2 The stochastic approach

The results of this chapter can be summarised as follows:

In the non-linear stochastic models (SN and SNM) the epidemic dies out almost surely, no matter what values the parameters take (Theorems 2.2 and 2.7). This behaviour is due to the finite number of individuals in the system. On the other hand, if we let the number of individuals tend to infinity in a way to be specified later, we found threshold results, depending on θ, which give us a better insight of the development of the epidemic (Theorems 2.3 and 2.8). In the linear stochastic models (SL and SLM) we have again threshold results depending on θ (Theorems 2.5 and 2.10). These results are analogous to Theorems 2.3 and 2.8 in the non-linear cases. Besides that we could compute the exact value of the expectation of the number of parasites at any time in the linear stochastic models (equations (2.5) and (2.6)). All these results must be compared with the results using the deterministic approach (chapter 4). All results in the stochastic environment have an analogue in the deterministic environment except Theorems 2.2 and 2.7.

2.1 The stochastic non-linear model without mortality of humans SN

We begin with the stochastic approach in model SN, that is the stochastic, non-linear model without mortality of humans. This model was analysed in Barbour and Kafetzaki (1993)). The first result is frequently used in analysing all eight models and has no obvious epidemiological interpretation.

Lemma 2.1 [Barbour and Kafetzaki (1993), Equation (2.2)] *For all $k \geq 0$ we have*

$$\sum_{i \geq 1} p_{ik} < \infty.$$

The next result shows that in the non-linear stochastic model without mortality of humans, the epidemic dies out with probability one no matter what values the parameters take.

Theorem 2.2 [Barbour and Kafetzaki (1993), Theorem 2.3] *In the model SN the infection dies out with probability one, that is*

$$\mathbb{P}[\lim_{t \to \infty} x^{(0)}(t) = e_0] = 1.$$

Remarks 1. There is no deterministic analogue of Theorem 2.2 (see Theorems 2.3 and 4.2 as a contrast).

2. As a consequence of Theorem 2.2 the process SN is in particular 'regular', in the sense that it makes only finitely many transitions in any finite time interval [0,T] almost surely.

Remarks on the basic reproduction ratios I Let us define $R_0^{(0)} := \lambda\theta/\mu$ and $R_1^{(0)} := (\lambda e \log\theta)/\mu$. These are quantities which emerge as being critical in determining the behaviour of the models without mortality, as, for instance, in Theorem 2.3 below. $R_0^{(0)}$ 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: A worm has an exponentially distributed lifetime with parameter μ which means that his expected lifetime is μ^{-1}. During such a life he makes contacts at a rate of λ per time unit and on average these contacts result in infections with θ worms. We have not yet found a good interpretation for $R_1^{(0)}$ (see "Remarks on the basic reproduction ratios II, III and IV" in this chapter for further discussions of these questions)!

By the expression **threshold behaviour** we *usually* denote general statements of the following type: If $R_0 > 1$ the epidemic develops in deterministic systems and if $R_0 < 1$ the epidemic dies out. In a stochastic environment statements are *usually* such that if $R_0 > 1$ the epidemic has a positive probability to develop and if $R_0 \leq 1$ the epidemic dies out almost surely. We are going to see in chapters 2 and 4 that the situation in our eight models is far more complex than that stated above.

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

Theorem 2.3 *Fix $y \in \mathbb{N}_0^\infty$, such that $0 < Y := \sum_{j\geq 1} y_j < \infty$, and suppose that for each $M > Y$ we have $x_j^{(M,0)}(0) = y_j/M$ for all $j \geq 1$. Then in model SN we have the following threshold behaviour:*

Case 1): $\theta \leq e$. Then

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

Case 2): $\theta > e$. Then

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

Explanation The initial number of infected individuals stays constant and equal to Y; as M tends to ∞, only the initial number of uninfected individuals $Mx_0^{(M,0)} = M - Y$ grows.

Remarks 1) The deterministic analogue of Theorem 2.3 is Theorem 4.2.

2) 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 $\lambda x_0^{(M,0)}$. As we increase M, we only increase the initial number of uninfected individuals. The initial proportion of uninfected tends to 1 as M tends to infinity. So we almost have a linear model (at least in the initial phase). So it is not too surprising, that we have analogous results to those in Theorem 2.5. Note that it is vital to let M converge to infinity first and then we let t converge to infinity. Otherwise these probabilities would be 0 in all cases because of Theorem 2.2.

Proof of Theorem 2.3 The idea of the proof is to show that for fixed M there exists a linear process $X^{(0)}/M$ which is *in all components* larger than our original $x^{(M,0)}$, and such that, the larger we choose M, the more $x^{(M,0)}$ behaves like $X^{(0)}/M$. Then we use Theorem 2.5 (we do *not* use Theorem 2.3 to prove Theorem 2.5).

1. First we have to find that linear process $X^{(0)}$: 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 $X_j^{(0)} := X_j^{(nl)} + X_j^{(r)}$ behaves like SL 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)

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

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

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

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

$$R'(u) \geq I[R(u) > 0]I[u < \tau]. \tag{2.1}$$

As we construct $X^{(r)}$ such that $X^{(0)}$ develops according to SL, 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 (2.1), 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

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

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

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

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

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

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

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

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

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

at rate $\lambda a^2/M$. With this construction (2.1) 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 (2.1).

As only the first two components of this process are important in part 1 of the proof, we repeat for better understanding the last part of the rate at which the first two co-ordinates change, neglecting R':

$$\big(X^{(nl)}, X^{(r)}\big) \to \big(X^{(nl)}, X^{(r)} + e_k\big) \text{ at rate}$$

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

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 SN. Let us look at the sum $X_j^{(0)} := (X^{(nl)} + X^{(r)})_j$ for $j \geq 1$. The development of $X^{(0)}$ is that of SL and it is independent of 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 $X^{(0)}$ does not include the zero co-ordinate, it is independent of M.

2. We now have to examine the limit

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

For all fixed M we introduce the notation $L(u) := \sum_{j\geq 1} X_j^{(0)}(u)$, 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 $R(u)$ depend on M, the distribution of $L(u)$ does not depend on M. We have

$$\mathbb{P}\big[\sum_{j\geq 1} x_j^{(M,0)}(t) = 0\big] = \mathbb{P}\big[\sum_{j\geq 1} X_j^{(nl)}(t) = 0\big] = \mathbb{P}\big[N = 0\big]. \tag{2.2}$$

As $L = N + R$ we have

$$\begin{aligned}\mathbb{P}\big[N = 0\big] &= \mathbb{P}\big[L - R = 0\big]\\ &= \mathbb{P}\big[L - R = 0|R = 0\big]\mathbb{P}\big[R = 0\big] + \mathbb{P}\big[L - R = 0|R > 0\big]\mathbb{P}\big[R > 0\big]\\ &= \mathbb{P}\big[L = 0\big] + \mathbb{P}\big[L - R = 0|R > 0\big]\mathbb{P}\big[R > 0\big]. \end{aligned} \tag{2.3}$$

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 $(X^{(0)}, B)$ such that $X^{(0)}$ is the SL process and behaves as before. Additionally we add a univariate random variable $B \geq 0$. The initial values are $X_j^{(0)}(0) = y_j$ for $j \geq 1$ and $B(0) = 0$ and let us recall that $Y := \sum_{j\geq 1} y_j$. The vector $(X^{(0)}, B)$ changes according to the following rates:

$$\begin{aligned}&\big(X^{(0)}, B\big) \to \big(X^{(0)} + (e_{j-1} - e_j), B\big) \text{ at rate } j\mu X_j^{(0)} \ ; \ j \geq 2,\\ &\big(X^{(0)}, B\big) \to \big(X^{(0)} - e_1, B + 1\big) \text{ at rate } \mu X_1^{(0)} \ ; \ (j = 1),\\ &\big(X^{(0)}, B\big) \to \big(X^{(0)} + e_k, B\big) \text{ at rate } \lambda \sum_{u\geq 1} X_u^{(0)} p_{uk} \ ; k \geq 1,\\ &\big(X^{(0)}, B\big) \to \big(X^{(0)}, B + 1\big) \text{ at rate } \lambda B + \lambda \sum_{u\geq 1} X_u^{(0)} p_{u0}.\end{aligned}$$

As is easily seen, $X^{(0)}$ is still our linear process constructed in step 1. B cancels almost surely every loss of an infected individual in the linear process $X^{(0)}$: an infected individual drops out of the system if a parasite dies in an individual with only one worm and additionally B cancels infections with zero parasites in the linear process $X^{(0)}$ through adding that rate in the fourth 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:

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

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

$$\begin{aligned} \mathbb{P}\big[R > 0\big] &= \mathbb{P}\big[RI_{\{t<\tau\}} + RI_{\{t\ge\tau\}} > 0\big] \\ &\le \mathbb{P}\big[RI_{\{t<\tau\}} > 0\big] + \mathbb{P}\big[RI_{\{t\ge\tau\}} > 0\big] \\ &\le \mathbb{P}\big[RI_{\{t<\tau\}} > 0\big] + \mathbb{P}\big[I_{\{t\ge\tau\}} > 0\big] \\ &\le \mathbb{P}\big[RI_{\{t<\tau\}} > 0\big] + \epsilon. \end{aligned} \tag{2.4}$$

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 (2.1)

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

as the probability that there is no event in the Poisson process until time t is $\exp(-t\lambda a^2/M)$. So, letting M tend to infinity, we have in (2.4), as $\epsilon > 0$ was chosen arbitrarily, that $\lim_{M\to\infty} \mathbb{P}[R > 0] = 0$. Hence, from (2.2) and (2.3) we have

$$\lim_{M\to\infty} \mathbb{P}\big[\sum_{j\ge1} x_j^{(M,0)}(t) = 0\big] = \mathbb{P}\big[L(t) = 0\big].$$

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 ($\theta \le e$ and $R_0^{(0)} \le 1$ or $\theta > e$ and $R_1^{(0)} \le 1$) follow immediately: We can use Theorem 2.5 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 ($\theta \leq e$ and $R_0^{(0)} > 1$ or $\theta > e$ and $R_1^{(0)} > 1$) 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 2.5 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.

□

2.2 The stochastic linear model without mortality of humans SL

In this section the model SL (stochastic, linear, without mortality of humans) is analysed. We first want to be sure that the process SL 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 2.4 *The process $X^{(0)}$ that evolves according to SL is regular.*

Proof of Lemma 2.4 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.

□

Next a result of Barbour (1994) is presented. In that paper the model SL (stochastic, linear, without mortality of humans) is analysed. Theorem 2.5 describes the threshold behaviour in the model SL and gives the expected number of parasites at time t:

Theorem 2.5 *Let us assume that $0 < \sum_{j\geq 1} X_j^{(0)}(0) < \infty$. Then the following result holds:*

Case 1): $\theta \leq e$. Then $\mathbb{P}[\lim_{t\to\infty} \sum_{j\geq 1} X_j^{(0)}(t) = 0] = 1$ if and only if $R_0^{(0)} \leq 1$.

Case 2): $\theta > e$. Then $\mathbb{P}[\lim_{t\to\infty} \sum_{j\geq 1} X_j^{(0)}(t) = 0] = 1$ if and only if $R_1^{(0)} \leq 1$.

In addition, the expected number of parasites in SL grows at an exponential rate $(\lambda\theta - \mu)$:

$$\mathbb{E}[\sum_{j\geq 1} jX_j^{(0)}(t)] = (\sum_{j\geq 1} jX_j^{(0)}(0))e^{(\lambda\theta-\mu)t}. \tag{2.5}$$

Remark The deterministic analogue of Theorem 2.5, cases 1) & 2) is Remark 1 following Theorem 4.8; that of equation (2.5) is equation (4.7).

Proof of Theorem 2.5 Cases 1) and 2) of this theorem have been proven in Barbour (1994) as Theorem 2.1.

(2.5) is proven as follows: Let us define $M(t) := \sum_{j\geq 1} jX_j^{(0)}(t)$. Further we define

$$\begin{aligned} c(X^{(0)}) :&= \sum_{j\geq 1} j\mu X_j^{(0)}\{(j-1)-j\} + \lambda \sum_{k\geq 1}\sum_{j\geq 1} X_j^{(0)} p_{jk} k \\ &= -\mu M + \lambda\theta M \end{aligned}$$

and

$$W(t) := M(t) - M(0) - \int_0^t c(X^{(0)}(u))du.$$

In Corollary A7 of the Appendix we prove that W is a martingale. We have:

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

Now we take the expectation, giving

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

since $W(0) = 0$. As $c(X^{(0)}(u)) = (\lambda\theta - \mu)M(u)$ we have the integral equation

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

where $y(t) := \mathbb{E}[M(t)]$. But this immediately leads to the equation (2.5), completing the proof of Theorem 2.5.

□

Remarks on the basic reproduction ratios II A first important remark that has to be made looking at Theorem 2.5 is as follows: Depending on the value of θ ($\theta > e$) it is possible that $R_0^{(0)} > 1$ and $R_1^{(0)} < 1$. Let us assume we are in such a situation and $\theta > e$. This implies that the epidemic dies out with probability one; but it means too that the expected number of parasites tends to infinity. It is clear that in the *stochastic* model the number of parasites goes to 0 too with probability 1.

Let us look at an analogous situation in model SN. If the number of individuals M is constant, the epidemic finally dies out with probability one (Theorem 2.2). We could ask ourselves whether if $R_0^{(0)} > 1$, then the expected number of parasites tends to infinity, that is $\mathbb{E}[\sum_{j\geq 1} j x_j^{(M,0)}(t)] \to \infty$ for $t \to \infty$? Such a behaviour is suggested through Remark 5) of Theorems 4.3 and 4.25. This question is open.

But instead, let us compare this result with the results of the deterministic approach in chapter 4: in both systems, DN and DL, we have an analogous behaviour (Remark 5 to Theorems 4.3 and 4.25 for DN and for DL it is equation (4.7) and Remark 1 to Theorem 4.8). But in DN and DL it is the number of parasites (and not an expectation as in chapter 2) that tends to infinity. This difference between the results of chapters 2 and 4 is due to the fact that in the deterministic models the number of individuals can be any nonnegative real number, possibly smaller than 1, while in the stochastic models we only have natural numbers.

The remainder of chapter 2.2 comes from Barbour (1994): When $\theta > e$ and $R_1^{(0)} < 1 < R_0^{(0)}$, the expected number of parasites $\mathbb{E}[\sum_{j\geq 1} j X_j^{(0)}(t)]$ increases with t, but, for $\beta = 1/\log\theta$, $\mathbb{E}[\sum_{j\geq 1} j^{\beta} X_j^{(0)}(t)]$ tends to zero (see proof of Theorem 2.5, Case 2, first direction in Barbour (1994)). This suggests that the expected number of parasites is in this case dominated by the possibility of having a few individuals with very large parasite burdens. Thus, to understand why $\lambda e \log\theta/\mu = 1$ emerges as a threshold, we consider what happens to individuals infected by large numbers of parasites. As time goes by, the number of parasites carried by such an individual decreases almost exactly exponentially at rate μ, and from time to time, at rate λ, he causes new infections, each of which starts with almost θ times as many parasites as he currently carries. Thus, on a *logarithmic* scale, his parasite burden decreases almost linearly towards zero at rate μ, and each of those he infects behaves in similar fashion, but with initial burden having a value almost $\log\theta$ greater than his current burden.

This motivates the following definition of a branching process Y with drift. $Y(t)$ describes the positions in $\mathbb{R}_+$ of a random number of particles. Each particle drifts steadily at rate μ towards 0, and is annihilated upon reaching 0. Until this time, it gives birth to further particles according to a Poisson process of rate λ, independently of all other particles. If a particle is born to a parent at position x, it is initially placed at position $x + \log\theta$, and it thereafter behaves according to the same rules governing drift, annihilation and reproduction, independently of all other particles. We are interested in the distribution of $N_Y \leq \infty$, the total number of particles ever in existence. By scaling, we can equivalently take $\lambda' = 1$ and $\mu' = 1$, then setting $d := \lambda\log\theta/\mu$

for the translation at birth. Clearly, the larger the value of d, the larger the values to be expected of N_Y. Let $\mathbb{P}_s$ denote the distribution conditional on starting with a single particle at position s.

Theorem 2.6 [Barbour (1994), Theorem 3.1] *If $d \leq 1/e$, then we have $\mathbb{P}_s[N_Y < \infty] = 1$ for all s, and $\mathbb{E}_d[N_Y] \leq e$. If $d > 1/e$, $\mathbb{P}_s[N_Y < \infty] < 1$ for all s.*

Remark The change at the critical value of d is quite abrupt. When d takes the value $1/e$, not only is N_Y almost surely finite, but its mean is also finite (and equal to e under $\mathbb{P}_d$), although, for any $d > 1/e$, there is a positive probability that $N_Y = \infty$. Note that $d = 1/e$ represents $\lambda e \log \theta/\mu = 1$ in the notation of the original problem. This suggests the interpretation that, for $R_1^{(0)} \leq 1$, the few individuals with large numbers of parasites are unable to support the growth of $X^{(0)}$, but that when $R_1^{(0)} > 1$ they can.

2.3 The stochastic non-linear model with mortality of humans SNM

In this section, the stochastic non-linear model including mortality of humans is analysed. The following theorem should be compared with Theorem 2.2. It shows that in the non-linear cases the epidemic dies out with probability one no matter what values the parameters take.

Theorem 2.7 *In the model SNM the infection dies out with probability one, that is*

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

Define $T_M^{ext} := \inf\{t : x(t) = e_0\}$, the time until the epidemic dies out (extinction). Then

$$\mathbb{E}[T_M^{ext}] \leq 1 + e^{rM}$$

with $r := \lambda - \log(1 - e^{-\kappa})$.

Remarks 1. There is no deterministic analogue of Theorem 2.7 (see Theorems 2.8 and 4.24 as a contrast).

2. As a consequence of Theorem 2.7 the process SNM is in particular 'regular', in the sense that it makes only finitely many transitions in any finite time interval [0,T] almost surely.

Proof of Theorem 2.7 First we find a lower bound for the probability that the epidemic dies out in an arbitrary, single time-interval of length 1, given it has not died out yet. The probability that a given person dies in the next time interval and 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 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 2.7. 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 first part of the proof.

We then immediately have an upper bound for the expected time until the epidemic dies out:

$$\mathbb{E}[T_M^{ext}] \leq 1 + \sum_{n\geq 0} \mathbb{P}[B_n^c] \leq 1 + e^{rM}$$

where $r := \lambda - \log(1-e^{-\kappa})$, completing the proof.

□

Let us define $R_0 := \lambda\theta/(\mu+\kappa)$, $R_1 := (\lambda e \log\theta)/(\mu\theta^{\frac{\kappa}{\mu}})$ and $R_2 := \lambda/\kappa$.

Looking at Theorem 2.7 we see that the epidemic *finally* dies out almost surely in SNM no matter what values the parameters take. But the behaviour of SNM 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 2.8 *Fix* $y \in \mathbb{N}_0^\infty$, *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* $j \geq 1$. *Then in model SNM we have the following threshold behaviour:*

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

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

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

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

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

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

Explanation The initial number of infected individuals stays constant and equal to Y; as M tends to ∞, only the initial number of uninfected individuals $Mx_0^{(M)} = M - Y$ grows.

Remarks 1) The deterministic analogue of Theorem 2.8 is Theorem 4.24.

2) 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. The initial proportion of uninfected tends to 1 as M tends to infinity. So we almost have a linear model (at least in the initial phase). So it is not too surprising, that we have analogous results to those in Theorem 2.10. Note that it is vital to let M converge to infinity first and then we let t converge to infinity. Otherwise these probabilities were 0 in all cases because of Theorem 2.7.

Proof of Theorem 2.8 The proof of this theorem is almost exactly the same as that of Theorem 2.3. The difficult part lies in the infection process (which is the same in both SN and SNM) and the possibility of humans to die (in SNM) does not add any mathematical problems. We use Theorem 2.10 instead of Theorem 2.5 and here again, we do *not* use Theorem 2.8 to prove Theorem 2.10.

□

Remarks on the basic reproduction ratios III Let us have a first look at the basic reproduction ratios R_i: $R_0 := \lambda\theta/(\mu+\kappa)$, $R_1 := (\lambda e \log\theta)/(\mu\theta^{\frac{\kappa}{\mu}})$ and $R_2 := \lambda/\kappa$. Again, as in SN, R_0 denotes the average number of offspring of a single parasite during his whole lifetime in the absence of density dependent constraints (the average lifetime is now $(\mu+\kappa)^{-1}$, since a worm dies too if the person he lives in dies). We do not have an obvious interpretation for R_1. R_2 is the average number of people an infected person infects during his whole lifetime in the absence of density dependent constraints. 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 has been seen, 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.

2.4 The stochastic linear model with mortality of humans SLM

In this section the model SLM (stochastic, linear, including mortality of humans) is analysed. We first want to be sure that the process SLM is ‘regular’, in the sence that it makes only finitely many transitions in any finite time interval [0,T], almost surely. This is shown in the following

Lemma 2.9 *The process X that evolves according to SLM is regular.*

Proof of Lemma 2.9 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.

□

The next result is the analogue of Theorem 2.5 when mortality of humans is included. The threshold behaviour in model SLM is as follows:

Theorem 2.10 *We assume that* $0 < \sum_{j\geq 1} X_j(0) < \infty$. *Then the following results hold:*

Case 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* $R_0 \leq 1$.

Case 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* $R_1 \leq 1$.

Case 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 SLM grows with an exponential rate $(\lambda\theta - \mu - \kappa)$:

$$\mathbb{E}[\sum_{j\geq 1} jX_j(t)] = (\sum_{j\geq 1} jX_j(0))e^{(\lambda\theta-\mu-\kappa)t}. \tag{2.6}$$

Remark The deterministic analogue of Theorem 2.10, cases 1), 2) and 3) is Remark 4 following Theorem 4.18, that of equation (2.6) is equation (4.15).

For the proof of Theorem 2.10 we first need three technical lemmas (Lemmas 2.11, 2.12 and 2.13).

Lemma 2.11 *Recall the definitions of the basic reproduction ratios* R_i *as follows:* $R_0 := \lambda\theta/(\mu+\kappa)$, $R_1 := \lambda e\log\theta/(\mu\theta^{\kappa/\mu})$ *and* $R_2 := \lambda/\kappa$. *Then*

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 2.11 a) This follows from part b) because

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

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 $a := \alpha(\mu/\kappa)$ and use $\log\theta > \mu/\kappa$ we only have to show that $e^a \geq a+1$ which is true.

□

For the following lemma we define

$$g_1(j) := \frac{1}{1+\delta j}$$
$$g_2(j) := \frac{1}{1+\delta j^{\alpha(j)}} \tag{2.7}$$

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 2.12 *$\alpha(j)$ 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 $K \geq e^{e^3}$.

Proof of Lemma 2.12 a)-d) These proofs are simple though partly tedious and only need elementary calculus.

e) In view of b) the first inequality is clear. For the second part we need d) and remember that for $x \leq K$ we have $\alpha'(x) = 0$. Then we can argue as follows:

$$x^{\alpha(cx)-\alpha(x)} = \exp\{(\alpha(cx) - \alpha(x))\log x\} \geq \exp\left\{\left(\frac{-2(c-1)x}{x \log x \log\log K}\right)\log x\right\}$$
$$\geq 1 - \frac{2(c-1)}{\log\log K},$$

which ends the proof.

f) If $x \leq K$, then $\alpha(x) = 1$ and so we have

$$g_2'(x) = -\frac{\delta}{(1+\delta x)^2}$$

and

$$g_2''(x) = \frac{2\delta^2}{(1+\delta x)^3}.$$

So we need to prove that

$$\frac{2\delta^2}{(1+\delta x)^3} \leq k\delta x^{-1},$$

as $\alpha(x) = 1$ for $x \leq K$. This is equivalent to finding a k such that

$$k \geq \frac{2\delta x}{(1+\delta x)^3}.$$

As is easily seen, choosing $k = 2.1$ already satisfies this equation uniformly in $x > 0$, $\delta \le 1$ and $K \ge e^{e^3}$. So we need to examine the second derivative of g_2 for $x > K$ and show that the left and the right limit of g_2, g_2' and g_2'' coincide at $x = K$.

For $x > K$, we first calculate the derivatives of g_2 and α:

$$g_2'(x) = -\frac{\delta x^\alpha}{(1+\delta x^\alpha)^2}\left[\frac{\alpha}{x} + \alpha' \log x\right],$$

where we used $\alpha := \alpha(x)$. Define $A := (\delta x^\alpha)/(1+\delta x^\alpha)^2$ and $B := \big[(\alpha/x) + \alpha' \log x\big]$. Then the second derivative of g_2 is $g_2'' = -[BA' + AB']$, that is

$$g_2''(x) = -\left\{B\left[\frac{\delta x^\alpha(1+\delta x^\alpha)^2 B - 2\delta^2 x^{2\alpha}(1+\delta x^\alpha)B}{(1+\delta x^\alpha)^4}\right]\right.$$
$$\left. + A\left[\frac{\alpha' x - \alpha}{x^2} + \alpha'' \log x + \frac{\alpha'}{x}\right]\right\}$$

We can write this in a slightly different way:

$$g_2''(x) = \frac{\delta x^\alpha}{(1+\delta x^\alpha)^2}\left[\frac{2\delta x^\alpha B^2}{(1+\delta x^\alpha)} - B^2 - \alpha'' \log x - 2\frac{\alpha'}{x} + \frac{\alpha}{x^2}\right]$$

We must show that $g_2''(x) \le k\delta x^{\alpha-2}$. This is equivalent to show that

$$\frac{2\delta x^{\alpha+2} B^2}{(1+\delta x^\alpha)} - (Bx)^2 - \alpha'' x^2 \log x - 2x\alpha' + \alpha \le k(1+\delta x^\alpha)^2. \tag{2.8}$$

The second term $(-(Bx)^2)$ on the left side of (2.8) is negative and therefore causes no problems. The last term (α) is bounded by 1 and therefore does not cause any problems either. Then, by Lemma 2.12 d), we know that $0 \le -\alpha'(x) \le 2/(x \log x \log\log K)$ which shows that the fourth term $(-2x\alpha')$ does not cause any problems either. We therefore only have to show that there exists a constant k such that

$$\frac{2\delta x^{\alpha+2} B^2}{(1+\delta x^\alpha)} - \alpha'' x^2 \log x \le k(1+\delta x^\alpha)^2$$

Using $B^2 = \frac{\alpha^2}{x^2} + 2\frac{\alpha\alpha' \log x}{x} + (\alpha' \log x)^2$, this leads to

$$\frac{2\delta x^{\alpha+2}\big(\frac{\alpha^2}{x^2} + 2\frac{\alpha\alpha' \log x}{x} + (\alpha' \log x)^2\big)}{(1+\delta x^\alpha)} - \alpha'' x^2 \log x \le k(1+\delta x^\alpha)^2. \tag{2.9}$$

Using Lemma 2.12 d) once again, we see that the first term of (2.9) is under control too. So we only need to prove that there is a $k > 2$ such that

$$-\alpha'' x^2 \log x \leq k(1 + \delta x^{\alpha})^2$$

Now we must calculate the second derivative of α. For the first derivative we have

$$\alpha'(x) = -2(1-\alpha_*)\left(1 - \frac{\log\log K}{\log\log x}\right)\frac{\log\log K}{(\log\log x)^2}\frac{1}{\log x}\frac{1}{x}.$$

The second derivative of α is

$$\alpha''(x) = -2(1-\alpha_*)\frac{(\log\log K)^2}{(\log\log x)^4}\frac{1}{(\log x)^2}\frac{1}{x^2}$$

$$+\frac{2(1-\alpha_*)}{(\log\log x)^2}\left(1 - \frac{\log\log K}{\log\log x}\right)\left\{\left(\frac{\log\log K}{x^2\log x} + \frac{\log\log K}{(x\log x)^2}\right) + \frac{2\log\log K}{(x\log x)^2(\log\log x)}\right\}$$

Only the first (negative) term of α'' is of interest for us as we need to find a $k > 2$ such that

$$-\alpha'' x^2 \log x \leq k(1 + \delta x^{\alpha})^2.$$

But such a k exists obviously.

The reader can easily check that the left and the right limit of g_2, g_2' and g_2'' coincide at $x = K$ finishing the proof.

□

Lemma 2.13 *a) Suppose S_j, $j \geq 1$, has the distribution F_j (see chapter 1), that is $S_j := \sum_{i=1}^{j} Y_i$ and the Y_i are independent and identically distributed with mean θ and variance σ^2 : $\mathbb{P}[S_j = k] = p_{jk}$. Define $g_1(j) := (1+\delta j)^{-1}$, $j \geq 0$, for $\delta > 0$. Then 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 2.12 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 2.12 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 $\alpha(j) \le 2\alpha_* < 1/3$ 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

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

Remark Lemma 2.13 allows us in four situations to replace $\mathbb{E}[g(S_j)]$ by $g(j\theta)$ with only small impact.

Proof of Lemma 2.13 a) We have to prove that

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

For any $x, y \ge 0$,

$$\begin{aligned}\frac{1}{1+\delta x} - \frac{1}{1+\delta y} &= \frac{\delta(y-x)}{(1+\delta x)^2} - \frac{\delta^2(y-x)^2}{(1+\delta x)^2(1+\delta y)} \\ &\ge \frac{\delta(y-x)}{(1+\delta x)^2} - \frac{\delta^2(y-x)^2}{(1+\delta x)^2},\end{aligned}$$

so that

$$1 - g_1(y) = 1 - \frac{1}{1+\delta x} + \frac{1}{1+\delta x} - \frac{1}{1+\delta y} \ge \frac{\delta x}{(1+\delta x)} + \frac{\delta(y-x)}{(1+\delta x)^2} - \frac{\delta^2(y-x)^2}{(1+\delta x)^2}.$$

Hence, taking $y = S_j$ and $x = j\theta$, it follows that

$$\begin{aligned}1 - \mathbb{E}[g_1(S_j)] &\ge \frac{\delta j\theta}{1+\delta j\theta} - \frac{\delta^2 j\sigma^2}{(1+\delta j\theta)^2} \\ &= \frac{\delta j\theta}{1+\delta j\theta}\left\{1 - \frac{\delta\sigma^2}{\theta(1+\delta j\theta)}\right\}.\end{aligned}$$

b) Take any $X \in (0, K)$, and consider the parabola

$$y(x) := (1 - g_2(X)) - (x - X)g_2'(X) - \frac{1}{2}(x-X)^2 k\delta/K,$$

for k as in Lemma 2.12 f). We show that $y(x) \le 1 - g_2(x)$ for all x, *independently* of the choice of X. It is immediate that $1 - g_2(x) \ge 0$ for all x. Then, as the leading term of the parabola has a negative sign, we can argue as follows: First, the smaller root of $y(x) = 0$ is at least as large as

$$X_2 := X - \sqrt{\frac{2(1 - g_2(X))}{k\delta/K}}.$$

So, for all $x \leq X_2$ we have $y(x) \leq 1 - g_2(x)$ because there $y(x) \leq 0$. Second, $y(X) = 1 - g_2(X)$ and $y'(X) = -g_2'(X)$. We thus have to check that $y''(x) \leq -g_2''(x)$ for all $x > X_2$ and we are finished.
For $x < K$ we have $g_2''(x) \leq 2\delta^2(1 + \delta x)^{-3} \leq 2\delta^2$. Now since $\delta \leq k/(2K)$ we therefore have $\sup_{x<K} g_2''(x) \leq k\delta/K$. By Lemma 2.12 f) we also have $\sup_{x\geq K} g_2''(x) \leq k\delta/K$, and so $-g_2''(x) \geq -k\delta/K = y''(x)$ for all $x > 0$. Hence

$$1 - g_2(x) \geq (1 - g_2(X)) - (x - X)g_2'(X) - \frac{1}{2}(x - X)^2 k\delta/K$$

for all x and X. Now choose $x = S_j$ and $X = j\theta$, giving

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

because $j\theta \leq K$ and $\delta \leq k/(2K)$, which ends the proof of b).

c) and d) For results c) and d) we need some preparation just as in b): Take any $X > 0$, and consider the parabola

$$z(x) := (1 - g_2(X)) - (x - X)g_2'(X) - \frac{1}{2}(x - X)^2 s(k)k\delta X^{\alpha(X)-2}$$

for $s(k)$ a constant yet to be determined and k as in Lemma 2.12 f). We show that $z(x) \leq 1 - g_2(x)$ for all x, if $s(k)$ is chosen large enough, *independently* of the choice of X. It is immediate that $1 - g_2(x) \geq 0$ for all x. Then, as the leading term of the parabola has a negative sign, we can argue as follows: First, the smaller root of $z(x) = 0$ is at least as large as

$$X_1 := X - \sqrt{\frac{2(1 - g_2(X))}{s(k)k\delta X^{\alpha(X)-2}}} \geq X - \sqrt{\frac{2X^2}{s(k)k}} = X\left(1 - \sqrt{\frac{2}{s(k)k}}\right).$$

So, for all $x \leq X_1$ we have $z(x) \leq 1 - g_2(x)$ because there $z(x) \leq 0$. Second, $z(X) = 1 - g_2(X)$ and $z'(X) = -g_2'(X)$. We thus have to check that $z''(x) \leq -g_2''(x)$ for all $x > X_1$ and we are finished. By Lemma 2.12 f) we have $g_2''(x) \leq k\delta X_1^{\alpha(X_1)-2}$ for all $x \geq X_1$ and so

$$z''(x) = -s(k)k\delta X^{\alpha(X)-2} \leq -k\delta X_1^{\alpha(X_1)-2} \leq -g_2''(x)$$

if we can show that

$$k\delta X_1^{\alpha(X_1)-2} \leq s(k)k\delta X^{\alpha(X)-2}. \tag{2.10}$$

So we need to prove (2.10). Pick any $s := s(k)$ such that $sk \geq 8$: then $X_1 \geq X/2$, and hence, from Lemma 2.12 b) and e),

$$1 \geq X^{\alpha(X)-\alpha(X_1)} \geq X^{\alpha(X)-\alpha(X/2)} \geq 1 - \frac{2}{\log\log K} \geq \frac{1}{3},$$

uniformly in $X > 0$, provided that $K > e^{e^3}$. Then, for (2.10) to hold, it is enough that

$$\left(1 - \sqrt{\frac{2}{sk}}\right)^{-2} \leq \frac{s}{3},$$

as can be arranged by picking $s = s(k)$ larger if necessary. For this choice of s,

$$1 - g_2(x) \geq (1 - g_2(X)) - (x - X)g_2'(X) - \frac{1}{2}(x - X)^2 s(k)k\delta X^{\alpha(X)-2}, \tag{2.11}$$

whatever the value of x and X.

After this preparation we can proceed to c), and d):
c) Now take $X = j\theta$ and $x = S_j$ in (2.11). This yields

$$1 - g_2(S_j) \geq (1 - g_2(j\theta)) - (S_j - j\theta)g_2'(j\theta) - \frac{1}{2}(S_j - j\theta)^2 s(k)k\delta(j\theta)^{\alpha(j\theta)-2},$$

and hence

$$1 - \mathbb{E}[g_2(S_j)] \geq (1 - g_2(j\theta)) - \frac{1}{2}j\sigma^2 s(k)k\delta(j\theta)^{\alpha(j\theta)-2}. \tag{2.12}$$

For $\delta(j\theta)^{\alpha(j\theta)} \leq 1$ we have

$$(1 - g_2(j\theta)) - \frac{1}{2}j\sigma^2 s(k)k\delta(j\theta)^{\alpha(j\theta)-2} \geq \frac{\delta(j\theta)^{\alpha(j\theta)}}{1 + \delta(j\theta)^{\alpha(j\theta)}}\left\{1 - \frac{ks(k)\sigma^2}{\theta^2 j}\right\}$$

and therefore c) follows from (2.12).

d) In the situation of d), using Lemma 2.12 e), we have

$$(1 - g_2(j\theta)) - \frac{1}{2}j\sigma^2 s(k)k\delta(j\theta)^{\alpha(j\theta)-2} \geq \frac{\delta(j\theta)^{\alpha(j\theta)}}{1 + \delta(j\theta)^{\alpha(j\theta)}}\left\{1 - O(j^{-2/3})\right\}$$

and therefore d) follows from (2.12).

□

Proof of Theorem 2.10 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) Let us define $M_\beta(t) := \sum_{j\geq 1} j^\beta X_j(t)$. Further we define

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

and

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

In Corollary A8 of the Appendix we prove that for each $0 < \beta \leq 1$, $W_\beta(t)$ is an $\mathcal{I}_t$-martingale where $\mathcal{I}_s := \sigma\{X(u), 0 \leq u \leq s\}$.

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 ($W(0) = 0$):

$$M(t) = W(t) + M(0) + \int_0^t c(X(u))du. \tag{2.13}$$

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

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

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

$$\mathbb{E}[M(t)|\mathcal{I}_s] = M(s) + \mathbb{E}[\int_s^t c(X(u))du|\mathcal{I}_s]. \tag{2.14}$$

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

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

So $\mathbb{E}[M(t)|\mathcal{I}_s] \leq M(s)$ 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), 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 (2.14). 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 (2.14). So we have

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

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

$$\begin{aligned} c_\beta(X) &= \sum_{j\geq 1} j\mu X_j\{(j-1)^\beta - j^\beta)\} \\ &\quad + \lambda \sum_{l\geq 1} \sum_{j\geq 1} X_j p_{jl} l^\beta - \kappa \sum_{j\geq 1} j^\beta X_j \\ &\leq (\lambda\theta^\beta - \beta\mu - \kappa) \sum_{j\geq 1} j^\beta X_j. \end{aligned}$$

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 $(\lambda\theta^{\beta} - \mu\beta - \kappa) \leq 0$ 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 $R_0 > 1$, $R_1 \leq 1$ and $(1+\kappa/\mu)^{-1} < \log\theta \leq \mu/\kappa$, 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 ($R_1 \leq 1$) 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 ($R_2 \leq 1$) 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 $\theta \leq 1$, and in B3) we treat the remaining case ($\theta > 1$).

B1) We think in terms of a discrete generation branching process with types $j = 1, 2, 3, \ldots$. 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.

We first want to explain why it is justified to examine this *discrete time* branching process with such a structure instead of our original process X. The final aim is to show that the process X does not die out with probability one in the cases where the relevant R_i's are greater than 1. We can easily see that each of the two processes, X and the discrete branching process, eventually becomes extinct whenever the other one does. Suppose X dies out at a time t_0. Now as X is a regular process by Lemma 2.9, with probability one there are only finitely many transitions in that process. So there can only be finitely many transitions in the discrete branching process and therefore it dies out too. On the other hand, we suppose the discrete branching process dies out by generation n. Then process X must eventually die out too, except if there is at least one individual that remains alive but makes no transitions. But this means that an exponentially distributed random variable with a rate of at least $(\lambda + \kappa + \mu)$ does not have a finite value with positive probability which is not possible. So we may examine the discrete process we constructed above.

Without loss of generality, we begin with only one infected individual with j parasites. This is justified because of the linearity of the process X; we show that even so the probability of extinction is smaller than 1. 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

$$(Tf)(j) = [j\mu/(\lambda + j\mu + \kappa)]f(j-1) + [\lambda/(\lambda + j\mu + \kappa)]f(j)\mathbb{E}[f(S_j)] \\ + [\kappa/(\lambda + j\mu + \kappa)],\ j \geq 1,$$

where S_j has the distribution F_j (see chapter 1), that is $S_j := \sum_{i=1}^{j} Y_i$ and the Y_i are independent and identically distributed with mean θ and variance $\sigma^2 : \mathbb{P}[S_j = k] = p_{jk}$. 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 2.11 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, where $f(j) := (1-(\kappa/\lambda))g(j) + (\kappa/\lambda)$. 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:

$$j\mu\big(g(j-1) - g(j)\big) + \kappa\big(1 - g(j)\big) \leq (1 - \mathbb{E}[g(S_j)])\big(\kappa - g(j)\kappa + \lambda g(j)\big), \quad (2.15)$$

and if we talk about a g satisfying condition (2.15), 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 (2.15) 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 (2.15) is satisfied. We try $g_1(j) := (1+\delta j)^{-1}$, as defined in (2.7), for $\delta > 0$ to be chosen later. With this choice of g and using Lemma 2.13 a) we see that (2.15) 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 term equivalent to

$$\mu\delta\frac{(\theta+\sigma^2-\delta\sigma^2)+j(\theta^2-\theta+\sigma^2\delta+\delta\theta^2)+j^2(\delta\theta^3-\delta\theta^2)}{(\theta-\delta\sigma^2-\delta\theta+\delta^2\sigma^2)+j(\delta\theta^2+\delta\theta-\sigma^2\delta^2-\delta^2\theta^2)+j^2\theta^2\delta^2} + \kappa\delta\frac{\sigma^2+j\delta\theta\sigma^2}{\theta+\delta j\theta^2-\delta\sigma^2}\leq c. \tag{2.16}$$

We now examine the first term of the left side of (2.16). 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 (2.16) 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, (2.16) 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 (2.15) is satisfied.

In this part we cannot choose the simple function $g_1(j) := (1 + \delta j)^{-1}$ as before, because (2.15) is not satisfied for all j no matter how we choose δ. Instead we define as in (2.7)

$$g_2(j) := \frac{1}{1 + \delta j^{\alpha(j)}}$$

for $\delta > 0$ to be chosen later 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 $\alpha_* < 1/6$ is made smaller if necessary later on. Besides that, in what follows, δ is always smaller than 1 and $K \geq e^{e^3}$ even if we do not mention it every time.

We frequently force δ to be small, depending on some parameters such as K. On the other hand we force K to be large, depending on various parameters. The reader can easily check that we *never force K to be large depending on δ* because this could lead to contradictions: in fact, we first choose an α_* and then construct α with a final K, then we choose a J (see B3.2)) *and then* choose δ appropriately, although these steps are mixed together in the proof!

This 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 2.12.

With this choice of g we see that (2.15) is satisfied if

$$j\mu\left[\frac{g_2(j-1)}{g_2(j)} - 1\right] + \kappa\delta j^{\alpha(j)} \leq (1 - \mathbb{E}[g(S_j)])(\kappa\delta j^{\alpha(j)} + \lambda) \tag{2.17}$$

is satisfied. Again, if we talk about a g satisfying condition (2.17), 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 2.13 b), it is enough to show that

$$\mu + \kappa \leq \frac{\theta}{1 + \delta j \theta} \left\{ 1 - \frac{k^2 \sigma^2}{\theta K} \right\} (\kappa \delta j + \lambda) \tag{2.18}$$

for (2.17) 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 $c := \lambda\theta - \mu - \kappa > 0$. (2.18) is then equivalent to

$$\mu \delta j \theta + \frac{\kappa \sigma^2 k^2 \delta j}{K} + \frac{\lambda \sigma^2 k^2}{K} \leq c. \tag{2.19}$$

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 (2.19) 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

$$\frac{g_2(j-1)}{g_2(j)} - 1 \leq \delta\big(j^{\alpha(j)} - (j-1)^{\alpha(j-1)}\big) \leq \delta \alpha(j-1)(j-1)^{\alpha(j-1)-1}. \tag{2.20}$$

We choose $\delta < \delta_3 := \min\big(\delta_2\ ,\ (KJ\theta)^{-1}\big)$. Then Lemma 2.13 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 2.13 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 (2.20) we see that (2.17) 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

$$j\mu\alpha(j-1)(j-1)^{\alpha(j-1)-1} + \kappa j^{\alpha(j)} \leq \lambda(j\theta)^{\alpha(j\theta)}(1 - O(K^{-1})) \tag{2.21}$$

is satisfied. If we can show that

$$\frac{\lambda(j\theta)^{\alpha(j\theta)}(1-O(K^{-1}))}{j\mu\alpha(j-1)(j-1)^{\alpha(j-1)-1}+\kappa j^{\alpha(j)}} \geq \frac{\lambda\theta^{\alpha(j)}}{\mu\alpha(j)+\kappa}(1-O((\log\log K)^{-1})) \geq 1, \tag{2.22}$$

then (2.21) is satisfied. The last inequality of (2.22) is surely true by Lemma 2.11 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:

$$\alpha(j)(j\theta)^{\alpha(j\theta)}(1-O(K^{-1})) \geq \theta^{\alpha(j)}(1-O((\log\log K)^{-1}))j\alpha(j-1)(j-1)^{\alpha(j-1)-1} \tag{2.23}$$

and

$$(j\theta)^{\alpha(j\theta)}(1-O(K^{-1})) \geq \theta^{\alpha(j)}(1-O((\log\log K)^{-1}))j^{\alpha(j)}. \tag{2.24}$$

Equation (2.23) is satisfied because the following three relations (2.25), (2.26) and (2.27) hold. Because of Lemma 2.12 d), we have

$$\begin{aligned}\frac{\alpha(j)}{\alpha(j-1)} &= 1-\big(\alpha(j-1)-\alpha(j)\big)\frac{1}{\alpha(j-1)}\\ &\geq 1-\frac{2}{\alpha(j-1)(j-1)\log(j-1)\log\log K}.\end{aligned} \tag{2.25}$$

Then, again by Lemma 2.12 d), we have

$$\begin{aligned}\theta^{\alpha(j\theta)-\alpha(j)} &= \exp([\alpha(j\theta)-\alpha(j)]\log\theta)\\ &\geq 1+[\alpha(j\theta)-\alpha(j)]\log\theta \geq 1-\log\theta\frac{2j(\theta-1)}{j\log j\log\log K}.\end{aligned} \tag{2.26}$$

Finally, again by Lemma 2.12 d) we can derive

$$\begin{aligned}\frac{j^{\alpha(j\theta)-1}}{(j-1)^{\alpha(j-1)-1}} &\geq (j-1)^{\alpha(j\theta)-\alpha(j-1)}\\ &= \exp\big(\log(j-1)[\alpha(j\theta)-\alpha(j-1)]\big)\\ &\geq 1+\log(j-1)[\alpha(j\theta)-\alpha(j-1)]\\ &\geq 1-\log(j-1)\frac{2(1+j(\theta-1))}{(j-1)\log(j-1)\log\log K}.\end{aligned} \tag{2.27}$$

Therefore (2.23) is satisfied. Furthermore, (2.26) and Lemma 2.12 e) show immediately that (2.24) 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 2.12 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 2.13 d) is weaker than the inequality of Lemma 2.13 c). So, after making δ even smaller if necessary, we may use

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

during the whole part of B3.3). Again, for the last time we want inequality (2.17) to be satisfied. All we need to show is therefore that

$$\begin{aligned} j\mu \left(\frac{(j-1)^{\alpha(j-1)}}{1+\delta(j-1)^{\alpha(j-1)}} \right) \frac{\alpha(j-1)}{(j-1)} + \kappa j^{\alpha(j)} \\ \leq \frac{(j\theta)^{\alpha(j\theta)}}{1+\delta(j\theta)^{\alpha(j\theta)}} \Big\{ 1 - O(j^{-2/3}) \Big\} (\kappa\delta j^{\alpha(j)} + \lambda). \end{aligned} \tag{2.28}$$

We want to get rid of the denominators: Equation (2.28) 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

$$\begin{aligned} & j\mu(j-1)^{\alpha(j-1)}\alpha(j-1) + j\mu(j-1)^{\alpha(j-1)}\alpha(j-1)\delta j^{\alpha(j\theta)}\theta^{\alpha(j\theta)} \\ & + \kappa j^{\alpha(j)}(j-1) + \kappa j^{\alpha(j)}\delta(j-1)^{\alpha(j-1)+1} \\ & \leq \big(1 - O(j^{-2/3})\big) \Big((j-1)j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}\lambda + j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}\lambda(j-1)^{\alpha(j-1)+1}\delta \Big) \\ & - O(j^{-2/3})\big(\kappa j^{\alpha(j)+\alpha(j\theta)}\delta\theta^{\alpha(j\theta)}(j-1) + \kappa j^{\alpha(j)+\alpha(j\theta)}(j-1)^{\alpha(j-1)+1}\delta^2\theta^{\alpha(j\theta)}\big). \end{aligned} \tag{2.29}$$

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

$$\begin{aligned} & j\mu\alpha(j-1)j^{\alpha(j\theta)}\theta^{\alpha(j\theta)} + \kappa j^{\alpha(j)}(j-1) \\ & \qquad \leq \left(1 - O(j^{-2/3})\right)\left(j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}\lambda(j-1)\right) \\ & \qquad\quad - O(j^{-2/3})\kappa j^{\alpha(j)+\alpha(j\theta)}(j-1)\delta\theta^{\alpha(j\theta)}, \end{aligned} \tag{2.30}$$

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

$$\begin{aligned} & j\mu(j-1)^{\alpha(j-1)}\alpha(j-1) + \kappa j^{\alpha(j)}(j-1) \\ & \qquad \leq \left(1 - O(j^{-2/3})\right)\left((j-1)j^{\alpha(j\theta)}\theta^{\alpha(j\theta)}\lambda\right) \\ & \qquad\quad - O(j^{-2/3})\kappa j^{\alpha(j)+\alpha(j\theta)}\delta\theta^{\alpha(j\theta)}(j-1). \end{aligned} \tag{2.31}$$

The separation of inequality (2.29) is such that in inequality (2.30) 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 (2.31) 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 (2.30) is satisfied. We divide inequality (2.30) by $j^{1+\alpha(j)}$. Then it is enough to show that the following inequality is satisfied:

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

We can apply Lemma 2.12 e) to the right hand side, showing that it is enough to have

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

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 (2.30) is satisfied also.
We now have to show that (2.31) is satisfied too. But (2.31) is almost the same as (2.30); 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

$$j^{\alpha(j-1)-\alpha(j\theta)} = \exp(\log j[\alpha(j-1) - \alpha(j\theta)])$$
$$\leq \exp\left(\frac{2\log j(1 + j(\theta - 1))}{(j-1)\log(j-1)\log\log K}\right)$$

which is near 1 for K large and is therefore finally smaller than $\theta^{\alpha(j\theta)}$. This shows that (2.31) is satisfied too. This ends the proof of B3.3) and therefore the proof of part B).

C) We can use equation (2.13) ($\beta = 1$) and take the expectation. We thus get

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

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 (2.6) which finishes the proof of Theorem 2.10.

□

Remarks on the basic reproduction ratios IV A first important remark that has to be made looking at Theorem 2.10 is as follows: Depending on the value of θ it is possible that $R_0 > 1$ and $R_1 < 1$ or $R_2 < 1$ respectively. Let us assume we are in such a situation and $\log\theta > (1 + \kappa/\mu)^{-1}$. This implies that the epidemic dies out with probability one; but it means too that the expected number of parasites tends to infinity. It is clear that in the *stochastic* model the number of parasites goes to 0 too with probability 1.

Let us look at an analogous situation in model SNM. If the number of individuals M is constant, the epidemic finally dies out with probability one (Theorem 2.7). We could ask ourselves whether if $R_0 > 1$, then the expected number of parasites tends to infinity, that is $\mathbb{E}[\sum_{j\geq1} jx_j^{(M)}(t)] \to \infty$ for $t \to \infty$? Such a behaviour is suggested through Remark 5) of Theorem 4.25. This question is open.

But instead, let us compare this result with the results of the deterministic approach in chapter 4: in both systems, DNM and DLM, we have an analogous behaviour (Remark 5 to Theorem 4.25 for DNM and for DLM it is equation (4.15) and Remark 4 to Theorem 4.18). But in DNM and DLM it is the explicit number of parasites (and not an expectation as in chapter 2) that tends to infinity. This difference between the results of chapters 2 and 4 is due to the fact that in the deterministic models the number of individuals can be any nonnegative real number while in the stochastic models we only have natural numbers.