

# 1 Theory

## 1.1 Models of infectious diseases

**Observations:** Large scale patterns of epidemics have been measured [?] The spread of infectious diseases is something that everyone is familiar with.

**The science inspired by observation.** One goal of epidemiology is to understand the principles behind the spreading process, i.e. the way how a disease is transmitted through a population. In this context, *conceptional* models are used. They make use of simple assumptions for the local (person-to-person) dynamics and focus on the big picture of the process. Conceptual models are very similar to models in theoretical physics, because they focus on the very essence of the problem (here: the macroscopic view, spreading patterns). However, they have to neglect many details of the real problem (here: physiology, symptoms, individual behavior, infection pathways and many more!).

Another important issue of epidemiology is the *forecast* of epidemic spreading processes. Forecast models incorporate as much information as possible and the main focus is not an understanding of the basic principles.

This section summarizes the mathematical framework that roughly reproduces the behavior of some infectious diseases and briefly discusses some important insights.

### 1.1.1 Development

### 1.1.2 Infection dynamics

The spread of infectious diseases can be modeled in terms of compartment models as described in section xx. We differentiate between *conceptional models* and *realistic disease models*. While the former class is used to provide conceptual results as for the computation of thresholds or to test theories [?], realistic disease models use as many aspects as possible to provide a forecast of the spreading process. Realistic disease models can be very complex and are beyond the scope of this work, thus the author focuses on the use of conceptional models. The following section is inspired by the Lecture notes of J. R. Chasnov [?].

### 1.1.3 SI model

Let us consider a population of  $N$  individuals. In the simplest case, the infection status of each individual is either susceptible or infected and there are no births and deaths on the population. Susceptible individuals become infected, if they are in contact with an infected. This mimics the behavior of an infectious disease without immunization, i.e. infected individuals stay permanently infected.

Provided that  $\alpha$  is the rate, under which new susceptible become infected, the SI-model is as follows:

$$\begin{aligned}\frac{dS}{dt} &= -\alpha SI \\ \frac{dI}{dt} &= \alpha SI,\end{aligned}\tag{1.1}$$

where  $S$  and  $I$  are the numbers of susceptible and infected individuals respectively. The total population is  $N = S + I$ . Thus, (1.1) can be rewritten as

$$\frac{dI}{dt} = \alpha(N - I)I,$$

i.e. a logistic differential equation. Hence, in the limit  $t \rightarrow \infty$  the whole population is infected ( $I(\infty) = N$ ).

### 1.1.4 SIR model

In contrast of the infection dynamics introduced in the previous section, many epidemics allow for an immunization of the individuals. Examples are measles or whooping cough [?] [?]. In this case, individuals recover from the disease after being infected for a certain time period. The infection scheme has to be extended to susceptible-infected-recovered (SIR) as in the following infection model [?]:

$$\begin{aligned}\frac{dS}{dt} &= -\alpha SI \\ \frac{dI}{dt} &= \alpha SI - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}\tag{1.2}$$

where  $\alpha$  is the infection rate and  $\gamma$  is the immunization or recovery rate. There is no analytic solution for the system (1.2), but some fundamental conclusions can be obtained analytically. We show a typical solution of (1.2) in figure 1.1.

The SIR model shows more sophisticated features than the SI model (1.1). To begin with, we analyze the fixed points of the system, i.e.  $(S_*, I_*, R_*)$  where

$$\frac{dS_*}{dt} = -\alpha S_* I_* = 0, \quad \frac{dI_*}{dt} = \alpha S_* I_* - \gamma I_* = 0, \quad \frac{dR_*}{dt} = \gamma I_* = 0.\tag{1.3}$$

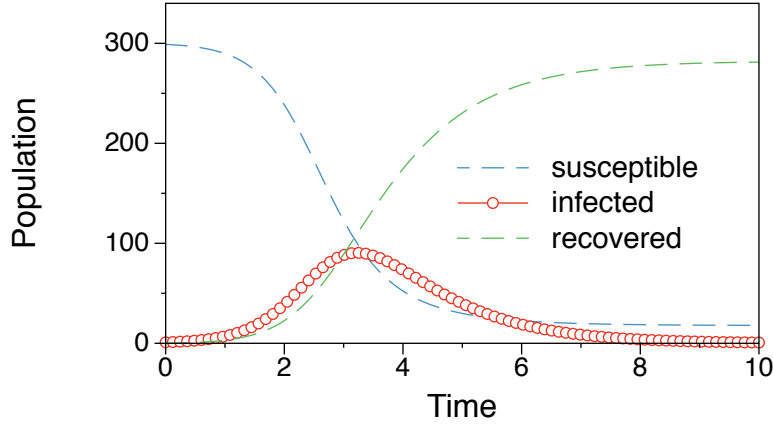


Figure 1.1: Typical solution of a susceptible-infected-recovered (SIR) model. Parameters:  $\alpha = 3$ ,  $\gamma = 1$ ,  $N = 300$ ,  $S_0 = 1$ .

It follows from the last equation that  $I_* = 0$  at the fixed point, where  $S_*$  and  $R_*$  can be arbitrary. Hence, a fixed point is  $(S_*, 0, R_*)$ .

Let us analyze the stability of the fixed point in the early phase of an infection. Almost all individuals are susceptible and consequently  $I_* = N - S_*$ . An outbreak occurs, if and only if  $dI/dt > 0$  in this phase, i.e.

$$\frac{dI}{dt} = \alpha S_*(N - S_*) - \gamma(N - S_*) = (N - S_*)(\alpha S_* - \gamma) > 0. \quad (1.4)$$

It follows from (1.4) that the number of infected grows, if

$$\alpha S_*/\gamma > 1. \quad (1.5)$$

Equation (1.5) is extremely important in epidemiology, because it defines a threshold for the unfolding of an infection spreading process. We call this fraction the *basic reproduction number*  $R_0$ . Recall that  $S_* \approx N$  in the fixed point. Thus it follows that the outbreak condition is

$$R_0 = N \frac{\alpha}{\gamma} > 1. \quad (1.6)$$

The basic reproduction number describes the average number of follow-up infections by each infected individual. It is one of the main goals in epidemiology to bring down the basic reproduction number of a disease below the critical value  $R_0 = 1$ . This is the reason for the implementation of mass vaccination. As one can immediately see from equation (1.6), this can be done by reducing the infection rate  $\alpha$  or by increasing the immunization

rate  $\gamma$ . In principle, one could also reduce the size of the initial population  $S_*$ . As an example, reducing the infection rate can be done by increasing hygiene standards or appropriate behavior, say wearing warm clothes in winter time to avoid common cold. The immunization rate can be increased by vaccination.

Let us now focus on the late phase of an SIR-infection. In contrast to the SI-model of section 1.1.3 an SIR like outbreak does not necessarily infect the whole population, even if  $R_0 > 1$ . The reason is that there has to be a critical mass of susceptible individuals in order to keep an infection alive (see equation (1.5)). The total number of infected during an infection given by the number of recovered at the end of the infection, since every recovered has to be in the infected state in the first place. A central measure throughout this work is therefore the *outbreak size*  $R_\infty$ .

To compute the outbreak size, we consider the second fixed point of (1.2), i.e. the fixed point for  $t \rightarrow \infty$ . At this point there are no infected and a fraction of the population is recovered. Hence the fixed point is  $(N - R_\infty, 0, R_\infty)$ . A simple way to obtain the outbreak size  $R_\infty$  is to use equations (1.2) and compute

$$\frac{dS}{dR} = -\frac{\alpha}{\gamma}S$$

and separate the variables [?]. This yields

$$\int_{S_*}^{N-R_\infty} \frac{dS}{S} = -\frac{\alpha}{\gamma} \int_{R_*}^{R_\infty} dR.$$

We integrate from the initial condition at  $t = 0$  to the final condition at  $t \rightarrow \infty$ , where  $S_\infty = N - R_\infty$ . Using that  $R_* = 0$  at  $t = 0$  gives

$$R_\infty = S_* - S_* e^{-\frac{\alpha}{\gamma} R_\infty}. \quad (1.7)$$

This transcendental equation can be solved numerically using a Newton-Raphson technique. The outbreak size  $R_\infty$  only takes finite values for  $\alpha/\gamma > 1$ . A solution of equation (1.7) is shown in figure 1.2

It should be noted that an SIR epidemic is a single event, i.e. it possesses a *characteristic time scale*. The analysis of the late phase of an epidemic also gives information about these time scales. Let us consider the second equation of (1.2).

$$\frac{dI}{dt} = \alpha SI - \gamma I \quad (1.8)$$

In the late phase of an SIR-type epidemic, the fraction of infected is small. Given sufficiently large values of  $R_0$ , the fraction of recovered is also small in this phase (see figure 1.2). Thus, we neglect the quadratic term in (1.8). This gives  $\frac{dI}{dt} = -\gamma I$ , which has the solution

$$I(t) = I(0)e^{-\gamma t}.$$

Hence, the infection decays exponentially for large  $t$  and the inverse recovery rate  $1/\gamma$  defines the characteristic time of the epidemic.

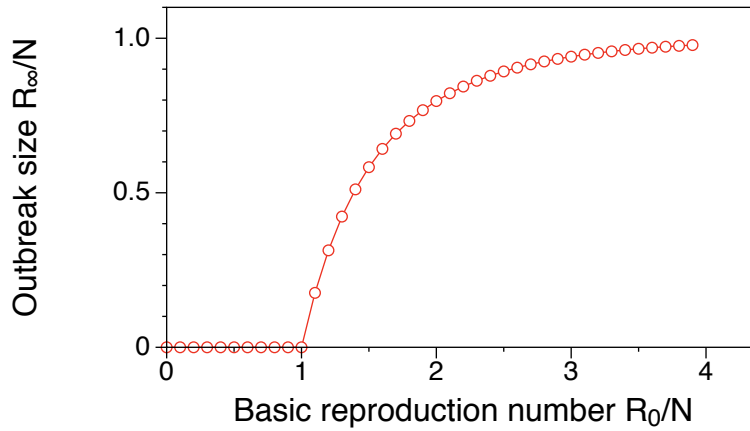


Figure 1.2: Relative outbreak size vs. basic reproduction number. The outbreak size takes finite values only for  $R_0/N > 1$ . Note that even for supercritical  $R_0$  the outbreak size is in general smaller than the total population.

### 1.1.5 Force of infection

The model presented in section 1.1.4 describes only the very basic behavior of epidemic dynamics, and is therefore a conceptual model. However, it is one of the main objectives in epidemiology to have an understanding of the exact infection rates in the process. Infection rates themselves can cause complex infection dynamics.

The term  $\alpha I$  used in section 1.1.4 is a special, very simple case of an infection rate. More generally, we have to replace  $\alpha I$  by an abstract infection rate  $\lambda$  containing more information about the interaction between susceptible and infected individuals [?]. Thus, the equation for the infected becomes

$$dI/dt = -\lambda S - \gamma I.$$

The rate  $\lambda$  is called the *force of infection*. In principle, this parameter can be arbitrarily complex, because it contains detailed information about the mixing properties of the population. This information could be given as contact networks, demographic contact structures, etc.

In most cases, detailed information about mixing is not available. Instead, we assume *random mixing* of the population, i.e. every individual can be in contact with every other individual. This yields a transmission rate [?]

$$\lambda = \tau n \frac{I}{N} \equiv \beta \frac{I}{N}, \quad (1.9)$$

where  $\tau$  is the transmission rate,  $n$  is the effective contact rate and  $I/N$  is the fraction of infectious contacts. It is therefore reasonable to replace the infection term  $\alpha$  in (1.2) by  $\beta/N$  to explicitly include the force of infection. Nevertheless, the results presented in section 1.1.4 remain qualitatively the same.

Although the force of infection gives a more reasonable description of the infection process, the assumption of random mixing remains inappropriate for many real world systems. Due to the availability of contact data, the random mixing assumption can be improved in terms of contact networks. Even if the exact data of an epidemic system is not available, research on complex networks allows us to give more realistic models about mixing. In the next section, we briefly report important results in complex network research and focus on the interplay between networks and epidemics in section 1.3.

## 1.2 Complex networks

## 1.3 Epidemics on networks

### 1.3.1 From random mixing to developed structure