nl)}, X^{(r)})$ are themselves infinite dimensional: the first component is an infinite vector $(X_j^{(nl)}(t))_{j\geq 0}$ and the second component is an infinite vector $(X_k^{(r)}(t))_{k\geq 1}$. We assume that $X_j^{(nl)}(t) \in \mathbb{N}_0$ and $X_k^{(r)}(t) \in \mathbb{N}_0$ for all t, j, k. We choose the initial values to be such that $X_0^{(nl)}(0) = M - Y$, $X_j^{(nl)}(0) = y_j$ for $j \geq 1$ and $X_k^{(r)}(0) = 0$ for $k \geq 1$. Our aim is to construct $X^{(nl)}$ and $X^{(r)}$ such that $X_j := X_j^{(nl)} + X_j^{(r)}$ behaves like L for $j \geq 1$. We define the univariate, random process $R'(t)$ to have values on the nonnegative integers and to have initial value $R'(0) = 0$. We let these processes develop according to the following rates:

$$\left(X^{(nl)}, X^{(r)}, R'\right) \to \left(X^{(nl)} + (e_{j-1} - e_j), X^{(r)}, R'\right)$$

at rate $j\mu X_j^{(nl)}$; $j \geq 1$, (death of a parasite in the non-linear process)

$$\left(X^{(nl)}, X^{(r)}, R'\right) \to \left(X^{(nl)} + (e_k - e_0), X^{(r)}, R'\right)$$

at rate $\lambda(X_0^{(nl)}/M) \sum_{u\geq 1} X_u^{(nl)} p_{uk}$; $k \geq 1$, (infection in the non-linear process)

$$\left(X^{(nl)}, X^{(r)}, R'\right) \to \left(X^{(nl)} + (e_0 - e_v), X^{(r)}, R'\right)$$

at rate $\kappa X_v^{(nl)}$; $v \geq 1$, (death of an individual in the non-linear process)

$$\left(X^{(nl)}, X^{(r)}, R'\right) \to \left(X^{(nl)}, X^{(r)} + (e_{j-1} - e_j), R'\right)$$

at rate $j\mu X_j^{(r)}$; $j \geq 2$, (death of a parasite in the residual process)

$$\left(X^{(nl)}, X^{(r)}, R'\right) \to \left(X^{(nl)}, X^{(r)} - e_1, R'\right)$$

at rate $\mu X_1^{(r)}$, (death of a parasite in the residual process when $j = 1$)

$$\big(X^{(nl)}, X^{(r)}, R'\big) \to \big(X^{(nl)}, X^{(r)} - e_v, R'\big)$$

at rate $\mu X_v^{(r)}$; $v \geq 1$, (death of an individual in the residual process). As can be seen, non of the above events change the state of R'.

Let us first motivate the rates to come. Define $R(u) := \sum_{j\geq 1} X_j^{(r)}(u)$, and $N(u) := \sum_{j\geq 1} X_j^{(nl)}(u)$. Then we define $\tau := \inf\{u : N(u) > a\}$ for a a (usually large) positive number to be chosen later. Our aim is to define a time-homogeneous Poisson process R' such that almost surely the following relation holds:

$$R'(u) \geq I[R(u) > 0]I[u < \tau]. \tag{3.21}$$

As we construct $X^{(r)}$ such that X develops according to L, we already know that the total rate at which infections take place in $X^{(r)}$ (and so in R) must be

$$\lambda \sum_{k\geq 1} \big(\sum_{j\geq 1} X_j^{(r)}(u) p_{jk} + (1 - X_0^{(nl)}(u)/M) \sum_{j\geq 1} X_j^{(nl)}(u) p_{jk}\big).$$

But in (3.21), the right side is 0 at time 0 and as long as $u < \tau$ increases to 1 as soon as a first infection takes place in $X^{(r)}$. This happens at rate

$$\lambda(1 - X_0^{(nl)}(u)/M) \sum_{k\geq 1} \sum_{j\geq 1} X_j^{(nl)}(u) p_{jk}$$

as until then $R = 0$. Let us have a closer look at this rate, as long as $u < \tau$:

$$\begin{aligned}\lambda(1 - X_0^{(nl)}(u)/M) \sum_{k\geq 1} \sum_{j\geq 1} X_j^{(nl)}(u) p_{jk} &\leq \lambda(1 - X_0^{(nl)}(u)/M) \sum_{j\geq 1} X_j^{(nl)}(u) \\ &\leq \lambda\big(1 - \frac{M-a}{M}\big)a = \lambda a^2/M\end{aligned}$$

So we define a time-homogeneous Poisson process R' of rate $\lambda a^2/M$ coupled to the development of R in the following way:

Define

$$\begin{aligned}b(u) := {} & a^2/M \\ & - \sum_{k\geq 1} \big(\sum_{j\geq 1} X_j^{(r)}(u) p_{jk} + (1 - X_0^{(nl)}(u)/M) \sum_{j\geq 1} X_j^{(nl)}(u) p_{jk}\big).\end{aligned}$$

Note that we have just shown that $b(u) \geq 0$ until the first infection takes place in the residual process and as long as $u < \tau$. Then, if $b(u) \geq 0$ we have the following rates

$$\big(X^{(nl)}, X^{(r)}, R'\big) \rightarrow \big(X^{(nl)}, X^{(r)} + e_k, R' + 1\big)$$

at rate

$$\lambda \sum_{l \geq 1} X_l^{(r)} p_{lk} + \lambda(1 - \frac{X_0^{(nl)}}{M}) \sum_{u \geq 1} X_u^{(nl)} p_{uk} \; ; \; k \geq 1,$$

this is an infection in the residual process. Additionally, we have the following changes

$$\big(X^{(nl)}, X^{(r)}, R'\big) \rightarrow \big(X^{(nl)}, X^{(r)}, R' + 1\big)$$

at rate

$$\lambda a^2/M - \sum_{k \geq 1} \big(\lambda \sum_{l \geq 1} X_l^{(r)} p_{lk} + \lambda(1 - \frac{X_0^{(nl)}}{M}) \sum_{u \geq 1} X_u^{(nl)} p_{uk}\big).$$

Now if $b < 0$, we have the following rates

$$\big(X^{(nl)}, X^{(r)}, R'\big) \rightarrow \big(X^{(nl)}, X^{(r)} + e_k, R'\big)$$

at rate

$$\lambda \sum_{l \geq 1} X_l^{(r)} p_{lk} + \lambda(1 - \frac{X_0^{(nl)}}{M}) \sum_{u \geq 1} X_u^{(nl)} p_{uk} \; ; \; k \geq 1,$$

this is again an infection in the residual process. Additionally, we have the following changes

$$\big(X^{(nl)}, X^{(r)}, R'\big) \rightarrow \big(X^{(nl)}, X^{(r)}, R' + 1\big)$$

at rate $\lambda a^2/M$. With this construction (3.21) holds almost surely for the following reasons: we showed that $b \geq 0$ until the first infection, R' increases too at the first infection but does not decrease any more, additionally, note that we look at $I_{\{R>0\}}$ and not R in (3.21). R' is a time-homogeneous Poisson process of rate $\lambda a^2/M$. The reader can easily check that $X^{(nl)}/M$ behaves according to N. Let us look at the sum $X_j := (X^{(nl)} + X^{(r)})_j$ for $j \geq 1$. The development of X is that of L and is the same for all M, as the rates involving M cancel. M also appears in the initial values, but there it only appears in the initial number of uninfected individuals; since X does not include the zero co-ordinate, it remains the same for all M.

2. We now have to examine the limit

$$\lim_{M\to\infty} \mathbb{P}\big[\sum_{j\geq 1} x_j^{(M)}(t) = 0\big].$$

For all fixed M we introduce the notation $L(u) := \sum_{j\geq 1} X_j(u)$, where we still have $N(u) := \sum_{j\geq 1} X_j^{(nl)}(u)$ and $R(u) := \sum_{j\geq 1} X_j^{(r)}(u)$.
Now we fix t and define $L := L(t), N := N(t)$ and $R := R(t)$. Note that while the distributions of $N(u)$ and $R(u)$ vary with M, the distribution of $L(u)$ is the same for all M. We have

$$\mathbb{P}\big[\sum_{j\geq 1} x_j^{(M)}(t) = 0\big] = \mathbb{P}\big[\sum_{j\geq 1} X_j^{(nl)}(t) = 0\big] = \mathbb{P}\big[N = 0\big]. \tag{3.22}$$

As $L = N + R$ we have

$$\begin{aligned}\mathbb{P}\big[N = 0\big] &= \mathbb{P}\big[L - R = 0\big]\\ &= \mathbb{P}\big[L - R = 0|R = 0\big]\mathbb{P}\big[R = 0\big] + \mathbb{P}\big[L - R = 0|R > 0\big]\mathbb{P}\big[R > 0\big]\\ &= \mathbb{P}\big[L = 0\big] + \mathbb{P}\big[L - R = 0|R > 0\big]\mathbb{P}\big[R > 0\big]. \end{aligned} \tag{3.23}$$

The last equality holds because if $L = 0$ then $R = 0$ too.

The next step is to show that $\mathbb{P}\big[R > 0\big]$ tends to 0 as M tends to infinity. Define a bivariate Markov process (X, B) such that X is the L process and behaves as before. Additionally we add a univariate random variable $B \geq 0$. The initial values are $X_j(0) = y_j$ for $j \geq 1$ and $B(0) = 0$ and let us recall that $Y := \sum_{j\geq 1} y_j$. The vector (X, B) changes according to the following rates:

$$\begin{aligned}
&\big(X, B\big) \to \big(X + (e_{j-1} - e_j), B\big) \text{ at rate } j\mu X_j \ ; \ j \geq 2,\\
&\big(X, B\big) \to \big(X - e_1, B + 1\big) \text{ at rate } \mu X_1 \ ; \ (j = 1),\\
&\big(X, B\big) \to \big(X + e_k, B\big) \text{ at rate } \lambda \sum_{u\geq 1} X_u p_{uk} \ ; k \geq 1,\\
&\big(X, B\big) \to \big(X - e_u, B + 1\big) \text{ at rate } \kappa X_u \ ; \ u \geq 1,\\
&\big(X, B\big) \to \big(X, B + 1\big) \text{ at rate } \lambda B + \lambda \sum_{u\geq 1} X_u p_{u0}.
\end{aligned}$$

As is easily seen, X is still our linear process constructed in step 1. B cancels almost surely every loss of an infected individual in the linear process X: an infected individual drops out of the system if a parasite dies in an individual with only one parasite and additionally B cancels infections with zero parasites in the linear process X through adding that rate in the fifth line of our rates. Hence, if we define $\tilde{L} := L + B$, then $\tilde{L}$ is almost surely a pure birth process of rate λ. If L increases, $\tilde{L}$ increases too, but $\tilde{L}$ does not decrease when L decreases; more, the growing part B of the sum $\tilde{L} = L + B$ contributes increasingly to the growth of $\tilde{L}$.

We can now argue as follows: for positive a, to be chosen later (the reader should think of a being much larger than Y), we have the following relations:

$$\mathbb{P}\big[N > a\big] \leq \mathbb{P}\big[\tilde{L} > a\big] \leq \frac{1}{a}\mathbb{E}\big[\tilde{L}\big] = \frac{1}{a} Y e^{\lambda t}.$$

If we choose a such that $a^{-1} Y e^{\lambda t} < \epsilon$, for an arbitrary $\epsilon > 0$, we can continue as follows: as $\tau := \inf\{u : N(u) > a\} \leq \infty$,

$$\begin{aligned}
\mathbb{P}\big[R > 0\big] &= \mathbb{P}\big[RI_{\{t<\tau\}} + RI_{\{t\geq\tau\}} > 0\big] \\
&\leq \mathbb{P}\big[RI_{\{t<\tau\}} > 0\big] + \mathbb{P}\big[RI_{\{t\geq\tau\}} > 0\big] \\
&\leq \mathbb{P}\big[RI_{\{t<\tau\}} > 0\big] + \mathbb{P}\big[I_{\{t\geq\tau\}} > 0\big] \\
&\leq \mathbb{P}\big[RI_{\{t<\tau\}} > 0\big] + \epsilon.
\end{aligned} \tag{3.24}$$

In the last inequality we used that N is dominated by $\tilde{L}$. We now have to show that $\mathbb{P}\big[RI_{\{t<\tau\}} > 0\big]$ tends to 0 as M tends to infinity. But by (3.21)

$$\mathbb{P}\big[RI_{\{t<\tau\}} > 0\big] = \mathbb{P}\big[I_{\{R>0\}} I_{\{t<\tau\}} > 0\big] \leq \mathbb{P}\big[R' > 0\big] = 1 - \exp(-t\lambda a^2/M),$$

as the probability that there is no event in the Poisson process until time t is $\exp(-t\lambda a^2/M)$. $\blacksquare$
So, letting M tend to infinity, we have in (3.24), as $\epsilon > 0$ was chosen arbitrarily, that $\lim_{M\to\infty} \mathbb{P}[R > 0] = 0$. Hence, from (3.22) and (3.23) we have

$$\lim_{M\to\infty} \mathbb{P}\big[\sum_{j\geq 1} x_j^{(M)}(t) = 0\big] = \mathbb{P}\big[L(t) = 0\big].$$

3. We now have to examine the expression

$$\lim_{t\to\infty} \mathbb{P}\big[L(t) = 0\big]$$

to finish the proof.

The first directions ($\log\theta \leq 1/(1+\kappa/\mu)^{-1}$ and $R_0 \leq 1$ or $1/(1+\kappa/\mu)^{-1} < \log\theta \leq \mu/\kappa$ and $R_1 \leq 1$ or $\log\theta > \mu/\kappa$ and $R_2 \leq 1$) follow immediately: we can use Theorem 2.1 because convergence to 0 a.s. implies convergence to 0 in probability (note that $\{L(t) = 0\} = \{L(t) > 1/2\}^c$).

The inverse directions ($\log\theta \leq 1/(1+\kappa/\mu)^{-1}$ and $R_0 > 1$ or $1/(1+\kappa/\mu)^{-1} < \log\theta \leq \mu/\kappa$ and $R_1 > 1$ or $\log\theta > \mu/\kappa$ and $R_2 > 1$) need the following reasoning: let us define the random process $I(t)$ in the following way:

$$I(t) := \begin{cases} 1 & \text{if } L(t) > 0 \\ 0 & \text{if } L(t) = 0. \end{cases}$$

As $I(t)(\omega)$ is a decreasing function in t for each ω, $\lim_{t\to\infty} I(t)$ exists a.s. and so we can define a.s. the limit-function I_∞ as follows:

$$I_\infty(\omega) := \lim_{t\to\infty} I(t)(\omega).$$

By Theorem 2.1 we have $\mathbb{P}\big[I_\infty = 0\big] =: q < 1$ under the above constraints. But as $I(t)$ is a decreasing function, we have $\mathbb{P}[I(t) = 0] \leq \mathbb{P}[I_\infty = 0] = q < 1$ completing the proof.

$\square$

Acknowledgements. C.J.L. wants to thank his PhD-advisor A.D. Barbour for his continued support and D. Mollison for his helpful comments.

References

[B1] Barbour, A.D. (1977). Breakpoint phenomena, and the initial stages of the transmission of bilharzia. *Bull. Int. Stat. Inst.*, **XLVII (2)**, 22-30

[BK] Barbour, A.D. and Kafetzaki, M. (1993). A host-parasite model yielding heterogeneous parasite loads. *J. Math. Biol.* **31** 157-176

[B2] Barbour, A.D. (1994). Threshold Phenomena in Epidemic Theory. In: F.P. Kelly (Ed.), *Probability, Statistics and Optimisation*, 101-116, Wiley, New York

[BM] Bradley, D.J. and McCullough, F.S. (1973). Egg output stability and the epidemiology of Schistosomiasis haematobium, Part II. An Analysis of the epidemiology of endemic S. haematobium. *Trans. Roy. Soc. Trop. Med. Hyg.*, **67**, 491-500

[HK] Hamza, K. and Klebaner, F.C. (1995). Conditions for integrability of Markov chains. *J. Appl. Prob.* **32** 541-547

[Ld] Luchsinger, C.J. (2001). Approximating the long term behaviour of a model for parasitic infection. *Preprint.*

[RY] Revuz, D. and Yor, M. (1991). *Continuous Martingales and Brownian Motion.* Springer, Berlin