University of Amsterdam

MSc Mathematics

Master Thesis

The SIR model: a queueing theoretic approach

Author:
Gijs Bijsterbosch

Supervisor:
prof. dr. R. Núñez Queija

Examination date:
December 20, 2021

Korteweg-de Vries Institute for
Mathematics

Abstract

In epidemic analysis the SIR model is a key tool used to understand the progression of a disease that spreads from person to person. The SIR model uses a system of partial differential equations (PDEs) to capture the dynamics of the epidemic spread. While the solution to the PDEs provides important understanding of the spread of the disease, the simplicity of the SIR model fails to accurately capture the discrete and stochastic nature of real-life epidemics. In this thesis we address this shortcoming by developing a stochastic and discrete model of disease spread and analyse its properties using a queueing theoretic framework. Such analysis defies easy results. We use simulations to build intuition of the inherent randomness in the way the disease spreads. Our analysis focuses on large populations and our mathematical results are formulated in limit theorems for a population size that grows to infinity. Our approach uses tools from the theory of fluid limits for stochastic processes. We will show that when scaling the population, two different regimes determine the behaviour of the system. The first regime can be referred to as the pre-fluid regime: it captures the stochastic dynamics until the scale of the epidemics is large enough for the second phase to manifest itself. In the second phase, the dynamics show much less variance and can be studied using the aforementioned fluid limits. Due to the randomness in the first phase, two important characteristics of the epidemic spread retain their stochastic nature for large population sizes. The first is whether or not the disease is able to spread to a large ("fluid") scale. If the disease spreads to a fluid scale, the time at which the outbreak peaks remains stochastic as well. Interestingly the height of the peak (the maximum number of simultaneous infections) is quite invariable.

Title: The SIR model: a queueing theoretic approach
Author: Gijs Bijsterbosch, gijs2b@gmail.com, 10550569
Supervisor: prof. dr. R. Núñez Queija
Second Examiner: dr. Michel Mandjes
Examination date: December 20, 2021

Korteweg-de Vries Institute for Mathematics
University of Amsterdam
Science Park 105-107, 1098 XG Amsterdam
`http://kdvi.uva.nl`

Preface

Initially I got the idea for this type of project hearing the words Covid and queue in the news repeatedly, there were queues to get tested, queues for the vaccin and so on. Thinking about a queue fed by an epidemic immediately spawned interesting ideas, for instance the total population was finite, and the people in the queue for a test would no longer participate in the infection process if we assume they would self quarantine well enough, then the test queue itself would maybe be a deterministic time queue feeding back into the general population to simulate the time one needed to wait for a result. A good start would be to just capture a reasonable model of the epidemic in queueing theoretic terms and then make the model more complex by allowing for a testing queue component, a vaccine queue component etc. This thesis is just an attempt at that first step.

This first step turned out to be difficult and interesting enough to occupy all the time allotted to me by the 34 EC one has for a master thesis. One of the initial inspirations was peer to peer models where the sharing of a file by many peers follows the basic dynamics of a virus: a peer starts sharing the file once he has received it, like an infected person a virus. While I couldn't directly apply many of these papers, wading through them has served as an inspiration for the ideas presented here. Thanks to this I have gained some understanding of the topology of this area of mathematics, which problems are easy, which results are powerful and which problems are still hard. Before we start I would like to give my thanks.

First of all I would like to thank the UvA and Mastermath for providing the high level of mathematics education that I have enjoyed over the past few years. I would like to especially thank a few teachers who have defined this master for me, Sonja Cox, Peter Spreij, Sindo Nuñez Queija, Michel Mandjes and Bas Kleijn. Furthermore I would like to thank all my peers who I studied and worked with throughout my Bachelor and Master, I certainly wouldn't have gotten here without your help. Then I would like to thank my supervisor whose endless ideas and support made this thesis possible. My parents and brother who were always there for me, and finally Maud Suzat whose infinite patience and support I could always count on.

Notations

$\mathbb{R}_+$	$[0, \infty)$
$\mathbb{N}$	$[0, 1, ...)$
$\mathcal{N}_\zeta$	Poisson process at rate ζ
λ	Arrival rate, or rate of infection
ν	Departure rate, or rate of curing
ρ	Utilisation, or effective utilisation
ρ_0	$\frac{\lambda}{\nu}$
$(a, ..., z)^t$	The vector $(a, ..., z)$ transposed

Contents

1 Introduction

> Rather, a top-down approach to the study of complex entities needs to be complemented with a bottom-up approach: analysis needs to go hand in hand with synthesis.
>
> Manuel Delanda

When modelling a certain phenomenon one typically has two approaches: one could investigate the data in a top-down fashion, summarised and gathered in a statistic, or one could attempt a bottom-up approach, to model the underlying phenomena of which the statistic is a summary. In this thesis we attempt to set up and study an underlying model that would be the bottom-up mirror of the top-down SIR model.

This thesis is concerned with the number of infections in a finite population as an epidemic spreads. In this area top-down modelling is the norm, specifically the SIR model is so widely used it is sometimes referred to as "the curve".

The SIR model (SIR is an acronym for susceptible, infected, recovered), models the outbreak of an epidemic by sorting a population into three compartments labeled susceptible, which is the category of people who haven't had the infection and might get infected; infected, which is the category currently infected and the source of any further infections; and recovered, which are people who have been infected and are no longer susceptible. Many variations of this type of model are possible by adding further compartments that sort the population further. We focus our attention on the most basic version with just the 3 categories, and further assume that the dynamics don't change during the course of the pandemic. When trying to incorporate changing dynamics as a result of natural conditions such as the weather or intentional (government) interventions our analysis can be applied to each stage of the pandemic separately. The population is taken to be constant, so letting $(s(t), i(t), r(t))$ be the amount of people in each category at time t we have $s(t) + i(t) + r(t) = N$.

The SIR model is a system of partial differential equations

$$s'(t) = -\lambda s(t)i(t)/N \tag{1.1}$$
$$i'(t) = \lambda s(t)i(t)/N - \nu i(t) \tag{1.2}$$
$$r'(t) = \nu i(t) \tag{1.3}$$

that uniquely solve a $(s(t), i(t), r(t))$, we in chapter 2 fully introduce this model. N is the population parameter, λ is the parameter for the rate of infection and ν the parameter for the rate of healing. The solution is continuous and deterministic. Any infection that spreads from person to person is discrete and stochastic: an infected person comes into contact with a random number of people, of which he infects a randomly distributed fraction.

To take these dynamics into account we introduce in detail the SIR queue in 2.2. We define a population sorted into the compartments susceptible, infected and recovered and a contact process that selects two people from the population, homogeneously. When an

infected person comes into contact with a susceptible person according to this process the susceptible person leaves the susceptible category and joins the infected category with a fixed probability. The instances of contact are Poisson distributed on the time parameter $t \geq 0$. A person infected remains infected for an exponentially distributed amount of time. The resulting process is a stochastic network given by the following dynamics

$$S(t) \to S(t) - 1 \text{ at rate } \lambda S(t)I(t)/N \tag{1.4}$$

$$I(t) \to \begin{cases} I(t) + 1 & \text{at rate } \lambda S(t)I(t)/N \\ I(t) - 1 & \text{at rate } \nu I(t) \end{cases} \tag{1.5}$$

$$R(t) \to R(t) + 1 \text{ at rate } \nu I(t). \tag{1.6}$$

We call this stochastic network the SIR *queue*. The equality $S(t) + I(t) + R(t) = N$ fixes the third component whenever the other two are known, this reduces the system to two dimensions.

We point out that these dynamics appear more frequent than it might seem. Take the example of customers buying an orange hat during a football tournament. If the behaviour of buying an orange hat spreads through seeing another person with a hat we are dealing with a finite amount of interested customers, whose hat buying is determined by a contact between people without a hat and with a hat. The growth dynamics as described above would be an appropriate model.

Investigating this stochastic network is the essential thrust of this thesis. We are initially agnostic about our method of investigation and attempt several elementary methods in chapter 2 with mixed succes. Then in chapter 3 we numerically calculate and run simulations to build an intuitive understanding of the SIR queue's behaviour especially as we scale the population N. In chapter 4 we build the most productive framework we've found for this system. This involves restating the SIR queue as a system of stochastic partial differential equations and studying its behaviour asymptotically. The key questions this thesis tries to answer are the following:

- How can we analyse the SIR queue;
- In what way can we link the SIR queue to the SIR model;
- What does the SIR queue add to the SIR model, or does the SIR model completely capture the dynamics of the SIR queue already?

This stochastic network is challenging to analyse with elementary techniques since at any time t the transition probabilities are dependent on the current state and are therefore themselves stochastic. Furthermore this system eventually gets stuck in the state $(\cdot, 0, \cdot)$ after some time and is thus transient. A paper, [4], that served as inspiration for this thesis reminds us that "Even in very simple queueing systems, the transient behaviour is delicate to analyse".

One method of analysis is to look for special values of the parameters λ, ν and N that simplify the SIR queue. The simplification we find is $\lambda > 0$ and $\nu = 0$, this corresponds to a model where an infection spreads from person to person and no one heals. An important inspiration for the construction of the SIR queue is [5]. In that paper a process with finite population and a contact process is made to analyse Peer-to-Peer file sharing, with an important difference: the possibility of extinction is disabled. In the special case that $\nu = 0$ however our models coincide and we are able to apply the techniques from [5]. In this case the system will always end with everyone infected. The expected time it takes

to end in this state is analysed in [5], we manage to fill in the details in the proof of this result and even sharpen the bound by a little bit. This will be done in section 2.3.

Section 2.4 will be devoted to writing down the probability of any particular state $(S(t), I(t)) = (s, i)$ being reached from starting position $(S(0), I(0)) = (s_0, i_0)$. While we arrive at a closed form expression for this probability, it is sufficiently complex that we did not succeed in proving further properties from this expression. We encounter in this section a counting problem regarding the number of paths that are possible from the initial position to a certain end state. Adopting the reflection principle as described in [6] to this system we come up with a simple closed form solution.

Then in section 2.5 we calculate the generating function of these hitting probabilities. When analysing the hitting probabilities we condition the process $(S(t), I(t))$ on its jump times $\{t_1, ...\}$ after which it is a random walk taking the steps

$$(S(t), I(t)) \to (S(t) - 1, I(t) + 1) \text{ with probability } p_t(S(t), I(t)) \tag{1.7}$$
$$(S(t), I(t)) \to (S(t), I(t) - 1) \text{ with probability } p_c(S(t), I(t)) \tag{1.8}$$

where $p_t(\cdot)$ and $p_c(\cdot)$ are the state dependent probabilities of a transmission of the infection, or a cure respectively. Therefore, there is a recurrence relation

$$h(s, i) = h(s + 1, i - 1)p_t(s + 1, i - 1) + h(s, i + 1)p_c(s, i + 1) \tag{1.9}$$

where $h(s, i)$ is the probability of reaching (s, i) starting in some given (s_0, i_0). The initial condition $h(s_0, i_0) = 1$ then fixes all the values of $h(s, i)$. We do not manage to satisfyingly solve the resulting generating function and consider this section a motivation to search for alternative methods of analysis.

In the third chapter, section 3.1, we numerically calculate the extinction distribution $h(s, 0)$, that is the size of the susceptible group at the end of the epidemic. This distribution has a clear intuitive meaning, it describes the probability that s people will never get the disease. We can calculate this distribution using the recurrence relation 1.9. We see that there are two events that largely determine the location of the absorbing state. The first event is that the process dies out before the infection grows large and the second event is that the infection grows big and 'group immunity' sets in: the susceptible group becomes small enough for the effective utilisation of the infected queue $\rho(t) = \frac{\lambda S(t) I(t)/N}{\nu I(t)}$ to go below 1. This system's dualistic behaviour will be a theme for the remainder of the thesis.

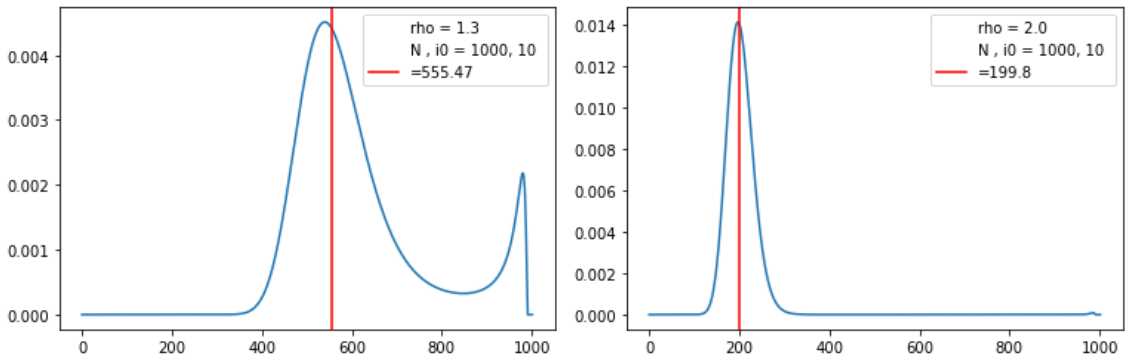

Figure 1.1: $\rho = \lambda/\nu$, on the x axis we read the size of the susceptible group when the epidemic terminates, on the y axis the probability density. In red we plot the location as predicted by the deterministic SIR model.

In section 3.2 we simulate the SIR queue fixing the parameters such that $\rho_0 = 2$ and increase the population N. We distinguish two ways to scale the initial position corresponding with what we will call the fluid regime and the initial regime. For the fluid regime we scale the initial position $(S(0), I(0)) = (N(1 - i_0), i_0 N)$ while $N \to \infty$, so the fraction of the initial infected group is constant, we pick $i_0 = 0.01$.

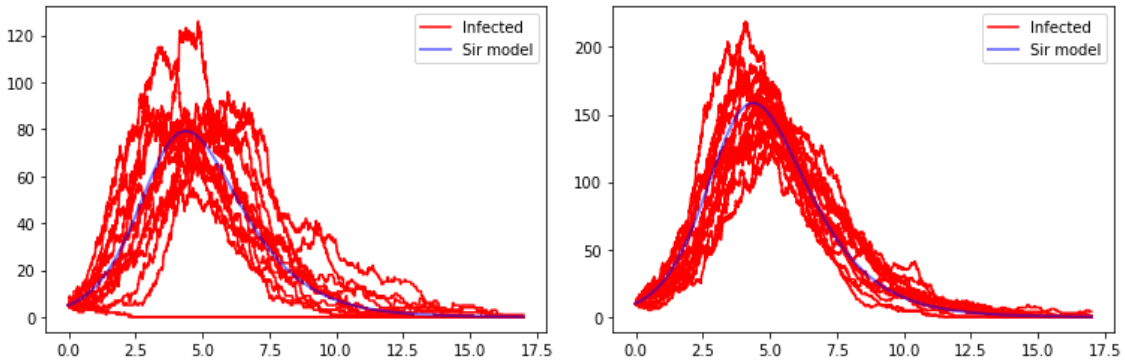

Figure 1.2: 15 simulations for $N = 500$ and $N = 1000$, in red the infected queue and in blue the deterministic SIR model.

As we can see, as the population increases the sample paths seemingly concentrate themselves increasingly around the deterministic SIR model. This is the behaviour characteristic of the **fluid regime**. However if we scale the initial position such that $(S(0), I(0)) = (N - i_0, i_0)$, that is we increase the population but leave the number of people initially infected constant, we see the behaviour characteristic of the initial regime. We pick $i_0 = 1$.

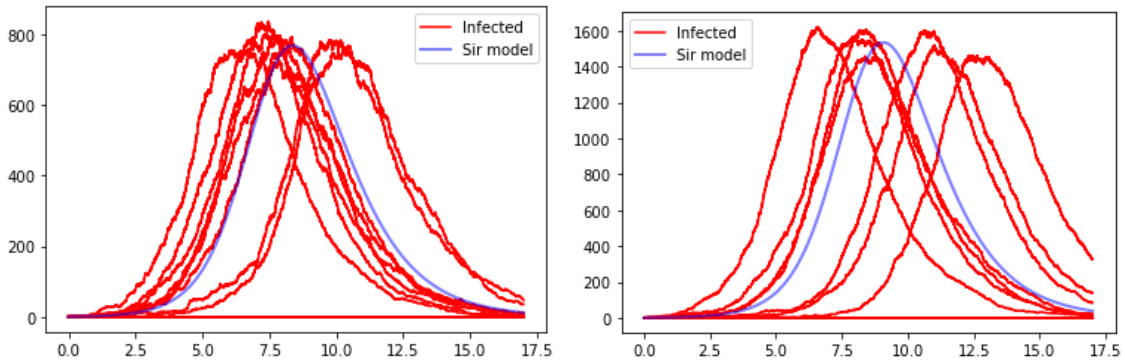

Figure 1.3: 15 simulations for $N = 5000$ and $N = 10000$. Red lines on the horizontal axis correspond with sample paths that go extinct almost immediately.

In this scaling we see the behaviour characteristic of the **initial regime**: the sample paths on the horizontal axis show that the probability of extinction does not disappear for large N, and that the location of the curve stays random for large N, when the size of the infected group is big enough the shape of the curve follows the prediction of the SIR model. The shape of the sample paths appears to converge to the SIR model once the infected group reaches a high enough level.

In chapter 4 we state and prove our these intuitions, these are our key results. We restate our model in the probabilistic framework found in [1]. This provides a powerful machinery for proving the convergence of queues to ordinary differential equations in one dimension, after a proper scaling is defined. We expand this method in order to apply

it to this more complicated, two dimensional stochastic network. The machinery works as follows: first we prove that the stochastic process corresponding to dynamics 5.38 - 5.40 starting in (s_0, i_0) is identical in distribution to the system of stochastic differential equations $L(t)$ given by

$$dL(t) = \begin{pmatrix} dS(t) \\ dI(t) \end{pmatrix} = \sum_{s=1}^{S(t-)} \sum_{i=1}^{I(t-)} (-1,1)^t \mathcal{N}_{\lambda/N}^{(s,i)}(dt) + \sum_{i=1}^{I(t-)} (0,-1)^t \mathcal{N}_{\nu}^{i}(dt) \tag{1.10}$$

where $\mathcal{N}_x^i$ are independent copies of the Poisson distribution with parameter x, with initial condition $L(0) = (s_0, i_0)^t$. The second step is to split up the process into a martingale part and a remaining part. Then finally by letting the population $N \to \infty$, scale the initial position by N and divide the resulting process by N, we prove that this scaling results in convergence in L_1 to the SIR model. We interpret this as the fluid regime, where the size of the infected group is proportional to N. Starting with a proportional initial position allows the process to skip over the stochastic nature the initial regime.

Here we find the answer to our second key question, we can link the SIR queue to the SIR model by the linear scaling of the initial position and the population. Doing so we obtain exactly the SIR model. This result justifies the study of the SIR queue as a meaningful model that is related to the SIR model, which itself has been empirically justified, see e.g. [17]. It also further justifies the SIR model as a meaningful model resulting from reasonable (but simplified) assumptions about an infection spreading through individuals meeting each other. Furthermore since the SIR queue is a bottom-up model we can (although this will not be the focus of this thesis) adjust more precisely the effects of different assumptions at the level of individuals. If for instance the contact process were not homogenous but had a more general probability distribution there would be no obvious way to implement this change in the SIR model. Using however the link between the SIR queue and the SIR model one could modify the SIR queue to reflect these adaptations and then link it to a higher level model to obtain a modified SIR model.

Then we investigate the scaling of the population keeping the initial infected group $i_0 = c$ for some constant c. This regime has the susceptible group large compared to the infected group. The initial regime is linked to the fluid regime through an exit time $\tau_{N,\varepsilon}$, defined as $\tau_{N,\varepsilon} = \inf_{t \geq 0}\{t : S(t) \leq (1-\varepsilon)N\}$. We define the initial regime then as the system at time $t \leq \tau_{N,\varepsilon}$. In this regime we can neglect somewhat the difference the $S(t)/N$ term makes, as $1 - \varepsilon \leq S(t)/N \leq 1$, so the dynamics of the resulting system are *roughly*

$$I(t) \to \begin{cases} I(t)+1 & \text{at rate } \lambda I(t) \\ I(t)-1 & \text{at rate } \nu I(t) \end{cases} \tag{1.11}$$

$$R(t) \to R(t) + 1 \text{ at rate } \nu I(t). \tag{1.12}$$

This is a linear birth-death process where we keep track of the deaths. To turn this rough estimation into precise results we define two linear birth-death processes $\bar{I}(t)$ and $\underline{I}(t)$ such that we have the stochastic ordering $\bar{I}(t) \geq I(t) \geq \underline{I}(t)$. Then as we let $N \to \infty$ and $\varepsilon \to 0$ we squeeze the $(I(t), R(t))$ system between these two linear birth-death processes. Results about the linear birth-death process can then be translated into results about the $(I(t), R(t))$ system in the initial regime. To do so some questions about the scaling of $N \to \infty$ and $\varepsilon \to 0$ need to be resolved this will be expanded on in chapter 4. We think of the fluid regime as starting after $\tau_{N,\varepsilon}$ when $I(t) + R(t)$ has become larger than εN and is on the same scale as $S(t)$. We prove that at time the time of this transition the ratio between $I(\cdot)$ and $R(\cdot)$ is fixed, $\lim_{N\to\infty} \frac{I(\tau_{N,\varepsilon})}{R(\tau_{N,\varepsilon})} = \frac{\lambda-\nu}{\nu} = \frac{\lambda}{\nu} - 1$ with probability $1 - (\frac{\nu}{\lambda})^{i_0}$

and 0 otherwise. The challenge in proving these types of results is keeping track of both the $I(t)$ process and $R(t)$ process simultaneously, and linking two different regimes without ignoring all the interesting properties, such as the state dependent transition rate.

While both the scaled behaviour of the SIR queue from $\tau_{N,\varepsilon}$ on is deterministic, and the state of $(I(\tau_{N,\varepsilon}), R(\tau_{N,\varepsilon}))$ is known, the actual time $\tau_{N,\varepsilon}$ remains stochastic. It is possible to approximate this $\tau_{N,\varepsilon}$ by the hitting time distribution of linear birth-death processes using a construction in which we squeeze $I(t)$ expanded upon in section 4.2. While the hitting time distribution of a linear birth-death process τ_M is generally known for fixed M, see [18], we could not prove any properties for scaling M from this distribution.

The third key question is thereby also answered, even when scaling the population the outbreak remains stochastic: it might exterminate before reaching a critical mass, and the time at which it either goes extinct or reaches fluid scale is distributed like the hitting time of a linear birth-death process, although we didn't manage a precise analysis of this phenomena. As for the first key question, regarding what method are appropriate for the analysis of the SIR queue, we have seen that the elementary methods have trouble extracting more than the most basic properties. The probabilistic framework expanded upon in chapter 4 is therefore to be the favoured framework.

2 The SIR model and the SIR queue

2.1 The SIR model

SIR models and their variants are widely used in epidemic modelling. The basic SIR model works by sorting the population into the groups Susceptible, Infected and Recovered. The Dutch research institute concerned with public health, the RIVM, uses a variant of the SIR model (see [2] p.13-14). The RIVM considers in the paper an SEIR model that sorts people into the compartments susceptible, exposed, infected and recovered, and then further distinguishes these people by age. We choose to work with a the most simple version of the SIR model that retains its mathematically interesting properties, the dynamics caused by a finite population and the interdependence of its variables.

For the mathematical statement of the SIR model we follow and simplify [7] p. 6-8. It models the spread of a disease by sorting the population into the following three compartments: the susceptible group consists of the population that has not previously been infected; the infected group is currently infected; and the recovered group used to be infected and are no longer infected or susceptible. We will use the variables $s(t), i(t), r(t)$ where the variable $t \in \mathbb{R}_+$ is the time parameter. The SIR model is a system of partial differential equations (PDEs), the solution of which is deterministic and with continuous values. We make the following assumptions in line with [7] :

1. The population is constant, so $s(t) + i(t) + r(t) = N$.
2. Only the infected group causes more infections.
3. Acquired immunity is permanent, the recovered group doesn't get reinfected.
4. We don't take into account dying from the infection. From a modelling perspective a death group would behave the same as a recovered group in that it doesn't get infected, nor does it infect the susceptible group.

The derivative of $i(t)$ is proportional with $s(t)i(t)$, this is intuitively justified since the growth of $i(t)$ is the result of contact between the susceptible group and the infected group. A justification directly from a modelling perspective through the SIR queue (see section 2.2 and this thesis) will follow by the convergence proven in section 4.1. These assumptions lead to the following model.

$$s'(t) = -\lambda s(t)i(t)/N \tag{2.1}$$
$$i'(t) = \lambda s(t)i(t)/N - \nu i(t) \tag{2.2}$$
$$r'(t) = \nu i(t). \tag{2.3}$$

The parameters λ and ν are respectively the rate of infection and the rate of recovery. We choose to use λ/N as this will arise in the discrete model naturally and is convention, see e.g. [7], noting that it is possible, when N is taken to be a constant, to choose a λ' such that division by N isn't necessary.

Definition 2.1.1. *The SIR model is the vector* $(s(t), i(t), r(t))$ *with the dynamics given by 5.35 - 5.37. The parameters are* λ, ν *and* N*. We have for all* $t \geq 0$ *that* $s(t)+i(t)+r(t) = N$*. We often consider the initial condition* $(s(0), i(0), r(0)) = (s_0, i_0, r_0) \in \mathbb{R}^3_+$ *as given.*

Definition 2.1.2. *We furthermore define* $\rho_0 = \frac{\lambda}{\nu}$ *and the effective utilisation of the* $i(t)$ *component* $\rho(t) = \frac{\lambda s(t) i(t)/N}{\nu i(t)} = \frac{\lambda}{\nu}\frac{s(t)}{N}\mathbb{1}_{\{i(t)\neq 0\}}$.

We have not yet shown that the solution of these PDEs exists, we will now.

Solution of the SIR model

We give a simple derivation for a solution following [9] where the simplifying assumption is made that $R_0 = 0$. This would mean that initially there are no people immune for the infection and is heuristically justified for modelling the outbreak of an epidemic.

We restate 5.35 - 5.37 as

$$\frac{ds(t)}{dt} = -\lambda s(t) i(t)/N \quad \frac{di(t)}{dt} = \lambda s(t) i(t)/N - \nu i(t) \quad \frac{dr(t)}{dt} = \nu i(t) \tag{2.4}$$

and note that $\frac{ds(t)}{dr(t)} = -\frac{\rho_0}{N} s(t)$. Solving and using $R_0 = 0$ we obtain $s(t) = s_0 e^{-\frac{\rho_0}{N} r(t)}$. Furthermore the assumption of a constant population allows us to write $i(t) = N - s(t) - r(t)$. Now 2.4 can be written as

$$s(t) = s_0 e^{-\frac{\rho_0}{N} r(t)} \quad i(t) = N - s(t) - r(t) \quad \frac{dr(t)}{dt} = \nu(N - s_0 e^{-\frac{\rho_0}{N} r(t)} - r(t)). \tag{2.5}$$

Now the only differential equation that remains to be solved is

$$\frac{dr(t)}{dt} = \nu(N - s_0 e^{-\frac{\rho_0}{N} r(t)} - r(t)) \tag{2.6}$$

which gives

$$t = \frac{1}{\nu}\int_0^r \frac{du}{N - s_0 e^{-\frac{\rho_0}{N} u} - u}. \tag{2.7}$$

This is an exact solution for the SIR model: we have $t(r)$ as a strictly increasing function, so the inverse is well defined and gives us $r(t)$. For further analysis of 2.6 and 2.7 we recommend [9]. Furthermore in [8] the model is analysed and solved for $r_0 > 0$, showing that the solution exists for all initial parameters.

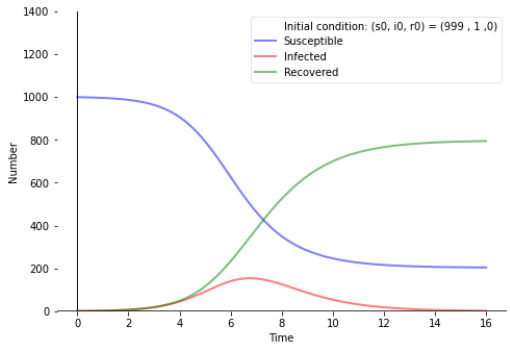

Figure 2.1: We plot the solutions taking the parameters $(\lambda, \nu) = (2, 1)$.

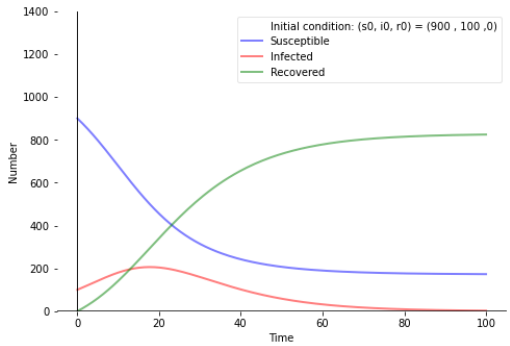

Figure 2.2: Different initial conditions and we take $(\lambda, \nu) = (0.2, 0.1)$.

McKendrick and Kermack [10] emphasise in their seminal paper that their model implies that it is possible for an epidemic to end before the the susceptible group reaches 0. This phenomena is sometimes referred to as group immunity. It follows from the solutions as follows: as the denominator in the integral of 2.7 goes to 0 the time t goes to infinity. Denoting the solution of

$$N = s_0 e^{-\frac{\rho_0}{N} r(t)} + r(t) \tag{2.8}$$

as r_{end} we have that $lim_{t\to\infty}(s(t), i(t), r(t)) = (N - r_{end}, 0, r_{end})$, this means that no matter how long the epidemic lasts no more than r_{end} people will be infected in total. We can solve equations like 2.8 using the Lambert W function see [3], or appendix section 5.1 for the definition and some properties of the Lambert W function.

Lemma 2.1.3. *The solution of 2.8 given as a fraction of the population is*

$$r_{end}/N = 1 + \frac{W(-\rho_0 \frac{s_0}{N} e^{-\rho_0})}{\rho_0} \tag{2.9}$$

and consequently

$$s_{end}/N = -\frac{W(-\rho_0 \frac{s_0}{N} e^{-\rho_0})}{\rho_0}. \tag{2.10}$$

Proof. See appendix 5.1. □

To obtain some intuition for this function we provide a table for some values of ρ_0 and $i_0/N = 1 - s_0/N$, rounded to 2 significant numbers.

	$i_0/N = 0.9$	0.5	0.25	0.1	0.05	0.01
ρ_0 =0.1	0.91	0.53	0.27	0.11	0.055	0.011
0.5	0.94	0.64	0.38	0.18	0.093	0.02
1	0.96	0.77	0.58	0.39	0.29	0.14
2	0.99	0.92	0.87	0.83	0.81	0.80
3	0.99	0.97	0.96	0.95	0.94	0.94
5	1.0	1.0	0.99	0.99	0.99	0.99
10	1.0	1.0	1.0	1.0	1.0	1.0

This table gives r_{end}/N, so the amount of people who will have been infected in the end, as a function of ρ_0 a constant related to the effective utilisation, and i_0/N, the initial amount of susceptible people. It is readily observed that r_{end} is increasing in ρ_0 and i_0/N.

2.2 Queuing model

In this section we will construct a model that takes into account the discrete and stochastic nature of an infection that spreads from person to person. An important inspiration for this type of construction is the literature on peer-to-peer (P2P) networks that use a queueing theoretic framework for the spread of a file among different servers that each contribute to the spread of a file see for example [4]. For our construction a paper [5] on P2P file-sharing was particularly important, as it takes into account a finite population and a contact process, although it doesn't take into account the event of extinction. We also use some basic properties of Poisson processes one could look up in e.g. [1] chapter 1.

We will construct a stochastic process with the same compartments as the SIR model. As with the SIR model we divide the population into three compartments, a susceptible

group, an infected group and a recovered group, the total population is denoted as before with N.

We are concerned with an infection that spreads from person to person, as they (randomly) encounter each other. We therefore have a stochastic process of people meeting each other and spreading the disease as they do. We call this the contact process at Poisson rate $N\lambda^*$ that works as follows:

- The times $t_1, t_2, ...$ are Poisson distributed with parameter $N\lambda^*$
- At each time t_i two[1] individuals will be selected from the population of N homogeneously. These two individuals are said to be in contact with each other.

Then if one of the people in this contact is in the susceptible group and one in the infected group there is a probability β of infection, which means that individual joins the infected group. Each infected individual joins the recovered group after a waiting time that is exponentially distributed with parameter ν.

At every contact t_i the probability of a match between someone in the susceptible group and someone in the infected group is $2\frac{S(t)}{N}\frac{I(t)}{N}$. So the contact process marked with probability $2\beta\frac{S(t)}{N}\frac{I(t)}{N}$ is the infection process, it has a Poisson dynamics with rate $2\beta\lambda^*\frac{S(t)}{N}I(t)$. For ease of notation we introduce $\lambda = 2\beta\lambda^*$. Since this λ is the only parameter that is relevant to the system we are able to choose $\beta = 1$ and will do so from now on.

The first time the system $(S(t), I(t), R(t))$ changes is the first moment either someone recovers or becomes infected. This is the minimum of exponentially distributed variables with parameters $\frac{\lambda}{N}S(t)I(t)$ and $\nu I(t)$, and therefore again an exponentially distributed random variable. We obtain these Poisson dynamics

$$S(t) \to S(t) - 1 \text{ at rate } \lambda S(t)I(t)/N \tag{2.11}$$

$$I(t) \to \begin{cases} I(t) + 1 & \text{at rate } \lambda S(t)I(t)/N \\ I(t) - 1 & \text{at rate } \nu I(t) \end{cases} \tag{2.12}$$

$$R(t) \to R(t) + 1 \text{ at rate } \nu I(t). \tag{2.13}$$

This system $(S(t), I(t), R(t))$ jumps at randomly distributed times $t_1, t_2...$ We call the time of the n-th jump $\tau(n)$.

Definition 2.2.1 (**The SIR queue**). *Let $(S(t), I(t), R(t))$ a process in $\mathbb{N}^3$ with dynamics 2.11 - 2.13 . Let furthermore $S(t) + I(t) + R(t) = N$. We call this process the SIRqueue with the parameters N, λ and ν. We generally consider the initial condition $(S(0), I(0), R(0)) = (s_0, i_0, r_0)$ to be given.*

Remark 2.2.2. *We denote the states of the SIR queue with $(S(t), I(t), R(t))$ and the SIR model from previous section with $(s(t), i(t), r(t))$.*

Remark 2.2.3. *Because of the $S(t)+I(t)+R(t) = N$ property any two out of $(S(t), I(t), R(t))$ completely determines the whole process. We will generally just focus on $(S(t), I(t))$, although in section 4.2 when we take $S(t)$ large we will change our focus to $(I(t), R(t))$.*

The state space of $(S(t), I(t))$ is $\mathcal{S} = \mathbb{N}^2$. We write down the Q-Matrix $Q = (q_{xy})$ associated with this process. For a definition of Q-matrices see the appendix.

[1] We allow for the same individual to be chosen twice, for our purposes this will result in a negligible chance as we take N to be large.

$$q_{(s,i),(s-1,i+1)} = \lambda si/N, \qquad s > 0, i > 0 \tag{2.14}$$

$$q_{(s,i),(s,i-1)} = \nu i, \qquad i > 0 \tag{2.15}$$

$$q_{xy} = 0, \qquad \text{otherwise if } x \neq y \tag{2.16}$$

Remark 2.2.4. *For all $m < n$ we have that $(S(\tau(n)), I(\tau(n))) \neq (S(\tau(m)), I(\tau(m)))$. This is because $S(t)$ is non-increasing in time so $S(\tau(n)) = S(\tau(m)) \implies I(\tau(m)) - I(\tau(n)) = m - n$. We call this the* ***non-returning property*** *of the SIR queue.*

The effective utilisation of the $I(t)$ queue at time t denoted by $\rho(t)$ is given by

$$\rho(t) = \frac{\lambda S(t) I(t)/N}{\nu I(t)} = \rho_0 \frac{S(t)}{N} \mathbb{1}_{\{I(t)\neq 0\}}. \tag{2.17}$$

Note that this only depends on the size of the susceptible group. As such $\rho(t)$ is a non-increasing function in time in this system. We are especially interested in parameters[2] such that $\rho(0) > 1$. We'd expect the queue $I(t)$ to be increasing as long as $\rho(t) > 1$, and then there will be a tipping point after which the queue will start to decrease in expectation. This is when $\rho(t) = \rho_0 \frac{S(t)}{N} \mathbb{1}_{\{I(t)\neq 0\}} = 1$ so when $S(t) = N/\rho_0$. Since $\rho(t)$ is decreasing in t there are no significant values for λ, ν that greatly simplify analysis other than $\nu = 0$ or $\lambda = 0$ (trivial). We will analyse the $\nu = 0$ case.

2.3 Mean time until termination with no recovery

In this section we will follow the analysis in [5]. This is possible because setting the parameter ν to 0 we don't have the eventuality of extinction and as such the SIR queue is similar to the model analysed in [5]. Some preciser calculations of the sum allows us to make the estimation slightly more precise than the estimation given in the paper. The $\nu = 0$ case is relevant when the epidemic spreads without recovery or as a limit case when $\nu \to 0$. Since there is no recovery the epidemic will spread until $S(t) = 0$. We will analyse $t_{end} = \inf_t\{t \in \mathbb{R}_+ : S(t) = 0\}$. We will start the process in the state $(S(0), I(0)) = (N - 1, 1)$, so starting from just 1 infection.

Let λ be the contact parameter, then the time between contacts is exponentially distributed with mean $(N\lambda)^{-1}$. Let $I(n)$ the number of people in the infected group after the i-th contact. We have the following recursive expression for $I(n)$

$$I(n+1) \to \begin{cases} I(n) + 1 & \text{w.p. } 1 - p(i) \\ I(n) & \text{w.p. } p(i) \end{cases} \tag{2.18}$$

where we let $p(i) = 1 - 2\frac{N-i}{N}\frac{i}{N}$ the probability of not matching a susceptible and infected person in the contact process (recall that we allow for picking the same person twice).

Let C_i be the number of required for a successful transmission when $I(t)$ is in state i. C_i is geometrically distributed with parameter $p(i)$:

$$\mathbb{P}(C_i = k) = (1 - p(i))p(i)^{k-1}. \tag{2.19}$$

[2] One of the reasons for this is that the behaviour of a SIR queue starting with $\rho(0) < 1$ is similar to the behaviour of a SIR queue starting with $\rho(0) > 1$ after some time t such that $\rho(t) < 1$.

Then let $D_j = \sum_{i=1}^{j} C_i$ be the amount of contacts needed to reach the state with $I(t) = j + 1$. Then $\mathbb{E}[D_{N-1}]$ is the expected number of contacts needed to infect the whole population. We obtain the following sum

$$\mathbb{E}[D_{N-1}] = \sum_{i=1}^{N-1} \mathbb{E}[C_i] \tag{2.20}$$

$$= \sum_{i=1}^{N-1} \frac{1}{1-p(i)} \tag{2.21}$$

$$= \sum_{i=1}^{N-1} \frac{N^2}{2(N-i)(i)}. \tag{2.22}$$

We give a precise solution to sums of these types in the following lemma.

Lemma 2.3.1.

$$\sum_{i=1}^{n} \frac{1}{(i+c)(N-i)} = \frac{1}{c+N}(\log(n+1+c) - \log(1+c) + \log(N) - \log(N-n) + O(1)) \tag{2.23}$$

where $c \in \mathbb{Z}_{\geq 0}$ *and* $n \leq N-1$.

Proof. The crux of this calculation is the relationship between harmonic numbers and the Euler-Mascheroni constant $\sum_{i=1}^{n} \frac{1}{i} = \int_1^{n+1} \frac{1}{x} dx + \varepsilon(n)$ where $\varepsilon(n)$ is the error term that converges monotonically to the Euler-Mascheroni constant $\gamma \approx 0.577$. We obtain

$$\sum_{i=1}^{n} \frac{1}{(i+c)(N-i)} = \frac{1}{c+N}(\sum_{i=1}^{n} \frac{1}{i+c} + \frac{1}{N-i}) \tag{2.24}$$

$$= \frac{1}{c+N}(\sum_{i=1}^{n} \frac{1}{i+c} + \sum_{i=N-n}^{N-1} \frac{1}{i}) \tag{2.25}$$

Now the trick $\sum_{i=1}^{n} \frac{1}{i+c} + \sum_{i=N-n}^{N-1} \frac{1}{i} - \int_1^{n+1} \frac{1}{x+c} dx - \int_{N-n}^{N} \frac{1}{x} dx = \varepsilon(n) + \varepsilon(N-1) - \varepsilon(N-n-1) = O(1)$.

$$= \frac{1}{c+N}(\int_1^{n+1} \frac{1}{x+c} dx + \int_{N-n}^{N} \frac{1}{x} dx + O(1)) \tag{2.26}$$

$$= \frac{1}{c+N}(\int_1^{n+1} \frac{1}{x+c} + \frac{1}{N-x+1} dx + O(1)) \tag{2.27}$$

$$= \frac{1}{c+N}(\log(n+1+c) - \log(1+c) + \log(N) - \log(N-n) + O(1)) \tag{2.28}$$

□

Plugging the lemma in for the sum 2.22, $n = N-1$, $c = 0$ gives the elegant result

$$\sum_{i=1}^{N-1} \frac{N^2}{2(N-i)(i)} = N(\log(N) + O(1)) \tag{2.29}$$

a slightly preciser approximation than stated in [5], where the bound $N(log(N)+o(log(N))$ is obtained.

Let T_j be the time that's needed to reach the state with j infected. We have

$$T_j = \sum_{k=1}^{D_{j-1}} \tau_k \tag{2.30}$$

where τ_k is the time between contacts in the contact process, so per construction $\tau_k \sim exp(N\lambda)$, i.i.d. exponentially distributed random variables with mean $1/(N\lambda)$. We use Wald's lemma to arrive at the following theorem.

theorem 2.3.2. *The mean time until the whole population is infected is given by*

$$\mathbb{E}[T_N] = \mathbb{E}[D_{j-1}]\mathbb{E}[\tau_1] \tag{2.31}$$

$$= \frac{1}{\lambda}\log(N) + O(1). \tag{2.32}$$

2.4 Hitting probabilities

In this section we investigate the hitting probabilities of states starting from a known initial condition and some parameters $\lambda > 0$, $\nu > 0$. The initial state (s_0, i_0) is considered known where s_0 denotes the number of healthy people at $t = 0$, and i_0 the number of infected people. We take the number of recovered people as 0 in the initial state.

Definition 2.4.1. *Let $h_{s_0,i_0}(s,i) := \mathbb{P}(\exists t : S(t) = s, I(t) = i | S(0) = s_0, I(0) = i_0)$. This is the probability that the SIR queue hits the state (s,i). We will often drop the subscript and simply write $h(s,i)$.*

The process stops once there are 0 infected people left, so the states $(s,0)$ are absorbing states, and if $s \leq s_0$ then $h(s,0) > 0$. We are especially interested in the *extinction distribution* $h(s,0)$ (note that this is a probability distribution). This is the distribution of the number of people that have never been infected during the epidemic.

We call the event of a transmission a transmission jump; the event of an individual curing a cure jump. Denote $p_t(\cdot)$ and $p_c(\cdot)$ as the probability of transmission or cure jumps respectively:

$$p_t(s,i) = \mathbb{P}((S(\tau(n+1), I(\tau(n+1)) = (s-1, i+1) | (S(\tau(n), I(\tau(n)) = (s,i)) \tag{2.33}$$

$$= \frac{s\lambda/N}{s\lambda/N + \nu}\mathbb{1}_{\{i\neq 0\}} \tag{2.34}$$

$$p_c(s,i) = \mathbb{P}((S(\tau(n+1), I(\tau(n+1)) = (s, i-1) | (S(\tau(n), I(\tau(n)) = (s,i)) \tag{2.35}$$

$$= \frac{\nu}{s\lambda/N + \nu}\mathbb{1}_{\{i\neq 0\}}. \tag{2.36}$$

By simply conditioning on the order of each possible sequence of transmission jumps and cure jumps leading to a certain end state we can calculate the probability of hitting that end state. Notice that since the SIR queue has the non-returning property we have that the hitting probability is also the expected times passing a state, $h(s,i) = \mathbb{E}[\#\text{ times in } (s,i)]$. We write down this idea formally.

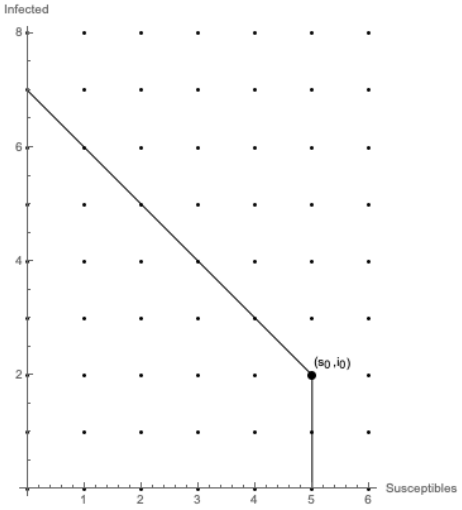

Figure 2.3: All reachable states are enclosed in the quadrilateral.

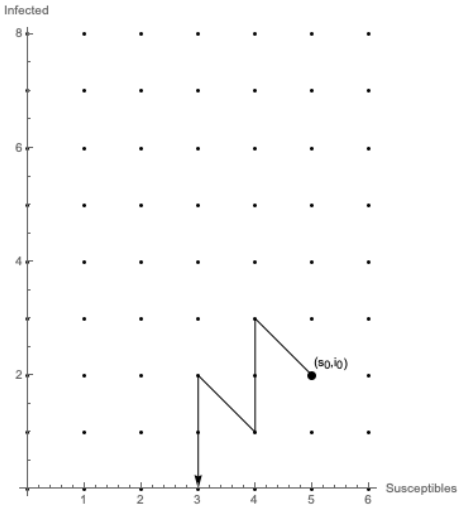

Figure 2.4: A sample path resulting in termination.

Note that $h(s,i) > 0 \iff 0 \leq s \leq s_0$ and $0 \leq i \leq i_0 + s_0 - s$. To reach the state (s,i), $s_0 - s = J_t$ transmission jumps are needed, and $i_0 - i + s_0 - s = J_c$ cure jumps. Let J the total number of jumps so $J := J_c + J_t = 2J_t + i_0 - i$. Let S be the set of all orders in which J_c cure jumps and J_t transmission jumps lead to the state (s,i) (this is not all combinations of J_c and J_t, we will expand on this point in the next section). Elements σ^a of S, $1 \leq a \leq \#S$ are J tuples $(s_1^a, ..., s_J^a) = \sigma^a$, with each position σ_j^a in $\{c,t\}$. For a $\sigma^a \in S$ we write $\mathbb{P}(\sigma^a) = \Pi_{j=1}^J p_{s_j^a}(s_{(j-1)}, i_{(j-1)})$, where $\mathbb{P}(\sigma^a)$ is the probability of the event that the process makes these jumps in the order σ^a. The states (s_j, i_j) are here recursively defined as $(s_j, i_j) = (s_{(j-1)}, i_{(j-1)}) + \mathbb{1}_{\{s_{j-1}^a = t\}}(-1,1) + \mathbb{1}_{\{s_{j-1}^a = c\}}(0,-1)$. Since each event σ^a is independent we can simply sum over all paths for a closed expression.

theorem 2.4.2. *Let $\sigma^a \in S$ all the permissible paths from (s_0, i_0) to (s,i). The hitting probability $h_{s_0,i_0}(s,i)$ has the following closed form*

$$h_{s_0,i_0}(s,i) = \sum_{\sigma^a \in S} \Pi_{j=1}^J p_{s_j^a}(s_{(j-1)}, i_{(j-1)}) \tag{2.37}$$

Remark 2.4.3. *Notice that in general the probabilities $h(s,i)$ do not form a distribution as $\sum_{s,i} h(s,i) > 1$. The most important example of a distribution in this family is the earlier mentioned extinction distribution $h(s,0)$.*

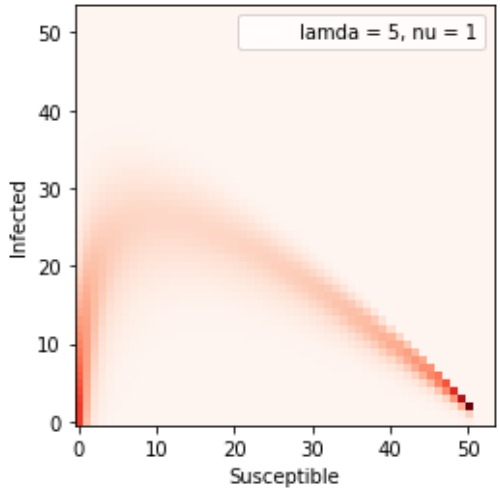

Figure 2.5: We plot the heat map of hitting probabilities starting in $(s_0, i_0) = (50, 2)$.

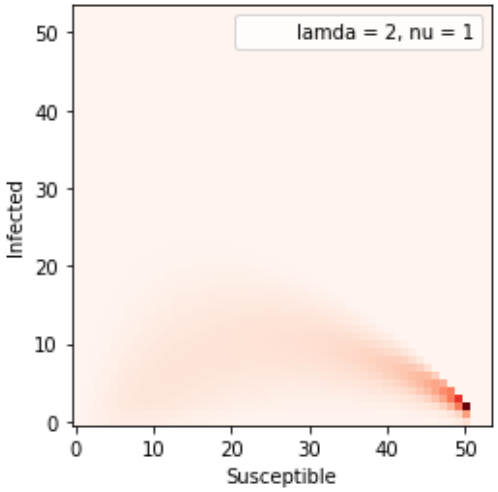

Figure 2.6: Starting in the same point $(s_0, i_0) = (50, 2)$ but with a lower ρ_0.

Note that the property $h(s,i) = h(s+1,i-1)p_t(s+1,i-1) + h(s,i+1)p_c(s,i+1)$ along with the starting condition $h(s_0, i_0) = 1$ determines all the hitting probabilities, this leads to a useful algorithm for numerical calculations. Using this we numerically calculate and plot the heat map of the hitting probabilities in figure 2.5 and 2.6. We see that there are two paths to extinction that dominate, one is extinction before $I(t)$ becomes big and the other is extinction after group immunity.

2.4.1 A counting problem

We are specifically interested in the extinction distribution $h(s, 0)$, so of interest are the paths that start in (s_0, i_0) and end in termination. A nice counting problem presents itself, how many elements are there in S that lead from (s_0, i_0) to $(s, 0)$? We know that every such path has $J_t = s_0 - s$ and $J_c = s_0 - s + i_0$. There are $\binom{J}{J_c}$ combinations of J_c cure jumps and J_t transmission jumps. However each order of jumps which hits 0 infected before the last jump is inadmissible because the process stops once there are 0 infected people. We use the reflection principle as in [6] to count these possibilities, however given that the jumps are both diagonal and horizontal we 'reflect' across a diagonal axis instead of a simple horizontal axis.

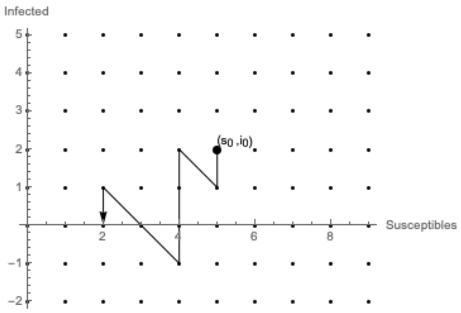

Figure 2.7: A defective path.

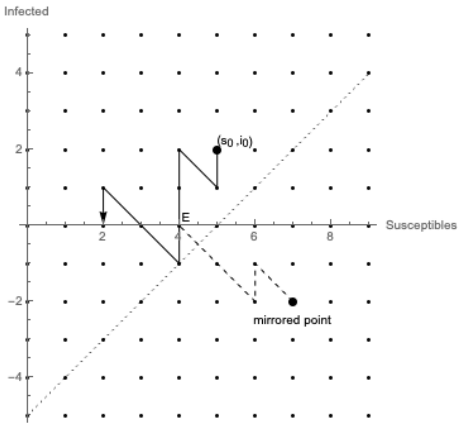

Figure 2.8: The same defective path with its mirrored path.

The number of elements in S are then the combination of J_c and J_t cure and transmission jumps, minus the number of defunct paths that hit some point $(s', 0)$, $s' < s$, where the process terminates before the endstate. We denote the set of all defunct paths by D. Since the last jump is a cure jump, and thus known, there are thus $\binom{J-1}{J_t}$ possible paths. As a formula this is

$$\#S = \binom{J-1}{J_t} - \#D. \tag{2.38}$$

The mirroring principle will be applied to find a bijection between D and a set of paths that is easier to count. Every path $d \in D$ hits the horizontal axis for the first time in a point $(s', 0)$ called E. We identify every such path with a reflected path that starts in the diagonally mirrored point $(s_0 + i_0, -i_0)$, and until reaching $(s', 0)$ does the opposite jump compared to d. These paths are constructed such that they intersect at the point E. The path d reaches E after some fixed number of J_c^0 (vertical) cure jumps and J_t^0 (diagonal) transmission jumps. This must solve $(s', 0) = (s_0, i_0) + J_t^0 \cdot (-1, 1) + J_c^0 \cdot (0, -1)$ so we have $J_c^0 = J_t^0 + i_0$ and $E = (s_0 - J_t^0, 0)$. Then for our mirrored path, d', we get $(s_0 + i_0, -i_0) + J_c^0 \cdot (-1, 1) + J_t^0 \cdot (0, -1) = (s_0 + i_0 - J_c^0, J_c^0 - i_0 - J_t^0) = (s_0 - J_t^0, 0) = E$. By symmetry this is the first time d' hits $(\cdot, 0)$ since it is the first time d reaches $(\cdot, 0)$.

Since every path between $(s_0+i_0, -i_0)$ and $(s, 1)$ constructed with the described modified jumps crosses the horizontal axis before $(s, 0)$ there is a bijunction between the set D of defunctive paths and D' the paths from $(s_0 + i_0, -i_0)$ to $(s, 1)$, so $\#D' = \binom{2s_0-2s+i_0-1}{s_0+i_0-s}$.

theorem 2.4.4. *The number of paths in S is given by*

$$\#S = \binom{J-1}{J_t} - \binom{2s_0 - 2s + i_0 - 1}{s_0 + i_0 - s} \tag{2.39}$$

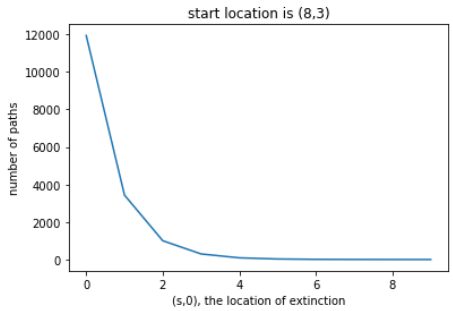

Figure 2.9: The number of paths calculated for a small population.

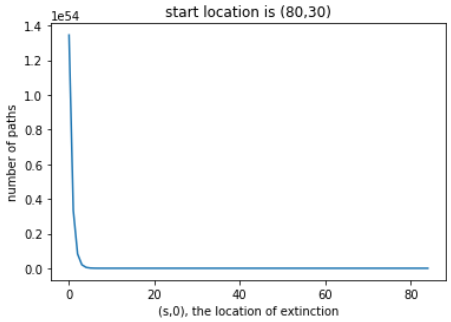

Figure 2.10: The number of paths for a slightly larger population.

It is readily observed that the the possible paths are largely concentrated near termination with zero susceptibles left.

2.5 Generating function of the hitting probabilities

As it is difficult to apply the closed form of the hitting probabilities, theorem 2.4.2, directly we explore other methods. In this section the hitting probabilities are investigated using generating functions. We will apply the recurrence relation $h(s,i) = h(s+1,i-1)p_t(s+1,i-1) + h(s,i+1)p_c(s,i+1)$ along with the starting condition $h(s_0,i_0) = 1$ and $h(s,i) \neq 0 \implies s \leq s_0, s+i \leq s_0+i_0$. We obtain

$$\tilde{h}(x,y) = \sum_{s=0}^{s_0} \sum_{i=0}^{s+i\leq s_0+i_0} h(s,i)x^s y^i = x^{s_0}y^{i_0} + \sum_{s=0}^{s_0-1} \sum_{i=2}^{s+i\leq s_0+i_0} h(s+1,i-1)p_t(s+1,i-1)x^s y^i \tag{2.40}$$

$$+ \sum_{s=0}^{s_0} \sum_{i=0}^{s+i\leq s_0+i_0-1} h(s,i+1)p_c(s,i+1)x^s y^i. \tag{2.41}$$

By carefully keeping track of the edge conditions we can ignore the indicator functions in the p_c and p_t functions. Let

$$g(s,i) = \frac{h(s,i)}{s\lambda/N + \nu}, \tag{2.42}$$

and

$$\tilde{g}(x,y) = \sum_{s=0}^{s_0} \sum_{i=0}^{s+i\leq s_0+i_0} g(s,i)x^s y^i \tag{2.43}$$

its generating function. Then we use that both simplify using that both $p_t(s,i)$ and $p_c(s,i)$ have the denominator $s\lambda/N + \nu$,

$$\tilde{h}(x,y) = \sum_{s=0}^{i_0} \sum_{i=0}^{s+i\leq s_0+i_0} (s\lambda/N) g(s,i) x^s y^i + \nu g(s,i) x^s y^i \tag{2.44}$$

$$= x\lambda/N \frac{d}{dx}\tilde{g}(x,y) + \nu\tilde{g}(x,y) \tag{2.45}$$

$$= x^{s_0}y^{i_0} + \frac{y}{x}\lambda/N \sum_{s=0}^{s_0-1} \sum_{i=2}^{s+i\leq s_0+i_0} \frac{d}{dx}(g(s+1,i-1)x^{s+1}y^{i-1}) \tag{2.46}$$

$$+ \frac{\nu}{y} \sum_{s=0}^{s_0} \sum_{i=0}^{s+i\leq s_0+i_0-1} g(s,i+1)x^s y^{i+1} \tag{2.47}$$

$$\tag{2.48}$$

For the first expression we note that the missing terms in the $\sum_{s=0}^{s_0-1}\sum_{i=2}^{s+i\leq s_0+i_0}\frac{d}{dx}(g(s+1,i-1)x^{s+1}y^{i-1})$ sum are the $g(0,i)$ terms and the $g(s,0)$ terms so we get

$$\sum_{s=0}^{s_0-1} \sum_{i=2}^{s+i\leq s_0+i_0} \frac{d}{dx}(g(s+1,i-1)x^{s+1}y^{i-1}) = \frac{d}{dx}\left(\tilde{g}(x,y) - \sum_{i=0}^{s_0+i_0} g(0,i)y^i - \sum_{s=0}^{s_0} g(s,0)x^s\right) \tag{2.49}$$

$$= \frac{d}{dx}\left(\tilde{g}(x,y) - \sum_{s=0}^{s_0} g(s,0)x^s\right). \tag{2.50}$$

Similarly for the second sum $\sum_{i=0}^{i_0}\sum_{j=0}^{i+j\leq i_0+j_0-1} g(i,j+1)x^i y^{j+1}$ we miss the $g(i,0)$ terms so

$$\sum_{i=0}^{i_0} \sum_{j=0}^{i+j\leq i_0+j_0-1} g(i,j+1)x^i y^{j+1} = \tilde{g}(x,y) - \sum_{i=0}^{i_0} g(i,0)x^i \tag{2.51}$$

Noting that $\sum_{i=0}^{i_0} g(i,0)x^i = \tilde{g}(x,0)$ we get the following differential equation

$$(x\frac{\lambda}{N} - \frac{y}{x}\frac{\lambda}{N})\frac{d}{dx}\tilde{g}(x,y) + \frac{y}{x}\frac{\lambda}{N}\frac{d}{dx}\tilde{g}(x,0) \tag{2.52}$$

$$= x^{i_0}y^{j_0} + (\frac{\nu}{y} - \nu)\tilde{g}(x,y) - \frac{\nu}{y}\tilde{g}(x,0) \tag{2.53}$$

We did not manage to solve this equation further and consider this section a negative result. We conclude chapter 2 after attempting different elementary techniques with mixed results. More sophisticated methods are necessary.

3 Simulations and Numerical calculations

3.1 Numerical calculations of the extinction distribution

We numerically calculate the extinction distribution $h(s,0)$, the probability of going extinct with s people remaining uninfected. We calculate these from the recurrence relation mentioned in section 2.4, $h(s,i) = h(s+1,i-1)p_t(s+1,i-1)+h(s,i+1)p_c(s,i+1)$. This fixes the distribution when the initial condition $h(s_0,i_0) = 1$ is given. We're especially interested in the impact of $\rho_0 = \frac{\lambda}{\nu}$, the initial fraction of infected i_0/N and the size of the population N. We also plot the value predicted according to lemma 2.1.3. We start with varying the ρ_0 and leaving $N = 1000$, $i_0 = 10$ constant.

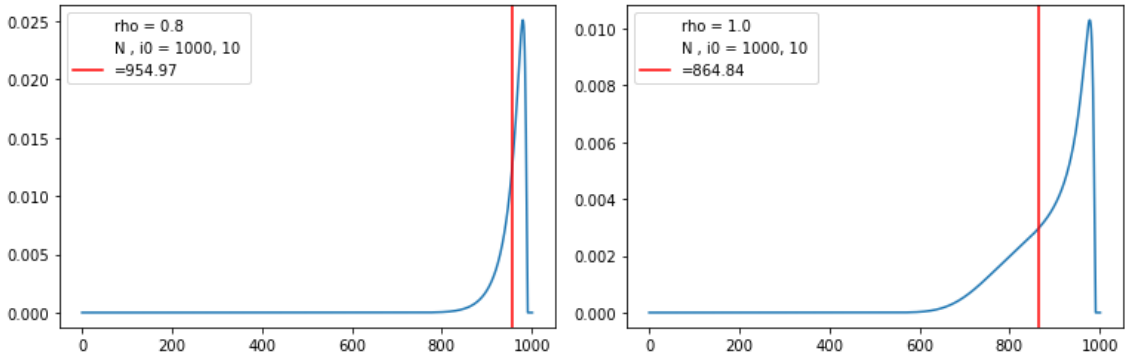

Figure 3.1: The distribution for low values of ρ_0.

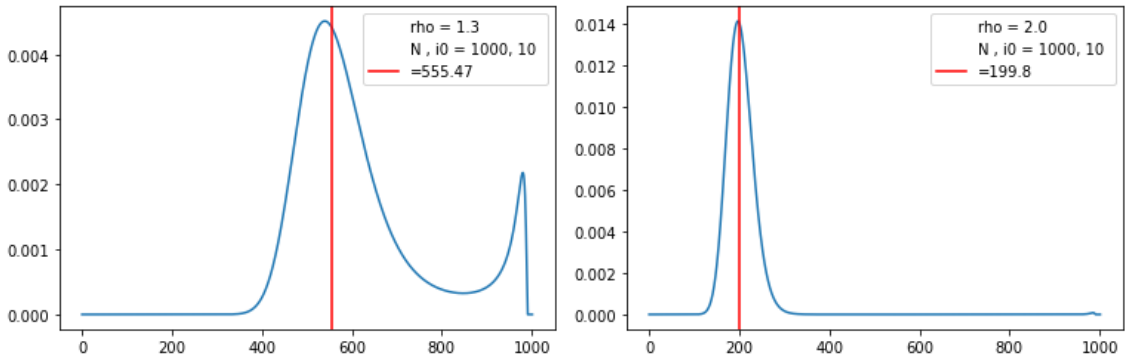

Figure 3.2: The distribution for medium values of ρ_0

We see that as far as extinction is concerned two events dominate: either the infection dies out before it spreads significantly, or it will continue to spread until the susceptible group becomes smaller and group immunity lets the process die out. These two moments of extinction can both play a significant role depending on the parameters. We now scale the population leaving $\rho_0 = 1.3$ and $i_0/N = 0.01$.

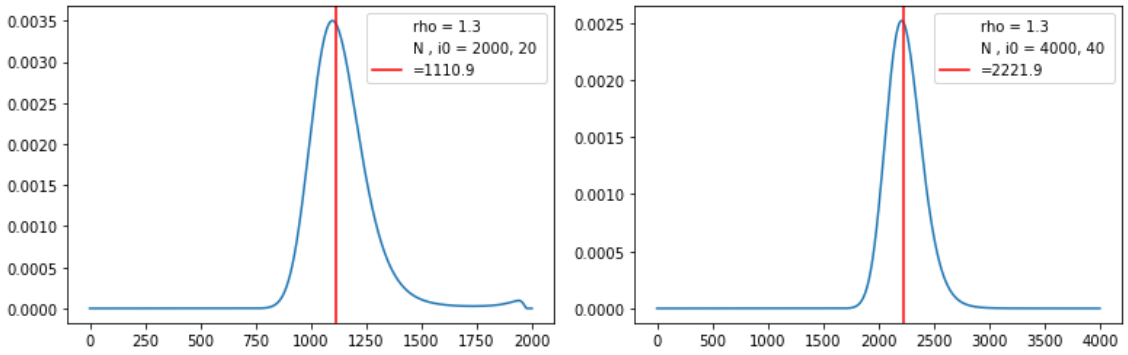

Figure 3.3: We double the population twice

We see that as we increase the population the probability mass concentrates around the same event, extinction on the location of lemma 2.1.3. We call this the second moment of extinction. Scaling the population N while leaving the initial fraction i_0/N constant will be further developed in section 4.1, where we see that this scaling does lead to a convergence. Finally we let the initial fraction vary.

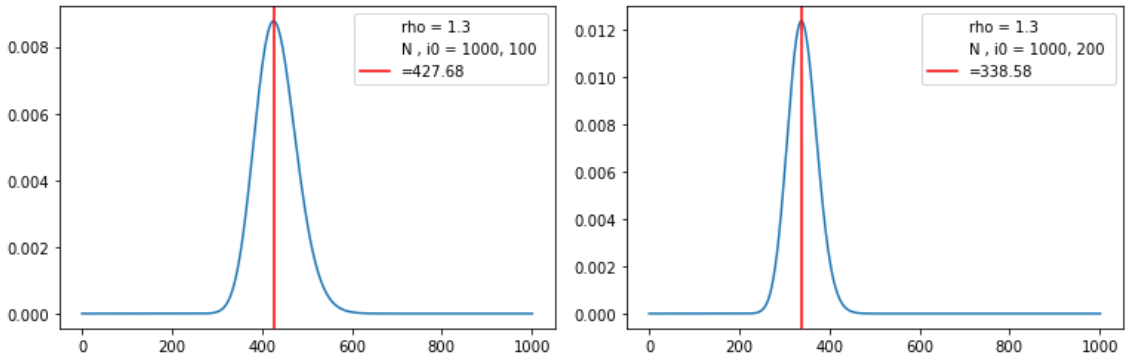

Figure 3.4: We let i_0/N equal 0.1 and then 0.2

We see that a higher fraction of initially infected leads to a smaller variance around the second moment of extinction. Since the initial infected population is high the first moment of extinction disappears.

3.2 Simulating the SIR queue

Looking at the SIR queue in its totality, instead of just the extinction distribution, it is preferable to use simulations rather than numerical calculations. We simulate the SIR queue starting in $i_0/N = 0.1$ with $\lambda = 2$, $\nu = 1$ at population size $N = 100$ and $N = 500$. The vertical axis shows the amount in each group while the horizontal axis represents the time parameter.

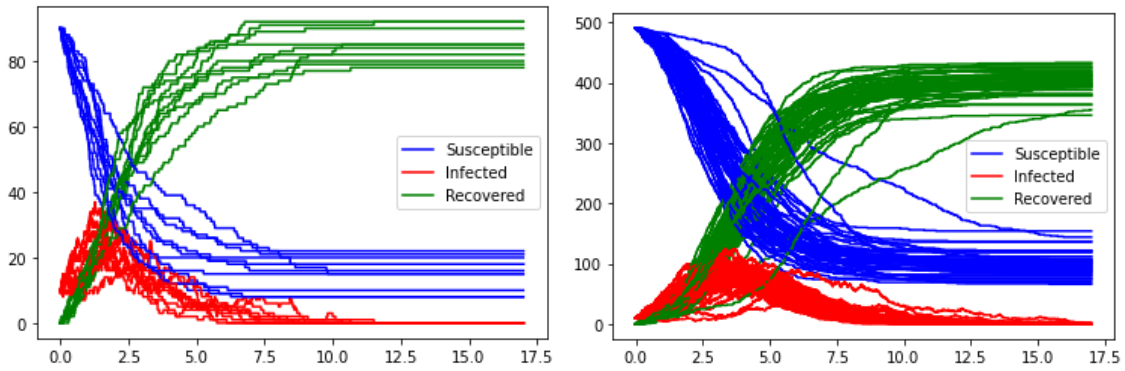

Figure 3.5: 50 simulations for $N = 100$ and $N = 500$.

When looking at the infected group we see the basic behaviour of the SIR queue is as follows: it grows until the tipping point discussed at the end of section 2.2 and then decreases until extinction. Since the infected group is the most clearly legible we won't show the Susceptible and the Recovered group in the next graphs, as to reduce cluttering.

We'll simulate the scaling the population in two ways. First we'll look at the behaviour as we increase N and increase the initial infected group proportionally. Then we will simulate a scaling of N leaving i_0 constant. We show 15 sample paths of the infected component of the SIR queue starting in $i_0/N = 0.01$ with $\lambda = 2$, $\nu = 1$ at population size $N = 100, 200, 500$ and then 1000. We also show the trajectory of the I component in the SIR model.

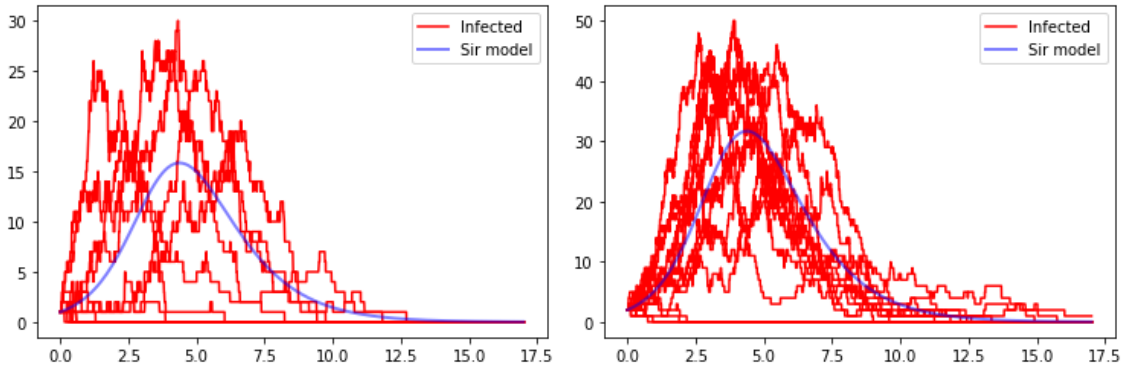

Figure 3.6: 15 simulations for $N = 100$ and $N = 200$.

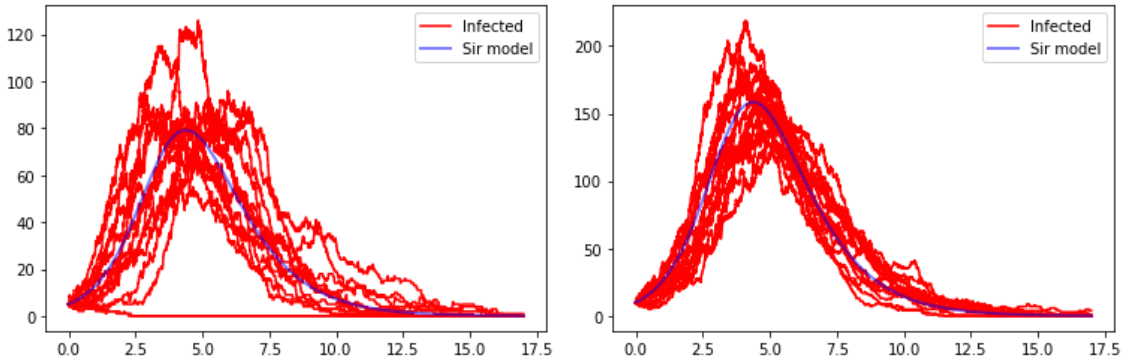

Figure 3.7: 15 simulations for $N = 500$ and $N = 1000$.

It seems that as N is scaled in this way that the SIR queue will behave increasingly like the SIR model, and its variance decreases when scaled. We will prove a formal statement of this in section 4.1.

We now consider the second aforementioned scaling, we let $i_0 = 1$ and consider $N = 500$, 1000, 5000 and 10000.

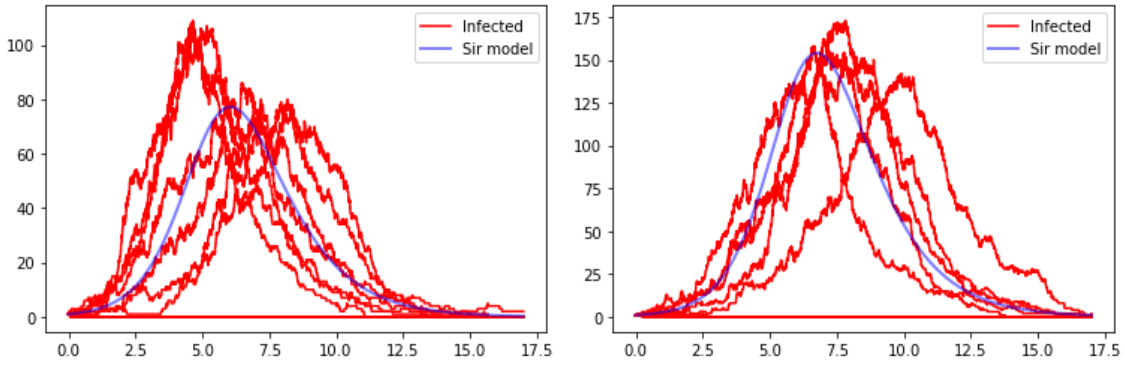

Figure 3.8: 15 simulations for $N = 500$ and $N = 1000$.

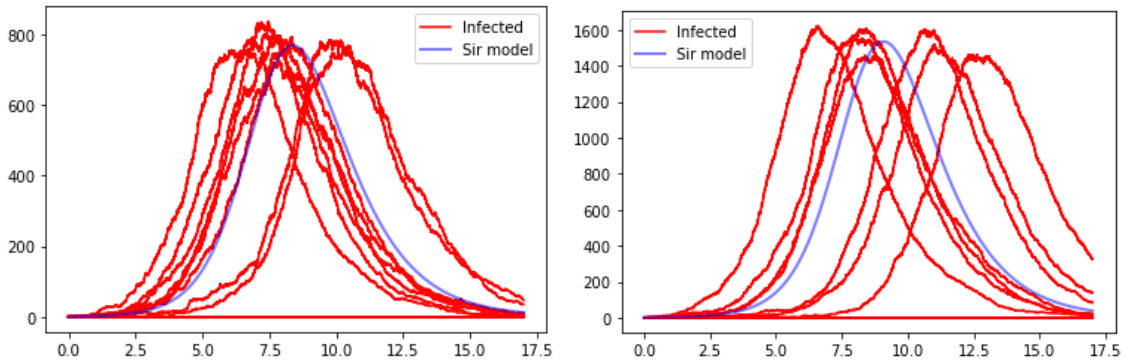

Figure 3.9: 15 simulations for $N = 5000$ and $N = 10000$.

We see that as we increase N the infected still regularly die out, the red lines on the horizontal axis correspond to sample paths that terminated early, and that while the shape of the sample paths becomes more similar to the SIR model, the location of the curve is still random. We will argue in the next chapter that there are two different regimes linked together in this type of scaling, where the population is large but the initial group of infected is small. In the initial regime, that is while $I(t)$ is small compared to the population, there are two effects that impact the overal course of epidemic:

- The infected group might go extinct before becoming large.
- The time until $I(t)$ becomes large and therefore enters the second regime has a significant variance and largely determines the location of the curve.

Once $I(t)$ becomes big enough it enters the second regime and starts to resemble the SIR model. We will try to formally state these ideas and prove them in the next chapter. To get some further insight in the distribution of the hitting times we simulated 300 sample paths with the parameters $\lambda = 1.5$ and $N = 10000$ and then $N = 15000$ but only until they hit $\tau_{N,\varepsilon} = \{\inf t | I(t) = n\varepsilon \text{ or } I(t) = 0\}$ for $\varepsilon = 0.015$.

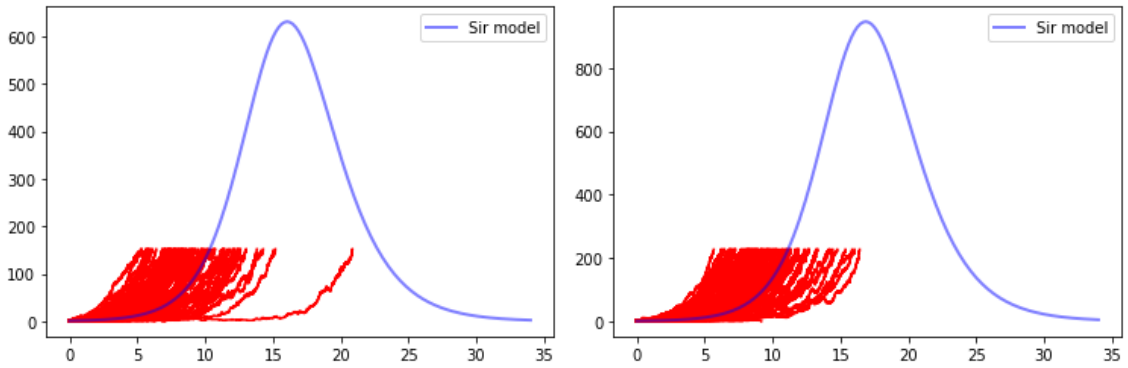

Figure 3.10: 300 simulations for $N = 10000$ and $N = 15000$.

We obtain for $N = 10000$ based on this sample a mean $\tau_{N,\varepsilon}$ of 9.04 and a variance of 6.30 and for $N = 15000$ a mean of 9.82 and a variance of 5.50.

3.3 A brief discussion on the problems of modelling with the SIR queue and SIR model

While real world modelling is not the focus of this thesis we have to adres a striking difference between the SIR model and SIR queue and for example the case count of Coronavirus in the Netherlands. The SIR queue and model will tend to peak once and then drift towards extinction, in fact one of the most readily obtained results is that the effective utilisation is non increasing in time, while the previously mentioned case count has many different peaks. There are three principle adaptions one could make to the SIR model or queue to allow for multiple peaks.

The first one is to allow the intensity of the contact process λ to vary in time. For both reasons that are independent of the levels of infection such as the weather and holidays, and reasons that depend on current levels of infection such as lockdowns, the λ parameter might wildly vary over time. Multiple peaks in the λ parameter might correspond to peaks in the size of the infected group.

We have furthermore assumed that there is a homogenous contact process, however in practice there might be densely connected clusters of people such as cities that are relatively isolated from other clusters. When modelling this could be simulated by assigning every densely clustered subpopulation a SIR queue, and considering the whole population the sum of these queues.

Finally it is possible that a different variation of the disease behaves, from the perspective of the SIR model, like a new disease: the recovered population being once again susceptible. This could also be simulated by summing different SIR queues or models together, one for each variant.

What we haven't mentioned here is the loss of immunity over time, which would correspond with the recovered group feeding into the susceptible group. While this might be a realistic assumption it would not necessarily lead to multiple peaks, but rather to convergence to an equilibrium.

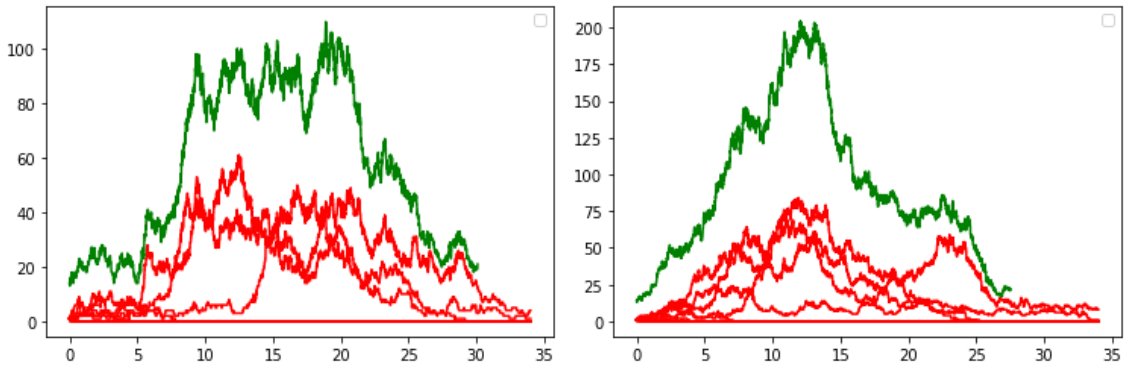

Figure 3.11: Two simulations where we sum (in green) over multiple sample paths of the infected queue.

4 Limiting behaviour

In this chapter we will discuss and prove the key results of our analysis. In section 4.1 we will apply a scaling that converges to the deterministic SIR model. The scaling used will be the one introduced in [16], and the technical framework from [1] ,especially chapter 6, that we will call the probabilistic framework. The key insight that allows us to use these methods is that the growth of the SIR queue behaves like a $M/M/\infty$ queue in that its utilisation at any time is dependent on the current state of the queue. The fact that this method could be expanded to the considerably more complex SIR queue shows that the probabilistic framework is a powerful framework. Proving the convergence to the SIR model justifies both the SIR model and the SIR queue as tools to model the spread of epidemics. The SIR model is justified because it is the limit of a model of randomised contact between individuals that infect each other, and the SIR queue is justified because its limit is the SIR-model, a model with widespread use and empirical justification, see e.g. [17]. Our second key question about linking the SIR model and the SIR queue will thus be answered.

For section 4.2 we separate the initial behaviour from the behaviour for large $I(t)$ into two regimes, and then link these regimes together. It is through this somewhat nuanced task that we accomplish a non-trivial addition to the SIR model. We scale the SIR queue in such a way that some stochastic behaviour is retained, even for large populations. This answers our third key question on what remains from our queueing theoretic approach once it is linked to the SIR model.

4.1 Global behaviour of the scaled SIR queue

We discuss our key convergence result, theorem 4.1.6, in this section. We build upon the framework found in [1] and expand it to a two dimensional system, doing so requires a familiarity with Q - matrices, Poisson processes and some martingale and stochastic integration theory. We briefly go through the necessary definitions and lemmas in appendix 5.2 and 5.3.

We state two technical lemmas here without proof, see [1] appendix A and B.

Lemma 4.1.1. *Let J a countable set and for $j \in J$ and*

- *$C_j \subseteq \mathbb{R}^d$.*
- *V_j is a marked Poisson process on $\mathbb{R}_+ \times \mathbb{Z}^d$ with intensity $\lambda_j dt \otimes \nu_j(dm)$, $\lambda_j \geq 0$ and ν_j a probability distribution on $\mathbb{Z}^d$. V_j, $j \in J$ are assumed to be independent.*
- *$\sum_{j\in J} \lambda_j < \infty$.*

Then the stochastic differential equation

$$dX(t) = \sum_{j\in J} \int_{\mathbb{Z}^d} \mathbb{1}_{\{X(t-)\in C_j\}} m V_j(dt, dm) \tag{4.1}$$

exists and has a unique solution s.t. $X(0) = x_0 \in \mathbb{Z}^d$.

Lemma 4.1.2. *Let (X(t)) a càdlàg Markov process, then if for any f on the statespace S*

$$f(X(t)) - f(X(0)) - \int_0^t Q(f)(X(s))ds \tag{4.2}$$

is a local martingale, then $X(t)$ has the same distribution as the Markov process starting from $X(0)$ with Q-matrix Q. See 5.2.2 for a definition of $Q(f)$.

We now begin applying the machinery of the probabilistic method. It is a method for proving the convergence between systems of stochastic partial differential equations (SPDEs) and PDEs. The first objective is to describe the SIR queue in terms of SPDEs, in theorem 4.1.3 we establish this link. We do this in order to then split up the SIR queue into a martingale part and a remaining process in 4.1.6.

theorem 4.1.3. ***Correspondence theorem*** *Let $\mathcal{N}_{\lambda/N}^{(s,i)}$, $\mathcal{N}_\nu^i$ be independent Poisson processes with the index (s,i) and i, and parameters λ/N and ν respectively. The SIR queue starting in $(s_0,i_0)^t$ with parameters N, λ has the same distribution as the unique solution of*

$$dL(t) = \begin{pmatrix} dS(t) \\ dI(t) \end{pmatrix} = \sum_{s=1}^{S(t-)} \sum_{i=1}^{I(t-)} (-1,1)^t \mathcal{N}_{\lambda/N}^{(s,i)}(dt) + \sum_{i=1}^{I(t-)} (0,-1)^t \mathcal{N}_\nu^i(dt) \tag{4.3}$$

with $L(0) = x_0 = \begin{pmatrix} s_0 \\ i_0 \end{pmatrix}$ or, equivalently

$$\begin{aligned} L(t) = x_0 &+ \sum_{s=1}^{\infty} \sum_{i=1}^{\infty} \int_{(0,t]} \mathbb{1}_{\{S(u-)\geq s\}} \mathbb{1}_{\{I(u-)\geq i\}} (-1,1)^t \mathcal{N}_{\lambda/N}^{(s,i)}(du) \\ &+ \sum_{i=1}^{\infty} \int_{(0,t]} \mathbb{1}_{\{I(u-)\geq i\}} (0,-1)^t \mathcal{N}_\nu^{(i)}(du) \end{aligned} \tag{4.4}$$

Furthermore the process

$$M(t) = L(t) - L(0) - \lambda/N \int_0^t S(u)I(u)(-1,1)^t du - \nu \int_0^t I(u)(0,-1)^t du \tag{4.5}$$

is a local martingale.

Proof. First, for the existence and uniqueness of 5.46 we apply lemma 4.1.1. We have a finite statespace so the $\sum \lambda_i < \infty$ condition is fulfilled. Next, to prove the correspondence we view the SIR queue as generated by its Q-matrix starting in x_0. We recall the Q-matrix of the SIR queue

$$q_{(s,i),(s-1,i+1)} = \lambda si/N, \qquad s > 0, i > 0 \tag{4.6}$$

$$q_{(s,i),(s,i-1)} = \nu i, \qquad i > 0 \tag{4.7}$$

$$q_{xy} = 0, \qquad \text{otherwise if } x \neq y. \tag{4.8}$$

So $Q(f)(s,i) = \lambda/N si[f(s-1,i+1) - f(s,i)] + \nu i[f(s,i-1) - f(s,i)]$ is the generator of Q, see appendix 5.2.

To prove the correspondence between the Markov process with Q-matrix Q and 5.46 we have by 4.1.2 that it is enough to show that for every function f on $\mathbb{N}^2$ that

$$f(L(t)) - f(L(0)) - \int_0^t Q(f)(L(s))ds \tag{4.9}$$

is a local martingale. The process $L(t)$ a.s. jumps by $(-1,1)^t$ or $(0,-1)^t$, we'll use this to find and expression for $df(L(t))$ which determines $f(L(t))$. So,

$$df(L(t)) = \lim_{s\uparrow t} f(L(t)) - f(L(s)) \tag{4.10}$$

$$= [f(L(t-) + (-1,1)^t) - f(L(t-))] \sum_{s=1}^{S(t-)} \sum_{i=1}^{I(t-)} \mathcal{N}_{\lambda/N}^{(s,i)}(dt) \tag{4.11}$$

$$+ [f(L(t-) + (0,-1^t) - f(L(t-)] \sum_{i=1}^{I(t-)} \mathcal{N}_{\nu}^{i}(dt). \tag{4.12}$$

So integrating over t gives

$$f(L(t)) = f(L(0)) + R(t) + \frac{\lambda}{N} \int_0^t [f(L(u-) + (-1,1)^t) - f(L(u-))] S(u-) I(u-) du \tag{4.13}$$

$$+ \nu \int_0^t [f(L(u-) + (0,-1)^t) - f(L(u-))] I(u-) du \tag{4.14}$$

$$= f(L(0)) + R(t) + \int_0^t Q(f)(L(u-)) du, \tag{4.15}$$

and

$$f(L(t)) = f(L(0)) + R(t) + \int_0^t Q(f)(L(u)) du, \tag{4.16}$$

where we define $R(t)$ as

$$R(t) = \sum_{s=1}^{\infty} \sum_{i=1}^{\infty} \int_{(0,t]} \mathbb{1}_{\{S(s-)\geq s\}} \mathbb{1}_{\{I(s-)\geq i\}} f(L(u-) + (-1,1)^t) - f(L(u-)(\mathcal{N}_{\lambda/N}^{(s,i)}(du) - \lambda/N du)$$

$$+ \sum_{i=1}^{\infty} \int_{(0,t]} \mathbb{1}_{\{I(s-)\geq i\}} f(L(u-) + (0,-1)^t) - f(L(u-))(\mathcal{N}_{\nu}^{(i)}(du) - \nu du). \tag{4.17}$$

Since we integrate with regards to the Poisson process minus the drift $R(t)$ is a local martingale, see 5.3.2. So

$$f(L(t)) - f(L(0)) - \int_0^t Q(f)(L(u)) du \tag{4.18}$$

is indeed a local martingale for any function f on the state space. We conclude that the process $L(t)$ has the same distribution as the SIR queue associated with the Q-matrix Q starting from x_0. Furthermore by taking $f(x) = x$ we see that $M(t)$ is a local martingale proving the theorem. □

In order to analyse the fluid limit of the SIR queue we first need an additional lemma on the $M(t)$ process, that allows us to bound the process under scaling later. In one dimension calculating the quadratic variation process suffices. Since $M(t)$ is a process in two dimensions we calculate a 'quadratic variation process' for a norm. The most important property for our use will be that $< M(t) >$ is such that $\mathbb{E}[\|M(t)\|^2] = \mathbb{E}[< M(t) >]$. Let $\|a\|^2 = \sum_i a_i^2$, the norm generated by the dot product on $\mathbb{R}^n$.

Lemma 4.1.4. *Let*

$$< M(t) >= 2\lambda/N \int_{(0,t]} S(u)I(u)du + \nu \int_{(0,t]} I(u)du \tag{4.19}$$

then this is the process such that $\|M(t)\|^2 - < M(t) >$ *is a martingale.*

Proof. We know from the correspondence theorem that equation 5.47 is a martingale, we rewrite $M(t)$ as follows

$$\begin{aligned} M(t) =& \sum_{s=1}^{\infty}\sum_{i=1}^{\infty} \int_{(0,t]} \mathbb{1}_{\{S(u-)\geq s\}}\mathbb{1}_{\{I(u-)\geq i\}}(-1,1)^t(\mathcal{N}^{(s,i)}_{\lambda/N}(du) - \lambda/Ndu) \\ &+ \sum_{i=1}^{\infty}\int_{(0,t]} \mathbb{1}_{\{I(u-)\geq i\}}(0,-1)^t(\mathcal{N}^{(i)}_{\nu}(du) - \nu du). \end{aligned} \tag{4.20}$$

A piecewise approach suffices: $\|M(t)\|^2 = M_1^2(t) + M_2^2(t)$ where M_i is the i-th component of $M(t)$, so if we find $< M_1(t) >$ and $< M_2(t) >$ then $< M(t) >=< M_1(t) > + < M_2(t) >$ will be our process. We'll do the second component first since it involves terms in two different sums.

$M_2^2(t)$ is the sum of the diagonal terms

$$\begin{aligned} &\sum_{s=1}^{\infty}\sum_{i=1}^{\infty}(\int_{(0,t]} \mathbb{1}_{\{S(u-)\geq s\}}\mathbb{1}_{\{I(u-)\geq i\}}(\mathcal{N}^{(s,i)}_{\lambda/N}(du) - \lambda/Ndu))^2 \\ &+ \sum_{i=1}^{\infty}(\int_{(0,t]} \mathbb{1}_{\{I(u-)\geq i\}}(\mathcal{N}^{(i)}_{\nu}(du) - \nu du))^2. \end{aligned} \tag{4.21}$$

the cross terms

$$\begin{aligned} \sum_{s=1}^{\infty}\sum_{i=1}^{\infty}\sum_{s'=1}^{\infty}\sum_{i'=1}^{\infty} &\int_{(0,t]} \mathbb{1}_{\{S(u-)\geq s\}}\mathbb{1}_{\{I(u-)\geq i\}}(\mathcal{N}^{(s,i)}_{\lambda/N}(du) - \lambda/Ndu) \\ &\int_{(0,t]} \mathbb{1}_{\{S(u-)\geq s'\}}\mathbb{1}_{\{I(u-)\geq i'\}}(\mathcal{N}^{(s',i')}_{\lambda/N}(du) - \lambda/Ndu) \end{aligned}$$

with $(s,i) \neq (s',i')$ and

$$\sum_{i=1}^{\infty}\sum_{i'=1}^{\infty}\int_{(0,t]} \mathbb{1}_{\{I(u-)\geq i\}}(\mathcal{N}^{(i)}_{\nu}(du) - \nu du)\int_{(0,t]} \mathbb{1}_{\{I(u-)\geq i'\}}(\mathcal{N}^{(i')}_{\nu}(du) - \nu du)$$

with $i \neq i'$, and finally

$$\begin{aligned} -2\sum_{s=1}^{\infty}\sum_{i=1}^{\infty}&\int_{(0,t]} \mathbb{1}_{\{S(u-)\geq s\}}\mathbb{1}_{\{I(u-)\geq i\}}(-1,1)^t(\mathcal{N}^{(s,i)}_{\lambda/N}(du) - \lambda/Ndu) \\ &\sum_{i'=1}^{\infty}\int_{(0,t]} \mathbb{1}_{\{I(u-)\geq i\}}(\mathcal{N}^{(i')}_{\nu}(du) - \nu du). \end{aligned}$$

Since all the cross terms are of the form $\int f(L(u-))(\mathcal{N}_\xi du - \xi du)\int g(L(u-)(\mathcal{N}_\zeta du - \zeta du)$, where the Poisson processes are independent, they are by 5.3.4 already martingales. The

increasing process is therefore the process that compensates for 4.21. Then by 5.3.2 we see that

$$< M_2(t) >= \lambda/N \sum_{s=1}^{\infty}\sum_{i=1}^{\infty} \int_{(0,t]} (\mathbb{1}_{\{S(u)\geq s\}}\mathbb{1}_{\{I(u)\geq i\}})^2 du \tag{4.22}$$

$$+ \nu \sum_{i=1}^{\infty} \int_{(0,t]} \mathbb{1}^2_{\{I(u)\geq i\}} du$$

$$= \lambda/N \int_{(0,t]} S(u)I(u)du + \nu \int_{(0,t]} I(u)du. \tag{4.23}$$

Using the same technique to calculate $< M_1(t) >$ we obtain

$$< M_1(t) >= \lambda/N \int_{(0,t]} S(u)I(u)du \tag{4.24}$$

so finally $< M(t) >= 2\lambda/N \int_{(0,t]} S(u)I(u)du + \nu \int_{(0,t]} I(u)du$.

□

We've now laid all the groundwork to scale the $L(t)$ process. The idea will be that we scale the population N, the initial position $L(0)$ and the λ/N parameter and then the resulting process is scaled by $1/N$. The need to scale λ/N can be understood as follows: as we scale the population the $S(u)I(u)$ term grows at rate N^2, so a scaling by λ/N and reduces this growth to order N. This scaling has been introduced in [16].

Let $L_N(t)$ the SIR queue process starting in $(s_0'N, i_0'N)$ with $(s_0', i_0') \in \mathbb{Q}^2 \cap (0,1)^2$ with parameters λ and ν[1]. Then define

$$\bar{L}_N(t) = \frac{L_N(t)}{N}. \tag{4.25}$$

the renormalised process with the components $(\bar{S}_N(t), \bar{I}_N(t))$.

We are confronted with a subtlety here, as we let $N \to \infty$ we have an $L_N(t)$ that doesn't always satisfy $L(0) \in \mathbb{N}^2$, since $(s_0'N, i_0'N)$ might not always be whole numbers. Therefore the resulting $L_N(t)$ won't be a SIR queue as we've constructed it. This can be solved by looking at the limit of the subsequence a_n instead, where a_n is the sequence $a_n = (n+1)lcm(m,p)$, $n+1$ times the least common multiple of the denominators m, p such that $(\frac{a}{m}, \frac{b}{p}) = (s_0', i_0')$, $a, b, m, p \in \mathbb{N}$. Since this won't affect the behaviour in the limit we can assume that $N \in \{a_0, a_1, ...\}$.

Now we have constructed $L_N(t)$ such that it is a SIR queue, we can therefore write it like the integral equation 5.46

$$L_N(t) = x_0 N + \sum_{s=1}^{\infty}\sum_{i=1}^{\infty} \int_{(0,t]} \mathbb{1}_{\{S_N(u-)\geq s\}}\mathbb{1}_{\{I_N(u-)\geq i\}}(-1,1)^t \mathcal{N}^{(s,i)}_{\lambda/N}(du) \tag{4.26}$$

$$+ \sum_{i=1}^{\infty} \int_{(0,t]} \mathbb{1}_{\{I_N(u-)\geq i\}}(0,-1)^t \mathcal{N}^{(i)}_{\nu}(du).$$

Lemma 4.1.5. *For $N \in \{a_0, a_1, ...\}$ the process*

$$\bar{M}_N(t) = \bar{L}(t) - \bar{L}(0) - \lambda/N^2 \int_0^t S_N(u)I_N(u)(-1,1)^t du - \nu/N \int_0^t I_N(u)(0,-1)^t du \tag{4.27}$$

[1] So, in order to fit $L_n(t)$ to a the parameters from a SIR queue with population M we'd pick (s_0', i_0') such that $(s_0'M, i_0'M)$.

is a martingale and furthermore

$$\mathbb{E}[\|\bar{M}_N(t)\|^2] = \frac{1}{N}(2\lambda \int_{(0,t]} \mathbb{E}[\bar{S}_N(u)\bar{I}_N(u)]du + \nu \int_{(0,t]} \mathbb{E}[\bar{I}_N(u)]du) \tag{4.28}$$

$$\leq \frac{1}{N}(\frac{1}{2}\lambda + \nu)t \tag{4.29}$$

Proof. The first part is seen simply by dividing 5.47 by N, using that $N\bar{M}_N(t) = M_N(t)$ is a martingale process corresponding to a SIR queue . Then we use theorem 4.1.4 to see that

$$\mathbb{E}[\|\bar{M}_N(t)\|^2] = \frac{1}{N^2}\mathbb{E}[\|M_N(t)\|^2] \tag{4.30}$$

$$= \frac{1}{N^2}\mathbb{E}[2\lambda/N \int_{(0,t]} S_N(u)I_N(u)du + \nu \int_{(0,t]} I_N(u)du] \tag{4.31}$$

$$= \frac{1}{N}(2\lambda \int_{(0,t]} \mathbb{E}[\bar{S}_N(u)\bar{I}_N(u)]du + \nu \int_{(0,t]} \mathbb{E}[\bar{I}_N(u)]du). \tag{4.32}$$

Next we use that $\bar{S}_N, \bar{I}_N$ are bounded by 1 and given that $\bar{S}_N + \bar{I}_N \leq 1$ we get the bound on the product $\bar{S}_N\bar{I}_N \leq \frac{1}{4}$,

$$\frac{1}{N}(2\lambda \int_{(0,t]} \mathbb{E}[\bar{S}_N(u)\bar{I}_N(u)]du + \nu \int_{(0,t]} \mathbb{E}[\bar{I}_N(u)]du) \tag{4.33}$$

$$\leq \frac{1}{N}(2\lambda \int_{(0,t]} \mathbb{E}[\frac{1}{4}]du + \nu \int_{(0,t]} \mathbb{E}[1]du) \tag{4.34}$$

$$= \frac{1}{N}(\frac{1}{2}\lambda + \nu)t \tag{4.35}$$

concluding the proof. □

We can write the scaled process as the following integral equation

$$\bar{L}_N(t) = \bar{L}_N(0) + \bar{M}_N(t) + \lambda \int_0^t \bar{S}_N(u)\bar{I}_N(u)(-1,1)^t du + \nu \int_0^t \bar{I}_N(u)(0,-1)^t du \tag{4.36}$$

with $\bar{M}_N(t)$ as in the last lemma, then using Jensen's inequality twice we see that our martingale goes to 0 of order $1/\sqrt{N}$

$$\|\mathbb{E}[\bar{M}_N(t)]\| \leq \sqrt{\mathbb{E}[\|\bar{M}_N(t)\|^2]} \leq \frac{f(t)}{\sqrt{N}} \tag{4.37}$$

where $f(t)$ is independent of N. Assuming that $\bar{L}_N(t)$ converges to a $l(t)$ and $\bar{M}_N(t)$ really disappears we'd expect $l(t)$ to satisfy

$$l(t) = l(0) + \lambda \int_{(0,t]} l_1(u)l_2(u)(-1,1)^t du + \nu \int_{(0,t]} l_2(u)(0,-1)^t du \tag{4.38}$$

which is exactly the integral equation of the SIR-model with population $N = 1$ and parameters λ and μ. We now prove this convergence.

theorem 4.1.6. ***Functional law of large numbers for the SIR queue*** *Let* $l(t) = \begin{pmatrix} l_1(t) \\ l_2(t) \end{pmatrix}$ *the unique solution to the SIR-model with parameters* $N = 1$, λ, ν *and initial condition* $l(0) = \begin{pmatrix} s_0' \\ i_0' \end{pmatrix}$. *Then* $\bar{L}_N(t)$ *converges to* $l(t)$ *in* L_1, *so for any* $\varepsilon > 0$,

$$\lim_{N\to\infty} \mathbb{P}(sup_{0\leq s\leq t}\|\bar{L}_N(t) - l(t)\|^2 > \varepsilon) = 0. \tag{4.39}$$

Proof. Note that $l(t)$ can be written as $l(t) = l(0)+\lambda\int_0^t l_1(u)l_2(u)(-1,1)^t du+\nu\int_0^t l_2(s)(0,-1)^t ds$. Let $Z_N(t) = \bar{L}_N(t) - l(t)$ then

$$\sup_{0\le s\le t}\|Z_N(t)\| \le \sup_{0\le s\le t}\|\bar{M}_N(t)\| \tag{4.40}$$

$$+\sqrt{2}\lambda\int_0^t \sup_{0\le s\le u}(\bar{S}_N(s)\bar{I}_N(s) - l_1(s)l_2(s)du \tag{4.41}$$

$$+\nu\int_0^t \sup_{0\le s\le u}(I(s) - l_2(s))du. \tag{4.42}$$

We want to upper bound 4.41 in terms of $\|Z_N(t)\|$ so that we can later apply Gronwall's lemma, define $\varepsilon(s), \delta(s)$ such that $l_1 = \bar{S}_N + \varepsilon$, $l_2 = \bar{I}_N + \delta$. Then

$$\begin{aligned}\bar{S}_N(s)\bar{I}_N(s) - (\bar{S}_N(s)+\varepsilon(s))(\bar{I}_N(s)+\delta(s)) &= -\delta(s)\varepsilon(s) - \bar{S}_N(s)\delta(s) - \bar{I}_N(s)\varepsilon(s)\\ &\le |\delta(s)\varepsilon(s)| + |\bar{S}_N(s)\delta(s)| + |\bar{I}_N(s)\varepsilon(s)|\end{aligned}$$

now we use the fact that $\bar{S}_N(s), \bar{I}_N(s) \in [0,1]$

$$\le |\delta(s)\varepsilon(s)| + |\delta(s)| + |\varepsilon(s)|$$

since also $\delta(s), \varepsilon(s) \in [0,1]$ we have that $\delta(s)\varepsilon(s) \le \delta(s) + \varepsilon(s)$ so

$$\begin{aligned}&\le 2(|\delta(s)| + |\varepsilon(s)|)\\ &= 2||Z_N(s)||_1\\ &\le 2\sqrt{2}||Z_N(s)||\end{aligned} \tag{4.43}$$

we write $||\cdot||_1$ for the manhattan distance, then use that $||x||_1 =< |x|, \bar{1} >\le \sqrt{n}||x||_2$ by Cauchy-Schwarz, here n is the dimension of x and $\bar{1}$ is the vector of ones. Also applying the $||x||_1 \le \sqrt{n}||x||_2$ trick to 4.42 we get

$$\sup_{0\le s\le t}\|Z_N(t)\| \le \sup_{0\le s\le t}\|\bar{M}_N(t)\| \tag{4.44}$$

$$+4\lambda\int_0^t \sup_{0\le s\le u}||Z_N(s)||du \tag{4.45}$$

$$+\nu\sqrt{2}\int_0^t \sup_{0\le s\le u}||Z_N(s)||du. \tag{4.46}$$

Next we find with Jensen's and Doob's inequality ($\mathbb{E}[\sup||X||^P]^{1/p} \le \frac{p}{p-1}\mathbb{E}[||X||^p]^{1/p} \implies \mathbb{E}[\sup||X||^2] \le 4\mathbb{E}[||X||]$) and lemma 4.1.5

$$\mathbb{E}[\sup\|\bar{M}_N(t)\|]^2 \le \mathbb{E}[\sup\|\bar{M}_N(t)\|^2] \tag{4.47}$$

$$\le 4\mathbb{E}[\bar{M}_N(t)] \le \frac{4}{N}(\frac{1}{2}\lambda + \nu)t \le \frac{C}{N} \tag{4.48}$$

where C is a constant w.r.t. N. So let

$$f_N(t) = \mathbb{E}[\sup||Z_n(s)||] \tag{4.49}$$

then we have

$$f_N(t) \le \sqrt{C/N} + (4\lambda + \nu\sqrt{2})\int_0^t f_N(s)ds \tag{4.50}$$

so by Gronwall's lemma (see the appendix 5.3.5)

$$f_N(t) \le \sqrt{C/N}e^{4\lambda+\nu\sqrt{2}t} \tag{4.51}$$

so f_N converges to 0 establishing L_1 convergence. □

We interpret this convergence to mean that as $N \to \infty$ as soon as the $I(t)$ component is at scale, that is there is some $\varepsilon > 0$ such that $I(t) > \varepsilon N$, the scaled path of the SIR queue is deterministic.

4.2 Initial behaviour of the scaled SIR queue

Crucial in the scaling of last section was that the initial condition $\frac{L_N(0)}{N}$ was kept constant as a fraction of the population N. It is however natural to investigate the behaviour of a large scaling population with an initial outbreak of constant size, as it is intuitive to think of an epidemic starting with just 1 case regardless of the total population size. We've seen in the last section that our stochastic system becomes deterministic as $N \to \infty$ when the initial condition is a fraction of the population, in this section we will investigate what stochasticity remains when we scale N when keeping the initial group of infected constant.

The simplest way to model this, taking the population and susceptible group to be infinite, simplifies the initial regime completely to a linear birth-death model, see [14] for a clear overview of some of its important properties. While this is a good way to model the initial regime it doesn't allow for an analysis of the transition from the initial regime to the fluid regime. We'll introduce the initial regime by scaling N, leaving the infected group at $t = 0$ constant, and then defining a hitting time $\tau_{N,\varepsilon}$ when $(I(t) + R(t))$ reaches the εN, a fraction of the total population. We interpret that moment as the end of the initial regime and beginning of the fluid regime. We will give a result for the ratio $I(\tau_{N,\varepsilon})/R(\tau_{N,\varepsilon})$. This ratio can be interpreted as a 'natural' starting point for a fluid regime, so that we pick as initial condition for the fluid regime for some small $\delta > 0$, $(1 - (\delta + \delta/C), \delta, \delta/C)$ where C is the calculated fraction. We'll discuss the significance of this C later in this chapter further.

4.2.1 Linear birth-death model

We start with some results for the Linear Birth Death model, which are necessary to give bounds on our scaled $(I(t), R(t))$ in the next section. We'll call the linear birth-death process $(\bar{I}(t), \bar{R}(t))$, the parameters are λ, ν and the initial condition is $(\bar{I}(0), \bar{R}(0)) = (i_0, 0)$. The dynamics are as follows

$$\bar{I}(t) \to \begin{cases} \bar{I}(t) + 1 & \text{at rate } \lambda \bar{I}(t) \\ \bar{I}(t) - 1 & \text{at rate } \nu \bar{I}(t) \end{cases} \tag{4.52}$$

$$\bar{R}(t) \to \bar{R}(t) + 1 \text{ at rate } \nu \bar{I}(t). \tag{4.53}$$

We are interested in the behaviour as t goes to infinity. When $\lambda \leq \nu$ this system will almost surely terminate in a state $(\bar{I}(t') = 0, \bar{R}(t') = r)$, where $r \in \mathbb{N}$ for some finite time t'. This follows from conditioning on the jump times and gambler's ruin for a simple random walk, see the appendix 5.4.1. For some properties of the time of extinction see [12] pages 8-10, or the section on the final regime at the end of this chapter. This means that for $\lambda \leq \nu$ the system will almost surely never reach the fluid scale. We switch our focus to the complementary case, let $\lambda > \nu$, $\bar{I}(0) = I_0, \bar{R}(0) = 0$.

Lemma 4.2.1. *The probability of extinction is given by*

$$\mathbb{P}(\lim_{t \to \infty} \bar{I}(t) = 0) = (\frac{\nu}{\lambda})^{i_0} \tag{4.54}$$

so the probability of escape is

$$\mathbb{P}(\lim_{t\to\infty} \bar{I}(t) = \infty) = 1 - (\frac{\nu}{\lambda})^{i_0} \tag{4.55}$$

Proof. This is once again seen by conditioning on the jump times and applying Gambler's ruin see the appendix 5.4.1. □

We see that the system either goes extinct or both $\bar{I}(t)$ and $\bar{R}(t)$ go to infinity as t goes to infinity, only in this case will the system reach fluid scale. In this case $\lim(\bar{I}(t), \bar{R}(t))$ itself is not interesting as they both tend to infinity, however it turns out that the ratio of the two converges, so we will give a result for $\lim_{t\to\infty} \frac{\bar{I}(t)}{\bar{R}(t)}$.

Intuitively we see that the derivative of $\bar{I}(t)$ should be $(\lambda - \nu)\bar{I}(t)$ so $\bar{I}(t) \sim I_0 e^{(\lambda-\nu)t}$ and the derivative of $\bar{R}(t)$ should be $\nu\bar{I}(t)$ so $R(t) \sim I_0 \frac{\nu}{(\lambda-\nu)} e^{(\lambda-\nu)t}$. This would imply that $I(t)/R(t) \sim \frac{\lambda-\nu}{\nu} = \frac{\lambda}{\nu} - 1$. Making the intuition that $\bar{I}(t) \sim I_0 e^{(\lambda-\nu)t}$ formal requires applying some fluid scaling. If we find some scaling such that $\lim \bar{I}_M(t)/M = I_0 e^{(\lambda-\nu)t}$ and $\lim \bar{R}_M(t)/M = I_0 \frac{\nu}{(\lambda-\nu)} e^{(\lambda-\nu)t}$ then the fraction of these fluid limits will certainly be the same as the fraction of the $\bar{I}(t)$ and $\bar{R}(t)$ processes themselves, since $\frac{\bar{I}(t)/M}{\bar{R}(t)/M} = \frac{\bar{I}(t)}{\bar{R}(t)}$. This turns out to be true but a subtle problem arrises in this line of reasoning. For the fluid limit the initial condition already needs to be at scale, and we are interested specifically in letting the initial condition constant. To bridge this gap we need to split up the proof into a theorem establishing the fluid limit, theorem 4.2.2, and then a theorem that couples a process that starts in the constant initial condition to the fluid limit, theorem 4.2.3. For the fluid limit we follow a scaling from [12], however we prove the results differently using the probabilistic method in [1], also used in the previous section, this allows us to prove the convergence in two dimensions simultaneously which is necessary to obtain a result for the ratio of the two. Because the proof is similar to the previous section's and lengthy we will state it in the appendix, section 5.5. For a different proof of a broader class of birth-death processes without keeping track of the deaths see [12] pages 14-16.

theorem 4.2.2. *For the scaled linear birth-death process* $L_M(t) = (I_M(t), R_M(t))^t = \frac{1}{M}\bar{L}_M(t)$ *where* $\bar{L}_M(t)$ *is a linear birth-death process starting in* $(i_{0,M}, j_{0,M})^t$ *such that* $\bar{L}_M(0) \to (i_0, j_0)^t \in \mathbb{R}^2$ *we have the following convergence*

$$\lim_{M\to\infty} \begin{pmatrix} I_M(t) \\ R_M(t) \end{pmatrix} = \begin{pmatrix} i_0 e^{(\lambda-\nu)t} \\ i_0 \frac{\nu}{(\lambda-\nu)}(e^{(\lambda-\nu)t} - 1) + j_0 \end{pmatrix} \tag{4.56}$$

in the sense that

$$\lim_M \mathbb{P}(\sup_{0\le s\le t} ||L_M(s) - \begin{pmatrix} i_0 e^{(\lambda-\nu)t} \\ i_0 \frac{\nu}{(\lambda-\nu)}(e^{(\lambda-\nu)t} - 1) + j_0 \end{pmatrix}|| > \varepsilon) = 0. \tag{4.57}$$

We now state a limiting result when the linear birth-death process starts in some constant i_0. Intuitively there is a probability the process goes extinct 'before' it reaches the fluid limit stage, so the initial condition for the scaling $\bar{I}(t) \sim i_0 M$ will only be reached with probability $1 - (\frac{\nu}{\lambda})^{i_0}$ as $M \to \infty$. However we need to show that while going to the fluid limit the $R(t)$ queue doesn't grow too large, and is negligible compared to the growth 'after' the system reached the fluid limit.

theorem 4.2.3. *Let* $(\bar{I}(t), \bar{R}(t))^t$ *a linear birth-death process starting in* $(i_0, 0)$ *then*

$$\lim_{t\to\infty} \frac{\bar{I}(t)}{\bar{R}(t)} = \begin{cases} \frac{\lambda}{\nu} - 1 & wp\ 1 - (\frac{\nu}{\lambda})^{i_0} \\ 0 & wp\ (\frac{\nu}{\lambda})^{i_0} \end{cases} \tag{4.58}$$

Proof. We couple the linear birth-death process to the fluid regime through a hitting time $\tau_M = \{\inf t : \bar{I}(t) = M \text{ or } \bar{I}(t) = 0\}$. We immediately see by lemma 4.2.1 that $\tau_M \to \infty$ with the probability that $\bar{I}(t)$ goes to infinity, so $1 - (\frac{\nu}{\lambda})^{i_0}$, and that in the other case $\bar{I}(\tau_M) = 0$. The scaled linear birth-death process will be constructed as follows: let $\frac{1}{M}V_M(t) = \frac{1}{M}(\bar{I}(\tau_M + t), \bar{R}(\tau_M + t))^t$. Then it is known by theorem 4.2.2 that $\lim \frac{1}{M}V_M(t) = (i_0 e^{(\lambda-\nu)t}, i_0 \frac{\nu}{(\lambda-\nu)}(e^{(\lambda-\nu)t-1)}) + j_0)^t$, if the initial value $\lim \frac{1}{M}V_M(0)$ is well behaved. We will prove the well behavedness later and assume this for now. Then, the limiting random variable $V'(t) = \lim_M \frac{1}{M}V_M(t)$ will be equal 0 for all t w.p. $(\frac{\nu}{\lambda})^{i_0}$ and equal $(i_0 e^{(\lambda-\nu)t}, i_0 \frac{\nu}{(\lambda-\nu)}(e^{(\lambda-\nu)t-1)}) + j_0)^t$ w.p. $1 - (\frac{\nu}{\lambda})^{i_0}$. Therefore $V_1'(t)/V_2'(t)$ behaves as 4.58.

In order to use the fluid limit obtained in theorem 4.2.2 we need that $\frac{1}{M}(\bar{I}(\tau_M), \bar{R}(\tau_M))^t \to (i_0, j_0)^t$. $\bar{I}(\tau_M)/M$ is by construction i_0 or 0 but $R(\tau_M)/M$ is a priori not clear. We need to exclude that $R(\tau_M)/M$ becomes arbitrarily large. As $\bar{R}(t)$ grows at rate $\nu\bar{I}(t)$ and $\bar{I}(t)$ grows at rate $(\lambda-\nu)\bar{I}(t)$, $\bar{R}(t)$ could get many times larger before τ_M. To simplify analysis somewhat we introduce $(\bar{I}'(t), \bar{R}'(t))$ that removes the event of extinction.

$$\bar{I}'(t) \to \begin{cases} \bar{I}'(t) + 1 & \text{at rate } \lambda(\bar{I}'(t) \vee 1) \\ \bar{I}'(t) - 1 & \text{at rate } \nu(\bar{I}'(t) \vee 1) \end{cases} \tag{4.59}$$

$$\bar{R}'(t) \to \bar{R}'(t) + 1 \text{ at rate } \nu(\bar{I}'(t) \vee 1). \tag{4.60}$$

We have that stochastically $\bar{R}(t) \leq \bar{R}'(t)$. We can look at the system $(\bar{I}'(t), \bar{R}'(t))$ given $\bar{I}'(t) + 2\bar{R}'(t)$ as a binomially distributed random variable with parameters $p = \frac{\lambda}{\lambda+\nu}$, where the state $(\bar{I}'(t), \bar{R}'(t)) = (a, b)$ corresponds with the the event of $a + b$ succeses after $a + 2b$ trials.

We want to show there is a constant C such that $\lim_M \mathbb{P}(\bar{R}(\tau_M) > CM) = 0$. For the process $\bar{R}'(t)$ we have that the event $E = \bar{R}'(\tau_M) > CM$ corresponds with

$$E = \{ \text{ At time } \tau_M, \bar{R}'(\tau_M) > CM\} \tag{4.61}$$

$$\overset{(1)}{\subseteq} \cup_{k \geq (2C+1)M} \{\text{After } k \text{ steps of the random walk there are } (C+1)M \text{ succeses for the first time.}\}$$

$$\subseteq \cup_{k \geq (2C+1)M} \{\text{After } k \text{ steps there are } (C+1)M \text{ or less successes.}\} \tag{4.62}$$

$$= E_2 \tag{4.63}$$

Inclusion (1) follows from seeing the event $\bar{I}'(\tau_M)$ as hitting M *more* succeses then there are failures at time τ_M. Since there are at least $(2C+1)M$ steps taken that means $\#succeses = (C+1)M$. We then allow E_2 to contain also all paths that hit level M earlier, and end lower than M. Let $X_k \sim \text{Binom}(k, \frac{\lambda}{\lambda+\nu})$ independent random variables distributed binomially with k trials. We will now show that

$$\mathbb{P}(E_2) = \sum_{k \geq M(2C+1)} \mathbb{P}(X_k \leq (C+1)M) \to 0. \tag{4.64}$$

First we note that $\mathbb{E}[X_k] = k\frac{\lambda}{\lambda+\nu} \geq M(2C+1)\frac{\lambda}{\lambda+\nu}$. Since $\lambda > \nu$ we have that $\frac{\lambda}{\lambda+\nu} > 0.5$, the most difficult situation being when the fraction is close to 0.5. With this in mind we write $\frac{\lambda}{\lambda+\nu} = 0.5 + \varepsilon$ for some $\varepsilon > 0$. We can then write $\mathbb{E}[X_k] \geq M(2C+1)(0.5+\varepsilon) = MC + 0.5M + 2MC\varepsilon + M\varepsilon = M(C+d)$ for $d = 0.5 + (C+1)\varepsilon$. By choosing our C dependent on the probability $\frac{\lambda}{\lambda+\nu}$ we can set our d arbitrarily large, it turns out we will need to set it such that $d > 1$. Let $\mu_{4,k} = kp(1 - p(1 + 3k - 6)p(1-p))$ with $p = \frac{\lambda}{\lambda+\nu}$ the fourth centralised moment of X_k.

We're going to use the Chebyshev inequality for fourth moments in order to estimate the probabilities of X_k deviating from its expectation. This chebyshev inequality is

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq c\sqrt[4]{\mu_4}) \leq \frac{1}{c^4}. \tag{4.65}$$

We are going to choose $c = \sqrt[3]{k}$. Then we calculate that for large enough M, picking C such that $d > 1$ we have

$$k\frac{\lambda}{\lambda+\nu} - \sqrt[3]{k}\sqrt[4]{\mu_4} \geq \tag{4.66}$$

$$M(C+d) - \sqrt[3]{M(2C+1)}\sqrt[4]{\mu_4} \geq (C+1)M. \tag{4.67}$$

We use here that $\sqrt[3]{k}\sqrt[4]{\mu_{4,k}} \leq C_1\sqrt[6]{K^5}$ for some constant C_1 so for large enough k the equation $k\frac{\lambda}{\lambda+\nu} - \sqrt[3]{k}\sqrt[4]{\mu_4}$ is growing in k, and that $\sqrt[3]{M(2C+1)}\sqrt[4]{\mu_{4,M(2C+1)}} \leq C_1\sqrt[6]{M^5}$. This equation shows that the region of values lower then $(C+1)M$ is far enough away from the mean value to apply equation 4.65.

$$\mathbb{P}(E_2) = \sum_{k \geq M(2C+1)} \mathbb{P}(X_k \leq (C+1)M) \tag{4.68}$$

$$\leq \sum_{k \geq M(2C+1)} \mathbb{P}(|X_k - \mathbb{E}[X_k]| \geq \sqrt[3]{M(2C+1)}\sqrt[4]{\mu_4}) \tag{4.69}$$

$$\leq \sum_{k \geq M(2C+1)} \frac{1}{k^{4/3}} \tag{4.70}$$

Where equation 4.70 is smaller then $\sum_i \frac{1}{i^{4/3}} = \zeta(4/3) = 3.60...$ so letting $M \to \infty$ gives

$$\lim_M \mathbb{P}(E_2) = 0. \tag{4.71}$$

This shows that the growth of R before τ_M will be insignificant compared to the growth after τ_M, and therefore negligible when taking $t \to \infty$. This concludes the proof. □

4.2.2 The initial regime, transition to the fluid regime

We will define the initial regime for a SIR queue $(S(t), I(t), R(t))$ starting in $(N - i_0, i_0, 0)$ as the system before $S(t) \leq (1-\varepsilon)N$ for some small $\varepsilon > 0$. We define the exit time $\tau_{N,\varepsilon}$.

Definition 4.2.4. ***Exit time*** *Let $(S(t), I(t), R(t))$ the SIR queue with initial condition as above and population N then*

$$\tau_{N,\varepsilon} = \inf_{t \geq 0}\{t : S(t) \leq (1-\varepsilon)N\} \tag{4.72}$$

is the exit time. The SIR queue with $t \leq \tau_{N,\varepsilon}$ is in the initial regime. If the SIR queue terminates before $S(t) \leq (1-\varepsilon)N$ we take $\tau_{N,\varepsilon}$ to be infinite.

So $t < \tau_{N,\varepsilon} \implies (1-\varepsilon)N \leq S(t) \leq N$. In this section we will assume $t \leq \tau_{N,\varepsilon}$. Our goal is to find a scaling of $N \to \infty$ and $\varepsilon \to 0$ such that we can give a non trivial result for $\lim_{(N,\varepsilon)\to(\infty,0)} I(\tau_{N,\varepsilon})/R(\tau_{N,\varepsilon})$ however this turns out to be a somewhat subtle question. We begin with giving an upper and lower bound for $(I(t), R(t))$ in terms of linear birth-death processes.

Let $\underline{I}(t)$, $\bar{I}(t)$ a linear birth-death processes starting with a population i_0 with parameters $(\lambda - \varepsilon\lambda, \mu)$ and (λ, μ) respectively. Then we have $\mathbb{P}(\underline{I}(t) \leq x) \leq \mathbb{P}(I(t) \leq x) \leq \mathbb{P}(\bar{I}(t) \leq x)$.

Note that since $I(t)$ will peak and then go down this stochastic ordering is only true if ε is picked such that this peak has not happened yet. $I(t)$ peaks when $\rho(t) = \frac{\lambda S(t)I(t)/N}{\nu I(t)} = \frac{\lambda}{\nu}\frac{S(t)}{N}\mathbb{1}_{\{I(t)\neq 0\}} = 1$ so when $S(t) = \frac{\nu}{\lambda}N$. So $\varepsilon < \frac{\nu}{\lambda}$ will suffice, and since we let $\varepsilon \to 0$ and this inequality doesn't depend on N this poses no problems.

Setting 4.2.5. *Let $t \leq \tau_{N,\varepsilon}$. We let $\lambda > \nu$ in order for the infection to not almost surely go extinct. $(S(t), I(t), R(t))$ is the SIR queue starting in $(N - i_0, i_0, 0)$. Let further more the processes $(\bar{I}(t), \bar{R}(t))^t$, $(I(t), R(t))^t$ and $(\underline{I}(t), \underline{R}(t))^t$ be defined on the same probability space such that for every sample path one has $\bar{I}(t)(\omega) \geq I(t)(\omega) \geq \underline{I}(t)(\omega)$. This is possible by taking the sigma algebra generated by $\bar{I}(t)$ and marking every jump with probability such that the marked process is I and then again for $\underline{I}$.*

Let $\underline{R}(t)$, $\bar{R}(t)$ the processes that keep track of the deaths in $\underline{I}(t)$, $\bar{I}(t)$ respectively (note they have the same stochastic ordering with regards to $R(t)$ as $\underline{I}(t)$, $\bar{I}(t)$ have with regards to $I(t)$). We then have the stochastic ordering (since we don't have the property that for all paths $\underline{R}(t)(\omega) \leq R(t)(\omega) \leq \bar{R}(t)(\omega)$ like we have for $I(t)$)

$$\frac{\underline{I}(\tau_{N,\varepsilon})}{\bar{R}(\tau_{N,\varepsilon})} \leq \frac{I(\tau_{N,\varepsilon})}{R(\tau_{N,\varepsilon})} \leq \frac{\bar{I}(\tau_{N,\varepsilon})}{\underline{R}(\tau_{N,\varepsilon})}. \tag{4.73}$$

In order to apply theorem 4.2.3, it is necessary for $\tau_{N,\varepsilon} \to \infty$. This is the **first property** our scaling needs to have. Since $\tau_{N,\varepsilon}$ is the moment such that $I(\tau_{N,\varepsilon}) + R(\tau_{N,\varepsilon}) = \varepsilon N$ it is stochastically bigger then $\bar{\tau}_{N,\varepsilon}$ defined as $\inf_t\{t : \bar{I}(t) + \bar{R}(t) \geq \varepsilon N\}$. This hitting time $\bar{\tau}_{N,\varepsilon}$ then clearly goes to infinity almost surely as $\varepsilon N \to \infty$. Therefore the first property comes down to picking a scaling such that $\varepsilon N \to \infty$.

As $\tau_{N,\varepsilon} \to \infty$ we have that

$$\frac{\underline{I}(\tau_{N,\varepsilon})}{\bar{R}(\tau_{N,\varepsilon})} \to \frac{\lambda - \nu}{\nu}\frac{e^{(\lambda(1-\varepsilon)-\mu)\tau_{N,\varepsilon}}}{e^{(\lambda-\mu)\tau_{N,\varepsilon}}} = \frac{\lambda - \nu}{\nu}e^{-\varepsilon\tau_{N,\varepsilon}\lambda} \tag{4.74}$$

and

$$\frac{\bar{I}(\tau_{N,\varepsilon})}{\underline{R}(\tau_{N,\varepsilon})} \to \frac{\lambda - \varepsilon\lambda - \nu}{\nu}\frac{e^{(\lambda-\mu)\tau_{N,\varepsilon}}}{e^{(\lambda(1-\varepsilon)-\mu)\tau_{N,\varepsilon}}} = \frac{\lambda - \varepsilon\lambda - \nu}{\nu}e^{\varepsilon\tau_{N,\varepsilon}\lambda}. \tag{4.75}$$

So to apply the squeeze lemma to 4.73 we need the **second property** of the scaling, that is that $\varepsilon\tau_{N,\varepsilon} \to 0$ almost surely. In order to find the correct scaling we need to understand $\varepsilon\tau_{N,\varepsilon}$ a bit better.

We define $\underline{\tau}_{N,\varepsilon}$ symmetrically to $\bar{\tau}_{N,\varepsilon}$ and notice the stochastic ordering $\underline{\tau}_{N,\varepsilon} \geq \tau_{N,\varepsilon} \geq \bar{\tau}_{N,\varepsilon}$. Conditional on the underlying processes not going extinct (we can ignore this case since then clearly $I(t)/R(t) \to 0$) we have that $\underline{\tau}_{N,\varepsilon}$ and $\bar{\tau}_{N,\varepsilon}$ are of order $\frac{log(\varepsilon N)}{\lambda-\nu}$. Letting $\varepsilon = f(N)$ for some f such that $f(N) \to 0$ we get that the second property means f has to solve $log(f(N)N)f(N) \to 0$ as $N \to \infty$. A simple class of functions that fulfils both properties is $f(N) = N^{-\delta}$ for some $\delta \in (0, 1)$.

theorem 4.2.6. *Assume setting 4.2.5. Let $\varepsilon = N^{-\delta}$ for some $\delta \in (0, 1)$ then*

$$\lim_{N\to\infty} \frac{I(\tau_{N,\varepsilon})}{R(\tau_{N,\varepsilon})} = \frac{\lambda - \nu}{\nu} = \frac{\lambda}{\nu} - 1 \tag{4.76}$$

with probability $1 - (\frac{\nu}{\lambda})^{i_0}$ and 0 otherwise.

Interpretation of equation 4.76

As discussed at the start of this section the constant $\frac{\lambda-\nu}{\nu}$ can be used to calculate a natural starting point for the fluid regime. As the fluid regime requires a somewhat arbitrary constant $\xi > 0$, the percentage of the population that is infected at $t = 0$, equation 4.76 leads to a natural choice for the recovered group at $t = 0$: $R(0) = \frac{\nu}{\lambda-\nu}\xi$. As it is a function of λ/ν it also leads to an estimator for λ/ν when at the beginning op an epidemic both $I(t)$ and $R(t)$ are known, this follows from the relation $\frac{\lambda}{\nu} = \frac{I(\tau_{N,\varepsilon})}{R(\tau_{N,\varepsilon})} + 1$.

On the distribution of $\tau_{N,\varepsilon}$

The relation $\tau_{N,\varepsilon} \geq \tau_{N,\varepsilon} \geq \bar{\tau}_{N,\varepsilon}$ implies that $\tau_{N,\varepsilon}$ behaves like the hitting time of a linear birth-death process $\bar{\tau}_{N,\varepsilon}$. One possible route of investigation is therefore to study the first passage time distribution of linear birth death processes conditioned on not going extinct. The distribution of first passage times when scaling N resists easy characterisation however. When M is fixed the distribution is known in terms of the eigenvalues of the Q-matrix, see [18], this could lead to some results when letting $N \to \infty$. Outside of the property that $\bar{\tau}_{N,\varepsilon}$ is of order $\log(N\varepsilon)$ we did not manage this. Simulations of $\tau_{N,\varepsilon}$ don't offer conclusive support for either our thesis that the variance remains after scaling or its complement that the variance goes to 0 as we scale N. We simulate the mean and variation of $\tau_{N,\varepsilon}$ from a sample of 300 paths starting in $i_0 = 1$, parameters $\lambda = 1.5$, $\nu = 1$, we leave $\varepsilon = 0.015$.

	$N = 15000$	30000	45000	60000
$\mathbb{E}[T_{N,\varepsilon}]$	10.18	11.83	12.91	13.36
$\mathbb{V}ar[\tau_{N,\varepsilon}]$	6.30	6.33	8.80	8.03
$\mathbb{V}ar[\tau_{N,\varepsilon}/\log(N\varepsilon)]$	0.21	0.17	0.20	0.17

4.2.3 Brief discussion on modelling a 'final regime'

We showed that the behaviour of the initial regime is still stochastic even when scaling the population, and that this model becomes deterministic in the 'middle regime' which we called the fluid regime. We can symmetrically to the initial regime model a final regime, investigating the behaviour of the SIR queue when it is close to going extinct, 'after' the fluid regime. When letting the SIR queue progress the $I(t)$ component will become small again at the end when it the SIR queue is almost terminated. We look at the SIR queue after the hitting time $\tau'_{N,\varepsilon} = \{\inf t : I(t)/N \leq \varepsilon \text{ and } S(t)/N\rho_0 \leq 1\}$ where we also demand that $S(t)/N\rho_0 \leq 1$ to guarantee this time is after the epidemic has peaked. We use the initial regime as a framework to understand the final regime: because $I(t)$ is small compared to $S(t)$ the $I(t)$ will be effectively approximated by linear birth-death processes $I_-(t)$ and $I_+(t)$. We also know that as the SIR model comes to an end in $S_{end}/N \to -W(-\rho_0 \frac{S_0}{N} e^{-\rho_0})$ by lemma 2.1.3. Therefore we can estimate

$$(-W(-\rho_0 \frac{S_0}{N} e^{-\rho_0}) - \delta)\lambda I(t) \leq \lambda I(t)S(t)/N \leq -W(-\rho_0 \frac{S_0}{N} e^{-\rho_0})\lambda I(t). \tag{4.77}$$

The parameters of the linear birth-death processes are thus $\lambda_- = -W(-\rho_0 \frac{S_0}{N} e^{-\rho_0}) - \delta$ and $\lambda_+ = -W(-\rho_0 \frac{S_0}{N} e^{-\rho_0})$ and ν, this leads to the stochastic ordering $I_+(t) \geq I(t) \geq I_-(t)$. These linear birth-death processes start at $\tau'_{N,\varepsilon}$ with $I(\tau'_{N,\varepsilon}) = I_+(\tau'_{N,\varepsilon}) = I_-(\tau'_{N,\varepsilon}) = N\varepsilon$. They have a negative effective utilisation so they will almost surely terminate in a state $(\cdot, 0, \cdot)$. Since this regime is the tail of the fluid regime we know that $(S(t)/N, I(t)/N, R(t)/N) \to$

$(-\frac{W(-\rho_0 \frac{S_0}{N} e^{-\rho_0})}{\rho_0}, 0, 1 + \frac{W(-\rho_0 \frac{S_0}{N} e^{-\rho_0})}{\rho_0})$. The most interesting question that remains is how the extinction time is distributed. Let τ_0^+, τ_0, τ_0^- the hitting time $\tau_0^{\cdot} = \inf_t\{t \geq \tau'_{N,\varepsilon} | I^{\cdot}(t) = 0\}$ belonging to the processes I_+, I, I_- respectively. Then the hitting times have the stochastic ordering $\tau_0^- \geq \tau_0 \geq \tau_0^+$. Since τ_0^- and τ_0^+ are extinction times that belong to linear birth-death processes with negative drift their first moment is finite and an explicit form is known. Proposition 2.6 in [12] gives a result for the moments of hitting times. We introduce $\mathbb{E}_M[T^-_{M-1}]$ the expected time hitting $N-1$ starting in N for the process I_-, and write $\xi_- = -W(-\rho_0 \frac{S_0}{N} e^{-\rho_0}) - \delta$, for its growth parameter. Then we have the following result.

Lemma 4.2.7. ***Expectation of the extinction time*** *For a linear birth-death process starting in M, with birth parameter ξ, death parameter ν and $\nu > \xi$ we have*

$$\mathbb{E}_M[T_{M-1}] = \frac{1}{\nu M} + \frac{\xi}{\nu^2} \Phi(\xi/\nu, 1, M+1) \tag{4.78}$$

where $\Phi(\xi/\nu, 1, M+1)$ is the Lerch transcendent defined as $\Phi(x, 1, c) = \sum_{i=0}^{\infty} \frac{x^k}{c+1}$. Then we have that the extinction time has expectation $\mathbb{E}_M[T_0] = \sum_{j=1}^{M} \mathbb{E}_j[T_{j-1}]$ so

$$\mathbb{E}_M[T_0] = \sum_{j=1}^{M} \frac{1}{\nu j} + \frac{\xi}{\nu^2} \Phi(\xi/\nu, 1, j+1). \tag{4.79}$$

In [12] also a result for the second moment of T_{M-1} is given.

Lemma 4.2.8. ***Second moment and variance of the extinction time*** *With the setup as in lemma 4.2.7 we obtain*

$$\mathbb{E}_M[T^2_{M-1}] = 2 \left(\frac{\nu}{\xi}\right)^M \sum_{j \geq M} \left(\frac{\xi}{\nu}\right)^j \mathbb{E}_{j+1}[T_j]^2 \tag{4.80}$$

so the variation of the extinction time is given by

$$\mathbb{V}ar_M(T_0) = \sum_{k=1}^{M} \mathbb{E}_M[T^2_{M-1}] - \mathbb{E}_M[T_{M-1}]^2. \tag{4.81}$$

Then by letting $\delta \to 0$ one obtains the moments for the SIR queue in the final regime by the squeeze lemma.

Discussion and further research

The most immediately productive unanswered question seems to us to be a further characterisation of the hitting time distribution $\tau_{N,\varepsilon}$. What makes the analysis difficult is that $\mathbb{E}[\tau_{N,\varepsilon}]$ is infinite for any linear birth-death process since it never hits $N\varepsilon$ with some positive probability (unless N, ε is lower than the starting state and $\rho_0 < 1$, the property of almost sure extinction), so we need to condition the underlying process on hitting that state to obtain finite moments for the hitting time. Doing this will transform the linear birth-death process into a process that isn't described in terms of simple Markov dynamics and some new framework is needed. Given the abundance of literature on population models and branching processes we are certain there is a way of productively engaging this problem. Specifically we want to prove some result about how the variance doesn't disappear against the time scaling of the SIR queue, which is of order $log(N)$. An example of such a result would be $\mathbb{V}ar(\tau_{N,\varepsilon}/log(N)|I(t)$ does not terminate early$) > c$ for some constant $c > 0$. We need the additional condition that $I(t)$ does not terminate early because then the variation would be unbounded as N grows.

In chapter 4 we managed to prove a functional law of large numbers for the SIR queue. However in [1] Robert manages to also prove a functional central limit theorem for the $M/M/\infty$ queue. Although this queue is substantially more simple than the SIR queue, from an analysis perspective they behave somewhat similar, that is to say the maximum rate of growth for both queues is similarly well behaved in expectation. We suspect for a certain scaling $f(N)$ a result is possible for $(f(N)[\bar{L}_N(t) - l(t)]) \to X$ where X is some non degenerate distribution. Such a result would greatly aid the understanding of the relation between the SIR model and the SIR queue in a very precise way. We did not manage because the methods [1] uses are one dimensional and already complex. Expanding upon the probabilistic framework will certainly be productive, since the proofs of the convergence of the linear birth death process and the SIR queue were so similar there might be a result possible that proves the convergence for a wide class of SPDEs to their PDEs. The category will need properties that govern the maximum rates of growth of the SPDEs, such that the offered load to queues can be no more then polynomials of the component queues.

Complications of the model are possible, such as adding testing, or allowing a more complex distribution in the contact process, but when the properties that make the SIR queue so difficult to analyse (chance of extinction, finite population, transience) are retained one would have to resort to computations or simulations. An interesting question is the loss of immunity, if the $R(t)$ queue feeds back into the $S(t)$ queue analysis of the system might actually be facilitated by stationary behaviour, and explicit results might be available.

Popular summary

In the modelling of disease spread the SIR model is often used. This model is broadly applicable and captures the initial exponential growth, the peak, and the exponential shrinking of an epidemic well. It does however have the downside that it is continuous: the model transitions from 1 ill person to 2 by crossing all the values between 1 and 2. The second downside is that it is deterministic, the time between 1 and 2 ill people will always be the same. In reality an epidemic moves discretely, that is to say it jumps from 1 ill person to 2, and stochastically, it might take a day to get 2 ill people, or a week, or it might not happen at all. In this thesis we investigate an alternative model that takes this into account and see what we can say about this new model. Then we compare the new model with the SIR model by zooming out enough that the discrete changes start behaving continuously, and we investigate if our zoomed out model behaves identically to the SIR model, or if there is still a difference.

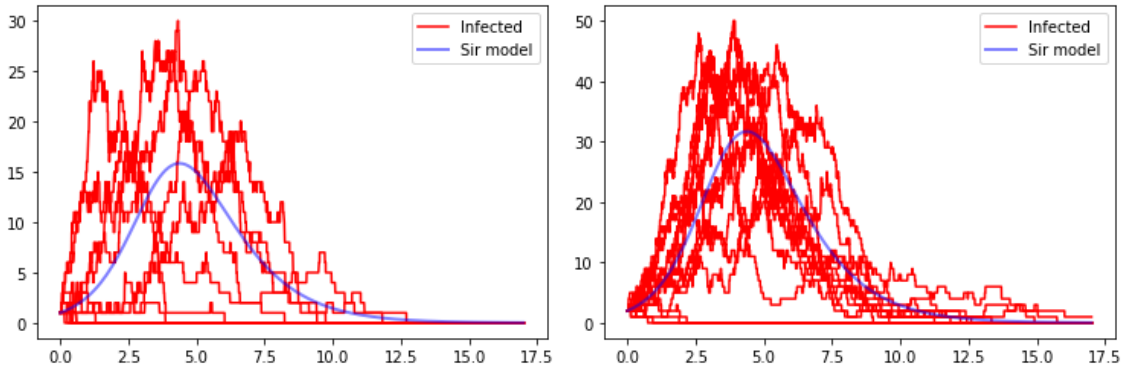

Figure 4.1: 15 simulations for $N = 100$ and $N = 200$.

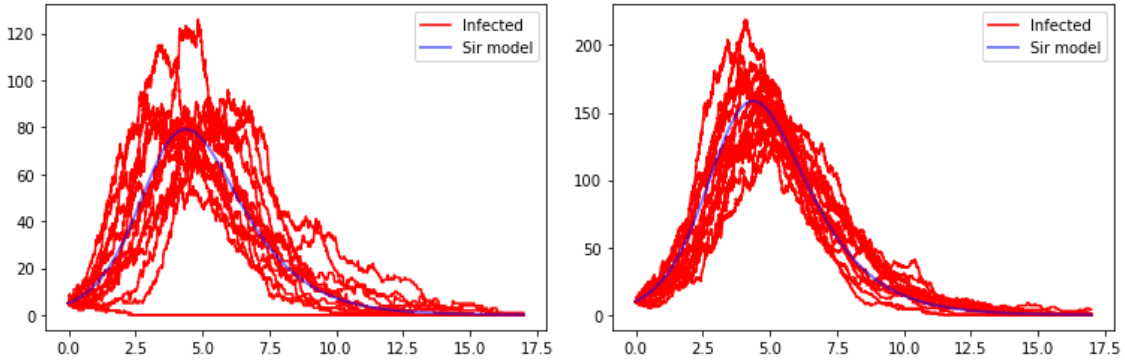

Figure 4.2: 15 simulations for $N = 500$ and $N = 1000$.

In these pictures we plot in red the sample paths for increasingly big populations, and in blue the deterministic SIR model. We see that the SIR model captures the shape of

the sample paths most of the time, and for bigger populations the SIR model the sample paths behave more nicely, and more often.

As it turns out there are two things that remain random when we approximate the SIR model using our techniques, this is the probability that the epidemic goes extinct before it reaches a lot of people, and the time when it breaks out. Once enough people are infected the difference between the discrete and stochastic model and the smooth SIR model becomes very small, however the moment the epidemic switches from just a couple of infected people to an outbreak remains random.

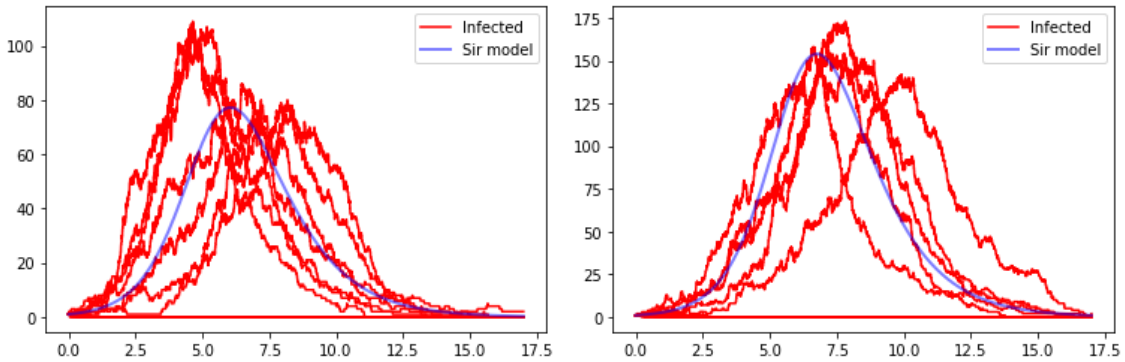

Figure 4.3: 15 simulations for $N = 500$ and $N = 1000$.

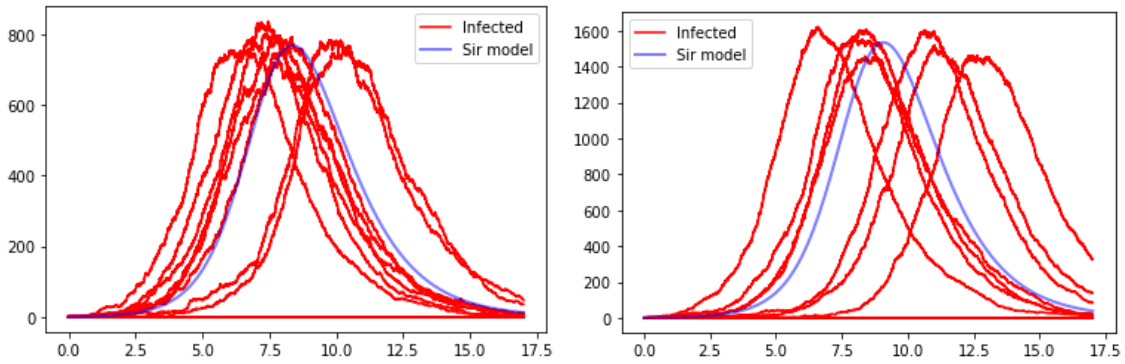

Figure 4.4: 15 simulations for $N = 5000$ and $N = 10000$.

When simulating very large population numbers but starting the outbreak every time at a single person, like we imagine an outbreak starting usually, we can see this phenomena with our own eyes. Notice that the red lines on the x axes, they are the sample paths of an infection that quickly exterminates.

We conclude that our random model is similar enough to the SIR model that investigation is justified, however proving theorems formally about it, such as the distribution of the time it hits the peak, turns out to be very difficult.

5 Appendix

5.1 Lambert W function, proof of lemma 2.1.3

The Lambert W function is the inverse of xe^x so the characterising equation is

$$x = W(x)e^{W(x)}. \tag{5.1}$$

This function has infinite branches, we only consider the principle branch which is defined on $x \geq -e^{-1}$. This is the only branch we'll consider in this text.

Lemma 5.1.1. *The equation $ax + b + ce^{dx} = 0$ is solved by $x = -\frac{b}{a} - \frac{1}{d}W(\frac{cde^{-bd/a}}{a})$*

Proof. We use that $x = W(x)e^{W(x)}$ implies $e^{-W(x)} = \frac{W(x)}{x}$ and then simply substitute $x = -\frac{b}{a} - \frac{1}{d}W(\frac{cde^{-bd/a}}{a})$ into $ax + b + ce^{dx}$. □

Proof of lemma 2.1.3

Proof. We simply apply 5.1.1 to obtain the formula $1+\frac{W(-\rho_0 \frac{s_0}{N} e^{-\rho_0})}{\rho_0}$. Since we're looking to calculate $W(-\rho_0 s_0/Ne^{-\rho_0})$ we have to show that $\rho_0 s_0/Ne^{-\rho_0} \leq e^{-1}$. Using $s_0/N \leq 1$ we obtain $\rho_0 s_0/Ne^{-\rho_0} \leq \rho_0 e^{-\rho_0}$. We calculate the maximum value of $\rho_0 e^{-\rho_0}$ by differentiating and finding $\frac{d}{d\rho_0}\rho_0 e^{-\rho_0} = e^{-\rho_0}(1 - \rho_0)$, which equals 0 only at $\rho_0 = 1$ and goes to $-\infty$, 0 as ρ_0 goes to $-\infty$ and ∞ respectively. Therefore $\rho_0 = 1$ is a global maximum, so $\rho_0 e^{-\rho_0} \leq e^{-1}$ □

5.2 Q-matrices

Definition 5.2.1. *A Q-matrix is a matrix associated with a Markov process $X(t)$ on the state space $\mathcal{S}$ as follows. For $Q = (q_{ij}|i, j \in \mathcal{S})$:*

$$q_{ij} = lim_{s\to 0}\frac{1}{s}\mathbb{P}(X(s) = j|X(0) = i), \quad i \neq j$$
$$q_{ii} = -\sum_{i\neq j} q_{ij}$$

Definition 5.2.2. *The functional operator of the Q-matrix Q, for a non-negative function f on S, is given by*

$$Q(f)(i) = \sum_{j\in S} q_{ij}f(j) = \sum_{j\neq i} qij(f(j) - f(i)) \tag{5.2}$$

5.3 Martingales and Stochastic Differential Equations

We briefly recall the definition of stochastic integral w.r.t. Poisson processes as given in the appendix of [1]. We will also use the notation and some results for Poisson processes given in chapter one of that book.

For $\lambda \geq 0$ let the point process $\mathcal{N}_\lambda$ a Poisson process on $\mathbb{R}_+$. Furthermore let $\mathcal{F}_t$, $t \geq 0$ the σ-algebra generated by the random variables $\mathcal{N}_\lambda((0,s])$ where $s \leq t$. Note that the process $\mathcal{N}_\lambda((0,t]) - \lambda t$ is a martingale. Let X an adapted càdlàg process.

Definition 5.3.1. *We define the the stochastic integral of X w.r.t. this martingale as*

$$Z(t) = \int_0^t X(s)[\mathcal{N}_\lambda(\omega, ds) - \lambda(ds)] = \sum_{n\in\mathbb{N}, 0<t\leq t_n} X(t_n) - \lambda \int_0^t X(s)ds, \tag{5.3}$$

or equivalently

$$dZ(t) = \lim_{s\uparrow t} Z(t) - Z(s) = X(t)[\mathcal{N}_\lambda(dt) - \lambda dt]. \tag{5.4}$$

Lemma 5.3.2. *If $X(t)$ is a càdlàg adapted process, the process*

$$\int_{(0,t]} X(s-)[\mathcal{N}_\lambda(\omega, ds) - \lambda ds] \tag{5.5}$$

is a martingale and

$$\lambda \int_0^t X^2(s)ds \tag{5.6}$$

its increasing process.

Definition 5.3.3. ***Increasing process*** *Let $M(t)$ a local martingale for which there exists a non-decreasing sequence of stopping times $\lim_{\tau_n\to\infty}$ such that $M(\tau_n \wedge t)$ is square integrable then there exists a unique deterministic non-decreasing process $< M(t) >$ such that*

$$M(t)^2 - < M(t) > \tag{5.7}$$

is a local martingale, we call $< M(t) >$ the increasing process of $M(t)$.

Lemma 5.3.4. *Let f, g bounded functions on $\mathbb{N}$ and $X(s)$ a càdlàg process, μ, $\lambda > 0$ and $\mathcal{N}_\lambda$, $\mathcal{N}_\mu$ independent Poisson processes, then the process*

$$\int_0^t f(X(s-))(\mathcal{N}_\lambda(ds) - \lambda ds) \int_0^t g(X(s-))(\mathcal{N}_\mu - \mu ds) \tag{5.8}$$

is a martingale.

Proof. See [1] page 356. □

Lemma 5.3.5. ***Gronwall's lemma*** *Let g and f non-negative measurable functions on $\mathbb{R}_+$ and let $\varepsilon > 0$ such that for all $s \leq t$*

$$f(s) \leq \varepsilon + \int_0^s f(u)g(u)du \tag{5.9}$$

then,

$$f(s) \leq \varepsilon \exp\left(\int_0^s h(u)du\right). \tag{5.10}$$

5.4 Gamblers ruin

Let $S_N = \mathbb{N}$ the state space for a Markov process R_n characterised by

$$R_{n+1} \to R_n \text{ if } R_n = 0$$
$$R_{n+1} \to \begin{cases} R_n + 1 & \text{w.p. } p \\ R_n - 1 & \text{w.p. } q \end{cases}$$

with $p + q = 1$.

theorem 5.4.1. ***Gambler's ruin*** *Let $R_0 = i \in \mathbb{N}$ then if $p \leq q$ then*

$$\mathbb{P}(\exists n \in \mathbb{N} : R_n = 0) = 1. \tag{5.11}$$

That is to say that almost surely the absorbing state 0 will be reached regardless of the starting position. Furthermore if $p > q$ we have

$$\mathbb{P}(\exists n \in \mathbb{N} : R_n = 0) = (q/p)^i. \tag{5.12}$$

Proof. We refer for a proof to [11] page 15. □

5.5 Proof of theorem 4.2.2

We'll use the acronym LBDP for linear birth-death process. Since these proofs will share a significant similarity with the proofs given in 4.1 we will summarise the proofs rather than give them in full detail. We split up the proof into 4 lemmas.

Lemma 5.5.1. ***Lemma 1, correspondence lemma.*** *The process $(\bar{I}(t), \bar{R}(t))^t$ has the same distribution as the solution to the stochastic differential equation $L(t) = (L_1(t), L_2(t))$*

$$dL(t) = \sum_{i=1}^{L_1(t)} (1,0)^t \mathcal{N}_\lambda^{(i,1)}(dt) + \sum_{i=1}^{L_1(t)} (-1,1)^t \mathcal{N}_\nu^{(i,2)}(dt) \tag{5.13}$$

starting in $L(0) = (\bar{I}(0), \bar{R}(0))^t$ where $\mathcal{N}_\lambda^{(a,b)}$ are independent Poisson distributions with parameter λ. Equivalently,

$$L(t) = L(0) + \sum_{i=1}^{\infty} \int_0^t \mathbb{1}_{\{L_1(u-) \geq i\}} (1,0)^t \mathcal{N}_\lambda^{(i,1)}(du) + \sum_{i=1}^{\infty} \int_0^t \mathbb{1}_{\{L_1(u-) \geq i\}} (-1,1)^t \mathcal{N}_\nu^{(i,2)}(du). \tag{5.14}$$

Furthermore the process

$$M(t) = L(t) - L(0) - (\lambda - \nu) \int_0^t (1,0)^t L_1(u) du - \nu \int_0^t (0,1)^t L_1(u) du \tag{5.15}$$

is a martingale.

Proof. Let Q the Q-matrix associated with the process $(\bar{I}(t), \bar{R}(t))^t$. Then Q is given by

$$q_{(i,j)(i+1,j)} = \lambda i, \quad i > 0 \tag{5.16}$$
$$q_{(i,j)(i-1,j+1)} = \nu i, \quad i > 0 \tag{5.17}$$

and $Q(f)(n_1,n_2) = \lambda n_1[f(n_1+1,n_2)-f(n_1,n_2)] + n_1\nu[f(n_1-1,n_2+1)-f(n_1,n_2)]$. We can now simply follow the proof of the correspondence theorem obtaining

$$f(L(t)) = f(L(0)) + R(t) + \int_0^t Q(f)(L(u))du \tag{5.18}$$

with

$$R(t) = \sum_{i=1}^{\infty} \int_0^t \mathbb{1}_{\{L_1(u-)\geq i\}}[f(L(u-)+(1,0)^t) - f(L(u-)](\mathcal{N}_\lambda^{(i,1)}(du) - \lambda du) \tag{5.19}$$

$$+ \sum_{i=1}^{\infty} \int_0^t \mathbb{1}_{\{L_1(u-)\geq i\}}[f(L(u-)+(-1,1)^t) - f(L(u-)](\mathcal{N}_\nu^{(i,2)}(du) - \nu du). \tag{5.20}$$

We apply lemma 4.1.2 and let $f(x) = x$ to conclude the proof. □

Lemma 5.5.2. ***Lemma 2, increasing process lemma.*** *The process*

$$< M(t) >= (\lambda + 2\nu)\int_0^t L_1(u)du \tag{5.21}$$

is the increasing process such that $||M(t)||^2 - < M(t) >$ *is a martingale.*

Proof. First we write down $M(t)$ in the form

$$M(t) = \sum_i \int_0^t (1,0)^t \mathbb{1}_{\{L_1(u-)\geq i\}}(\mathcal{N}_\lambda^{i,1}(du) - \lambda du) \tag{5.22}$$

$$+ \sum_i \int_0^t (-1,1)^t \mathbb{1}_{\{L_1(u-)\geq i\}}(\mathcal{N}_\nu^{i,2}(du) - \nu du). \tag{5.23}$$

Then using similar calculations as in the proof of lemma 4.1.4 we find that the only non martingale terms in $||M(t)||^2$ are the diagonal terms

$$M_1^2(t) = \sum_i \left(\int_0^t \mathbb{1}_{\{L_1(u-)\geq i\}}(\mathcal{N}_\lambda^{i,1}(du) - \lambda du)\right)^2 \tag{5.24}$$

$$- \sum_i \left(\int_0^t \mathbb{1}_{\{L_1(u-)\geq i\}}(\mathcal{N}_\nu^{i,2}(du) - \nu du)\right)^2. \tag{5.25}$$

and

$$M_2^2(t) = \sum_i \left(\int_0^t \mathbb{1}_{\{L_1(u-)\geq i\}}(\mathcal{N}_\nu^{i,2}(du) - \nu du)\right)^2. \tag{5.26}$$

Then by lemma 5.3.2 we and that $< M(T) >=< M_1(t) > + < M_2(t) >$ we obtain

$$< M(t) >= (\lambda + 2\nu)\int_0^t L_1(u)du \tag{5.27}$$

□

We scale the LBDP $L(t)$ with M such that $\bar{L}^M(t)/M = \frac{1}{M}(\bar{I}_M(t), \bar{R}_M(t))^t$ where $(\bar{I}_M(t), \bar{R}_M(t))^t$ is a LBDP starting in $(i_0^M, j_0^M)^t\mathbb{R}^2$ such that $\frac{1}{M}(i_{0,M}, j_{0,M})^t$ converges to some $(i_0, j_0)^t \in \mathbb{R}^2$. We call $L_M = (I_M(t), R_M(t)) = \bar{L}^M(t)/M$ a scaled LBDP.

Lemma 5.5.3. ***Lemma 3 upper bound lemma.*** *Let L_M a scaled LBDP such that $L_M(0) \to (i_0, j_0)^t$, then*

$$M_M(t) = L_M(t) - L_M(0) - \lambda \int_0^t (1,0)^t I_M(u)du - \nu \int_0^t (-1,1)^t I_M(u)du \tag{5.28}$$

is a martingale such that

$$\mathbb{E}[||M_M(t)||^2] \leq \frac{1}{M}(\lambda + 2\nu)f(t) \tag{5.29}$$

for some $f(t)$ independent from M.

Proof. Equation 5.5.3 follows just from dividing $\bar{L}^M(t)$ by M. For equation 5.29 we can apply the proof of lemma 4.1.5 if we know that a LBDP starting in $\bar{I}(0) = i_{0,M}/M$ is in expectation independent of M. Consider a LBDP starting in $i_{0,M}$, it is in distribution the same as the convolution of $i_{0,M}$ LBDPs starting in 1. Let X_1 a $LBDP$ starting in 1. Knowing that $i_{0,M}/M \to i_0$ we know there exists a C such that $i_{0,M}/M < C$ for all M. So $\mathbb{E}(I_M(t)) < C\mathbb{E}[X_1(t)] = Ce^{(\lambda-\nu)t}$ concludes the proof. □

We can now write $L_M(t) = L_M(0) + M_M(t) + \lambda \int_0^t (1,-1)^t I_M(t)du + \nu \int_0^t (-1,1)^t I_M(u)du$ and since $M_M(t) \to 0$ we expect the solution $l(t)$ to satisfy $y(t) = y(0) + \lambda \int_0^t (1,-1)^t l_1(t)du + \nu \int_0^t (-1,1)^t l_1(u)du$, so

$$\frac{d}{dt}l_1 = (\lambda - \nu)l_1 \tag{5.30}$$

$$\frac{d}{dt}l_2 = \nu l_1 \tag{5.31}$$

starting in $(i_0, j_0)^t$. The solution is $l(t) = (i_0 e^{(\lambda-\nu)t}, i_0 \frac{\nu}{\lambda-\nu}(e^{(\lambda-\nu)t} - 1) + j_0)^t$.

Lemma 5.5.4. ***Lemma 4 the functional law of large numbers for the LBDP.*** *Let $L_M(t)$ a scaled LBDP such that $\lim L_M(0) = (i_0, j_0)^t$ and let $l(t) = (i_0 e^{(\lambda-\nu)t}, i_0 \frac{\nu}{\lambda-\nu}(e^{(\lambda-\nu)t} - 1) + j_0)^t$. Then $L_M(t)$ converges to $l(t)$ in the sense that*

$$\lim_M \mathbb{P}(\sup_{0\leq s\leq t} ||L_M(s) - l(s)|| > \varepsilon) = 0. \tag{5.32}$$

Proof. We write $Z_M(t) = L_M(t) - l(t)$ and since $l(t)$ satisfies $l(t) = (i_0, j_0)^t + \lambda \int_0^t (1,-1)^t l_1(t)du + \nu \int_0^t (-1,1)^t l_1(u)du$ we obtain

$$Z_M(t) = (L_M(0) - (i_0, j_0)^t) + M_M(t) + \lambda \int_0^t (1,0)^t Z_{M,1}(u)du + \nu \int_0^t (-1,1)^t Z_{M,1}(u)du \tag{5.33}$$

where $Z_{M,1}$ is the first coordinate Z_M. Then we use the same proving technique as in theorem 4.1.6.

$$\begin{aligned}\sup_{0\leq s\leq t} ||Z_M(t)|| \leq & (L_M(0) - (i_0, j_0)^t) + \sup_{0\leq u\leq t} ||M_M(u)|| \\ & + \lambda \int_0^t \sup_{0\leq u\leq t} Z_{M,1}(u)du + \sqrt{2}\nu \int_0^t \sup_{0\leq u\leq t} Z_{M,1}(u)du.\end{aligned} \tag{5.34}$$

Then defining $f_M(t) = \mathbb{E}\left(\sup_{0\leq u\leq t} ||Z_M(u)||\right)$ we can apply Gronwall's lemma to $f_N(t) \leq ||(L_M(0) - (i_0, j_0)^t)|| + \sqrt{C_1/M} + (\lambda + 2\nu) \int_0^t f_M(u)du$ to prove the convergence. □

Formula sheet

SIR model PDEs:

$$s'(t) = -\lambda s(t)i(t)/N \tag{5.35}$$
$$i'(t) = \lambda s(t)i(t)/N - \nu i(t) \tag{5.36}$$
$$r'(t) = \nu i(t) \tag{5.37}$$

SIR queue dynamics:

$$S(t) \to S(t) - 1 \text{ at rate } \lambda S(t)I(t)/N \tag{5.38}$$
$$I(t) \to \begin{cases} I(t)+1 & \text{at rate } \lambda S(t)I(t)/N \\ I(t)-1 & \text{at rate } \nu I(t) \end{cases} \tag{5.39}$$
$$R(t) \to R(t) + 1 \text{ at rate } \nu I(t). \tag{5.40}$$

$$\rho_0 = \frac{\lambda}{\nu} \tag{5.41}$$

Effective utilisation:

$$\rho(t) = \rho_0 \frac{S(t)}{N} \mathbb{1}_{\{I(t)\neq 0\}}. \tag{5.42}$$

Linear birth death process:

$$\bar{I}(t) \to \begin{cases} \bar{I}(t)+1 & \text{at rate } \lambda \bar{I}(t) \\ \bar{I}(t)-1 & \text{at rate } \nu \bar{I}(t) \end{cases} \tag{5.43}$$
$$\bar{R}(t) \to \bar{R}(t) + 1 \text{ at rate } \nu \bar{I}(t). \tag{5.44}$$

The SDE and martingale

$$dL(t) = \begin{pmatrix} dS(t) \\ dI(t) \end{pmatrix} = \sum_{s=1}^{S(t-)} \sum_{i=1}^{I(t-)} (-1,1)^t \mathcal{N}^{(s,i)}_{\lambda/N}(dt) + \sum_{i=1}^{I(t-)} (0,-1)^t \mathcal{N}^{i}_{\nu}(dt) \tag{5.45}$$

with $L(0) = x_0 = \begin{pmatrix} s_0 \\ i_0 \end{pmatrix}$ or, equivalently

$$\begin{aligned} L(t) = x_0 &+ \sum_{s=1}^{\infty} \sum_{i=1}^{\infty} \int_{(0,t]} \mathbb{1}_{\{S(u-)\geq s\}} \mathbb{1}_{\{I(u-)\geq i\}} (-1,1)^t \mathcal{N}^{(s,i)}_{\lambda/N}(du) \\ &+ \sum_{i=1}^{\infty} \int_{(0,t]} \mathbb{1}_{\{I(u-)\geq i\}} (0,-1)^t \mathcal{N}^{(i)}_{\nu}(du) \end{aligned} \tag{5.46}$$

Furthermore the process

$$M(t) = L(t) - L(0) - \lambda/N \int_0^t S(u)I(u)(-1,1)^t du - \nu \int_0^t I(u)(0,-1)^t du \tag{5.47}$$

is a local martingale.

Bibliography

[1] Phillipe Robert. 2003. Stochastic Networks and Queues. Springer-Verlag Berlin Heidelberg.

[2] RIVM. 2021. Beschrijving transmissiemodel berekening zorgbelasting (version 2 April). www.Rivm.nl

[3] Frank Wang. 2010. Application of the Lambert W Function to the SIR Epidemic Model. VOL. 41, NO. 2, MARCH 2010 THE COLLEGE MATHEMATICS JOURNAL. pp. 156-159

[4] Florian Simatos, Phillipe Robert, Fabrice Guillemin. 2008. A Queueing System for modelling a File Sharing Principle.

[5] Rudesindo Núñez-Queija and Balakrishna Prabhu. 2008. Scaling Laws for File Dissemination in P2P Networks with Random Contacts. 16th Interntional Workshop on Quality of Service (2008): 75-79.

[6] Ray Hill. 1991. Counting and the reflection principle. The Mathematical Gazette , Oct., 1991, Vol. 75, No. 473 (Oct., 1991), pp. 308-312

[7] Muhammad Umar Niazi, Alain Kibangou, Carlos Canudas-de-Wit, Denis Nikitin, Liudmila Tumash and Pierre-Alexandre Bliman. 2020. Modeling and Control of COVID-19 Epidemic through Testing Policies. arXiv.org

[8] Tiberiu Harko, Francisco Lobo, Man Kwong Mak. 2014. Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. 2014. Applied Mathematics and Computation 236 , 184-194.

[9] Piet van Mieghem. 2020. Exact solution of the Kermack and McKendrick SIR differential equations. Delf University of Technology.

[10] William Ogilvy Kermack, Anderson Gray McKendrick. 1927. A Contribution to the Mathematical Theory of Epidemics. Proceedings of the Royal Society London A, vol. 115 No 772 , pp 700-721

[11] James Norris. 1997. Markov chains (Cambridge Series in Statistical and Probabilistic Mathematics). Cambridge: Cambridge University Press.

[12] Vincent Bansaye, Sylvie Méléard. 2017. Some stochastic models for structured populations : scaling limits and long time behavior. hal-01162547v2

[13] Krishna Athreya, Peter Ney. 1972. Branching processes. Springer-Verlag Berlin Heidelberg New York.

[14] Simon Tavaré. 2010. The Linear Birth-Death Process: An Inferential Retrospective. Advances in Applied Probability , Volume 50 , Issue A: Branching and Applied Probability , December 2018 , pp. 253 - 269

[15] Thomas Dence. 2013. A Brief Look into the Lambert W Function.

[16] Francis Patrick Kelly. 1986. “Blocking Probabilities in Large Circuit-Switched Networks.” Advances in Applied Probability, vol. 18, no. 2, Applied Probability Trust, pp. 473–505.

[17] Mark Shapiro, Fazle Karim, Guido Muscioni, Abel Augustine. 2021. Adaptive Susceptible-Infectious-Removed Model for Continuous Estimation of the COVID-19 Infection Rate and Reproduction Number in the United States: Modeling Study

[18] James Fill. 2009. The passage time distribution for a birth-and-death chain: Strong stationary duality gives a first stochastic proof