# Stochastic Epidemic Modeling

Priscilla E. Greenwood and Luis F. Gordillo

**Abstract** We review the topic of stochastic epidemic modeling with emphasis on compartmental stochastic models. A main theme is the usefulness of the correspondence between these and their large population deterministic limits, which describe dynamical systems. The dynamics of an ODE system informs us of the deterministic skeleton upon which the behavior of corresponding stochastic systems are built. In this chapter we present a number of examples, mostly in the context of susceptible-infected-recovered (SIR) models, and point out how this way of thinking may be useful in understanding other stochastic models. In particular we discuss the distribution of final epidemic size, the effect of different patterns of infectiousness, and the quantification of stochastically sustained oscillations.

## 1 Introduction

The topic of stochastic epidemic modeling is huge. There are many possible types of stochastic epidemic model. The decision of which type of model to choose, or to invent a new one, depends on the specific question to be explored and the data which is at hand or can be obtained. This chapter is a brief guide for newcomers, to the literature and to the construction of compartmental stochastic models. We will indicate some of the history of the subject in the next section. The one class of stochastic models which we will describe in some detail, compartmental models, is introduced in Section 3. Their natural form is multivariate Markov jump processes. When populations are large, they correspond, in the sense of non-limit approxima-

Priscilla E. Greenwood
Department of Mathematics and Statistics, Arizona State University, Tempe, AZ 85287-1804, e-mail: pgreenw@math.asu.edu

Luis F. Gordillo
Department of Mathematics, University of Puerto Rico-RUM, Mayaguez, PR 00681-9018 e-mail: gordillo@math.uprm.edu

tion, to systems of stochastic differential equations. Their large population limits are systems of ordinary differential equations.

Following the section on stochastic compartmental models we will describe three stochastic phenomena that illustrate some of the questions to which these models can yield answers. They happen to be questions to which we have recently contributed. The first, in Section 4, concerns the form of the distribution of the final size of an epidemic. In the context of a susceptible-infected-removed (SIR) epidemic, the final size distribution is bimodal, quite strikingly if the reproduction number is just slightly larger than one. Hence a prediction of epidemic size based on deterministic modeling may be meaningless.

Section 5 concerns stochastically sustained oscillations, which occur if the corresponding dynamical system has damped oscillations. Such sustained oscillations may help to explain the semi-regular recurrence of infectious disease outbreaks. Multiscale analysis has allowed the phenomenon to be interpreted in terms of stochastic process behavior, so that the role it plays in oscillatory disease phenomena can be quantified.

Another interesting stochastic effect, in Section 6, concerns a class of stochastic models in which nearly all homogeneity is abandoned. Still it is possible to say something about the distribution of epidemic size. It depends on the infectiousness of infected participants only through their total, or integrated, infectiousness.

A concluding section contains general observations about the essential role of dynamical systems analysis in the understanding of stochastic dynamic effects in epidemic models and additional examples.

## 2 History

An early stochastic epidemic model was proposed by A.G. McKendrick in 1926, [38], which precedes his work with Kermack on deterministic models, [28]. An account of McKendrick's paper can be found in [25]. In 1928 and 1931, Reed and Frost, and Greenwood proposed discrete time stochastic models, which proceeded by generations of infectives, [18]. The Reed-Frost model was not published at the time, but was presented in lectures in 1928. Bartlett, [14], studied a continuous time stochastic SIR model, and this began a large literature of which we mention only a few highlights. The book of Bailey [6] (first printed in 1957) is about both deterministic and stochastic epidemic models and the estimation of their parameters. In 1993 the Isaac Newton Institute in Cambridge held a semester-long workshop on stochastic epidemic modeling. Three collections of papers edited by Mollison [40], by Isham and Medley [26], and by Grenfell and Dobson [21] resulted. The monograph of Daley and Gani [18] is probably the best general source on this subject. In particular their Chapter 1 is a fine account of the early history of epidemic modeling in general. In fact Daley and Gani are two of the outstanding figures on this area. Another authoritative survey, including maximum likelihood estimation and Monte Carlo Markov Chain (MCMC) methods is by Anderson and Britton, [3]. A revised edition is currently under preparation. Among further prominent contributors to the area are Frank Ball [7, 8, 9, 11] Andrew Barbour [12, 13], Niels Becker [15, 16], David Kendall [27], Donald Ludwig [36], Anders Martin-Löf [39], Ingemar Nåsell [41, 42], Gian Paulo Scalia-Tomba [46, 47, 11] , Tom Sellke [48].

## 3 Stochastic Compartmental Models

Before looking at particular epidemic models, let us become familiar with some notation and ideas about stochastic compartmental models. Later the compartments will become disease states and their members, which we refer to here as particles, will be individuals. We will represent compartments, or classes of individuals by boxes, for example three of them as in Figure 1, and define a vector-valued process which describes the movement of particles into, and out of each box. Time is continuous. For each time, $t \geq 0$, $(X_1(t), X_2(t), X_3(t))$ is the number of particles in boxes 1, 2, and 3 respectively, where these three numbers sum up to $N(t)$, the total number of particles at that time.

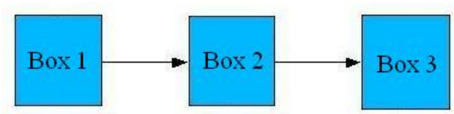


**Fig. 1** A schematic compartmental representation. Particles ”move” betwen different categories.

An underlying structure, basic to the class of stochastic compartmental models, and indeed to all Markov jump processes, is the Poisson process. Suppose there is just one compartment, and just one process, $X(t)$, representing the number of particles in the box at time $t$. Particles enter the box at random times. The initial value, $X(0)$, is fixed and for some $\lambda > 0$,

$$P(X(t+\Delta t) - X(t) = 1) = \lambda \Delta t + o(\Delta t), \tag{1}$$
$$P(X(t+\Delta t) - X(t) = 0) = 1 - \lambda \Delta t + o(\Delta t). \tag{2}$$

The increments of $X(t)$ in disjoint time intervals are independent. Then $X(t), t \geq 0$ is called a Poisson process. The number $\lambda$ is called the *intensity* or the *stochastic rate* of the process. The times between successive jumps of the process are exponentially distributed with parameter $\lambda$. Instead of being constant, $\lambda = \lambda(t)$ may depend on $t$ and may also depend on the value of the process at time $t$. For example, if

$$P(X(t+\Delta t) - X(t) = 1) = aX(t)\Delta t + o(\Delta t), \tag{3}$$

then we say $aX(t)$ is the *conditional instantaneous stochastic rate* of the process at time $t$, where the conditioning is on the value of $X(t)$, or, once this is understood, we shorten this to : $aX(t)$ is the stochastic rate, or simply the rate. The rate per particle is $a$. This process is a *pure birth process*. If our model has several compartments, so that our stochastic process of counts has several components, then the conditional instantaneous stochastic rates of particles entering or leaving each compartment at time $t$ may depend on the sizes of any of the components.

A Markov jump process necessarily has exponentially distributed times between the jumps. We saw this above in the description of the pure birth process, which involves just one component process and one type of jump. A compartmental epidemic model is a vector-valued process with a vector component for each compartment. There are several types of jumps, one type of jump for each arrow in the diagram. If the resulting multi-component process is Markov, each type of jump will occur according to a *locally Poisson* probability as in (1) above, and the times between jumps of any one type, given that nothing occurs in the interim to alter the rate, will be exponentially distributed with parameter given in terms of the states of the component processes at the beginning of the interval. Each component of the Markov jump process can be regarded as a birth and death process, with instantaneous stochastic rates depending on all the components.

The requirement that the inter-jump times be exponentially distributed is not essentially restrictive. There are ways to generalize without losing the advantages of Markov modeling. For example, additional stages can be introduced so that, for instance, one infective step occurs in a sum of independent exponential times, and the result will be a gamma-distributed time. What is important to the resulting stochastic process is how the conditional mean of each increment, each change between $t$ and $t+\Delta t$, relates to the conditional variance of the increment. In the case of conditionally Poisson increments, which are often used in this type of modeling, and yield a Markov structure, the mean and variance are equal. The paper of Lloyd in this volume discusses this point further.

**Example 1.** *A simple stochastic epidemic*. In this example we will consider two compartments corresponding to susceptible and infective individuals. We will use the letters $S$ and $I$ respectively to refer to the compartments and also, without confusion we hope, to the number of individuals in each class. We will assume that the total number of individuals is constant and equal to $N$, that is, $S+I=N$. An individual who belongs to the class $S$ may be contacted by an individual in $I$, who can transfer the infection. If that is the case, the susceptible individual changes his classification and belongs now to the class $I$, where he will remain indefinitely. Assume that individuals in each compartment are interchangeable, that the classes are homogeneously mixed, and that contacts between susceptible and infective individuals, or equivalently the movement of individuals from the class $S$ to the class $I$, occur at random times. If $\beta$ is the average number of contacts made by an average infective per unit of time that leads to an infection, the probability of a susceptible individual moving from class $S$ to class $I$ in the time interval $[t,t+\Delta t]$, that is, $S \to S-1$ and $I \to I+1$, is $\beta \frac{SI}{N}\Delta t + o(\Delta t)$.

This stochastic infection rate has come to be widely used, with various possible interpretations of the $N$ in the denominator. One can think of each susceptible contacting everyone in the population with a rate $\beta$ and encountering a proportion $I/N$ of infectives. Or one may think of each infective contacting everyone in the population with a rate $\beta$ and encountering a proportion $S/N$ of susceptibles. Or one may think of the $N$ in the denominator as a reduction of the infection rate due to incomplete mixing in population. From this last point of view, the denominator might be a different power of $N$ or some other function of $N$. This point is discussed in [49].

The process $(S_t, I_t)$, will represent the number of susceptible and infective individuals at time $t$. The probability of an infection during the time interval $[t, t+\Delta t]$ is

$$P((S_{t+\Delta t}, I_{t+\Delta t}) - (S_t, I_t) = (-1,1)) = \beta \frac{S_t I_t}{N} \Delta t + o(\Delta t), \tag{4}$$

with the complementary probability

$$P((S_{t+\Delta t}, I_{t+\Delta t}) - (S_t, I_t) = (0,0)) = 1 - \beta \frac{S_t I_t}{N} \Delta t + o(\Delta t). \tag{5}$$

**Example 2.** *The stochastic SIR model*. Consider three classes of individuals: susceptible, infected, and removed (by recovery or death). As in the previous example, we will use $S$, $I$ and $R$ to represent the compartments themselves, as well as the numbers of individuals in each compartment, and assume $S+I+R=N$, a constant. Thus, in the time interval $[t, t+\Delta t]$, the probability of an infection, that is, the simultaneous transitions $S \to S-1$ and $I \to I+1$ occur, is $\beta \frac{SI}{N} \Delta t + o(\Delta t)$, as in Example 1. If it is assumed that infected individuals recover with rate $\gamma$, the probability for a recovery, $I \to I-1$ and $R \to R+1$, in the interval $[t, t+\Delta t]$, is $\gamma I \Delta t + o(\Delta t)$. Because $R = N - S - I$, it is enough to consider the process $(S_t, I_t)$. Thus, the probabilities of an infection and of a recovery during the time interval $[t, t+\Delta t]$ are

$$P((S_{t+\Delta t}, I_{t+\Delta t}) - (S_t, I_t) = (-1,1)) = \beta \frac{S_t I_t}{N} \Delta t + o(\Delta t), \tag{6}$$

$$P((S_{t+\Delta t}, I_{t+\Delta t}) - (S_t, I_t) = (0,-1)) = \gamma I_t \Delta t + o(\Delta t), \tag{7}$$

with the complementary probability

$$P((S_{t+\Delta t}, I_{t+\Delta t}) - (S_t, I_t) = (0,0)) = 1 - \left(\beta \frac{S_t}{N} + \gamma\right) I_t \Delta t + o(\Delta t). \tag{8}$$

This model, widely known as the *general stochastic epidemic*, was introduced by Barlett in 1949, [14]. An extensive study can be found, for instance, in [3, 18]. The *stochastic equations* describing this process, which are going to be used in Section 4, are obtained by adding and subtracting, to each increment of $S_t$ and $I_t$, the conditional expectations, given the value of the process at the beginning of the corresponding time increment, say, of length $\Delta t$, [6]. Each increment of the process can be expressed as the expected value of the increment plus a sum of centered increments. In our example, the expected values of the increments $\Delta S = S_{t+\Delta t} - S_t$ and $\Delta I = I_{t+\Delta t} - I_t$ are $(-\beta \frac{S_t I_t}{N})\Delta t$ and $(\beta \frac{S_t I_t}{N} - \gamma I_t)\Delta t$ respectively, so the increments can be written as

$$\Delta S = \left(-\beta \frac{S_t I_t}{N}\right) \Delta t + \Delta Z_1 \tag{9}$$

$$\Delta I = \left(\beta \frac{S_t I_t}{N} - \gamma I_t\right) \Delta t - \Delta Z_1 + \Delta Z_2, \tag{10}$$

where $\Delta Z_1$ and $\Delta Z_2$ are conditionally centered Poisson increments with mean zero and conditional variances $\beta(S_t I_t/N)\Delta t$ and $\gamma I_t \Delta t$.

Now let us consider what happens if we drop the terms $\Delta Z_i$ from equations (9) and (10), and let $\Delta t$ go to zero. The resulting ordinary differential equations,

$$\frac{dS}{dt} = -\beta \frac{S_t I_t}{N},$$
$$\frac{dI}{dt} = \beta \frac{S_t I_t}{N} - \gamma I_t,$$

define a deterministic model. If $\hat{\beta} S_t I_t$ is used instead of $\beta S_t I_t/N$, with $\hat{\beta} = \beta/N$, we have, after dropping the hats, the so called Kermack and McKendrick ODE model, [19],

$$\frac{dS}{dt} = -\beta S_t I_t, \tag{11}$$

$$\frac{dI}{dt} = \beta S_t I_t - \gamma I_t. \tag{12}$$

In these first two examples, many aspects of a real contagion process have been put aside, for instance latent periods, varying infection and recovery rates, partial immunity, behavioral changes. The inclusion of such features would make the model more realistic, but would complicate the analysis. The strategy in modeling a particular system is first to consider the simplest model, even though some of the aspects one might eventually wish to include are absent. One looks at the analysis and then one may add, one step at a time, additional features. Adding compartments rapidly complicates the analysis. It may be necessary to evaluate the effect of an additional feature by numerical methods or simulation. In the next example, to which we return in Section 5, the effect of births and deaths is taken into account.

**Example 3.** *Stochastic SIR with demography*. The stochastic SIR presented in Example 2 might be appropriate when the rates of movement between compartments, and hence the evolution of the disease, are fast enough so that the life span of an individual does not need to be taken into account. This is often acceptable as an idealization when one is interested in looking at functionals of a particular epidemic outbreak such as epidemic size, which we discuss in the next section. However, we may be interested in the longer term recurrent or endemic aspects of a disease, such as the childhood diseases mumps, measles, smallpox, chickenpox, polio or rubella.

In this case, demography, meaning births and deaths of individuals, is often included in the model. A scheme including demography is shown in Figure 2, where births occur only in the susceptible class and deaths occur, at the same rate per individual, in the three compartments. The transition rates for this model are shown in Table 1 below.

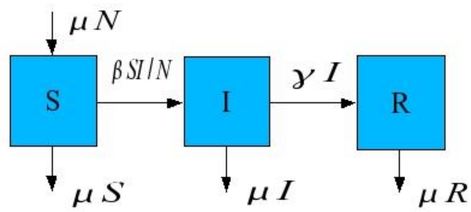


**Fig. 2** Schematic compartmental representation of SIR including ”demography”.

**Table 1**

| Transition | Rate |
|---|---|
| $S \to S+1$ | $\mu N$ |
| $S \to S-1$ | $\beta \frac{SI}{N} \Delta t + \mu N$ |
| $I \to I+1$ | $\beta \frac{SI}{N} \Delta t$ |
| $I \to I-1$ | $(\gamma + \mu) I$ |
| $R \to R+1$ | $\gamma I$ |
| $R \to R-1$ | $\mu I$ |

The stochastic rates of birth and death of individuals are assumed constant and equal to $\mu$. This makes the expected value of the total population constant and equal to $N$. The corresponding probabilities of the events are

$$\begin{aligned} P((S_{t+\Delta t}, I_{t+\Delta t}) - (S_t, I_t) = (1,0)) &= P((S_{t+\Delta t}, I_{t+\Delta t}) - (S_t, I_t) = (-1,0)) \\ &= \mu N \Delta t + o(\Delta t), \qquad (13) \\ P((S_{t+\Delta t}, I_{t+\Delta t}) - (S_t, I_t) = (-1,1)) &= \beta \frac{S_t I_t}{N} \Delta t + o(\Delta t), \qquad (14) \\ P((S_{t+\Delta t}, I_{t+\Delta t}) - (S_t, I_t) = (0,-1)) &= (\gamma + \mu) I_t \Delta t + o(\Delta t). \qquad (15) \end{aligned}$$

As in Example 2, it is enough to consider the process $(S_t, I_t)$, even though the total population at time t has become a stochastic process. The equations for $(S_t, I_t)$ form a closed system if we take $N$ to be the constant $EN$, the expected value of $N_t$. The stochastic equations describing this process, which will be used in Section 5, are obtained similarly, by adding and subtracting to each increment of $S_t$ and $I_t$, the conditional expectations, given the value of the process at the beginning of the corresponding time increment. For this example, the expected values of the increments $\Delta S = S_{t+\Delta t} - S_t$ and $\Delta I = I_{t+\Delta t} - I_t$ are $(\mu(N - S_t) - \beta \frac{S_t I_t}{N})\Delta t$ and $(\beta \frac{S_t I_t}{N} - (\gamma + \mu) I_t)\Delta t$ respectively, so the increments can be written as

$$\Delta S = \left(\mu(N - S_t) - \beta \frac{S_t I_t}{N}\right) \Delta t + \Delta Z_1 + \Delta Z_2, \tag{16}$$

$$\Delta I = \left(\beta \frac{S_t I_t}{N} - (\gamma + \mu) I_t\right) \Delta t - \Delta Z_2 + \Delta Z_3, \tag{17}$$

where $\Delta Z_1$ is the difference of the centered Poisson increments corresponding to births and deaths in the susceptible class with mean zero and variance $\mu(N + S_t)\Delta t$. Similarly the centered Poisson increments corresponding to the infections and removals are $\Delta Z_2$ and $\Delta Z_3$ respectively, both with conditional mean zero and with conditional variances $\beta(S_t I_t/N)\Delta t$ and $(\gamma + \mu)\Delta t$. If we drop the terms $\Delta Z_i$ from equations (16) and (17), and let $\Delta t$ go to zero. The resulting ordinary differential equations,

$$\frac{dS}{dt} = \mu(N - S_t) - \beta \frac{S_t I_t}{N},$$
$$\frac{dI}{dt} = \beta \frac{S_t I_t}{N} - (\gamma + \mu) I_t,$$

define a deterministic model, namely the deterministic SIR with demography, which has deterministic rates the same as the stochastic rates which yield the stochastic model defined by equations (16) and (17).

When $N$ is large, there is often a diffusion approximation to a stochastic compartmental model. We illustrate this in the case of Example 3. Let us normalize by dividing each of the stochastic processes in our model by $N$ so that the state variables are the proportions of the total expected population in the susceptible and infective classes at each time, $t$ and their jumps are of size $1/N$. Suppose we replace the conditionally centered Poisson increments, $\Delta Z_i/N$, by increments of Brownian motion, appropriate multiples of $\Delta W_i$, with the same standard deviations as the Poisson increments they replace. We obtain a diffusion approximation to our Markov jump process model which can be written as

$$\begin{aligned} ds &= (\mu(1 - s) - \beta s i)\, dt + G_1 dW_1(t) - G_2 dW_2(t), \\ di &= (\beta s i - (\gamma + \mu) i)\, dt + G_2 dW_2(t) - G_3 dW_3(t), \\ G_1 &= \sqrt{\mu(1 + s)}, \quad G_2 = \sqrt{\beta s i}, \quad G_3 = \sqrt{(\gamma + \mu) i}, \end{aligned} \tag{18}$$

where $s = S/N$ and $i = I/N$. Kurtz, [29, 30], showed that the normalized Markov jump process and the approximating diffusion (18) can be constructed on the same probability space in such a way that the maximum pointwise distance between their sample paths on a fixed finite interval of time is of order $\log N/N$. It is important to note that the diffusion approximation is good for $N$ large but becomes less useful if $N$ is too large. The limit of the solution of the stochastic system, as $N$ goes to infinity, is in fact the solution of the deterministic model, where the states are the fractions of the total population in each class (18).

For epidemic models, a main concern is to find conditions under which a disease introduced into a community will develop into a large outbreak, and if it does, conditions under which the disease may become endemic. For stochastic models, all such questions are in terms of probabilities. A useful parameter in this regard, called the *basic reproductive number*, $R_0$, is defined as the expected number of secondary infective cases per primary case in a completely susceptible population, [19]. In Examples 2 and 3 above, the basic reproductive number is $\beta/\gamma$ and $\beta/(\gamma+\mu)$, respectively. If the basic reproductive number is smaller than or equal to one, with a high probability the disease outbreak is relatively small. For this reason most studies of these examples concentrate on the complementary case. Arguably, the most important and interesting case is where $R_0$ is near one, as we shall see.

If the basic reproductive number is greater than one, the stochastic behavior of Examples 2 and 3 are very different. In Example 2, with no demography, the number of infected individuals generally increases, reaches a maximum and then generally decreases to zero. In Example 3, with demography, the solutions of the corresponding deterministic equations (see Section 5 below), will approach a nontrivial equilibrium as $t$ increases, called the endemic equilibrium. Simulations of the stochastic model show almost periodic oscillations of the process around this equilibrium. We return to this striking phenomenon in Section 5.

Other stochastic epidemic models can be defined along lines similar to these three examples. Compartments may be added corresponding to latent, asymptomatic, quarantined, or other disease-associated states. In this chapter we will confine ourselves mostly to questions pertaining to Examples 2 and 3. It will be clear that these and similar questions about other compartmental stochastic models might be pursued using, in part, similar methods. The relation of stochastic compartmental models to limiting deterministic models, and their approximations by diffusions, are illustrated for the SIR model in this section. Essentially all stochastic compartmental models have deterministic limits and diffusion approximations which can be obtained by the arguments analogous to those indicated here and given in detail and in great generality by Kurtz [30]. Often it is useful, and justified, to work with the diffusion approximation to the jump Markov chain model. The deterministic large-$N$ limit is also often of value for understanding the behavior of the stochastic dynamical system defined by the Markov chain model. One example is the information contained in the basic reproductive number, $R_0$, which can be regarded as a property of the deterministic limit. We see another example in Section 5 and discuss this point more generally in Section 7.

## 4 Distribution of the Final Epidemic Size

Public health policy may be influenced by predictions of how large an epidemic might be, that is, how many individuals ultimately become infected during the entire time an epidemic lasts. This involves the assumption that the disease in question does not become endemic and persist at a positive level indefinitely. In view of the nature of the dynamics of Examples 2 and 3 in Section 3, the assumption that there is a finite epidemic size pushes us in the direction of assuming that the population is fixed and finite, with no *demography*. In this case, the number of infectives will eventually reach zero, with probability one, so that the total number of individuals that are infected during the infectious process is almost surely finite, and the distribution of the final size of the epidemic can, in principle, be computed.

It was first observed by Bailey in 1953, [5], that the final size distribution for the stochastic SIR is bimodal, that is, there are two maxima. He provided formuli that allow the computation of the distribution of the final size if the population is rather small. Since then, the final size distribution has been investigated for various models, by Lefèvre and Picard [34, 35, 44], Ball [8, 9], Ball and Nåsell [10], Scalia-Tomba [46, 47] , Martin-Löf [39], Ludwig [36] among others. For large populations, the computer storage needed for computation of the final size distribution, together with numerical precision, have been important issues.

For the SIR, if the basic reproductive number is greater than one, the general shape of the epidemic size distribution can be deduced intuitively as follows. With a large enough population, during the first stages of an epidemic, the number of infectives evolves approximately like a branching process. If the probability of zero offspring in any family is positive, then the branching process goes extinct with positive probability. The event of zero offspring corresponds to the event that an infective infects no-one else, and this has positive probability. If extinction does occur, it is likely to occur early. Correspondingly there are several sample paths of the process of infectives, $I_t$, which reach zero relatively soon. On the other hand, if early extinction does not happen then the finite number of susceptible individuals begins to be depleted, so that the process no longer behaves as a branching process. In this case the size of the epidemic may be approximately normally distributed.

There have been attempts to produce a rigorous argument for bimodality of the epidemic size distribution along these lines, but apparently without success. On the other hand, careful observation of simulations shows that some degree of bimodality of this distribution is present for any combination of parameters. The most striking biomodality occurs when the basic reproductive number is just slightly larger than one, as in Figure 3.

Martin Löf, [39], found a normalization and relative rates, under which the process of infectives has a diffusion limit when, simultaneously, the total population, $N$, and the basic reproductive number approach infinity and one, respectively. Also, the distribution of the time until the epidemic stops converges to the time it takes a Brownian motion to hit a parabolic boundary. Martin Löf used an elegant approximation, using Airy functions, to produce the shape of the limiting epidemic size distribution by computation. Marion and Greenwood [37] found a way of computing the final size distribution for very large $N$, from which one can see the degree of agreement between Martin-Löf's limiting distributions and the pre-limit for large $N$. In this section we describe these results and others found in [22]. We look at the questions:

- How does the epidemic size distribution depend on the parameters of the model?
- Is there a way to incorporate a process of vaccination in the stochastic SIR that depends on the activity of the disease?
- How does the final epidemic size distribution change in the presence of vaccination?

In order to simplify notation we will re-scale the time to $\gamma t$, and define $\lambda = \beta/\gamma$, which is the basic reproductive number. Then, in the notation of Example 2, $\beta = \lambda$, and $\gamma = 1$. We observe how the shape of the final size distribution depends on the parameters of the model by accurately computing the distribution in a memory-efficient fashion, which we indicate here. For this we need to look only at the times at which transitions occur, that is, points in time where an event, contagion or removal, happens. Ignoring the waiting times between events, we obtain a discrete-time Markovian structure from the continuous time SIR that has the same epidemic size as the SIR process. This is called the *discrete time embedded Markov chain of jumps*. Let us number the consecutive jumps of the continuous time Markov chain by $j$, so that $j$ becomes the time parameter of the discrete time embedded Markov chain. The transition probabilities of this discrete time chain are given by

$$(\Delta S_j, \Delta I_j) = \begin{cases} (-1,1) & \text{with probability } \frac{\lambda S_j I_j/N}{\lambda S_j I_j/N + I_j} = \frac{\lambda S_j}{\lambda S_j + N}, \\ (0,-1) & \text{with probability } \frac{I_j}{\lambda S_j I_j/N + I_j} = \frac{N}{\lambda S_j + N}. \end{cases} \tag{19}$$

We introduce the possibility of immunization through an additional type of jump, of the form (-1,0). An individual is removed from the susceptible class through immunization, which may occur at each time step of the embedded chain. We denote by $\theta$ the average ratio of the number of vaccinations to the number of jumps. Thus, for instance, if $\theta = 0.5$, there is one vaccination every two steps, on the average; if $\theta = 2$, there are two vaccinations per step, on the average. This entry into the model of the vaccination procedure can be modified in various ways. For example, immunization can be considered only at the times when someone is infected, or only when a recovery occurs. The parameter $\theta$ can also depend on time or it may even be random with its distribution being time-dependent, and/or dependent on the current state of the process. Partial effectiveness can be modeled by multiplying $\theta$ by the probability of successful vaccination. Note that in this model the number of vaccinated persons is not directly related to the number of susceptibles, but is tied to the intensity of the epidemic as it evolves.

Let $U_j$ count the total number of infections which occur up to, and including, time $j$, disregarding the initial number of infected individuals. The probability that an individual gets infected in the time step from $j-1$ to $j$, given the value $U_{j-1}$, and if there were initially $n$ and $m$ susceptible and infected individuals, is

$$p_{k,j} \equiv P(U_j = k+1 | U_{j-1} = k) = \frac{\lambda(n-k-\theta(j-1))}{\lambda(n-k-\theta(j-1)) + (n+m)}, \tag{20}$$

with the complementary probability

$$q_{k,j} \equiv P(U_j = k | U_{j-1} = k) = 1 - p_{k,j}. \tag{21}$$

Although it does not appear in the notation, this probability is conditional on $S_t$ being positive.

Let $T$ denote the time at which the epidemic stops. Then $U_T$ is the number of individuals ultimately infected, in addition to the original m infectives. At each time step an infection or a recovery happens. The process $U_j$ is a random walk starting at 0, with a positive step when there is an infection and a zero step when there is a recovery. The epidemic stops when

$$U_t + m - (\text{number of recoveries}) = 0,$$

and

$$T = U_T + (\text{number of recoveries}).$$

Therefore, $U_T = (T-m)/2$. To obtain the distribution of $U_T$, it is enough to compute the distribution of the hitting time $T$. To compute this distribution we will use the following recursion. First define $W_j(k) = P(U_j = k, T > j)$ for non-negative integers, $j$. Notice that $W_0(k) = \delta_{k,0}$, $P(T=0) = 0$ and, if $j-m$ is even,

$$P(T = j) = W_{j-1}\left(\frac{j-m}{2}\right) q_{\frac{j-m}{2},j}. \tag{22}$$

If $j-m$ is odd, then $P(T=j) = 0$. The defective distribution $W_j(\cdot)$ is computed as

$$W_j(k) = \begin{cases} W_{j-1}(k-1)p_{k-1,j} + W_{j-1}(k)q_{k,j}, & \text{if } k > \frac{j-m}{2}; \\ 0, & \text{if } k \leq \frac{j-m}{2}. \end{cases} \tag{23}$$

This recursion allows us to compute the distribution $P(T=j)$, $j = 1,2,\ldots$, and the defective distributions $W_j(k)$, $k = 1,\ldots,n$, $j = 1,2,\ldots$, while storing only the values of $W_j(\cdot)$ at stage $j$. As $j$ increases, $W_j(\cdot)$ loses mass.

This algorithm can be used for any finite number $N$ of individuals in the population. If $N$ approaches infinity and simultaneously $\lambda$ approaches one, with suitable related rates, the distribution can be found using a diffusion approximation. For this, we define new random variables $X_t^N$ and $Y_t^N$ by

$$\frac{S_t}{N} = 1 - \frac{X_t^N}{N^\alpha},$$
$$I_t = Y_t^N N^\beta,$$

and let $\lambda_N = 1 + a/N^\gamma$. Martin Löf [39] found exponents $\alpha$, $\beta$ and $\gamma$ such that $X_t^N$ and $Y_t^N$ converge weakly to a limiting diffusion, see [39] or [22] for details.

Appropriate values for these exponents are $\alpha = \beta = \gamma = 1/3$. After re-scaling time as $s = tN^{-2/3}$ and letting $N \to \infty$, $X_s^N$ and $Y_s^N$ converge weakly to diffusions $X_s$ and $Y_s$, which satisfy the stochastic differential equations

$$dX_s = (1+2\theta)ds,$$
$$dY_s = (a - X_s)ds + \sqrt{2}\,dW_s,$$

where $a = \lim_{N\to\infty} N^{1/3}(\lambda_N - 1)$.

The process $X_s$ is deterministic linear drift, $X_s = (1+2\theta)s$, so that $Y_s$ is defined by

$$dY_s = (a - (1+2\theta)s)ds + \sqrt{2}\,dW_s.$$

After integration,

$$Y_s = b + as - (1+2\theta)s^2/2 + \sqrt{2}\,W_s, \tag{24}$$

a diffusion with parabolic drift starting at

$$b = \lim_{N,m\to\infty} m/N^{1/3}. \tag{25}$$

The limiting epidemic stops when the right hand side of (24) is equal to zero, or in other words, when the Brownian motion $\sqrt{2}W_s$ hits the parabola $b + as - (1+2\theta)s^2/2$ for the first time.

Epidemic size defines a continuous functional with respect to the topology of weak convergence of stochastic processes. Hence the weak convergence of the pre-limit processes to $Y$ implies convergence of the distribution of epidemic size for finite $N$ to a distribution associated with the time Brownian motion hits a parabola.

Pre-limit and limiting distribution curves are shown in Figure 3. In the figure we can observe the convergence of epidemic size distributions and get an idea of the size of $N$ necessary for the pre-limit to approximate the limit to a certain degree of accuracy. We compare the distributions obtained using the algorithm (22) and (23) for the pre-limit and the limiting diffusion obtained in [39] for different values of parameters $a$ and $b$, with $\theta = 0$. For the pre-limit, $\lambda$ and $m$ are chosen according to $\lambda = 1 + a/N^{1/3}$, $b = m/N^{1/3}$. We see that the shape of the distribution is highly sensitive to the values of the parameters $a$ and $b$ when the model is slightly super-critical, that is, when $\lambda > 1$ is very close to 1. The degree of agreement between the limit distribution and the pre-limit for $N = 50,000$ and for $N = 500,000$ is not as precise as one might have expected.

The effect of vaccination in our model can be observed in Figure 4. Vaccination pushes the mass of the distribution in the direction of smaller epidemic size, but the bimodality of the distribution persists. Increasing $\theta$ pushes the distribution towards zero, as one would expect.

A scaling parameter less than 1/3 can also be used and leads to the limit one would obtain from a branching process model. The limiting epidemic size distribution using this scaling is not bimodal, [20].

The embedded chain in the algorithmic scheme for computing the distribution of the final size, as presented in this section, ignores the amount of time between

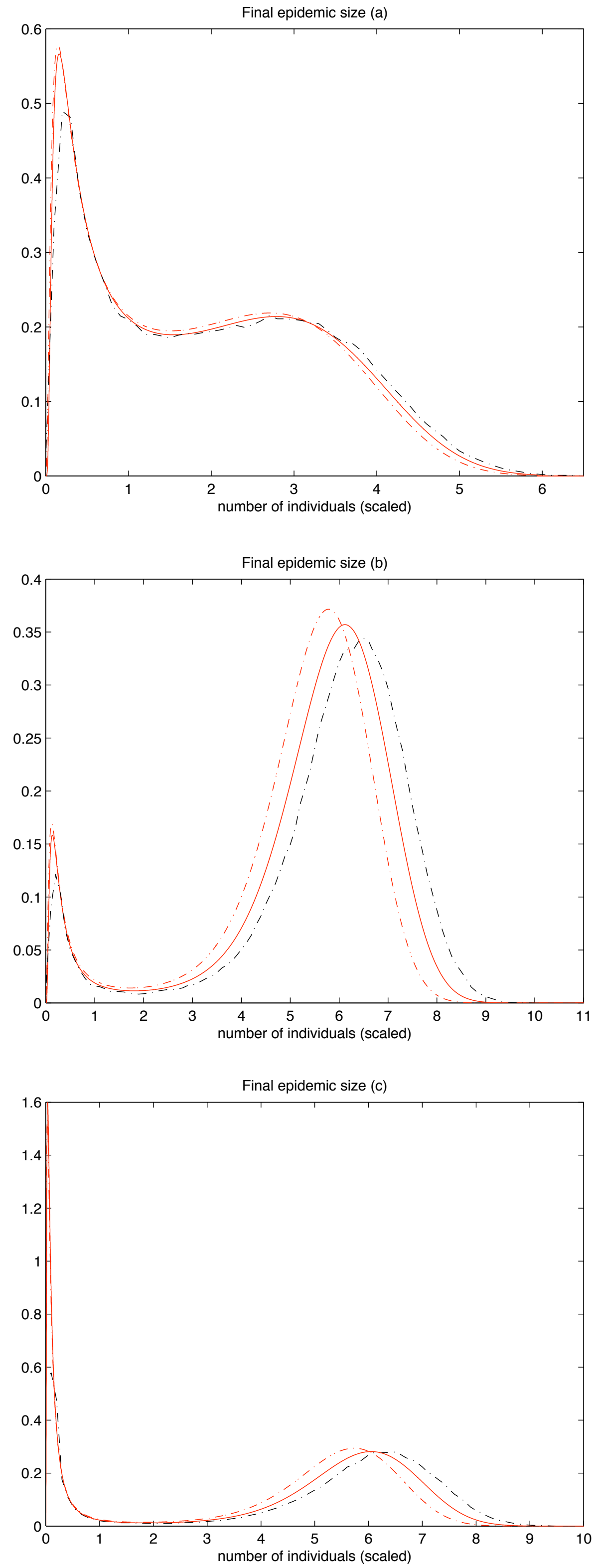


**Fig. 3** Distribution of the final epidemic size for the pre-limit (red): population $N = 50,000$ (dash-dots) and $N = 500,000$ (solid). The limit distribution (black) is in dash-dots. The parameters are (a) $a = 1$, $b = 1$, (b) $a = 3$, $b = 1$ (c) $a = 3$, $b = 0.5$. For all the figures $\theta = 0$. Notice that the three figures have different scales. Epidemic size is about $N^{2/3}$ times the scaled value.

events. This suggests that latency periods in infected individuals might not affect the distribution of the final size. In Section 6 we see that this is true in a very general setting, where homogeneity of the susceptible and infectious classes is abandoned.

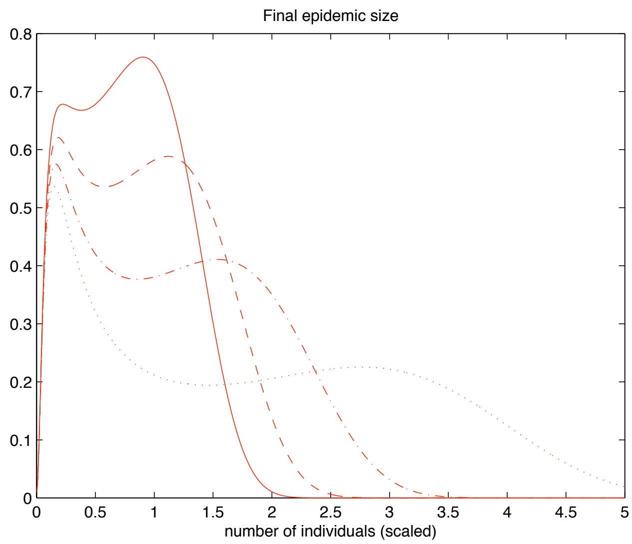


**Fig. 4** Distribution of the final epidemic size for $N = 30,000$. The vaccination levels are $\theta = 0$ (dots), $\theta = 0.5$ (dash-dot), $\theta = 1$ (dashes) and $\theta = 1.5$ (solid). The other parameters are $a = 1$ and $b = 1$. Vaccination pushes the mass of the distribution to the region of short outbreaks. Similar behavior can be observed for other values of the parameters $a$ and $b$, [22].

## 5 Stochastic Sustained Oscillations

In this section we will consider the model defined in Section 3, Example 3. As mentioned before, when the basic reproductive number is greater than one, the deterministic system has two equilibrium points, one unstable with no infectives, and a stable endemic point with a positive number of infectives. It can be observed, and shown analytically, that the solutions of the deterministic model oscillate around the endemic point and rapidly damp to the endemic equilibrium as time increases. However time-course data of diseases like measles or chickenpox show periodic oscillations that do not damp as time goes by. It is possible [24, 32] to produce deterministic models which have more slowly damped, or even sustained, oscillations by including, for instance, age structure, quarantine, multiple strains of infectious agents or delays. Of course seasonal periodic forcing produces seasonal oscillations in a deterministic epidemic.

Simulations of a stochastic SIR model follow the damped deterministic trajectory for a certain time after which the stochastic path remains oscillatory, with a varying amplitude, as can be seen in Figure 5 below.

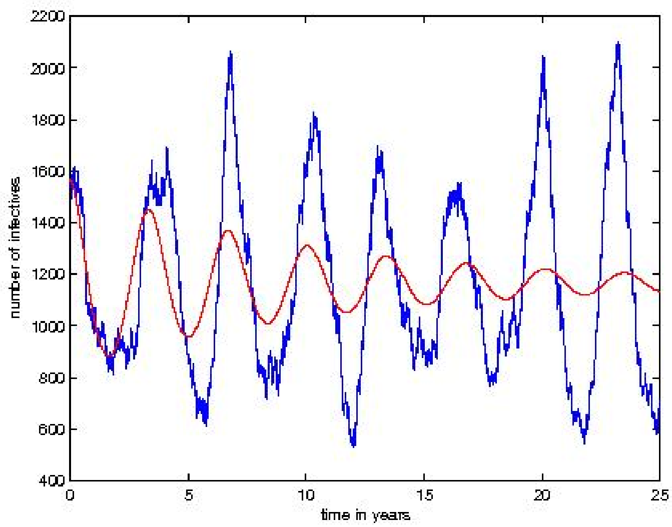


**Fig. 5** A sample path of the infective process of stochastic SIR with demography contrasted to the damped oscillations of the corresponding deterministic system with the same initial point. The parameters are $N = 2,000,000$, average life span $1/\mu = 80$ years, $R_0 = 15$, average time of infectiousness $1/\gamma = 15$ days. Notice that the stochastic path initially follows the deterministic trajectory.

The oscillations of stochatic SIR paths have a frequency distribution, evidenced by the power spectral density of the process of infectives, and a stochastically varying amplitude. This phenomenon, in which random fluctuations sustain nearly periodic oscillations in a system which has a stable constant equilibrium in the deterministic limit, has been called *coherence resonance* or *autonomous stochastic resonance*. Coherence resonance has been observed in a number of experimental studies of electrical, chemical and physiological phenomena, [43].

Although coherence resonance has been recognized as a possible occurrence in the presence of noise when a dynamical system has a small or hidden inherent periodicity, the phenomenon is only beginning to be understood quantitatively, from the stochastic process viewpoint. In [4], Aparicio and Solari give a convincing explanation of the non-damping of stochastic SIR in terms of the average change in a Liapunov function as the process moves inside and outside a parabola in phase space. In [31] we use multi-scale analysis to show that a stationary version of the system (16) and (17) can be rather closely approximated, in a neighborhood of the endemic equilibrium, for a suitable range of the parameters $\beta$, $\gamma$ and $\mu$, by a linear combination of sinusoids, where the coefficient processes in this approximation are Ornstein-Uhlenbeck processes running on a slower time scale.

The accuracy of the multiscale approximation to (16), (17) for a particular choice of parameters can be evaluated by comparing the power spectral densities of the two processes as in Figure 6.

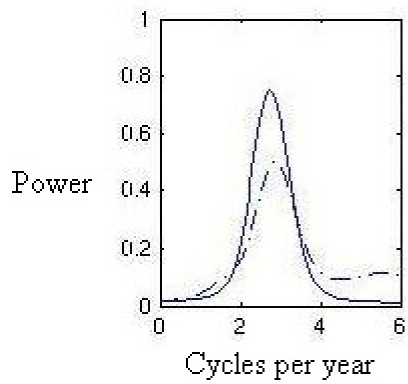


**Fig. 6** The power spectral density of the process of infectives in the stochastic SIR model (dot-dash) and the multi-scale approximation (solid). The values used are $R_0 = 15, N = 500,000, 1/\mu =$ 55 years and $1/\gamma = 25$ days.

## 6 Effects of Varying Infectiousness

If an individual becomes infected at time $t$, he may not become infectious immediately. There may be a latent period during which he is asymptomatic and/or remains uninfectious to others. More generally, infectiousness may vary in a variety of patterns following the event of disease transmission. In [23] we defined for each individual, a function which quantifies how infectious he is at time $t$ following the event of his infection. This function may depend on many factors, including a latent period, the response of the individual's immune system, the effects of medical treatments, etc. In epidemic modeling, it is natural to question how the pattern of infectiousness affects disease dynamics. In particular we are interested to see how the final size of an epidemic may be affected by how infectiousness varies.

In fact, a latency period does not influence the distribution of the final size of the stochastic SIR. This result has been shown in various contexts, see for example [8], [3], [1]. In [23] we proposed a more general formulation and showed that the final size depends only on the *integrated infectiousness.* Here are the main points of the argument.

The individuals in a population of constant size are labelled with the values $1,\ldots,N$, for identification. We think of individuals as distributed in space and related by a network of social or other connections, or as moving in space and encountering one another with pair independent frequencies. The numbers $c_{ij} \geq 0$ measure the rates of contact from individual $i$ to individual $j$ for all pairs $(i, j)$ with $1 \leq i, j \leq N$. Notice that $c_{ji} \neq c_{ij}$ in general. Our model assumption is that the time $T_{ij}$ of first infectious contact from $i$ to $j$ happens in the time interval $[t, t+\Delta t]$ with probability

$$P(T_{ij} \in [t, t+\Delta t] | T_{ij} > t, \mathscr{F}_i) = c_{ij} X_i(t) \Delta t + o(\Delta t). \tag{26}$$

Time in (26) runs according to a clock which starts at the first infectious contact made to individual $i$. The *infectiousness* process of individual $i$, $X_i(t)$, measures, at each time $t$, the probability that a contact made by $i$ at time $t$ is effective in transmitting the disease. The collection of sets $\mathscr{F}_i$ represents the information generated by the entire history of the random infectiousness process $X_i$, not including its start time. The infectiousness clock of individual $i$ may start at time 0, but it may be that $X_i(0) = 0$, so that $i$ is not actually infectious at time 0. Nevertheless we refer to such individuals as *initial infectives.* The product in (26) should be read as the probability of contact from $i$ to $j$ in the time increment $[t, t+\Delta t]$, $c_{ij}\Delta t + o(\Delta t)$, times the conditional probability of transmission of disease, given that contact is made, $X_i(t)$. The random function $c_{ij} X_i(t)$ is a random hazard function, which satisfies

$$P(T_{ij} > t | \mathscr{F}_i) = e^{-\int_0^t c_{ij} X_i(s) ds}, \tag{27}$$

for every $t \geq 0$. Notice that conditioning on $\mathscr{F}_i$ allows us to write a specific sample path, $X_i(s)$. If we let

$$D_i = \int_0^\infty X_i(s) ds < \infty, \tag{28}$$

the probability that an individual $j$ has no infectious contact from individual $i$, given a sample path of the process $X_i$, is

$$P(\text{no infection from } i \text{ to } j \,| \mathscr{F}_i) = e^{-c_{ij} D_i}.$$

We say that an individual $i$ is *nominally contacted* when an infectious contact to $i$ occurs. This may not be the first infectious contact. Let $\mathscr{F} = \bigcup_i \mathscr{F}_i$ be the $\sigma$-algebra generated by the infectiousness processes of all individuals in the population. $\mathscr{F}$ contains the information of the patterns of infectiousness of the entire population. Let $\mathscr{P}$ denote the set of all individuals in the population and let $\mathscr{X}_k$ and $\mathscr{Y}_k$, $k = 0, 1, 2, \ldots$ be

$$\begin{aligned}
&\mathscr{X}_0 = \{\text{initial infectives}\}, \\
&\mathscr{Y}_0 = \mathscr{P} - \mathscr{X}_0, \\
&\mathscr{X}_1 = \{j : j \in \mathscr{Y}_0, \exists i \in \mathscr{X}_0 \text{ such that } j \text{ is nominally contacted by } i\}, \\
&\mathscr{Y}_1 = \mathscr{Y}_0 - \mathscr{X}_1, \\
&\cdots .
\end{aligned}$$

We can see that $\mathscr{Y}_0 \supset \mathscr{Y}_1 \supset \ldots$ and that the set of all nominally contacted individuals, $\bigcup_{k=0}^\infty \mathscr{X}_k$, where $\mathscr{X}_i \cap \mathscr{X}_j = \emptyset$ if $i \neq j$, is equal to the set of all individuals who become infected. The size of this set is the total epidemic size. Thus, if we let $\mathscr{X} \subset$

$\mathscr{Y}_0$, the probability that the random set $\mathscr{X}_1$ is exactly $\mathscr{X}$, given $\mathscr{F}$, is

$$P(\mathscr{X}_1 = \mathscr{X} | \mathscr{F}) = \prod_{j \in \mathscr{X}} \left( 1 - \prod_{i \in \mathscr{X}_0} e^{-D_i c_{ij}} \right) \cdot \prod_{j \in \mathscr{Y}_0 - \mathscr{X}} \prod_{i \in \mathscr{X}_0} e^{-D_i c_{ij}}.$$

The distributions of $\mathscr{X}_k$, given $\mathscr{X}_{k-1}$, $k = 2, 3, ...$ can be computed similarly. Therefore, the probability distribution of the number of individuals that have nominal contacts, given all the patterns of infectiousness, is

$$P\left(|\cup_k \mathscr{X}_k| = n|\mathscr{F}\right) = \sum_{\substack{\mathscr{X} \subset \mathscr{P} \\ |\mathscr{X}|=n}} P\left(\cup_k \mathscr{X}_k = \mathscr{X} | \mathscr{F}\right),$$

which depends only on the random variables $D_i$, $i = 1, ..., N$ and the $c_{ij}$'s.

## 7 Stochastic and Deterministic Dynamics are Complementary

In this final section of the chapter we point out some useful general relationships between compartmental stochastic and deterministic epidemic models. The two types of model are alternate viewpoints on the same phenomenon, offering complementary insights.

The class of stochastic epidemic models of this chapter is defined by two properties: first, the dynamics can be described by a compartmental diagram such as Figure 2, with inputs and outputs, and second, the process is a vector-valued continuous time Markov process. This class of models is extremely large. For instance, to Examples 1, 2, and 3 can be added compartments which correspond to the latent, the asymptomatic, those quarantined, those vaccinated, the presence of multiple diseases, or classes of vectors such as mosquitoes which carry infectious agents.

We have indicated in Section 3, using an SIR example, how each such model corresponds to a deterministic model. One can write the stochastic increment equations as in (16) and (17), and then take the conditional expected value of each increment given the process at the beginning of that increment to obtain deterministic increment equations. Or one can divide each state variable by $N$ to obtain equations for the proportion of individuals in each class at time $t$, and apply the law of large numbers to these equations [29].

On the other hand, starting with a family of ordinary differential equations, which describes a dynamical system, one can arrive at a variety of corresponding stochastic models by interpreting some or all the the deterministic rates as stochastic rates in the sense of (1), (2).

Each of these model formulations, the stochastic and the deterministic dynamical system, augments our understanding of the other. The ODE model can be thought of as a deterministic skeleton of any corresponding stochastic model. An indispensable step in understanding the behavior of a stochastic model is the analysis of the dynamics of the ODE model. A dramatic example is the one described in Section

5. In fact, the analysis of stochastically sustained oscillations involves the details of the damped deterministic oscillations. Here are two additional examples.

In epidemic theory deterministic analysis often starts with the basic reproductive number, $R_0$. In Example 3, if $R_0$ is less that one, the unique equilibrium has no infectives. However in certain other models [17, 2], there is a more complex bifurcation structure in which, for a range of $R_0$ below one, there are two locally stable equilibria, one with no infectives and one with a positive number of infectives, separated by an unstable equilibrium. This structure is sometimes called a *backward bifurcation* because of the shape of the bifurcation diagram, Figure 7. An example is the model with susceptible, infected, and vaccinated individuals, defined by Brauer [17]. The dynamics tell us that a deterministic path is attracted to the equilibrium which is on the same side of the unstable equilibrium as the initial point of the path. However, a stochastic path started in a neighborhood of the unstable equilibrium will have probability about one half of being attracted to either equilibrium in the stochastic version of Brauer's model [33]. The authors of [33] continue to study the details of this problem. Of interest, for example, is the function which describes the probability of attraction of the process of infectives to each locally stable equilibrium as the distance of the starting point from the unstable equilibrium increases.

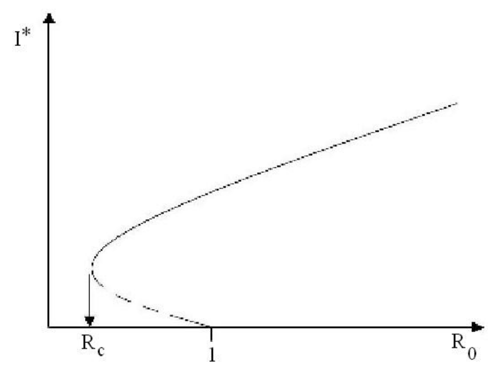


**Fig. 7** A typical backward bifurcation diagram. $I^*$ is the value of the infectives at the equilibrium. The solid lines stand for stability while dashed lines for unstability. Taken with permission from [45].

The deterministic susceptible-infected-susceptible (SIS) models are a bit simpler than those of Examples 2 and 3. For $R_0 > 1$, there is a stable equilibrium of the ODE, and the convergence of the path of susceptibles to this equilibrium is monotone rather than by damped oscillations as in Example 3. In fact this is a logistic model, whose susceptible class converges to a saturation point. The paths of the stochastic model lie near the deterministic path and continue to vary randomly around the deterministic equilibrium for a rather long time with high probability. The existence of an absorbing state, $I_t = 0$, in the finite state space means, according to general Markov chain theory, that ultimately the Markov process goes to the absorbing state.

Even in the simple SIS model the deterministic dynamic skeleton shows us a great deal about the behavior of the stochastic paths, and brings to our attention questions which pertain to the stochastic model: What is the nature of the stochastic path of $I_t$ as it varies near the deterministic equilibrium? Starting from the deterministic equilibrium, what is the distribution of the time until the stochastic path hits 0? These questions have been studied by Nåsell [41] for the SIS and other processes.

We should point out, in closing, that the stochastic models we have discussed here are simple ones, involving no more than two linked stochastic equations. The difficulty of a stochastic model grows closely in step with the difficulty of its companion ODE model. Additional stochastic models related to systems of ODE's result from the introduction of stochastic structure to parameters.

## References


1. C.L. Addy, I.M. Longini, M. Harber, A generalized stochastic model for the analysis of infectious disease final size data. Biometrics, **47**, No 3, 961–974 (1991).
2. L. Allen, P. van den Driessche, Stochastic epidemic models with a backward bifurcation, Mathematical Biosciences, **3**, 445–458 (2006).
3. H. Andersson, T. Britton, *Stochastic epidemic models and their statistical analysis*. Lecture Notes in Statistics, 151. Springer-Verlag, New York, 2000.
4. J.P. Aparicio, H.G. Solari, Sustained oscillations in stochastic systems. Mathematical Biosciences, **169**, 15–25 (2001).
5. N.T.J. Bailey, The total size of a general stochastic epidemic. Biometrika, **40**, No. 1/2, 177–185 (1953).
6. N.J.T. Bailey, *The mathematical theory of infectious diseases and its applications, Third ed.*, Oxford University Press, Oxford, 1975.
7. F. Ball, Deterministic and stochastic epidemics with several kinds of susceptibles. Advances in Applied Probability, **17**, 1–22 (1985).
8. F. Ball, A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models. Advances in Applied Probability **18**, 289–310 (1986).
9. F. Ball, D. Clancy, The final size and severity of a generalized stochastic multitype epidemic model, Advances in Applied Probability **25**, 721–736 (1993).
10. F. Ball, I. Nåsell, The shape of the size distribution of an epidemic in a finite population. Mathematical Biosciences, **123**, 167–181 (1994).
11. F. Ball, D. Mollison, G. Scalia-Tomba, Epidemics with two levels of mixing. Annals of Applied Probability, **7**, 46–89 (1997).
12. A.D. Barbour, The principle of diffusion of arbitrary constants. Journal of Applied Probability, **9**, 519–541 (1972).
13. A.D. Barbour, On a functional central limit theorem for Markov population processes. Advances in Applied Probability, **6**, 21–39 (1974).
14. M.S. Bartlett, Some evolutionary stochastic processes. J. Roy. Statist. Soc. B, **11**, 211–229 (1949).
15. N. Becker, *Analysis of infectious disease data.* Chapman and Hall, London, 1989.
16. N. Becker, K. Dietz, The effect of the household distribution on transmission and control of highly infectious diseases. Mathematical Biosciences, **127**, 207–219 (1995).
17. F. Brauer, Backward bifurcations in simple vaccination models, Journal of Mathematical Analysis and Applications, **298**, 418–431 (2004).
18. D.J. Daley, J. Gani, *Epidemic modeling: an introduction.* Cambridge Studies in Mathematical Biology, 15. Cambridge University Press, Cambridge, 1999.
19. O. Diekmann, J.A.P. Heesterbeek, *Mathematical epidemiology of infectious diseases. Model building, analysis and interpretation*. Wiley Series in Mathematical and Computational Biology. John Wiley & Sons, Ltd., New York, 2000.

20. R.G. Dolgoarshinnykh, S.P. Lalley, Critical scaling for the simple SIS stochastic epidemic. J. Appl. Probab. **43** 892–898 (2006).
21. B.T. Grenfell, A. P. Dobson, *Ecology of infectious diseases in natural populations.* Cambridge University Press, Cambridge, 1996.
22. L.F. Gordillo, S.A. Marion, A. Martin Löf, P.E. Greenwood, Bimodal epidemic size with vaccination. Bulletin of Mathematical Biology, **70**, 589–602 (2008).
23. L.F. Gordillo, S.A. Marion, P.E. Greenwood, The effects of patterns of infectiousness on epidemic size. Submitted.
24. H.W. Hethcote, S.A. Levin, Periodicity in epidemiological models. Applied Mathematical Ecology, Biomathematics, **18**, Springer, Berlin, 193–211, 1989.
25. J.O. Irwin, The place of mathematics in medical and biological statistics. Journal of the Royal Statistical Society. Series A (General), **126**, 1, 1–45 (1963).
26. V. Isham, G. Medley (eds.), *Models for infectious human diseases: their structure and relation to data.* Cambridge University Press, Cambridge, 1996.
27. D.G. Kendall, Deterministic and stochastic epidemics in closed populations. Proc. Third Berkeley Symp. Math. Statist. Probab. **4**, 149–165. University of California Press, Berkeley, 1956.
28. W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, Proceedings of the Royal society of London. Series A, **115**, no. 772, 700–721 (1927).
29. T.G. Kurtz, Strong approximation theorems for density dependent Markov chains. Stochastic Processes and their Applications, **6**, 223–240 (1978).
30. T.G. Kurtz, *Approximation of population processes.* CBMS-NSF Regional Conference Series in Applied Mathematics, 36. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa., 1981.
31. R. Kuske, L.F. Gordillo, P.E. Greenwood, Sustained oscillations via coherence resonance in SIR, Journal of Theoretical Biology, **245**, 459–469 (2007).
32. S.A. Levin, J. Dushoff, J. Plotkin, Evolution and persistence of influenza A and other diseases. Mathematical Biosciences, **188**, 17–28 (2004).
33. R.Lopez, B.Dembele, Stochasticity in Vaccination, manuscript, (2007).
34. C. Lefèvre, P. Picard, A Non-Standard Family of Polynomials and the Final Size Distribution of Reed-Frost Epidemic Processes. Advances in Applied Probability, **22**, no. 1, 25–48 (1990).
35. C. Lefèvre, P. Picard, Collective epidemic models. Mathematical Biosciences **134**, 51–70 (1996) .
36. D. Ludwig, *Stochastic population theories.* Lecture Notes in Biomathematics, 3. Springer-Verlag, New York, 1974.
37. S. Marion, P.E. Greenwood, Computation of the size of an epidemic in a finite heterogeneous population. *Second European Conference on Highly Structured Stochastic Systems,* 183–185 (1999).
38. A.G. McKendrick, Applications of mathematics to medical problems. Proceedings of the Edinburgh Mathematical Society, **14**, 98–130 (1926).
39. A. Martin-Löf, The final size of a nearly critical epidemic and the first passage time of a Wiener process to a parabolic barrier. J. Appl. Prob. **35**, No 3, 671–682 (1998).
40. D. Mollison (ed.), *Epidemic Models: Their Structure and Relation to Data.* Cambridge University Press, Cambridge, 1995.
41. I. Nåsell, The quasi-stationary distribution of the closed endemic SIS model. Adv. in Appl. Probab. **28**, no. 3, 895–932 (1996).
42. I. Nåsell, Endemicity, persistence, and quasi-stationarity. In: *Mathematical approaches for emerging and reemerging infectious diseases: an introduction (Minneapolis, MN, 1999)*, 199–227, IMA Vol. Math. Appl., 125, Springer, New York, 2002.
43. A. Neiman, Coherence Resonance, Scholarpedia, Art. 1442, (2007).
44. P. Picard, C. Lefèvre, A Unified Analysis of the Final Size and Severity Distribution in Collective Reed-Frost Epidemic Processes. Advances in Applied Probability, **22**, no. 2, 269–294 (1990).

45. F. Sánchez, X. Wang, C. Castillo-Chavez, D. Gorman, P. Gruenewald, Drinking as an epidemica simple mathematical model with recovery and relapse. In Therapists Guide to Evidence-Based Relapse Prevention: Practical Resources for the Mental Health Professional, Katie A. Witkiwitz G. Alan Marlatt (eds.), Academic Press (2006).
46. G. Scalia-Tomba, Asymptotic final size distribution for some chain-binomial processes. Adv. in Appl. Probab. **17**, 3, 477–495 (1985).
47. G. Scalia-Tomba, Asymptotic final size distribution of the multitype Reed-Frost process. J. Appl. Probab. **23**, 3, 563–584 (1986).
48. T. Sellke, On the asymptotic distribution of the size of the stochastic epidemic. Journal of Applied Probability, **20**, 390–394 (1983).
49. P. D. Stroud, S.J. Sydoriak, J.M. Riese, J.P. Smith, S.M. Mniszewski, P.R. Romero, Semi-empirical power-law scaling of new infection rate to model epidemic dynamics with inhomogeneous mixing. Mathematical Biosciences, **203**, 301-318 (2006).