THRESHOLDS AND INITIAL GROWTH RATES IN A MODEL OF PARASITIC INFECTION

BY A. D. BARBOUR, J. A. P. HEESTERBEEK AND C. J. LUCHSINGER

Universität Zürich, GLW–DLO Wageningen and Universität Zürich

We investigate the initial behavior of a deterministic model of parasitic infection, 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 this model, 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; nor need the growth rates of parasite numbers and numbers of infected hosts be the same. Similar phenomena are observed if real time is replaced by generation number, and the overlap of generations as time passes need not correspond to that found, for instance, in the supercritical linear birth and death process. The proofs involve martingale methods, applied to a Markov chain associated with the deterministic differential equation system.

1. Introduction. In many infections, the main mechanism of increase in the population of the causal agent in a host is replication within the host. Thus, for these infections, the infectivity of a currently infected host is largely independent of any new influx of parasites from outside. Examples of such infections are those caused by many viral, bacterial and protozoan agents. For this type of parasitic infection, there has long been an established theory to characterize invasion thresholds and growth rates; see Heesterbeek and Dietz (1996) for a recent review.

There is, however, a large class of infections in which multiplication of the causal agent within the host plays only a minor role, or indeed does not occur at all. For these diseases, the development of host infectivity is determined almost exclusively by repeated reinfection of the infected host. Examples are infections caused by helminths and other parasitic worms. Unfortunately, the characterization of growth rates and invasion thresholds for this class of infections is still an open problem.

Up to now, general definitions of threshold quantities have been borrowed from elsewhere. Typically, the threshold has been defined in terms of the mean number of adult female worms that are produced by one adult female worm during her entire life, in the absence of density-dependent constraints [see, e.g., Anderson and May (1991)], which should be greater than 1 for

infection to become established. There are a number of problems with this definition. One has been pointed out by MacDonald and Nåsell [see Heesterbeek and Dietz (1996)]: in the invasion limit, where the number of parasites converges to 0 in backward time, it can no longer be guaranteed that two parasites are present in the initial infected hosts, posing a conceptual problem for parasites with sexual reproduction in the definitive host.

There is a further problem with this definition. It makes no provision for the distribution among the available hosts of the new adult worms produced by a female. It could be the case that they all end up in just a few hosts, so that, even if the number of parasites increased, the number of newly infected hosts might grow much more slowly, or even not at all. In this paper, we give an example which shows that this situation can easily occur in models of parasitic infections, even when hosts mix homogeneously. The implication is that one cannot hope to arrive at a general characterization of invasion thresholds and growth rates for parasitic infections, either by looking only at infected hosts generating new infected hosts or by looking only at parasites generating new parasites.

Our example is of a parasitic infection in which it is possible for R_0, defined in the "usual" way as above, to exceed 1, and yet for the infection to be certain to die out. This was first shown in Barbour (1994) in a stochastic formulation of the model. Here we pursue the causes and consequences of the phenomenon, trying to elucidate why the behavior of this class of models runs counter to existing beliefs. We now work in terms of a deterministic version of the model, which has the advantage of showing that the unusual behavior is not just a product of a stochastic approach.

2. The model and its behavior. The model that we investigate was proposed by Barbour and Kafetzaki (1993) to describe the transmission of certain parasitic diseases. In the model, the infectivity of a definitive host is assumed to depend on his parasite burden, in such a way that his rate λ of making potentially infectious contacts always remains the same, but, at any given contact with an uninfected host, each of his parasites passes on a random number of "offspring" to the new host, drawn independently from a fixed distribution with mean θ and finite variance σ^2. This particular form of contact process serves as a simple model for parasites which are released in localized groups into the environment, and may then be ingested together into a new host. It was originally incorporated into a transmission model for schistosomiasis, where the real infection process, although somewhat of this general form, is more indirect, involving an aquatic snail as intermediate host. The model was nonetheless successful in generating at equilibrium the highly overdispersed distributions of parasites among definitive hosts which are characteristic of this disease. However, its threshold behavior was found to be somewhat unusual [Barbour (1994)]. Here, we consider the initial behavior of the model in more detail, with main emphasis on the rate of growth.

Since the initial development is our object of interest, we study the Whittle (Markov branching process) approximation to the initial stages of such an infection process. Apart from the infection mechanism described above, all that we need assume is that parasites have independent negative exponentially distributed lifetimes with mean $1/\mu$ and that the process is time homogeneous. For each $j \geq 1$, we let $X_j \in \mathbb{Z}_+$ denote the number of definitive hosts with j parasites, so that there are infinitely many "types" of hosts, one for each possible parasite burden: this specification is used to accommodate the different infectivity of hosts with different parasite burdens. The model is then specified by the (Markovian) transitions

(2.1)
$$\{X_j \to X_j - 1, X_{j-1} \to X_{j-1} + 1\} \quad \text{at rate } j\mu X_j, \, j \geq 2,$$
$$\{X_1 \to X_1 - 1\} \quad \text{at rate } \mu X_1,$$
$$\{X_k \to X_k + 1\} \quad \text{at rate } \lambda \sum_{j \geq 1} X_j p_{jk}, \, k \geq 1.$$

Here,

$$p_{jk} = \mathbb{P}\left[\sum_{l=1}^{j} Z_l = k\right],$$

and the $(Z_l, l \geq 1)$ are independent and identically distributed, with $\mathbb{P}[Z_1 = k] = p_{1k}$, so that

$$\theta = \sum_{k \geq 1} k p_{1k} \quad \text{and} \quad \sigma^2 = \sum_{k \geq 0} (k - \theta)^2 p_{1k} < \infty;$$

the common distribution of the Z_l is that of the number of "offspring" of a parasite at a single infectious contact. The first two transitons correspond to deaths of parasites, the third to infections: note that the latter is in a form consistent with the infected hosts making contacts independently of one another. The initial values $(X_j(0), j \geq 1)$ are taken to be such that $X_j(0) = 0$ for all but finitely many j, so that the initial number of infectives is finite, and this in turn implies that

(2.2)
$$\sum_{j \geq 1} X_j(t) < \infty \quad \text{and} \quad \sup_{0 \leq s \leq t} \sum_{j \geq 1} \mathbb{E}X_j(s) < \infty \quad \text{for all } t \geq 0.$$

In the usual time-homogeneous models of epidemics in populations of independently mixing individuals, the epidemic threshold theorem takes the form that a "large" outbreak is impossible if $R_0 < 1$ and possible (certain in deterministic formulations) if $R_0 > 1$, where R_0 is the basic reproduction number. Note that R_0 can typically be interpreted as the dominant eigenvalue of the mean matrix in the multitype Galton–Watson process, with the time step being one "generation" of infection in the original epidemic model, and the threshold theorem in these terms becomes the criticality theorem for branching processes [Athreya and Ney (1972), Theorem 2, page 186]. In our model, whatever the initial parasite burden of a newly infected host, each of his parasites has an average of $\lambda\theta/\mu$ offspring over its whole lifetime. This

suggests the definition of $R_0 = \lambda\theta/\mu$ as the basic reproduction number for the parasite population. The definition is supported by looking at transmission from the host's point of view; the $\mathbb{N} \times \mathbb{N}$ mean rates matrix associated with the Markov branching process (2.1) has a positive right eigenvector with eigenvalue R_0. However, it is shown in Barbour (1994), Theorem 2.1, that, with this definition, $R_0 = 1$ is only critical if $\theta \leq e$, and that for $\theta > e$ the threshold occurs when $\lambda e \log \theta/\mu = 1$. In the remainder of this section, we outline the behavior of the model in greater detail; the proofs are deferred to the later section.

We concentrate on the deterministic analog of (2.1), partly for simplicity, and partly to emphasize that the phenomena we describe do not arise purely because of a "stochastic" formulation of the model. The deterministic version is given by the monotone system of linear differential equations

$$(2.3) \qquad \frac{dx_j}{dt} = (j+1)\mu x_{j+1} - j\mu x_j + \lambda \sum_{l \geq 1} x_l p_{lj}, \qquad j \geq 1.$$

where $x_j(t) \in \mathbb{R}_+$ for each j and for all $t \geq 0$, and where the initial values $(x_j(0), j \geq 1)$ are such that $\sum_{j \geq 1} x_j(0) < \infty$. That these equations are indeed a deterministic analog of (2.1) is shown in the following theorems. In order to state them, we need a deterministic condition reflecting (2.2):

$$(C) \qquad \sup_{0 \leq s \leq t} \sum_{j \geq 1} x_j(s) < \infty \quad \text{for all } t \geq 0.$$

THEOREM 2.1. *Let $(X^{(M)}, M \geq 1)$ be a sequence of Markov branching processes as specified in (2.1), having initial values $X^{(M)}(0)$ such that $\sum_{j \geq 1} X_j^{(M)}(0) < \infty$ and such that $M^{-1}X^{(M)}(0) \to x^{(0)}$, where $0 < \sum_{j \geq 1} x_j^{(0)} < \infty$. Then $M^{-1}X^{(M)}$ converges weakly in $D^\infty[0, T]$ for each $T > 0$ to a nonrandom process x, which evolves according to the differential equations (2.3) with initial state $x(0) = x^{(0)}$ and which satisfies Condition C.*

THEOREM 2.2. *Let X be a Markov process with rates given in (2.1) and with initial state $X(0)$ satisfying $\sum_{j \geq 1} X_j(0) = M < \infty$, and set $x^{(0)} = M^{-1}X(0)$. Then ξ, defined by*

$$\xi_j(t) = M^{-1}\mathbb{E}\{X_j(t)|X(0) = Mx^{(0)}\},$$

satisfies the differential equations (2.3) with $\xi(0) = x^{(0)}$, as well as Condition C.

Thus the initial values $x_j(0)$ in (2.3) can be taken to represent the initial proportions of the different types of infectives in a large mixing population, in which case the initial condition $\sum_{j \geq 1} x_j(0) = 1$ is appropriate. The $x_j(t)$ then represent the numbers of the different types of infectives at time t, expressed relative to the total number of initial infectives. Since the total number of infectives can grow or decline with time, it is not reasonable to suppose that $\sum_{j \geq 1} x_j(t) = 1$ for all t. However, the fact that the solutions of (2.3), which are

interesting as approximations to the behavior of (2.1), satisfy Condition C is reassuring, in view of (2.2).

The differential equations (2.3) can be written in vector form $\dot{x} = xR$, with

$$(2.4) \qquad R_{ij} = i\mu(\delta_{i-1,j} - \delta_{ij}) + \lambda p_{ij}, \qquad i,j \geq 1,$$

where δ_{kl} is the Kronecker delta. Since $x(t) \in \mathbb{R}_+^\infty$ and all the off-diagonal elements of R are nonnegative, the notation is unambiguous. The matrix R has a positive right eigenvector v with components $v_j = j$, $j \geq 1$, since

$$(2.5) \quad (Rv)_i = i\mu\{(i-1) - i\} + \lambda \sum_{j \geq 1} jp_{ij} = -i\mu + \lambda i\theta = (\lambda\theta - \mu)v_i$$

which apparently suggests that the solution of (2.3) grows if $R_0 > 1$, $\lambda\theta - \mu > 0$, and declines otherwise, but the fact that there are infinitely many coordinates complicates matters. Nonetheless, it follows from (2.5) that the matrix S defined by

$$(2.6) \qquad S_{ij} = \frac{j}{i}\left(R_{ij} - (\lambda\theta - \mu)\delta_{ij}\right), \qquad i,j \geq 1,$$

has nonnegative off-diagonal elements and satisfies

$$\sum_{j \geq 1} S_{ij} = \frac{1}{i} \sum_{j \geq 1} jR_{ij} - (\lambda\theta - \mu) = 0$$

for all $i \geq 1$, and is thus a "Q-matrix" in the sense of Markov chain theory [Chung (1967)]. We write

$$S_i = -S_{ii} = \lambda\theta + \mu(i-1) - \lambda p_{ii}, \qquad i \geq 1.$$

Furthermore, if $y(t)$ is a nonnegative solution of $\dot{y} = yS$ with $y(0) = y^{(0)}$, then setting

$$(2.7) \qquad x_j(t) = j^{-1}y_j(t)e^{(\lambda\theta - \mu)t}, \qquad t \geq 0, j \geq 1,$$

yields a nonnegative solution to $\dot{x} = xR$ with $x_j(0) = j^{-1}y_j^{(0)}$, since

$$\dot{x}_j = j^{-1}\left(\dot{y}_j + (\lambda\theta - \mu)y_j\right)e^{(\lambda\theta - \mu)t}$$

$$= j^{-1}\left[\sum_{i \geq 1} y_i S_{ij} + (\lambda\theta - \mu)y_j\right]e^{(\lambda\theta - \mu)t}$$

$$= j^{-1} \sum_{i \geq 1} i^{-1}y_i jR_{ij}e^{(\lambda\theta - \mu)t} = \sum_{i \geq 1} x_i R_{ij}.$$

Similarly, starting with a nonnegative solution of $\dot{x} = xR$, one obtains a solution of $\dot{y} = yS$ from (2.7). Thus the solutions of (2.3) are simply related to the solutions of $\dot{y} = yS$, for a specific Q-matrix S. This enables one to prove the following theorem.

THEOREM 2.3. *The equations $\dot{x} = xR$, $x(0) = x^{(0)}$, with $x^{(0)}$ such that $0 < \sum_{j \geq 1} x_j^{(0)} < \infty$, have a unique nonnegative solution satisfying Condition* C. *The solution is given by*

$$(2.8) \qquad x_j(t) = j^{-1}\left(\sum_{l \geq 1} lx_l^{(0)}\mathbb{P}_l[Y(t) = j]\right)e^{(\lambda\theta - \mu)t},$$

where Y is the unique pure jump Markov process with Q-matrix S and $\mathbb{P}_l$ denotes probability conditional on $Y(0) = l$.

REMARK 2.4. With the more restrictive initial condition $\sum_{j \geq 1} j x_j^{(0)} < \infty$, corresponding to a finite initial number of *parasites*, rather than infectives, the solution can be even more simply expressed as

$$(2.9) \qquad x_j(t) = j^{-1} \left(\sum_{l \geq 1} l x_l^{(0)} \right) e^{(\lambda \theta - \mu)t} \mathbb{P}^{(0)} [Y(t) = j],$$

where $\mathbb{P}^{(0)}$ denotes probability conditional on the initial distribution

$$\mathbb{P}^{(0)}[Y(0) = j] = j x_j^{(0)} \Big/ \left(\sum_{l \geq 1} l x_l^{(0)} \right)$$

for Y. From now on, we shall always assume that $\sum_{j \geq 1} j x_j^{(0)} < \infty$ is satisfied.

For x given by (2.9), we always have

$$(2.10) \qquad \sum_{j \geq 1} j x_j(t) = \left(\sum_{l \geq 1} l x_l^{(0)} \right) e^{(\lambda \theta - \mu)t},$$

growing with t if and only if $R_0 > 1$. Thus R_0 is appropriate for characterizing the asymptotic behavior of $\sum_{j \geq 1} j x_j(t)$. However, for a "large outbreak" in the epidemiological sense, it is the behavior of $\sum_{j \geq 1} x_j(t)$ that is relevant, and this may be rather different. In fact, if Y is irreducible and positive recurrent, there is a probability distribution π on $\mathbb{N}$ such that $\mathbb{P}[Y(t) = j] \to \pi_j$ as $t \to \infty$, irrespective of the value of $x^{(0)}$, and hence

$$\sum_{j \geq 1} x_j(t) \sim \left(\sum_{j \geq 1} j^{-1} \pi_j \right) \left(\sum_{l \geq 1} l x_l^{(0)} \right) e^{(\lambda \theta - \mu)t}$$

has its growth or decay characterized by R_0. However, if Y is null recurrent or transient,

$$(2.11) \qquad \lim_{t \to \infty} \sum_{j \geq 1} j^{-1} \mathbb{P}[Y(t) = j] = 0,$$

and the "natural" asymptotic order $e^{(\lambda \theta - \mu)t}$ always overestimates that of $\sum_{j \geq 1} x_j(t)$. Thus the recurrence classification of Y is a first important step in understanding how $\sum_{j \geq 1} x_j(t)$ behaves. This is the substance of the next result.

THEOREM 2.5. *The Markov process Y is irreducible if $p_{11} + p_{10} < 1$. Assuming this, Y is positive recurrent if $R_0 \log \theta < 1$, null recurrent if $R_0 \log \theta = 1$ and transient if $R_0 \log \theta > 1$. If $p_{11} + p_{10} = 1$, Y is eventually absorbed in state 1.*

As a consequence of Theorem 2.5, if $p_{11} + p_{10} < 1$,

$$x_j(t) \sim j^{-1} \pi_j \left(\sum_{l \geq 1} l x_l^{(0)} \right) e^{(\lambda \theta - \mu)t}$$

exactly when $R_0 \log \theta < 1$. Under these circumstances, growth or decay is determined by whether $R_0 > 1$ or $R_0 < 1$ holds (with $R_0 = 1$ implying asymptotic stability), the exponential rate is exactly $(\lambda\theta - \mu)$ and, whether growing or declining, the proportions of hosts with j parasites, $j \geq 1$, converges as $t \to \infty$:

$$(2.12) \qquad \lim_{t\to\infty} e^{-(\lambda\theta-\mu)t} x_j(t) = j^{-1}\pi_j\left(\sum_{l\geq 1} l x_l^{(0)}\right)$$

and

$$(2.13) \qquad \lim_{t\to\infty}\left\{x_j(t)\Big/\sum_{l\geq 1} x_l(t)\right\} = j^{-1}\pi_j\Big/\left(\sum_{l\geq 1} l^{-1}\pi_l\right).$$

In particular, if $\theta \leq 1$, the behavior is as in (2.12) and (2.13), whatever the values of λ and μ. This includes the case $p_{11} + p_{10} = 1$, for which one takes $\pi_1 = 1$ and $\pi_j = 0$ otherwise.

If $1 < \theta < e$, then $R_0 \leq 1$ entails $R_0 \log \theta < 1$, so that, for fixed θ in this range, (2.12) and (2.13) are true for all choices of λ and μ such that $R_0 \leq 1$, that is, such that $\lambda/\mu \leq 1/\theta$, and indeed for all values of λ/μ up to, but not including, $1/\{\theta \log \theta\}$. Since increasing λ while leaving θ and μ unchanged increases $x(t)$ for all t, it follows that $R_0 = 1$ is the critical value separating growth from decay for θ in the range $1 < \theta < e$ also. However, if $\lambda/\mu \geq 1/\{\theta \log \theta\}$, (2.12) no longer holds, and all that can immediately be guaranteed is an exponential growth rate of at least

$$\frac{\mu}{\log\theta}(1 - \log\theta) - \varepsilon < \lambda\theta - \mu$$

for any $\varepsilon > 0$. More precise statements about the behavior when $1 < \theta < e$, and even the most elementary properties when $\theta > e$, require a more detailed analysis of the process Y. Our main result in this direction is the following.

THEOREM 2.6. *The solution* (2.10) *to the equations* $\dot{x} = xR$, $x(0) = x^{(0)}$, *for initial conditions satisfying* $\sum_{j\geq 1} j x_j^{(0)} < \infty$, *is such that the limit*

$$\lim_{t\to\infty} t^{-1}\log\sum_{j\geq 1} x_j(t) =: c(\lambda, \mu, \theta)$$

exists and is given by

$$c(\lambda, \mu, \theta) = \begin{cases} \lambda\theta - \mu, & \text{if } R_0\log\theta \leq 1, \\ \dfrac{\lambda\theta}{R_0\log\theta}(1 + \log(R_0\log\theta)) - \mu, & \text{if } 1 < R_0\log\theta \leq \theta, \\ \lambda, & \text{if } R_0\log\theta > \theta. \end{cases}$$

REMARK 2.7. In particular, $c < 0$ if $\theta \leq e$ and $R_0 < 1$, or if $\theta > e$ and $R_0 < \theta/(e\log\theta)$, these being the same threshold conditions as for the stochastic model (2.1). In addition, Theorem 2.6 also gives an exact exponential rate of growth or decay. Note, however, that the conclusion of Theo-

rem 2.6 is not quite strong enough to justify a result analogous to (2.12), nor would it, in general, be correct: for $\theta < e$ and $R_0 \log \theta = 1$, we have $c(\lambda, \mu, \theta) = \lambda \theta - \mu$, but, from Theorem 2.5,

$$e^{-(\lambda\theta - \mu)t} x_j(t) = j^{-1} \left(\sum_{l \geq 1} l x_l^{(0)} \right) \mathbb{P}^{(0)}[Y(t) = j] \to 0.$$

The form of the exponential rate $c(\lambda, \mu, \theta)$ depends on the properties of the auxiliary Markov process Y, which emerged purely as an analytical aid in (2.6). It also has a biological interpretation. Suppose, in an infinite population setting, that $q_i = q_i(t)$ denotes the proportion of parasites at time t which are living in hosts who have exactly i parasites; thus $q_i = i x_i / \sum_{j \geq 1} j x_j$. Using (2.3) and (2.10), it follows that

$$(2.14) \qquad \dot{q}_i = i \mu q_{i+1} - [(i-1)\mu + \lambda\theta] q_i + \lambda \sum_{l \geq 1} q_l l^{-1} i p_{li},$$

so that the $q_i(t)$ solve the Kolmogorov forward differential equations for the Markov process Y; $q_i(t) = \mathbb{P}[Y(t) = i]$, with $\mathbb{P}[Y(0) = i]$ assigned according to the initial constitution of the population. Thus the distribution of $Y(t)$ is that of the number of parasites in a host which is selected at time t by choosing a *parasite* uniformly at random and then fixing on its host.

Our discussion so far has been in terms of the evolution of the model (2.1) in real time. However, the basic reproduction number is usually defined by considering the reproductive success of an individual in terms of its offspring in the next generation; see, for example, Diekmann, Heesterbeek and Metz (1990). What does the infection process look like, if it is described in terms of its evolution in time measured by generations of infected hosts?

As before, it is important to distinguish hosts according to their infective potential. Here, a type i host denotes a host who was *initially* infected by i parasites. The expected number T_{ij} of type j "offspring" in the next generation of infection arising from a single type i host is then given by

$$T_{ij} = \lambda \int_0^\infty \sum_{l=1}^{i} P_{il}(t) p_{lj} \, dt,$$

where $P_{il}(t)$ is the probability that, at time t after its infection, a type i host has exactly l surviving parasites: $P_{il}(t) = \mathbb{P}[\mathrm{Bi}(i, e^{-\mu t}) = l]$. Since $\int_0^\infty P_{il}(t) \, dt = 1/l\mu$, the expected length of time for which exactly l parasites are alive in the host, we thus find that

$$(2.15) \qquad\qquad T_{ij} = \frac{\lambda}{\mu} \sum_{l=1}^{i} l^{-1} p_{lj}.$$

Note that $\sum_{j \geq 1} j T_{ij} = i \lambda \theta / \mu = R_0 i$; on average, parasite numbers are multiplied by R_0 in each generation.

The expected evolution, now with generation as the time parameter, is described by the discrete difference equations

$$(2.16) \qquad\qquad x^{(n)} = x^{(n-1)} T,$$

where $x^{(n)}$ denotes the expected numbers of hosts of the different types in generation n. The operator T (or more precisely its adjoint) corresponds to the next-generation operator as introduced by Diekmann, Heesterbeek and Metz (1990). These equations are analogous to (2.3), with $T - I$ corresponding to R, and have solution $x^{(n)} = x^{(0)}T^n$. To find its behavior, note that the matrix L defined by

$$(2.17) \qquad L_{ij} = \frac{j}{iR_0}T_{ij} = \frac{1}{i\theta}\sum_{l=1}^{i} l^{-1}jp_{lj}$$

is stochastic ($L - I$ corresponds to S), and that

$$(2.18) \qquad T_{ij}^n = R_0^n iL_{ij}^n/j.$$

Hence the development of $x^{(n)}$ depends on the properties of the Markov chain $\hat{Y}$ with transition matrix L. In contrast to the evolution in real time, the matrix L does not involve λ or μ, so that critical values for determining the form of the growth exponent of $x^{(n)}$ have to be different from those in Theorem 2.6. It turns out that the situation is rather simpler.

THEOREM 2.8. *The Markov chain $\hat{Y}$ with transition matrix L defined in (2.17) is positive recurrent if $\theta < e$, null recurrent if $\theta = e$ and transient if $\theta > e$. Furthermore, the limit*

$$\lim_{n \to \infty} \left\{ \sum_{j \geq 1} x_j^{(n)} \right\}^{1/n} = \hat{c}(R_0, \theta)$$

exists, and is given by

$$\hat{c}(R_0, \theta) = \begin{cases} R_0, & \text{if } \theta \leq e, \\ \lambda e \log \theta/\mu, & \text{if } \theta > e. \end{cases}$$

The contrast between Theorems 2.6 and 2.8 indicates that the way in which generations of infection overlap in real time also changes in character, depending on the parameter values. As an aid to understanding this, let $q_{ik} = q_{ik}(t)$ denote the proportion of *parasites* at time t which are living in hosts who have exactly i parasites and belong to the kth generation of infection. Then the analog of (2.14) yields

$$(2.19) \qquad \dot{q}_{ik} = i\mu q_{i+1, k} - \left[(i-1)\mu + \lambda\theta\right]q_{ik} + \lambda \sum_{l \geq 1} q_{l, k-1}l^{-1}ip_{li}.$$

Thus the $q_{ik}(t)$ solve the Kolmogorov forward equations for a two-dimensional Markov process (Y, Z) with transitions

$$(2.20) \qquad \begin{aligned} (j, k) &\to (j-1, k) \quad \text{at rate } (j-1)\mu, \\ (j, k) &\to (l, k+1) \quad \text{at rate } \lambda j^{-1}lp_{jl}, \end{aligned}$$

for $j, l \geq 1$, $k \geq 0$. Hence the distribution of $(Y(t), Z(t))$ can be interpreted as the status, in terms of number of parasites and generation, of a host selected

at time t by choosing a parasite uniformly at random and then taking its host. The first component is the same as the original auxiliary Markov process Y. The second component Z is a Poisson process of rate $\lambda\theta$, showing that the proportion of parasites at time t which are living in generation k hosts is given by $e^{-\lambda\theta t}(\lambda\theta t)^k/k!$. Hence also, using (2.10), the proportion of the parasites of generation k which are alive at time t is given by $e^{-\mu t}(\mu t)^k/k!$. These are the same proportions as expected in a linear birth and death process with per capita birth rate $\lambda\theta$ and death rate μ, suggesting that the overlap of generations in terms of parasite numbers remains entirely normal. However, host type and generation are not independent of one another; instead, we have the following result.

THEOREM 2.9. *For the Markov process (Y, Z) with transitions given in (2.20), we have*

$$\mathbb{E}\{y^{Y(t)}z^{Z(t)}|Y(0) = i,\, Z(0) = 0\} = e^{\lambda\theta t(z-1)}\mathbb{E}\{y^{Y_z(t)}|Y_z(0) = i\},$$

where Y_z is the auxiliary Y-process obtained when λ is replaced by λz. In particular, because the distribution of Y_z is not the same as that of Y, $Y(t)$ and $Z(t)$ are not independent.

Despite Theorem 2.9, it still makes sense to ask whether the distributions of $Y(t)$ and $Z(t)$ are in some sense asymptotically independent. If $R_0 \log\theta < 1$, the Y-process is positive recurrent, and so $Y(t)$ has a limiting distribution π as $t \to \infty$. The distribution of $Z(t)$ never converges, but that of $\hat{Z}(t) = t^{-1/2}\{Z(t) - \lambda\theta t\}$ has limit $\mathcal{N}(0, \lambda\theta)$. In the case when $R_0 \log\theta < 1$, it follows from Theorem 2.9, by setting $z = \exp\{-st^{-1/2}\}$ and letting $t \to \infty$, that the pair $(Y(t), \hat{Z}(t))$ has $\pi \times \mathcal{N}(0, \lambda\theta)$ as limiting distribution, with independence between the components.

Thus, sampling in real time, the distribution of host type settles asymptotically to a fixed distribution which is independent of generation number, whenever $R_0 \log\theta < 1$. This remains true even when $\theta > e$, although, under these circumstances, the number of hosts in generation n decreases like $\{R_0 e(\log\theta)/\theta\}^n$ instead of like R_0^n, and so the *average* number of parasites per host at infection in generation n grows like $(\theta/e(\log\theta))^n$. That these apparently different kinds of behavior can coexist seems surprising. However, there are two factors which could help to account for it. First, even if the average number of parasites per host tends to ∞ with generation number, it need not be the case that the "typical" number of parasites per host, which is described by a probability distribution, also tends to ∞—a proper probability distribution can well have infinite mean. Second, a host with a large initial parasite load spends a longer time infected than one with a small initial load, but for most of this time the bulk of his initial parasites are already dead. However, from an epidemiological point of view, this combination of parameters is uninteresting, since then $R_0 < 1$ also, and the infection dies out.

The more interesting case is that in which $R_0 \log\theta > 1$ and $\theta < e$. This implies geometric growth of host and parasite numbers like R_0^n in terms of generations, but a slower exponential rate of growth in real time for the number of infected hosts than the rate $e^{(\lambda\theta - \mu)t}$ for the number of parasites.

Thus, in real time, the average number of parasites per host tends to ∞ exponentially fast, although it remains steady when time is expressed in terms of number generations. Once again, a proper distribution can have infinite mean; also, it is now not obviously the case from Theorem 2.9 that generation number and host type are asymptotically independent, so that the way in which the generations overlap may be biased by parasite burden. In both of these curious cases, the difficulties in interpreting the results arise because the hosts are sampled according to a scheme weighted by parasite number. Such a scheme has no meaning when the mean parasite burden is infinite, as may asymptotically be the case.

3. Proofs.

PROOF OF THEOREM 2.3.   The proof of Theorem 2.1 actually uses Theorem 2.3, which we therefore start by proving.

LEMMA 3.1.   *The matrix S is the Q-matrix corresponding to a regular* (*nonexplosive*) *pure jump Markov process* $(Y(t),\ t \geq 0)$.

PROOF.   Since Y is positive integer valued, it can only make infinitely many jumps if it makes infinitely many upward jumps. However, upward jumps occur according to a Poisson process of rate $\lambda\theta$, so that only finitely many can occur in finite time, with probability 1. $\square$

LEMMA 3.2.   *Let* $(g(j),\ j \geq 1)$ *be a bounded nonnegative sequence. Then if,* *for some* $c \in \mathbb{R}$,

$$\sum_{j \geq 1} S_{ij} g(j) \leq (\geq) cg(i) \quad \text{for all } i \geq 1,$$

it follows that

$$e^{-ct} \mathbb{E} g(Y(t)) \leq (\geq) \mathbb{E} g(Y(0)) \quad \text{for all } t > 0.$$

PROOF.   Let the jump times of Y be denoted by $(\tau_n,\ n \geq 0)$, with $\tau_0 = 0$, and set $M_n = \exp(-c(t \wedge \tau_n)) g(Y(t \wedge \tau_n))$. Then, assuming the first direction in the inequalities, the sequence M_n is a supermartingale with respect to the σ-algebras $\mathscr{F}_{\tau_n}$, where $\mathscr{F}_s = \sigma(Y(u),\ 0 \leq u \leq s)$. To see this, take any $n \geq 0$, $0 < s < t$ and $i \geq 1$, and compute

$$\mathbb{E}\big(M_{n+1} | \mathscr{F}_{\tau_n} \cap \{\tau_n = s,\, Y(\tau_n) = i\}\big)$$

$$= \exp(-S_i(t - s)) g(i) e^{-ct}$$

$$\quad + \int_s^t \exp(-S_i(u - s)) \sum_{j \neq i} S_{ij} g(j) e^{-cu}\, du$$

$$\leq \exp(-S_i(t - s)) g(i) e^{-ct}$$

$$\quad + g(i) e^{-cs}(c + S_i) \int_s^t \exp(-(c + S_i)(u - s))\, du$$

$$= g(i) e^{-cs} = M_n.$$

Hence, for all $n \geq 0$,

$$\mathbb{E}\{\exp(-c(t \wedge \tau_n))g(Y(t \wedge \tau_n))\}$$
$$= \mathbb{E}\{e^{-ct}g(Y(t))I[t \leq \tau_n]\} + \mathbb{E}\{\exp(-c\tau_n)g(Y(\tau_n))I[t > \tau_n]\}$$
$$\leq \mathbb{E}g(Y(0)).$$

Letting $n \to \infty$, monotone convergence and Lemma 3.1 show that

$$\mathbb{E}\{e^{-ct}g(Y(t))I[t \leq \tau_n]\} \uparrow e^{-ct}\mathbb{E}g(Y(t)),$$

whereas

$$0 \leq \mathbb{E}\{\exp(-c\tau_n)g(Y(\tau_n))I[t > \tau_n]\}$$
$$\leq \exp\{t \max(0, -c)\} \sup_j g(j)\mathbb{P}[t > \tau_n] \to 0.$$

The argument with the inequalities in the other direction is similar. $\square$

LEMMA 3.3.   *If* $\sum_{j \geq 1} x_j^{(0)} < \infty$, *the equations* $\dot{x} = xR$, $x(0) = x^{(0)}$, *with* $x^{(0)} \geq 0$ (*componentwise*), *have a nonnegative solution*

$$(3.1) \qquad x_j(t) = j^{-1}\left(\sum_{l \geq 1} lx_l^{(0)}\mathbb{P}_l[Y(t) = j]\right)e^{(\lambda\theta - \mu)t},$$

which satisfies Condition C.

PROOF.   Suppose first that $x_l^{(0)} = \delta_{lk}$ for some $k \geq 1$, so that the corresponding $y^{(0)}$ is given by $y_l^{(0)} = k\delta_{lk}$. Then, from the standard Markov theory, $\dot{y} = yS$ has a solution given by

$$y_j(t) = k\mathbb{P}_k[Y(t) = j], \qquad j \geq 1,$$

and this solution is the minimal nonnegative solution with the given initial condition. Hence, from (2.7), $\dot{x} = xR$ has a solution

$$x_j(t) = j^{-1}k\mathbb{P}_k[Y(t) = j]e^{(\lambda\theta - \mu)t}$$
$$= j^{-1}\left(\sum_{l \geq 1} l\,\delta_{lk}\mathbb{P}_l[Y(t) = j]\right)e^{(\lambda\theta - \mu)t},$$

in agreement with (3.1). Linearity now implies that (3.1) is a solution for all initial vectors $x^{(0)}$ which have only finitely many nonzero components.

To solve the differential equations $\dot{x} = xR$ for an arbitrary initial condition $x^{(0)} \geq 0$ with $\sum_{j \geq 1} x_j^{(0)} < \infty$, use (2.7) and instead solve the equations $\dot{y} = yS$ with an arbitrary initial condition $y^{(0)} \geq 0$ satisfying $\sum_{j \geq 1} j^{-1}y_j^{(0)} < \infty$. The differential equations $\dot{y} = yS$ can be rewritten in integral form as

$$(3.2) \quad y_j(t) = y_j(0)\exp(-S_j t) + \int_0^t \sum_{i \neq j} y_i(u)S_{i,j}\exp(-S_j(t - u))\, du.$$

Thus, if $y^{(0)} \geq 0$ is an arbitrary vector with $\sum_{j \geq 1} j^{-1}y_j^{(0)} < \infty$ and if $y_j^{[M]}$ solves $\dot{y} = yS$ with initial condition $y_j^{(0)}1_{j \leq M}$, it is immediate that $y_j^{[M]}(t)$ is

increasing in M for each fixed j and t. Hence it follows from monotone convergence in (3.2) that $y = \lim_{M \to \infty} y^{[M]}$ satisfies (3.2) with initial condition $y^{(0)}$, this y being explicitly given by

$$y_j(t) = \sum_{l \geq 1} y_l^{(0)} \mathbb{P}_l[Y(t) = j].$$

Hence, from (2.7), the solution x proposed in (3.1) indeed satisfies the equations $\dot{x} = xR$. That this solution is not infinite, and indeed satisfies Condition C, follows by taking $g(j) = j^{-1}$ in Lemma 3.2, for which we have

$$\sum_{j \geq 1} S_{ij} g(j) = \mu - i^{-1}[\mu(i-1) + \lambda\theta] + i^{-1}\lambda \sum_{j \geq 1} p_{ij} \leq cg(i)$$

for $c = -(\lambda\theta - \mu) + \lambda$. Thus

$$\sum_{j \geq 1} j^{-1} \mathbb{P}_l[Y(t) = j] \leq l^{-1} e^{ct},$$

from which it follows that, for x as in (3.1),

$$\sum_{j \geq 1} x_j(t) \leq \sum_{l \geq 1} l x_l^{(0)} l^{-1} e^{\lambda t} < \infty. \qquad \square$$

LEMMA 3.4. *The solution given in* (3.1) *is the only nonnegative solution which satisfies Condition* C.

PROOF. The differential equations $\dot{x} = xR$ can be rewritten in integral form as

$$(3.3) \quad x_j(t) = x_j(0) + \int_0^t \Big\{ (j+1)\mu x_{j+1}(u) - j\mu x_j(u)$$
$$+ \lambda \sum_{l \geq 1} x_l(u) p_{lj} \Big\} \, du, \qquad j \geq 1.$$

To simplify the algebra, we also introduce an $x_0(t)$, which does not influence the other $x_j(t)$ in the system of differential equations. We let $x_0(t)$ develop according to the equation

$$(3.4) \qquad x_0(t) = x_0(0) + \int_0^t \Big\{ \mu x_1(u) + \lambda \sum_{l \geq 0} x_l(u) p_{l0} \Big\} \, du,$$

with $x_0(0) = 0$.

We prove Lemma 3.4 by showing that each nonnegative solution of $\dot{x} = xR$ which satisfies Condition C has the same Laplace transform, and then applying the uniqueness theorem. To obtain the Laplace transform, we multiply the j equation in (3.3) [or (3.4) if $j = 0$] by e^{-js}, for any fixed $s > 0$, and add over $j \geq 0$, obtaining

$$\phi(s,t) = \phi(s,0) + \int_0^t \Big\{ \mu(1 - e^s) \frac{\partial\phi(s,u)}{\partial s} + \lambda\phi(-\log\psi(s), u) \Big\} \, du,$$

where $\phi(s,t) = \sum_{j \geq 0} x_j(t) e^{-js}$ and $\psi(s) = \sum_{j \geq 0} e^{-sj} p_{1j}$. Differentiating with respect to t then leads to the partial differential equation

$$\frac{\partial \phi(s,t)}{\partial t} = \mu(1 - e^s) \frac{\partial \phi(s,t)}{\partial s} + \lambda \phi(-\log \psi(s), t).$$

This can be integrated in $s > 0$, $t \geq 0$, using the method of characteristics, in the form

$$(3.5) \qquad \phi(s,t) = \phi(\sigma_{s,t}(v), v) + \int_v^t \lambda \phi(-\log \psi(\sigma_{s,t}(u)), u) \, du$$

for any v, and, in particular, for $v = 0$, where

$$\sigma_{s,t}(u) = -\log\{1 - (1 - e^{-s}) e^{-\mu(t-u)}\}.$$

Now if $\xi^{(1)}$ and $\xi^{(2)}$ are two nonnegative solutions of $\dot{x} = xR$ which satisfy Condition C, they give rise to functions ϕ_1 and ϕ_2 satisfying (3.5). Under Condition C, it follows from (3.4) and Gronwall's inequality that

$$M_w = \max_{i=1,2} \sup_{0 \leq r \leq w} \sum_{j \geq 0} \xi_j^{(i)}(r) < \infty$$

for all $w \geq 0$. For all t, $0 \leq t \leq w$, we thus have $0 \leq \phi_i(s,t) \leq M_w$ for $i = 1, 2$. Suppose now that, for any $v \geq 0$, $\phi_1(s,v) = \phi_2(s,v)$ for all s, as is certainly the case for $v = 0$. Let

$$d_{vw} = \sup_{v \leq t \leq w} \sup_{s > 0} |\phi_1(s,t) - \phi_2(s,t)| \leq M_w < \infty.$$

Then, from (3.5), for $v \leq t \leq w$,

$$|\phi_1(s,t) - \phi_2(s,t)|$$

$$\leq \lambda \int_v^t |\phi_1(-\log \psi(\sigma_{s,t}(u)), u) - \phi_2(-\log \psi(\sigma_{s,t}(u)), u)| \, du$$

$$\leq \lambda(t - v) d_{vw} \leq \lambda(w - v) d_{vw},$$

from which it follows that $d_{vw} \leq \lambda(w - v) d_{vw}$, implying in turn that $d_{vw} = 0$ for $w < v + 1/\lambda$. Iterating this procedure, starting with $v = 0$ and continuing in steps of $1/2\lambda$, shows that $\phi_1(s,t) = \phi_2(s,t)$ for all $s > 0$ and $t \geq 0$, which completes the proof of the lemma. $\square$

Lemmas 3.3 and 3.4 comprise the proof of Theorem 2.3. $\square$

PROOF OF THEOREM 2.1. For all $j \geq 1$ and $x \in \mathbb{R}_+^\infty$, define the functions

$$a_j(x) = (j + 1)\mu x_{j+1} - j\mu x_j + \lambda \sum_{l \geq 1} x_l p_{lj},$$

$$b_j(x) = (j + 1)\mu x_{j+1} + j\mu x_j + \lambda \sum_{l \geq 1} x_l p_{lj},$$

and the random processes

$$U_j^M(t) = x_j^M(t) - x_j^M(0) - \int_0^t a_j(x^M(u))\, du,$$

$$V_j^M(t) = U_j^M(t)^2 - \frac{1}{M} \int_0^t b_j(x^M(u))\, du,$$

where $x^M(t) = M^{-1} X^{(M)}(t)$. Further, let $\mathscr{G}_t^M$ denote $\sigma\{x^M(s), 0 \leq s \leq t\}$.

LEMMA 3.5. *The processes $U_j^M(t)$ and $V_j^M(t)$ are $\mathscr{G}_t^M$-martingales.*

PROOF. It is enough to modify the arguments in Hamza and Klebaner (1995), Theorem 2 and Corollary 3, replacing their Condition C with the bound

$$|a_j(x)| \leq b_j(x) \leq 2(\lambda \vee \mu)(j+1) \sum_{l \geq 1} x_l,$$

whenever it is used, and observing that, for any $T > 0$,

$$(3.6) \qquad \mathbb{E}\left\{ \sup_{0 \leq t \leq T} \sum_{l \geq 1} x_l^M(t) \right\} \leq e^{\lambda T}.$$

The latter is true, because $\sum_{j \geq 1} X_j^{(M)}(t)$ only increases at an infection, and infections occur at a total rate of $\lambda \sum_{j \geq 1} \sum_{k \geq 1} X_j^{(M)} p_{jk} \leq \lambda \sum_{j \geq 1} X_j^{(M)}$; thus, by comparison with a pure birth process with per capita birth rate λ, (3.6) follows. $\square$

LEMMA 3.6. *For each j, the sequence $\{x_j^M\}_{M \geq 1}$ is tight in $D[0, T]$ and any weak limit belongs to $C[0, T]$.*

PROOF. We apply Billingsley (1968), Theorem 15.5, for which we need only to check that, given any $\varepsilon, \eta > 0$, we can find $\delta, M_0 > 0$ such that, for all $M \geq M_0$,

$$(3.7) \qquad \mathbb{P}\left[\sup_{0 \leq s < t \leq T; \, t-s < \delta} |x_j^M(t) - x_j^M(s)| > \varepsilon \right] < \eta.$$

From the definition of U_j^M, it follows that

$$|x_j^M(t) - x_j^M(s)|$$

$$(3.8) \qquad \leq |U_j^M(t) - U_j^M(s)| + \int_s^t |a_j(x^M(u))|\, du$$

$$\leq |U_j^M(t) - U_j^M(s)| + \int_s^t (\lambda \vee \mu) 2(j+1) \sum_{l \geq 1} x_l^M(u)\, du.$$

Now by the Doob–Kolmogorov inequality for martingales, for s arbitrary but fixed,

$$
\mathbb{P}\left[\sup_{s < t \leq s+\delta} |U_j^M(t) - U_j^M(s)| \geq \frac{\varepsilon}{6}\right]
$$

(3.9)

$$
\leq \frac{36}{\varepsilon^2} \mathbb{E}\left[\left(U_j^M(s+\delta) - U_j^M(s)\right)^2\right]
$$

$$
= \frac{36}{\varepsilon^2 M} \mathbb{E}\left[\int_s^{s+\delta} b_j(x^M(u)) \, du\right]
$$

$$
\leq \frac{36}{\varepsilon^2 M} \mathbb{E}\left[\int_s^{s+\delta} (\lambda \vee \mu) 2(j+1) \sum_{l \geq 1} x_l^M(u) \, du\right]
$$

$$
= \frac{72(j+1)(\lambda \vee \mu)}{\varepsilon^2 M} \int_s^{s+\delta} \mathbb{E}\left[\sum_{l \geq 1} x_l^M(u)\right] du
$$

$$
\leq \frac{72(j+1)(\lambda \vee \mu)\delta e^{\lambda T}}{\varepsilon^2 M},
$$

where the last inequality follows from (3.6). For the second term in (3.8), comparison with the pure birth process with rate λ immediately gives an estimate which is uniform in $s \in [0, T]$:

(3.10)

$$
\mathbb{P}\left[\sup_{0 \leq s < t \leq T; \, t-s \leq \delta} \int_s^t (\lambda \vee \mu) 2(j+1) \sum_{l \geq 1} x_l^M(u) \, du \geq \frac{\varepsilon}{2}\right]
$$

$$
\leq \frac{4\delta(j+1)(\lambda \vee \mu)e^{\lambda T}}{\varepsilon}.
$$

Hence, given $\varepsilon, \eta > 0$, pick δ so small that $4\delta(j+1)(\lambda \vee \mu)e^{\lambda T} < \varepsilon\eta/2$ with $\delta = T/r$ for some integer r, so that the estimate in (3.10) is at most $\eta/2$. We then choose M_0 so large that, for all $M \geq M_0$,

$$
72(j+1)(\lambda \vee \mu)e^{\lambda T} < \varepsilon^2 M \eta/2T,
$$

so that, from (3.9),

$$
\mathbb{P}[A_s^M] < \frac{\eta\delta}{2T}
$$

for any $s \in [0, T]$ and $M \geq M_0$, where $A_s^M := \{\sup_{s < t \leq s+\delta} |U_j^M(t) - U_j^M(s)| \geq \varepsilon/6\}$. With these choices, we have

$$
\mathbb{P}\left[\sup_{0 \leq s < t \leq T; \, t-s < \delta} |U_j^M(t) - U_j^M(s)| > \frac{\varepsilon}{2}\right]
$$

$$
\leq \mathbb{P}\left[\bigcup_{i=0}^{r-1} A_{i\delta}^M\right] \leq \sum_{i=0}^{r-1} \mathbb{P}[A_{i\delta}^M] \leq \frac{r\eta\delta}{2T} = \frac{\eta}{2}
$$

for all $M \geq M_0$. This completes the proof of (3.7). $\square$

LEMMA 3.7.   *Given any infinite subsequence N_1 of $\mathbb{N}$, there exists a subsequence $N_2 \subset N_1$ such that x^M converges weakly in $D^\infty[0,T]$ along N_2. We denote the limit by $x^* = x^*(N_2)$.*

PROOF.   It is enough by Prohorov's theorem to show that the sequence x^M is tight in $D^\infty[0,T]$. Given $\varepsilon > 0$, let K_j be a compact set in $D[0,T]$ such that $\mathbb{P}[\,x_j^M \in K_j\,] > 1 - 2^{-j}\varepsilon$: such a K_j exists, by Lemma 3.6. Then $K = \prod_{j \geq 1} K_j$ is compact in $D^\infty[0,T]$, and $\mathbb{P}[\,x^M \in K\,] > 1 - \varepsilon$. $\square$

LEMMA 3.8.   *The condition $x^M \Rightarrow x^*(N_2)$ implies that $U_j^M \Rightarrow U_j^*$ in $D[0,T]$ along N_2, for any $j \geq 1$. Here,*

$$U_j^* = x_j^*(t) - x_j^*(0) - \int_0^t a_j(x^*(u))\,du.$$

PROOF.   The index N_2 is suppressed throughout the proof. Define the functions

$$h(x)(t) = x_j(t) - x_j(0) - \int_0^t \Bigg[(j+1)\mu x_{j+1}(u) - j\mu x_j(u)$$
$$+ \lambda \sum_{l \geq 1} x_l(u) p_{lj} \Bigg]\,du$$

and, for any $k > 0$,

$$h_k(x)(t) = x_j(t) - x_j(0) - \int_0^t \Bigg[(j+1)\mu x_{j+1}(u) - j\mu x_j(u)$$
$$+ \lambda \sum_{l \geq 1} (x_l(u) \wedge k) p_{lj} \Bigg]\,du.$$

Then, if $x \in D^\infty[0,T]$ satisfies $\sup_{0 \leq t \leq T} \sum_{j \geq 1} |x_j(t)| < \infty$, both $h(x)$ and $h_k(x)$ are elements of $D[0,T]$, and $U_j^M(t) = h(x^M)(t)$ and $U_j^*(t) = h(x^*)(t)$. We thus need to prove that

(3.11)
$$\lim_{M \to \infty} \mathbb{E}\big[\,f(h(x^M))\,\big] = \mathbb{E}\big[\,f(h(x^*))\,\big]$$

for each $f \in C^b(D[0,T])$.

Observe that, for any such f and any $k > 0$, we have

(3.12)
$$\big| \mathbb{E}\big[\,f(h(x^M))\,\big] - \mathbb{E}\big[\,f(h(x^*))\,\big] \big|$$
$$\leq \big| \mathbb{E}\big[\,f(h(x^M))\,\big] - \mathbb{E}\big[\,f(h_k(x^M))\,\big] \big|$$
$$+ \big| \mathbb{E}\big[\,f(h_k(x^M))\,\big] - \mathbb{E}\big[\,f(h_k(x^*))\,\big] \big|$$
$$+ \big| \mathbb{E}\big[\,f(h_k(x^*))\,\big] - \mathbb{E}\big[\,f(h(x^*))\,\big] \big|.$$

For the first term in (3.12), it follows from (3.6) that

$$\left| \mathbb{E}\big[f(h(x^M)) \big] - \mathbb{E}\big[f(h_k(x^M)) \big] \right| \le 2\|f\|\mathbb{P}\left[\sup_{0 \le t \le T} \sum_{j \ge 1} x_j^M(t) > k \right]$$

$$\le 2\|f\|e^{\lambda T}/k.$$

A similar argument can be used for $\|\mathbb{E}[f(h_k(x^*))] - \mathbb{E}[f(h(x^*))]\|$, since

$$\mathbb{P}\left[\sup_{0 \le t \le T} \sum_{j \ge 1} x_j^*(t) > k \right] = \lim_{J \to \infty} \mathbb{P}\left[\sup_{0 \le t \le T} \sum_{j=1}^{J} x_j^*(t) > k \right]$$

(3.13)
$$\le \lim_{J \to \infty} \liminf_{M \to \infty} \mathbb{P}\left[\sup_{0 \le t \le T} \sum_{j=1}^{J} x_j^M(t) > k \right]$$

$$\le e^{\lambda T}/k,$$

because $x^M \Rightarrow x^*$ and from (3.6). Finally, h_k is continuous at all points of $C^\infty[0,T]$ and $\mathbb{P}[x^* \in C^\infty[0,T]] = 1$ in view of Lemmas 3.6 and 3.7; thus $h_k(x^M) \Rightarrow h_k(x^*)$, and so the remaining term $\|\mathbb{E}[f(h_k(x^M))] - \mathbb{E}[f(h_k(x^*))]\|$ in (3.12) is small as $M \to \infty$. Because k was chosen arbitrarily, this proves (3.11). $\square$

LEMMA 3.9. *The condition $x^M \Rightarrow x^*(N_2)$ implies that*

$$\mathbb{P}\left[\sup_{0 \le t \le T} |U_j^M(t)| > \varepsilon \right] \to 0$$

along N_2, for each $j \ge 1$ and $\varepsilon > 0$.

PROOF.   By the Doob–Kolmogorov inequality applied to U_j^M and from the martingale property of V_j^M, it follows that

$$\mathbb{P}\left[\sup_{0 \le t \le T} |U_j^M(t)| > \varepsilon \right] \le \frac{1}{\varepsilon^2}\mathbb{E}\big[U_j^M(T)^2 \big]$$

$$\le \frac{1}{M\varepsilon^2}\int_0^T \mathbb{E}\big[b_j(x^M(u)) \big]\, du$$

$$\le \frac{1}{M\varepsilon^2}\int_0^T 2(j+1)(\lambda \vee \mu)e^{\lambda T}\, du$$

$$= \frac{2(j+1)(\lambda \vee \mu)Te^{\lambda T}}{M\varepsilon^2},$$

which converges to 0 as $M \to \infty$. $\square$

As a consequence of Lemma 3.9,

$$\mathbb{P}\left[x_j^*(t) = x_j^*(0) + \int_0^t a_j(x^*(u))\, du \ \text{ for all } 0 \le t \le T \right] = 1$$

for any weak limit $x^* = x^*(N_2)$. Thus x^* satisfies (2.3) with $x(0) = x^{(0)}$, as required. By Theorem 2.3, there is only one solution of (2.3) that satisfies Condition C. It thus simply remains to be shown that any $x^*(N_2)$ satisfies Condition C; but this follows from (3.13). This completes the proof of Theorem 2.1. $\square$

PROOF OF THEOREM 2.2. Let $m_{ij}(t) = \mathbb{E}\{X_j(t) | X(0) = e_i\}$, where e_i denotes the ith coordinate vector in $\mathbb{R}^\infty$, and note that

$$\xi_j(t) = \sum_{i \ge 1} x_i^{(0)} m_{ij}(t).$$

By the comparison with a pure birth process in (3.6), we have

$$0 \le \sum_{j \ge 1} m_{ij}(t) \le e^{\lambda t} < \infty$$

for all $i \ge 1$ and $t \ge 0$. Thus $\sum_{j \ge 1} \xi_j(t) \le e^{\lambda t} < \infty$ for all t also, and Condition C is satisfied.

By the Markov property, since

$$\mathbb{E}\{X_j(t + h) | \mathscr{G}_t\} = \sum_{i \ge 1} X_i(t) m_{ij}(h),$$

where $\mathscr{G}_t$ denotes $\sigma(X(s), 0 \le s \le t)$ and $h > 0$, it follows that

(3.14)
$$\begin{aligned}
h^{-1}\{\xi_j(t + h) - \xi_j(t)\} \\
= \sum_{i \ne j} \xi_i(t) h^{-1} m_{ij}(h) + \xi_j(t) h^{-1}(m_{jj}(h) - 1).
\end{aligned}$$

Using the branching structure of X and conditioning on the time and outcome of the first transition, we then have

$$\begin{aligned}
m_{ij}(h) = \delta_{ij} e^{-(\lambda + i\mu)h} \\
+ \int_0^h e^{-(\lambda + i\mu)u}\Big\{ i\mu m_{i-1,j}(h - u) \\
+ \lambda \sum_{l \ge 0} p_{il}\{m_{lj}(h - u) + m_{ij}(h - u)\}\Big\}\, du,
\end{aligned}$$

with $m_{0j}(t)$ taken to be 0 for all t, so that

$$\begin{aligned}
e^{(\lambda + i\mu)h} m_{ij}(h) \\
= \delta_{ij} + \int_0^h e^{(\lambda + i\mu)v}\Big\{ i\mu m_{i-1,j}(v) + \lambda m_{ij}(v) + \lambda \sum_{l \ge 0} p_{il} m_{lj}(v)\Big\}\, dv;
\end{aligned}$$

hence each m_{ij} is differentiable, and

$$(3.15) \qquad m_{ij}(0) = \delta_{ij}, \qquad \dot{m}_{ij}(0) = i\mu\delta_{i,j+1} - i\mu\delta_{ij} + \lambda p_{ij}.$$

Again by comparison with the pure birth process, we have

$$0 \le h^{-1}m_{ij}(h) \le h^{-1}(e^{\lambda h} - 1) \le \lambda e^{\lambda} \quad \text{in } h \le 1,$$

uniformly for all $i \ne j$, and hence, since $\sum_{j \ge 1} \xi_j(t) < \infty$, dominated convergence as $h \to 0$ in (3.14) shows that

$$\dot{\xi}_j(t) = \sum_{i \ne j} \xi_i(t)\{i\mu\delta_{i,j+1} + \lambda p_{ij}\} + \xi_j(t)\{-j\mu + \lambda p_{jj}\},$$

so that ξ satisfies (2.3) with $\xi(0) = x^{(0)}$. $\square$

PROOF OF THEOREM 2.5.   Observe that $S_{i,i-1} > 0$ for all $i \ge 2$. If $p_{11} + p_{10} = 1$, it follows, in addition, that $S_{ij} = 0$ for all $j > i$, and hence that state 1 is absorbing and that Y eventually reaches it. On the other hand, irreducibility under the condition $p_{11} + p_{10} < 1$ is immediate, because then, for any i, there is a $j > i$ with $S_{ij} > 0$.

For the remainder of the proof, we therefore have $p_{11} + p_{10} < 1$. We argue using the criteria of Foster (1953) and Tweedie (1975, 1976) throughout.

We begin by showing that Y is positive recurrent if and only if $R_0 \log \theta < 1$. If $R_0 \log \theta < 1$, we verify the conditions for positive recurrence given in Tweedie (1975), Theorem 2.3(i), for which it is sufficient to show that $\sum_{j \ge 1} S_{ij} \log j$ is finite for all i and bounded above by $-\frac{1}{2}(\mu - \lambda\theta \log \theta) < 0$ for all i sufficiently large. Now

$$\sum_{j \ge 1} S_{ij} \log j = \mu(i-1)(\log(i-1) - \log i) + \lambda i^{-1} \sum_{j \ge 1} jp_{ij} \log j - \lambda\theta \log i$$

$$\le -\mu i^{-1}(i-1) + \lambda\theta \log(i\theta + \sigma^2/\theta) - \lambda\theta \log i$$

by Jensen's inequality applied to the distribution with probabilities $jp_{ij}/(i\theta)$, $j \ge 1$, with the function $\log x$; for $i = 1$, define $(i-1)\log(i-1) = 0$. Hence

$$\sum_{j \ge 1} S_{ij} \log j \le \lambda\theta \log \theta - \mu + i^{-1}(\mu + \lambda\sigma^2/\theta),$$

and the result follows.

If $R_0 \log \theta \ge 1$, we use instead a theorem for continuous-time pure-jump Markov processes which is analogous to Tweedie (1976), Theorem 9.1(ii); to prove that Y is null recurrent or transient, it is enough to check that

$$(3.16) \qquad \sum_{j \ge 1} S_{ij} \log j > \lambda\theta \log \theta - \mu \ge 0$$

and

$$(3.17) \qquad \sum_{j \ge 1} S_{ij}|\log j - \log i| \le \mu + \lambda(\theta^2 + \sigma^2 + 1) < \infty$$

for all $i \geq 1$, while observing that $\inf_{i \geq 1} S_i > 0$; the latter follows from

$$\inf_{i \geq 1} S_i \geq \min\left\{\lambda(\theta - p_{11}), \mu + \lambda\left(\theta - \sup_{i \geq 2} p_{ii}\right)\right\},$$

because $\theta = i^{-1}\sum_{j \geq 1} j p_{ij} \geq p_{ii}$ for all i and $p_{11} < \theta$ if $p_{11} + p_{10} < 1$. A proof of the theorem is given in the Appendix.

For (3.16), use Jensen's inequality applied to the distribution with probabilities p_{ij}, $j \geq 0$, and the function $x \log x$, to give

$$\sum_{j \geq 1} S_{ij}\log j > -\mu + \lambda\theta(\log(i\theta) - \log i) = \lambda\theta \log\theta - \mu$$

for any $i \geq 2$. Note that, for such i, $\log(i-1) - \log i > -(i-1)^{-1}$. In the case $i = 1$, (3.16) is immediate.

For (3.17), simply observe that

$$\sum_{j \geq 1} S_{ij}|\log j - \log i| \leq \mu + \frac{\lambda}{i}\left[\sum_{j > i} j p_{ij}\left(\frac{j-i}{i}\right) + \sum_{j < i} p_{ij}(i-j)\right]$$

$$\leq \mu + \lambda i^{-1}\left[i^{-1}(i^2\theta^2 + i\sigma^2) + i\right].$$

We now show that Y is recurrent if $R_0 \log\theta = 1$. To do this, apply Theorem 3.3 of Tweedie (1975), showing that $\sum_{j \geq 1} S_{ij}(\log j)^{1/2} \leq 0$ for all i sufficiently large. This is equivalent to $\sum_{j \geq 1} P_{ij}(\log j)^{1/2} \leq (\log i)^{1/2}$ for P the "jump" matrix defined by $P_{ij} = (1 - \delta_{ij})S_{ij}/S_i$. The argument follows from the two estimates

$$\mu(i - 1)\left\{(\log(i-1))^{1/2} - (\log i)^{1/2}\right\}$$

$$\leq -\tfrac{1}{2}\mu(\log i)^{-1/2} + O\left(i^{-1}(\log i)^{-1/2}\right),$$

and, using Jensen's inequality,

$$\lambda i^{-1}\sum_{j \geq 1} j p_{ij}(\log j)^{1/2} - \lambda\theta(\log i)^{1/2}$$

$$\leq \lambda\theta\left\{\left(\log\left(i\theta + \frac{\sigma^2}{\theta}\right)\right)^{1/2} - (\log i)^{1/2}\right\}$$

$$= \lambda\theta(\log i)^{1/2}\left(\left[1 + \frac{\log\theta + O(i^{-1})}{\log i}\right]^{1/2} - 1\right)$$

$$\leq \lambda\theta(\log i)^{1/2}\left\{\left(1 + \frac{\log\theta}{\log i}\right)^{1/2} - 1 + O\left((i\log i)^{-1}\right)\right\}$$

$$\leq \frac{1}{2}\lambda\theta\log\theta(\log i)^{-1/2} - \frac{\lambda\theta}{8}\frac{(\log\theta)^2}{(\log i)^{3/2}}\left(1 + \frac{\log\theta}{\log i}\right)^{-3/2}$$

$$+ O\left(i^{-1}(\log i)^{-1/2}\right).$$

Finally, to show that Y is transient if $R_0 \log \theta > 1$, apply Theorem 6 of Foster (1953), by showing that there is an α, $0 < \alpha < 1$, such that

$$\sum_{j \geq 1} S_{ij} j^{-\alpha} \leq 0 \quad \text{for all } i.$$

In fact, by Jensen's inequality,

$$\sum_{j \geq 1} S_{ij} j^{-\alpha} \leq \mu\{(i-1)^{1-\alpha} - i^{1-\alpha}\} + \lambda i^{-1}(i\theta)^{1-\alpha} - (\lambda\theta - \mu)i^{-\alpha}$$

$$\leq i^{-\alpha}\{(\lambda\theta^{1-\alpha} - \mu(1-\alpha)) - (\lambda\theta - \mu)\},$$

and so an α with the required properties exists provided that the function $f(x) = \lambda\theta^x - \mu x$ satisfies $f(1) > \min_{0 \leq x \leq 1} f(x)$. However, if $R_0 \log \theta > 1$, $f'(x) = \lambda\theta^x \log \theta - \mu > 0$ for x near 1 and so this is indeed the case. $\square$

PROOF OF THEOREM 2.6.   Since Theorem 2.5 is enough when $R_0 \log \theta < 1$, we concentrate on the case $R_0 \log \theta > 1$; the case $R_0 \log \theta = 1$ can then be dealt with, using the monotonicity in λ.

As in the proof of Theorem 2.5, for any $0 \leq \alpha \leq 1$,

$$\sum_{j \geq 1} S_{ij} j^{-\alpha} \leq i^{-\alpha} c_\alpha,$$

with

$$c_\alpha = (\lambda\theta^{1-\alpha} - \mu(1-\alpha)) - (\lambda\theta - \mu).$$

If $R_0 \log \theta > 1$, $c_\alpha < 0$ for α near 0, taking its minimum at

$$1 - \log\left(\frac{\mu}{\lambda \log \theta}\right)\bigg/\log \theta = 1 - \log\left(\frac{\theta}{R_0 \log \theta}\right)\bigg/\log \theta > 0$$

if $R_0 \log \theta \leq \theta$, and at 1 if $R_0 \log \theta > \theta$. So, for $1 < R_0 \log \theta$, $\sum_{j \geq 1} S_{ij} j^{-\alpha} \leq -di^{-\alpha}$, with

$$\alpha = 1 - \log\left(\frac{\theta}{R_0 \log \theta}\right)\bigg/\log \theta$$

and

$$d = \frac{\mu}{\log \theta}(R_0 \log \theta - 1 - \log(R_0 \log \theta))$$

if $1 < R_0 \log \theta \leq \theta$ and with $\alpha = 1$ and $d = \lambda(\theta - 1) - \mu$ if $R_0 \log \theta \geq \theta$. Applying Lemma 3.2, it thus follows that, for each $l \geq 1$ and for all $t \geq 0$,

$$e^{dt}\mathbb{E}_l\{Y(t)^{-1}\} \leq e^{dt}\mathbb{E}_l\{Y(t)^{-\alpha}\} \leq \mathbb{E}_l\{Y(0)^{-\alpha}\} = l^{-\alpha}.$$

Thus, from (3.1),

$$\sum_{j \geq 1} x_j(t) = \left(\sum_{l \geq 1} l x_l^{(0)} \mathbb{E}_l\{Y(t)^{-1}\}\right) e^{(\lambda\theta - \mu)t}$$

$$\leq \sum_{j \geq 1} j^{1-\alpha} x_j^{(0)} e^{(\lambda\theta - \mu - d)t},$$

and hence, for $R_0 \log \theta > 1$,

$$(3.18) \qquad \limsup_{t \to \infty} t^{-1} \log \sum_{j \geq 1} x_j(t) \leq \lambda \theta - \mu - d = c(\lambda, \mu, \theta),$$

since now, by assumption, $\sum_{j \geq 1} j x_j^{(0)} < \infty$.

For the reverse inequality, the argument is more delicate. For any $K > e$, we define $g_K(j) = j^{-\alpha_K(j)}$ for

$$\alpha_K(j) = \begin{cases} 0, & \text{for } 1 \leq j \leq K, \\ \left(1 - \dfrac{\log \log K}{\log \log j} \right)^2, & \text{for } j > K, \end{cases}$$

and show that

$$(3.19) \qquad \sum_{j \geq 1} S_{ij} g_K(j) \geq -(d + \varepsilon(K)) g_K(i),$$

where $\varepsilon(K) = O((\log \log K)^{-1/2})$ and d is as above. From Lemma 3.2, it then follows that

$$\mathbb{E}_l \left\{ Y(t)^{-\alpha_K(Y(t))} \right\} \geq e^{-(d + \varepsilon(K))t} \mathbb{E}_l \left\{ Y(0)^{-\alpha_K(Y(0))} \right\}$$

$$= e^{-(d + \varepsilon(K))t} l^{-\alpha_K(l)}.$$

Now, for α_K so defined, $\sup_{x \geq 1} x^{1 - \alpha_K(x)} = C(K) < \infty$, and hence

$$C(K) \mathbb{E}_l \left\{ Y(t)^{-1} \right\} \geq e^{-(d + \varepsilon(K))t} l^{-\alpha_K(l)}$$

also. Hence, from (3.1),

$$
\begin{aligned}
(3.20) \qquad \liminf_{t \to \infty} t^{-1} \log \sum_{j \geq 1} x_j(t) &\geq \lambda \theta - \mu - d - \varepsilon(K) \\
&= c(\lambda, \mu, \theta) - \varepsilon(K),
\end{aligned}
$$

and since, by the choice of K, $\varepsilon(K)$ can be made arbitrarily small, the theorem follows.

It remains to prove (3.19). Suppressing the index K from now on, we note that

$$
\begin{aligned}
(3.21) \qquad \sum_{j \geq 1} S_{ij} g(j) &= \mu(i - 1)(g(i - 1) - g(i)) \\
&\quad + \lambda \theta \frac{1}{i\theta} \sum_{j \geq 1} j p_{ij} (g(j) - g(i)).
\end{aligned}
$$

Now, for $i \geq 2$,

$$
\begin{aligned}
(3.22) \qquad \frac{g(i - 1) - g(i)}{g(i)} &= i^{\alpha(i)}(i - 1)^{-\alpha(i-1)} - 1 \\
&= (1 - i^{-1})^{-\alpha(i)}(i - 1)^{\alpha(i) - \alpha(i-1)} - 1 \\
&\geq 1 + i^{-1}\alpha(i) - 1 = i^{-1}\alpha(i).
\end{aligned}
$$

Then, for any $j < i$,

$$\frac{g(j)}{g(i)} = \frac{i^{\alpha(i)}}{j^{\alpha(j)}} \geq \left(\frac{i}{j}\right)^{\alpha(i)},$$

whereas, for $j > i$,

$$\frac{g(j)}{g(i)} = \left(\frac{i}{j}\right)^{\alpha(i)} \exp\{\log j(\alpha(i) - \alpha(j))\}$$

$$\geq \left(\frac{i}{j}\right)^{\alpha(i)} \exp\left\{\frac{-4(j-i)\log j}{((K\log K) \vee (i\log i))\log\log K}\right\}$$

$$\geq \left(\frac{i}{j}\right)^{\alpha(i)} \exp\left\{\frac{-8(j-i)}{(K \vee i)\log\log K}\right\}$$

if also $j \leq (K \vee i)^2$. Thus, for all j,

(3.23)
$$\frac{g(j)}{g(i)} \geq \left(\frac{i}{j}\right)^{\alpha(i)} - 1_{\{j > (K \vee i)^2\}} - 8(\log\log K)^{-1/2}$$
$$- 1_{\{j > (K \vee i)(\log\log K)^{1/2}\}},$$

and hence, using the Jensen and Markov inequalities and the fact that $0 \leq \alpha(j) \leq 1$, it follows that

(3.24)
$$\frac{1}{i\theta}\sum_{j\geq 1}\frac{jp_{ij}g(j)}{g(i)} \geq \left(\frac{i}{i\theta + \sigma^2/\theta}\right)^{\alpha(i)} - \frac{i\theta + \sigma^2/\theta}{(K \vee i)^2}$$
$$- \frac{8}{\sqrt{\log\log k}} - \frac{i\theta + \sigma^2/\theta}{(K \vee i)\sqrt{\log\log k}}.$$

Combining (3.22) and (3.24) and recalling that $\alpha(i) = 0$ for $i \leq K$, we thus find that, for all $i \geq 1$,

$$\frac{1}{g(i)}\sum_{j\geq 1}S_{ij}g(j)$$

$$\geq \mu\alpha(i) + \lambda\theta^{1-\alpha(i)} - \lambda\theta - \mu i^{-1}\alpha(i)$$

$$- \lambda\theta\left\{\theta^{-\alpha(i)}\left[1 - \left(1 + \frac{\sigma^2}{i\theta^2}\right)^{-\alpha(i)}\right] + (K \vee i)^{-1}\theta\left(1 + \frac{\sigma^2}{i\theta^2}\right)\right.$$

$$\left. + \frac{8}{\sqrt{\log\log K}} + \frac{\theta(1 + \sigma^2/i\theta^2)}{\sqrt{\log\log K}}\right\}$$

$$= c_{\alpha(i)} - \eta(i, K),$$

where $|\eta(i, K)| = O((\log \log K)^{-1/2})$ uniformly in i. Hence, for all $i \geq 1$,

$$\frac{1}{g(i)} \sum_{j \geq 1} S_{ij} g(j) \geq -(d + \varepsilon(K)),$$

where $\varepsilon(K) = \sup_{i \geq 1} |\eta(i, K)| = O((\log \log K)^{-1/2})$, as required. $\square$

PROOF OF THEOREM 2.8.   The proofs of the recurrence properties of $\hat{Y}$ are discrete-time analogs of those used for Theorem 2.5, so we only mention the essential steps. To prove that $\hat{Y}$ is positive recurrent if $\theta < e$, use Tweedie (1976), Theorem 9.1(i), showing that

$$\sum_{j \geq 1} L_{ij} \log j \leq \log i + (\log \theta - 1) + \varepsilon_i,$$

where $\lim_{i \to \infty} \varepsilon_i = 0$. Then $\hat{Y}$ is null recurrent or transient when $\theta \geq e$ from Tweedie (1976), Theorem 9.1(ii), because

$$\sum_{j \geq 1} L_{ij} \log j \geq \log i + (\log \theta - 1) \geq \log i,$$

and because $\sum_{j \geq 1} L_{ij} |\log j - \log i|$ is uniformly bounded in i. That $\hat{Y}$ is recurrent if $\theta = e$ follows, as for Y when $R_0 \log \theta = 1$, from Tweedie (1975), Theorem 3.3, since $\sum_{j \geq 1} L_{ij} (\log j)^{1/2} \leq (\log i)^{1/2}$ for all i large enough, here, we use Jensen's inequality to show that

$$\sum_{j \geq 1} L_{ij} (\log j)^{1/2} \leq i^{-1} \sum_{l=1}^{i} \left\{ \log \left(l\theta + \frac{\sigma^2}{\theta} \right) \right\}^{1/2},$$

and then calculate that

$$i^{-1} \sum_{l=1}^{i} \left\{ \log \left(l\theta + \frac{\sigma^2}{\theta} \right) \right\}^{1/2} = i^{-1} \sum_{l=1}^{i} \left\{ \log(l\theta) \right\}^{1/2} + O\left(i^{-1} (\log i)^{1/2} \right)$$

$$= \frac{1}{i\theta} \int_{1}^{i\theta} (\log x)^{1/2} \, dx + O\left(i^{-1} (\log i)^{1/2} \right)$$

$$\leq (\log i)^{1/2} - \frac{1}{8(2 \log i)^{3/2}} + O\left(i^{-1} (\log i)^{1/2} \right)$$

for $\theta = e$. That $\hat{Y}$ is transient if $\theta > e$ follows from Foster (1953), Theorem 6, since

$$\sum_{j \geq 1} L_{ij} j^{-\alpha} \leq i^{-\alpha}$$

for all i, if $\alpha \in (0, 1)$ is such that $\theta^{\alpha}(1 - \alpha) > 1$; for $\theta > e$, any $\alpha > 0$ which is small enough does the trick.

To determine the asymptotic growth rate of $\hat{Y}$, the arguments are much as for Theorem 2.6. If $\theta < e$, $\hat{Y}$ is positive recurrent, and the growth rate follows using (2.18). For $\theta > e$, it follows easily that

$$\sum_{j \geq 1} L_{ij} j^{-\alpha} \leq i^{-\alpha} c_{\alpha},$$

where $c_\alpha = [\theta^\alpha(1 - \alpha)]^{-1}$ has a minimum value of $e \log \theta/\theta$, attained when $1 > \alpha = 1 - 1/\log \theta > 0$. Let α henceforth denote this value. Since $c_\alpha^{-n}\hat{Y}_n^{-\alpha}$ is a supermartingale, it follows that

$$c_\alpha^{-n}\mathbb{E}\left\{\hat{Y}_n^{-1}|\hat{Y}_0 = l\right\} \leq c_\alpha^{-n}\mathbb{E}\left\{\hat{Y}_n^{-\alpha}|\hat{Y}_0 = l\right\}$$

$$\leq \mathbb{E}\left\{\hat{Y}_0^{-\alpha}|\hat{Y}_0 = l\right\} = l^{-\alpha},$$

and thence, from (2.18), that

$$\limsup_{n \to \infty} \left\{\sum_{j \geq 1} x_j^{(n)}\right\}^{1/n} = \limsup_{n \to \infty} \left\{\sum_{j \geq 1}\sum_{i \geq 1} x_i^{(0)}T_{ij}^n\right\}^{1/n}$$

$$\leq R_0 c_\alpha = \lambda e \log \theta/\mu.$$

For the reverse inequality, we consider $g_K(j) = j^{-\alpha_K(j)}$, with $\alpha_K(j)$ as defined in the proof of Theorem 2.6, showing that

(3.25) $$\sum_{j \geq 1} L_{ij}g_K(j) \geq [c_\alpha - \eta(K)]g_K(i)$$

for all $i \geq 1$, where $\eta(K) = O((\log\log K)^{-1/2})$ and $\alpha = 1 - 1/\log \theta$. It follows from (3.25), much as in the argument preceding (3.20), that

$$\mathbb{E}\left\{\hat{Y}_n^{-\alpha_K(\hat{Y}_n)}|\hat{Y}_0 = l\right\} \geq [c_\alpha - \eta(K)]^n\mathbb{E}\left\{\hat{Y}_0^{-\alpha_K(\hat{Y}_0)}|\hat{Y}_0 = l\right\}$$

$$= [c_\alpha - \eta(K)]^n l^{-\alpha_K(l)}.$$

Since also

$$\mathbb{E}\left\{\hat{Y}_n^{-\alpha_K(\hat{Y}_n)}|\hat{Y}_0 = l\right\} = \mathbb{E}\left\{\hat{Y}_n^{-1}\hat{Y}_n^{1-\alpha_K(\hat{Y}_n)}|\hat{Y}_0 = l\right\}$$

$$\leq C(K)\mathbb{E}\left\{\hat{Y}_n^{-1}|\hat{Y}_0 = l\right\},$$

where $C(K) = \sup_{x \geq 1} x^{1-\alpha_K(x)}$ as before, we then have

$$\sum_{j \geq 1}\sum_{i \geq 1} x_i^{(0)}T_{ij}^n \geq R_0^n \sum_{i \geq 1} x_i^{(0)}i^{1-\alpha_K(i)}[c_\alpha - \eta(K)]^n/C(K),$$

and hence that

$$\liminf_{n \to \infty} \left\{\sum_{j \geq 1} x_j^{(n)}\right\}^{1/n} = \liminf_{n \to \infty} \left\{\sum_{j \geq 1}\sum_{i \geq 1} x_i^{(0)}T_{ij}^n\right\}^{1/n}$$

$$\geq R_0 c_\alpha = \lambda e \log \theta/\mu.$$

It thus remains to prove (3.25); we suppress the index K from now on. Using (3.23) and the Jensen and Markov inequalities, we find, analogously to (3.24), that

$$\sum_{j \geq 1} \frac{L_{ij}g(j)}{g(i)} \geq i^{\alpha(i)-1} \sum_{l=1}^{i}\left(l\theta + \frac{\sigma^2}{\theta}\right)^{-\alpha(i)} - i^{-1}\sum_{l=1}^{i}\frac{l\theta + \sigma^2/\theta}{(K \vee i)^2}$$

$$- \frac{8}{\sqrt{\log\log K}} - i^{-1}\sum_{l=1}^{i}\frac{l\theta + \sigma^2/\theta}{(K \vee i)\sqrt{\log\log K}}.$$

The last three terms in the lower bound are uniformly of order $(\log\log K)^{-1/2}$, so that only the first need be considered further: call it U_i. Then, for $i \le K$, we have $\alpha(i) = 0$, and thus $U_i = 1 > c_\alpha$. We also have

$$U_i \ge \frac{i^{\alpha(i)-1}}{\theta + \sigma^2/\theta} \sum_{l=1}^{i} l^{-\alpha(i)} \ge \frac{1}{\theta + \sigma^2/\theta} \int_{1/i}^{1} y^{-\alpha(i)}\, dy.$$

If we pick an α^* such that $\int_0^1 y^{-\alpha^*}\, dy > \theta + \sigma^2/\theta$, and then i^* such that $\int_{1/i^*}^1 y^{-\alpha^*}\, dy \ge \theta + \sigma^2/\theta$, it follows that $U_i \ge 1 > c_\alpha$ for all i so large that $\alpha(i) \ge \alpha^*$ and $i \ge i^*$; for K chosen large enough, $\alpha(i) \ge \alpha^*$ already implies that $i \ge i^*$, so that, then, $U_i \ge 1 > c_\alpha$ whenever $\alpha(i) \ge \alpha^*$. Finally, if $i > K$ and $\alpha(i) \le \alpha^*$,

$$U_i = \frac{i^{\alpha(i)-1}}{\theta^{\alpha(i)}} \sum_{l=1}^{i} l^{-\alpha(i)} - \frac{i^{\alpha(i)-1}}{\theta^{\alpha(i)}} \sum_{l=1}^{i} l^{-\alpha(i)}\left\{1 - \left(1 + \frac{\sigma^2}{l\theta^2}\right)^{-\alpha(i)}\right\}$$

$$\ge \frac{i^{\alpha(i)-1}}{\theta^{\alpha(i)}} \int_1^i x^{-\alpha(i)}\, dx - O\big((\log\log K)^{-1/2}\big)$$

$$= c_{\alpha(i)} - O\big((\log\log K)^{-1/2}\big),$$

where the errors are uniform in the stated range, and $c_{\alpha(i)} \ge c_\alpha$. □

PROOF OF THEOREM 2.9.   To analyze the process (Y, Z), let

$$\chi(i,t) = \chi_{y,z}(i,t) = \mathbb{E}\{y^{Y(t)} z^{Z(t)} | Y(0) = i, Z(0) = 0\}$$

for $0 \le y, z \le 1$. Then, conditioning on the time and outcome of the first transition, we have

$$\chi(i,t) = \exp\{-[(i-1)\mu + \lambda\theta]t\} y^i + \int_0^t \exp\{-[(i-1)\mu + \lambda\theta]u\}$$

$$\times \left\{(i-1)\mu\chi(i-1,t-u) + \sum_{j\ge1} \frac{\lambda}{i} jp_{ij} z\chi(j,t-u)\right\} du.$$

Changing variable to $v = t - u$ in the integral, multiplying both sides by $\exp\{[(i-1)\mu + \lambda\theta]t\}$ and differentiating gives

$$\frac{\partial\chi}{\partial t}(i,t) + [(i-1)\mu + \lambda\theta]\chi(i,t)$$

$$= (i-1)\mu\chi(i-1,t) + \sum_{j\ge1} i^{-1}\lambda zjp_{ij}\chi(j,t).$$

Rewriting in terms of $\psi_i(t) = \chi(i,t)e^{-\lambda\theta t(z-1)}$, we thus obtain

$$\frac{\partial\psi_i}{\partial t} + [(i-1)\mu + \lambda z\theta]\psi_i = (i-1)\mu\psi_{i-1} + \sum_{j\ge1} \frac{\lambda z}{i} jp_{ij}\psi_j,$$

the equation for the generating function of the process Y, when the infection rate is λz instead of λ. □

APPENDIX

Let X be the minimal pure-jump Markov process associated with an irreducible $\mathbb{N} \times \mathbb{N}$ intensity matrix Q for which $\inf_i q_i > 0$, where, as usual, $q_i = -q_{ii} = \sum_{j \neq i} q_{ij}$. Suppose that there exists a function $g : \mathbb{N} \rightarrow \mathbb{R}_+$ which satisfies:

(i) $(Qg)_i = \sum_j q_{ij} g(j) \geq 0$ *for all $i > i_0$,*

(ii) $U(i) = \sum_j q_{ij} |g(j) - g(i)| \leq \beta$ *for all $i \geq 1$,*

(iii) $\max_{i \leq i_0} g(i) < \inf_{i > i_0} g(i)$

for some $i_0 \geq 1$, $\beta < \infty$. Then X is either null recurrent or transient.

PROOF.   (a) *Without loss of generality, we may take g to satisfy* (i) *for all i, and to be such that $(Qg)_{j_0} > 0$ for at least one j_0.*

If not, define $g'(j) = g(j) \vee \max_{l \leq i_0} g(l)$; g' then has the required properties. For (i), if $i \leq i_0$, we have

$$(Qg')_i = \sum_{j \neq i} q_{ij} \left(g'(j) - \min_l g'(l) \right) \geq 0,$$

whereas, for $i > i_0$, we have $(Qg')_i \geq (Qg)_i \geq 0$; (ii) is immediate, since $|g'(j) - g'(i)| \leq |g(j) - g(i)|$ for all i, j; and, if $j \leq i_0$, we have

$$(Qg')_j = \sum_{l > i_0} q_{jl} \left(g(l) - \max_{i \leq i_0} g(i) \right),$$

where all terms of the sum with $q_{jl} > 0$ are positive and the rest are 0, so that the sum is positive for any j for which $q_{jl} > 0$ for some $l > i_0$, and one such j at least must exist, because Q is irreducible.

From now on, we assume if possible that X is positive recurrent.

(b) *Let T_0 denote the time at which j_0 is first reached. Then $\mathbb{E}_i T_0 \geq \beta^{-1} |g(i) - g(j_0)|$.*

Let $\tau_0 = 0$ and $(\tau_j, j \geq 1)$ denote the jump times of X, write $x_j = X(\tau_j)$ and set

$$M_0 = 0, \qquad M_n = \sum_{j=1}^n \left\{ |g(x_j) - g(x_{j-1})| - (\tau_j - \tau_{j-1}) U(x_{j-1}) \right\}.$$

Then $\mathbb{E}|M_n| \leq 2n\beta \sup_i q_i^{-1} < \infty$, and a simple calculation now shows that $(M_n, n \geq 0)$ is a martingale with respect to the σ-fields $\Sigma_n = \sigma\{(x_j, \tau_j),$

$0 \le j \le n$}. Hence, if $N_0 < \infty$ denotes the index of the jump at which j_0 is first reached, it follows that $\mathbb{E}_i(M_{n \wedge N_0}) = 0$ for any $n \ge 0$ and any i, so that

$$|g(i) - g(j_0)|\mathbb{P}[N_0 \le n] \le \mathbb{E}_i\left\{\sum_{j=1}^{N_0} |g(x_j) - g(x_{j-1})|I[N_0 \le n]\right\}$$

$$\le \mathbb{E}_i\left\{\sum_{j=1}^{n \wedge N_0} (\tau_j - \tau_{j-1})U(x_{j-1})\right\}$$

$$\le \beta\mathbb{E}_i(\tau_n \wedge T_0) \le \beta\mathbb{E}_i T_0$$

for any $n \ge 0$, from which claim (b) follows.

(c) *For all t, $\mathbb{E}_i g(X(t)) - g(i) = \mathbb{E}_i \int_0^t (Qg)_{X(u)}\, du$.*

This follows from Hamza and Klebaner (1995), Theorem 2, and the remark following it.

Let $T_0^{(t)} = \inf\{s \ge t: X(s) = j_0\}$, and set $e_t = \mathbb{E}_{j_0} T_0^{(t)} - t$.

(d) $\limsup_{t \to \infty} t^{-1}e_t = 0$.

We use a last entrance renewal formula to write

$$e_t = h(t) + \int_0^t h(t - u)\, dm(u),$$

where m is the renewal function corresponding to the distribution function F of the time $\tilde{T}_0$ of the first *return* to j_0 when starting form j_0, and

$$h(t) = \mathbb{E}_{j_0}\left\{(\tilde{T}_0 - t)I[\tilde{T}_0 > t]\right\} \le \mathbb{E}_{j_0}\left(\tilde{T}_0 I[\tilde{T}_0 > t]\right).$$

Note that $h^*(t) = \sup_{s \ge t} h(s) \to 0$ as $t \to \infty$ if X is taken to be positive. Hence, in particular, we have

$$e_t \le h(t) + m(t - s)h^*(s) + (m(t) - m(t - s))h^*(0),$$

so that, letting $s > 0$ be fixed but arbitrary, it follows that

$$\limsup_{t \to \infty} t^{-1}e_t \le \mu^{-1}h^*(s),$$

where $\mu = \int_0^\infty u\, dF(u)$, proving (d).

(e) $\liminf_{t \to \infty} t^{-1}e_t > 0$, *the desired contradiction.*

From (b) and (c), we have

$$e_t = \mathbb{E}_{j_0}\left\{\mathbb{E}(T_0^{(t)} - t|X(t))\right\} \ge \beta^{-1}\mathbb{E}_{j_0}|g(X(t)) - g(j_0)|$$

$$\ge \beta^{-1}|\mathbb{E}_{j_0} g(X(t)) - g(j_0)| \ge \beta^{-1}(Qg)_{j_0}\mathbb{E}_{j_0} H_t(j_0),$$

where

$$H_t(j_0) = \int_0^t I[X(s) = j_0]\, ds.$$

However, $(Qg)_{j_0} > 0$ and $\lim_{t \to \infty} t^{-1} \mathbb{E}_{j_0} H_t(j_0)$ exists and is positive if X is positive, proving the contradiction. $\square$

REFERENCES

ANDERSON R. M. and MAY, R. M. (1991). *Infectious Diseases of Humans: Dynamics and Control*. Oxford Univ. Press.

ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. Springer, Berlin.

BARBOUR, A. D. (1994). Threshold phenomena in epidemic theory. In *Probability, Statistics and Optimization* (F. P. Kelly, ed.) 101–116. Wiley, New York.

BARBOUR, A. D. and KAFETZAKI, M. (1993). A host–parasite model yielding heterogeneous parasite loads. *J. Math. Biol.* **31** 157–176.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

CHUNG, K. L. (1967). *Markov Chains with Stationary Transition Probabilities*, 2nd ed. Springer, Berlin.

DIEKMANN, O., HEESTERBEEK, J. A. P. and METZ, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28** 365–382.

FOSTER, F. G. (1953). On the stochastic matrices associated with certain queuing processes. *Ann. Math. Statist.* **24** 355–360.

HAMZA, K. and KLEBANER, F. C. (1995). Conditions for integrability of Markov chains. *J. Appl. Probab.* **32** 541–547.

HEESTERBEEK, J. A. P. and DIETZ, K. (1996). The concept of R_0 in epidemic theory. *Statist. Neerlandica* **50**. To appear.

TWEEDIE, R. L. (1975). Sufficient conditions for regularity, recurrence and ergodicity of Markov processes. *Math. Proc. Cambridge Philos. Soc.* **78** 125–136.

TWEEDIE, R. L. (1976). Criteria for classifying general Markov chains. *Adv. in Appl. Probab.* **8** 737–771.

A. D. BARBOUR
ABTEILUNG FÜR ANGEWANDTE MATHEMATIK
UNIVERSITÄT ZÜRICH
WINTERTHURERSTRASSE 190
CH-8057 ZÜRICH
SWITZERLAND
E-MAIL: adb@amath.unizh.ch

J. A. P. HEESTERBEEK
DLO–AGRICULTURAL MATHEMATICS GROUP
P.O. BOX 100
NL-6700 AC WAGENINGEN
THE NETHERLANDS
E-MAIL: j.a.p.heesterbeek@glw.dlo.nl

C. J. LUCHSINGER
ABTEILUNG FÜR ANGEWANDTE MATHEMATIK
UNIVERSITÄT ZÜRICH
WINTERTHURERSTRASSE 190
CH-8057 ZÜRICH
SWITZERLAND
E-MAIL: lynx@amath.unizh.ch