

Chapter 2

Stochastic Modelling in Life Sciences

The dynamics of natural phenomena such as the growth of populations of species, the spread of epidemics, changes in gene frequencies or the course of chemical reactions are all subject to random variation. Their evolution is not exactly predictable. However, the application of mathematical models enables insight into such complex processes.

This chapter motivates and reviews representative application fields from life sciences and appropriate mathematical models. These applications and models will recur throughout the entire book. They give rise to the model constructions in Chaps. 3–5 and the investigation and development of estimation procedures in Chaps. 6 and 7. Moreover, they form the basis for the application studies in Chaps. 8 and 9.

The emphasis of this and the following chapters is on the important role of chance. In the literature, there is a vast number of works for modelling the mentioned dynamics where randomness is not taken into account. Such deterministic models provide a convenient and sometimes also appropriate way to represent a situation of interest. For comparison purposes, this deterministic approach is also introduced here. In general, however, deterministic models are not able to capture the natural stochastic behaviour of a real-world phenomenon. For instance, a deterministic model for the spread of an infectious disease may predict a major outbreak in a marginal situation and possibly prove wrong (cf. Sect. 2.2). Deterministic models for the dynamics of chemical reactions typically fail when the number of reactants is small (e.g. [McQuarrie 1967](#)). As another example, [Lande et al. \(2003\)](#) invoke harvest strategies, say in fishery, which may do harm to small populations of endangered species when they are developed based on deterministic models. For that reason, this book particularly focuses on the application of stochastic models. These account for random fluctuations of the considered processes and assign probabilities to critical events.

The structure of the present chapter is as follows: Sect. 2.1 introduces the very general class of compartment models. From such a model, both deterministic and stochastic processes can be derived. Sections 2.2 and 2.3 provide introductions to two emerging fields of life sciences, namely to models for the spread of infectious

diseases and to models for processes in molecular biology, biochemistry and genetics. Both sections start from a compartmental representation and then consider three types of models. These are stochastic jump processes, deterministic continuous processes and stochastic diffusion processes. The first type of process mirrors the exact dynamics of the compartmental system, whereas the second and third can be considered as approximations of the first. The development of an exact simulation algorithm for the jump process in 1976 hence meant a considerable advancement in the field of statistical modelling. This algorithm is presented in Sect. 2.4. In many situations, however, its application is computationally costly. Hence, numerical approximation algorithms for the second and third type of process are outlined as well. Section 2.5 concludes this chapter.

2.1 Compartment Models

In a *compartment model*, all objects involved in a system of interest are arranged in a finite number of *compartments*, i.e. in groups of objects that are defined through certain specified properties (Jacquez 1972). The compartments are mutually disjoint, and the assignment of each object to a compartment is unambiguous. The elements of each compartment are assumed to be homogeneous and well-mixed. Interaction between different compartments happens through the exchange of objects which is described by transition equations. Such passages are assigned with some rate that typically depends on the concentrations of objects from the distinct compartments. In this book, the considered compartmental systems are usually *closed*, i.e. there is no flow of objects to and from the environment.

The classification of objects into different compartments may, for example, be due to the location of animals or humans in a geographical region, the kinetic properties of molecules, or the age or physical conditions of individuals that are susceptible to a disease. Figures 2.1 and 2.2 display two compartment models from the fields of applications that are considered in Sects. 2.2 and 2.3.

A compartment model is a convenient fundament for a dynamical system one wishes to represent. From this model, different types of processes can be derived, all of them standing for the same considered phenomenon. This book will consider

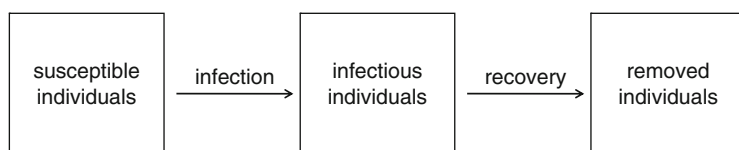


Fig. 2.1 Compartmental representation of the susceptible–infectious–removed (SIR) model that will be investigated in Sect. 2.2.2. In this model, a population of interest is classified into susceptible, infectious and removed individuals. Transitions between these three groups are due to infections and recoveries

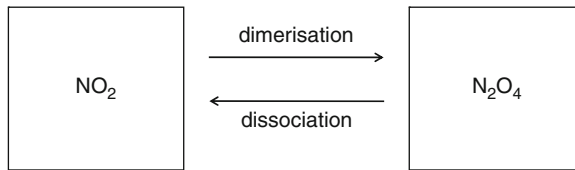


Fig. 2.2 Compartmental representation of the dimerisation of nitrogen dioxide. In this model, all nitrogen dioxide (NO_2) and dinitrogen tetroxide (N_2O_4) molecules in a gas are summarised in two compartments. Depending on the temperature and pressure, two of the NO_2 molecules may dimerise, yielding one N_2O_4 molecule. The other way round, an N_2O_4 molecule may dissociate into two NO_2 molecules

the following three kinds: pure Markov jump process, deterministic processes with continuous sample paths, and diffusion processes. First examples are shown in the next two sections.

2.2 Modelling the Spread of Infectious Diseases

Epidemics of infectious diseases have shaped the history of humankind. They have directly affected economy, politics and demography, the course of wars, social behaviour and religious beliefs (McNeill 1976; Cunha 2004; Smallman-Raynor and Cliff 2004; Sherman 2006; Oldstone 2010).

Devastating historic epidemics and pandemics include the Black Death in 1347–1350 with 25 million deaths in Europe, where there was up to 50% mortality of the urban population in England and Southern Europe; outbreaks of smallpox, measles and typhus in Mexico in 1518–1520 with 2–15 million deaths out of a population of 20 million; several cholera epidemics in India during the seventeenth century with more than 20 million deaths; and the Spanish influenza pandemic in 1918–1920 with estimated numbers of worldwide deaths lying between 25 and 50 million (Dobson and Carper 1996; Smallman-Raynor and Cliff 2004; Vasold 2008).

Present-day pandemics comprise for instance the acquired immunodeficiency syndrome (AIDS) caused by the human immunodeficiency virus (HIV) which was identified in the 1980s. It is assumed that in 2008 there were 2.7 million new infections, 2 million AIDS-related deaths and 33.4 million people living with the virus worldwide (UNAIDS 2009). Quite recently, in 2009, an influenza pandemic spread from Mexico over the whole world within a few months. It possibly affected between 11 and 21% of the global population (Kelly et al. 2011) and caused more than 18,000 deaths (WHO 2010). During the early stages of the epidemic, one even feared much higher mortality. Hence, the spread of diseases is still a serious concern in both the developed and developing world.

The elimination of infectious disease epidemics is desirable not only from a humane viewpoint but also regarding economic factors such as manpower and

public health costs. Even for diseases with relatively mild courses it is generally favourable to invest in prevention rather than cure. Considerable progress in understanding the propagation of infectious diseases from a medical point of view has been achieved by Louis Pasteur (1822–1895) and Robert Koch (1843–1910), who discovered the cause of infections by microorganisms. Targeted intervention against the spread of diseases, such as vaccination or isolation, however requires an overall comprehension of the typically complex dynamics of an epidemic. This is achieved by application of mathematical modelling (Brauer 2009).

The objectives of this section are the following: First, to introduce basic models for the spread of infectious diseases, and second, to motivate the utilisation of stochastic rather than deterministic models. This presentation is oriented towards the needs of subsequent chapters. For further information, the reader is referred to Bailey (1975), Anderson (1982), Becker (1989), Anderson and May (1991), Daley and Gani (1999), Andersson and Britton (2000), Diekmann and Heesterbeek (2000) and Keeling and Rohani (2008).

2.2.1 History of Epidemic Modelling

Detailed statistics on disease counts go back to John Graunt (1620–1674) who recorded weekly death counts in London together with their causes. The first mathematical model for the spread of infectious diseases, however, is generally accredited to Daniel Bernoulli (1700–1782), but epidemic modelling has not received much attention until the beginning of the twentieth century. Early works include En'ko (1889), Hamer (1906), Ross (1915) and Kermack and McKendrick (1927). Detailed historical accounts on the development of mathematical epidemiology can be found in Bailey (1975), Dietz (1967), Anderson and May (1991) and Daley and Gani (1999).

In the early stages of epidemic modelling, the spread of diseases was generally formulated as a deterministic process. According to Bailey (1975), the first author who included a random component in an epidemic model was McKendrick (1926), but that particular approach was only continued 20 years later. Instead, the class of *chain binomial models*, independently introduced by Lowell Reed and Wade Hampton Frost (see Abbey 1952 or Costa Maia 1952) and Greenwood (1931), established itself. A model of this type considers the evolution of an epidemic at discrete time points. To that end, the number of susceptible and infectious individuals in a population is assumed to be binomially distributed, conditioned on the state of the epidemic at the previous time point. An overview about chain binomial models is given in Becker (1989) and Daley and Gani (1999).

In subsequent years, both stochastic and deterministic models were refined and their mathematical analysis was extended; see e.g. Isham (2004) for a review. The class of *susceptible–infective–removed (SIR) models*, which is introduced in the next section, emerged as the most prominent description of the spread of infectious disease epidemics.

While the comprehension of disease dynamics and the development of mathematical tools progresses, the general framework of modelling the spread of diseases changes as well: First of all, the increased mobility of humans raises the risk of fast spreading pandemics. On the other hand, detailed medical knowledge of infection processes and improved hygienic conditions in many countries help prevent transmission of diseases. Modern epidemiological models take into account travel, social behaviour, the effect of intervention such as vaccination or isolation, and many other aspects.

The following section introduces a standard model from epidemiology which serves as the basis for many extensions as indicated in Sect. 2.2.3 and implemented in Chap. 5. This section concentrates on infectious diseases for humans. The considered diseases are assumed to be directly transmittable rather than vector-borne, i.e. transmitted for example by insects.

2.2.2 SIR Model

An *SIR model* (Kermack and McKendrick 1927; Bartlett 1949) classifies a population of fixed size N into susceptible (S), infectious (I) and removed (R) individuals. Transitions between these classes are

$$\textcircled{S} + \textcircled{I} \xrightarrow{\alpha} 2 \textcircled{I} \quad \text{and} \quad \textcircled{I} \xrightarrow{\beta} \textcircled{R}. \quad (2.1)$$

The first transition means that each contact between a susceptible and an infectious individual will cause an infection with rate $\alpha \in \mathbb{R}_+$, resulting in two infectious individuals. The second transition denotes that each of these infectious individuals will be removed with rate $\beta \in \mathbb{R}_+$ due to being recovered and immune, or quarantined, or dead. The parameter α is the *contact rate* of an infectious individual for spreading the disease, and β is the *reciprocal average infectious period*. Some authors also refer to α and β as the *infection rate* and *removal rate*, respectively.

Modifications of the SIR model e.g. disregard recovery (SI), allow a return to the susceptible status (SIS, SIRS), or incorporate a latent/exposed period (SLIR/SEIR). For simplicity, we assume in this section that an individual is infectious as soon as it is infected. The terms *infected*, *infectious* and *infective* are considered interchangeable.

The SIR model is conveniently described as a time-homogeneous Markov process. Unless otherwise stated, we assume the population closed during the time of consideration, ignoring births, non-related deaths, and migration. Furthermore, the population is presumed to mix homogeneously.

Different constructions of the SIR model can be found in the literature, see for example Andersson and Britton (2000) for an overview. The following paragraphs present three of the most common descriptions.

Representation as Pure Markov Jump Process

Denote by S and I the absolute numbers of susceptible and infectious individuals in the population under consideration. Due to the fixed population size N , the current state of an SIR process is completely described by the tuple $(S, I)'$, which is an element of the state space $\mathcal{D} = \{(S, I)' \in [0, N]^2 \cap \mathbb{N}_0^2 \mid S + I \leq N\}$; the number of removed individuals can be calculated as $R = N - S - I$.

Hence, let $(S, I)' \in \mathcal{D}$ be the state of the process at time $t \in \mathbb{R}_0$. Assuming that at most one event can occur within a small time interval of length Δt , there are three possibilities for the state of the process at time $t + \Delta t$:

1. $(S - 1, I + 1)'$ in case one infection occurs,
2. $(S, I - 1)'$ in case one recovery occurs,
3. $(S, I)'$ in case nothing happens.

These transitions come up with probabilities

$$p_1 = \alpha SI/N \Delta t + o(\Delta t), \quad p_2 = \beta I \Delta t + o(\Delta t) \quad \text{and} \quad p_3 = 1 - p_1 - p_2, \quad (2.2)$$

respectively, where $o(\Delta t)/\Delta t \rightarrow 0$ as $\Delta t \rightarrow 0$. See Sect. 5.1.2 for the derivation of (2.2). For $(S, I)' \notin ([0, N-1] \times [1, N-1]) \cap \mathcal{D}$, the above target states may not be an element of \mathcal{D} . In those cases, however, the respective transition probabilities leading to them are $o(\Delta t)$. For an initial condition $(S_0, I_0)' \in \mathcal{D}$, the process can therefore never leave the admissible state space.

A Markov process with the above described dynamics is also termed the *general stochastic epidemic*. Section 2.4.1 describes how an according Markov chain can exactly be simulated. Figure 2.3a shows a realisation of such a Markov chain.

A notable insight into the dynamics of the general stochastic epidemic is the following stochastic threshold result: Let $(S_0, I_0)' \in \mathcal{D}$ denote the initial state of the process and define $\mathcal{R}_0 = \alpha/\beta$. Then, in large populations, a major outbreak will occur with probability tending to

$$1 - \left(\min \left\{ 1, \frac{N}{S_0} \mathcal{R}_0^{-1} \right\} \right)^{I_0}$$

as N and $S_0 = N - I_0$ grow to infinity for fixed I_0 (Whittle 1955; Williams 1971; Ball 1983). This probability is positive if and only if the relative removal rate \mathcal{R}_0^{-1} is smaller than the initial fraction of susceptibles S_0/N . In this formulation, the term *major outbreak* means that the fraction S/N of susceptibles will fall below \mathcal{R}_0^{-1} roughly as far as it was above this threshold before, provided that the difference between S_0/N and \mathcal{R}_0^{-1} is not too large. For more details, see for example Daley and Gani (1999, Chap. 3.4). \mathcal{R}_0 is called the *basic reproductive ratio* and interpreted as the average number of infections caused by an infectious individual during its entire infectious period, provided that the infective enters a totally susceptible population.

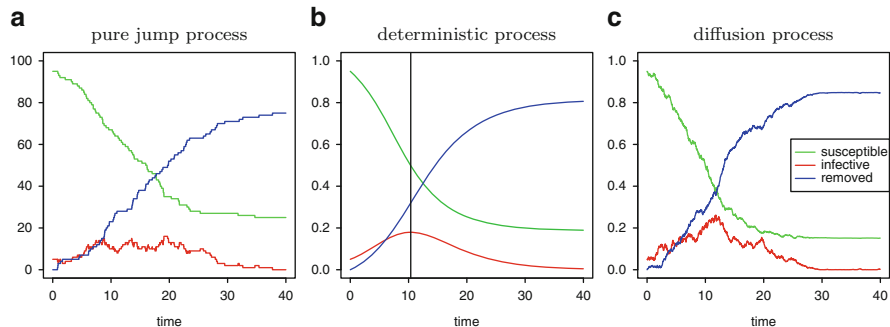


Fig. 2.3 Illustration of SIR model for parameters $\alpha = 0.5$, $\beta = 0.25$ and population size $N = 100$. **(a)** Temporal evolution of numbers of susceptible, infective and removed individuals in the stochastic SIR model with transition probabilities (2.2) for initial value $(S_0, I_0)' = (95, 5)'$. The graphs have been simulated by application of Gillespie's Algorithm, i.e. Algorithm 2.1 on p. 26. **(b)** Temporal evolution of fractions of susceptible, infective and removed individuals in the standard deterministic SIR model (2.3) for initial value $(s_0, i_0)' = (0.95, 0.05)'$. The graphs have been obtained by application of the standard Euler scheme with step length 0.025. The vertical line marks the instant at which the fraction of susceptibles falls below $\mathcal{R}_0^{-1} = \beta/\alpha = 0.5$. The fraction of infectives reaches its maximum at this point. **(c)** Temporal evolution of fractions of susceptible, infective and removed individuals in the SIR diffusion model (2.4) for initial value $(s_0, i_0)' = (0.95, 0.05)'$. The graphs have been obtained by application of the Euler-Maruyama scheme from Sect. 6.3.2 with step length 0.025

Representation Through a System of Ordinary Differential Equations

Another possibility to describe the infection dynamics in the SIR model is a deterministic representation via the set of ordinary differential equations (ODEs)

$$ds/dt = -\alpha si, \quad di/dt = \alpha si - \beta i, \quad (2.3)$$

where $s = S/N$ and $i = I/N$ denote the fractions of susceptible and infectious individuals. In this description, the state space $\mathcal{C} = \{(s, i)' \in [0, 1]^2 \cap \mathbb{R}_0^2 \mid s + i \leq 1\}$ is considered continuous, which is an eligible assumption for large populations. The remaining fraction $r = R/N$ can again be obtained as $r = 1 - s - i$. The ODEs (2.3) are subject to an initial condition $(s_0, i_0)' \in \mathcal{C}$. See Sect. 5.1.4 for their formal derivation.

Figure 2.3b shows the typical evolution of an epidemic following the deterministic description (2.3). While recovery follows a linear process, infections occur at high rate only when both the fractions of susceptibles and infectives are sufficiently large. As the ODEs are not explicitly solvable, the trajectories have been obtained numerically by application of the standard Euler scheme (cf. Sect. 2.4.2). Figure 2.4 displays the course of the deterministic SIR process for different values of α and β .

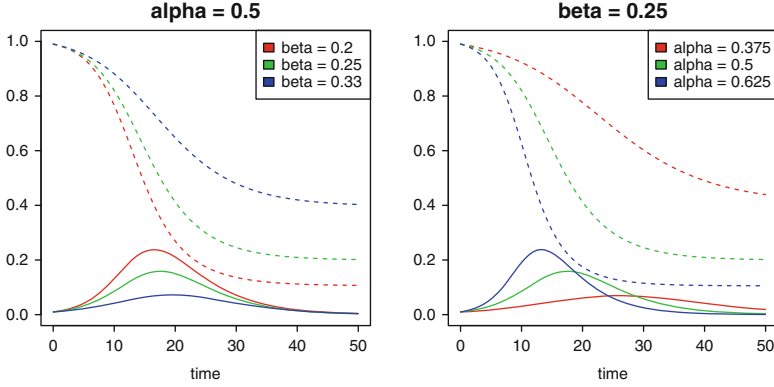


Fig. 2.4 Fractions of susceptibles (*dashed*) and infectives (*solid lines*) in an SIR epidemic following the deterministic model (2.3) for different values of α and β and initial value $(s_0, i_0)'$ equal to $(0.99, 0.01)'$. The graphs have been obtained by application of the standard Euler scheme with step length 0.025 for solving the ODE system. In both graphics, the parameters correspond to $\mathcal{R}_0 = \alpha/\beta \in \{1.5, 2.0, 2.5\}$

The first equation in (2.3) implies that the fraction of susceptibles is strictly decreasing as long as both s and i are nonzero. Solving $di/dt < 0$ leads to $s < \beta/\alpha$. That means, when $\mathcal{R}_0^{-1} := \beta/\alpha$ is greater than the initial fraction of susceptibles s_0 , no epidemic will develop. Otherwise, the epidemic will rise first but fall off as soon as the fraction s drops below this threshold. This is the famous *threshold theorem* by [Kermack and McKendrick \(1927\)](#). An obvious strategy to eradicate an epidemic is hence to vaccinate the population until the latter requirement is met. The vertical line in Fig. 2.3b indicates the first time point at which the fraction of susceptibles falls below \mathcal{R}_0^{-1} . Apparently, this mark agrees with the time point at which the epidemic reaches its maximum with respect to the number of infected individuals.

Representation Through a System of Stochastic Differential Equations

A third variant to express the SIR dynamics (2.1) as a mathematical process is by a stochastic differential equation (SDE)

$$\begin{pmatrix} ds \\ di \end{pmatrix} = \begin{pmatrix} -\alpha si \\ \alpha si - \beta i \end{pmatrix} dt + \frac{1}{\sqrt{N}} \begin{pmatrix} \sqrt{\alpha si} & 0 \\ -\sqrt{\alpha si} & \sqrt{\beta i} \end{pmatrix} \begin{pmatrix} dB_1 \\ dB_2 \end{pmatrix}. \quad (2.4)$$

In this equation, s and i denote again the fractions of susceptible and infectious individuals in the population. The right hand side of the differential equation (2.4) consists of a deterministic and a stochastic component, that is the first and the second summand, respectively. B_1 and B_2 are independent Brownian motions, representing stochasticity in disease transmission and recovery. As for the multivariate ODE (2.3), an appropriate initial condition has to be specified for the SDE (2.4).

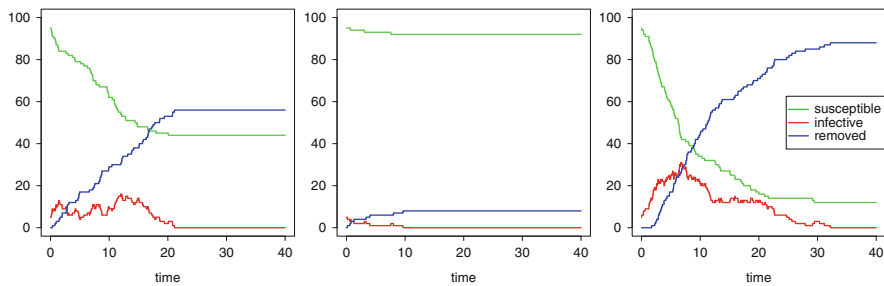


Fig. 2.5 Different courses of stochastic SIR model with transition probabilities (2.2). The simulations base on parameters $\alpha = 0.5$, $\beta = 0.25$, population size $N = 100$ and initial value $(S_0, I_0)' = (95, 5)'$. The graphs have been obtained by application of Gillespie's Algorithm (Algorithm 2.1)

Stochastic differential equations and their solutions, which are typically diffusion processes, will be formally introduced in Chap. 3. Diffusion processes possess extremely wiggly but almost surely continuous trajectories. Figure 2.3c displays the course of an SIR epidemic as described by Eq. (2.4).

Concluding Remarks

This section introduced three different representations of the standard SIR model. There naturally arises the question which type of process is the most appropriate one. The pure Markov jump process, considered first, mirrors the exact dynamics following the transitions (2.1). In many cases, however, this type of process is rather inconvenient for the purpose of simulation and statistical inference. The ODE representation, considered next, has the advantage of a non-individual-based view point. It facilitates interpretation and mathematical analysis, but unfortunately ignores possible variation by chance. In particular, the ODEs (2.3) do not even take into account the population size N and hence unrealistically predict identical fractions of infectives and susceptibles in small and large populations. Finally, the representation of the SIR model in terms of a multivariate SDE consists of both a deterministic and a stochastic component and this way compromises on the former two processes. For this reason, the utilisation of SDEs is favourable in many contexts. Their statistical analysis is the subject of this book. A more elaborate discussion concerning the three above representations is the topic of Chap. 4.

In order to further elucidate the impact of random events in the SIR model, recall the above deterministic and stochastic threshold results. Both the stochastic model with transition probabilities (2.2) and the deterministic model following the ODEs (2.3) possess the same threshold $\mathcal{R}_0^{-1} = s_0$. The interpretation of this threshold, however, differs substantially in these two models: In the deterministic case, a major epidemic will always occur whenever $\mathcal{R}_0^{-1} < s_0$. In the stochastic case, a major outbreak does not necessarily happen if $\mathcal{R}_0^{-1} < s_0$. The probability for this event lies strictly between zero and one. Figure 2.5 illustrates that different

realisations of the course of an epidemic may clearly differ in a stochastic framework. A deterministic simulation for the same model parameters is displayed in Fig. 2.3b. A further investigation of the SDE (2.4) requires its formal definition, which is the subject of Chap. 3. An illustration of this model is for example given in Sect. 5.1.5.

Epidemics will usually terminate due to a lack of infectives, not due to a lack of susceptibles, i.e. at the end of an epidemic outbreak not all individuals will typically have suffered from the disease. According to the above thresholds, major epidemics occur or have positive probability, respectively, when $\mathcal{R}_0 < s$. Suppose that this is the case. Then, there are three general measures to weaken the strength of an epidemic: First, to reduce the number of susceptibles, typically by vaccination, i.e. to decrease the fraction s . Second, to reduce the number of potentially infectious contacts, possibly by closing schools or simply invoking caution, i.e. to decrease α . Third, to reduce the time until an infectious individuals goes over to the removed class, for example by isolation, i.e. to reduce the average infectious period β^{-1} . Each of these three strategies aims at lowering the difference between $\mathcal{R}_0 = \alpha/\beta$ and s , at best accomplishing $\mathcal{R}_0 > s$. The fact that an epidemic does not start or fades out after sufficiently many individuals have left the susceptible state is known as *herd immunity*. The subject of herd immunity, including many examples, is discussed by [Anderson and May \(1985\)](#) and [Fine \(1993\)](#), corresponding control strategies by [Morton and Wickwire \(1974\)](#).

2.2.3 Model Extensions

So far, the SIR model considered in the previous section is fairly simplistic, assuming a homogeneously mixing population, homogeneity of individuals and a time-homogeneous course of an epidemic. In most contexts, some modifications are necessary in order to adapt the mathematical model to a real life situation in which an epidemic develops. Some of these aspects are outlined in the following.

First of all, one very often experiences heterogeneity in contacts among the population. In those cases, individuals typically mix homogeneously in certain subgroups but not with respect to the entire population. It is then meaningful to incorporate patterns into the model such as the age structure of the population e.g. for childhood diseases, a risk structure e.g. for sexually transmitted infections, a geographical structure like an assignment of individuals to different cities or countries, or social structures such as households, schools or circles of friends.

Moreover, there is typically heterogeneity among individuals in the population. For example, susceptible persons may differ in their degree of susceptibility, such as children or elderly people that possibly have a weaker immune system, or individuals that have acquired partial immunity to a disease due to previous epidemics.

In some cases, it is also appropriate to extend an epidemic model such that it accounts for time-varying background conditions. For example, the weather

and temperature may well have an effect on the susceptibility of individuals. Furthermore, there may be changes in social behaviour, either independently or dependently on the course of an ongoing epidemic, leading to a variation of contact rates. When observing the spread of a disease over a long period of time, demographic changes such as births and non-related deaths may be included in the model. Other models consider endemic components, i.e. the sustained presence of a certain number of infectious cases in the population, or the presence of carriers that are apparently healthy but infective.

Ample examples and references for the above model extensions are given by [Isham \(2004\)](#) and [Keeling and Rohani \(2008\)](#). In order to mention just a few of them, multipopulation epidemics are for example investigated by [Rushton and Mautner \(1955\)](#), [Ball \(1986\)](#), [Sattenspiel \(1987\)](#), [Sattenspiel and Dietz \(1995\)](#) and [Ball et al. \(1997\)](#). Such models can often be applied to any kind of contact heterogeneity but are in most cases described for the division of a population into several communities in distinct geographical areas. Chapter 5 in this book introduces a multitype SIR model for arbitrary contact heterogeneities as well. Concerning the remaining model modifications mentioned above, [Hethcote \(2000\)](#) takes into account the age of individuals, and [Hethcote \(1994\)](#) gives many references for models which take into account varying population sizes. [Neal \(2007\)](#) analyses an epidemic model where individuals differ with respect to both their susceptibility and infectivity. [Ireland et al. \(2007\)](#) consider seasonality in birth-rates of hosts. [Riley \(2007\)](#) reviews some recent approaches for spatial modelling. Finally, [Lloyd-Smith et al. \(2005\)](#) and [Galvani and May \(2005\)](#) investigate the impact of the presence of superspreaders, that are individuals that communicate a disease in a substantially greater extent than other individuals.

Appropriate modifications of the basic SIR model improve the compatibility between the model assumptions and reality and hence increase the applicability of the model. On the other hand, each extension automatically requires additional information such as community sizes or contact patterns between groups. One should hence balance carefully between complex and oversimplistic models. Stochastic models typically get along with fewer details as minor aspects can be covered by random fluctuations. Chapters 5 and 8 in this book derive and statistically infer on a probabilistic multitype model for the spread of an infectious disease.

2.3 Modelling Processes in Molecular Biology, Biochemistry and Genetics

Understanding the mechanisms of heredity and variation of living organisms, senescence and the emergence of diseases such as cancer has fascinated mankind within living memory. Nowadays one knows that these phenomena are based on chemical processes in living organisms and the structures and functions of living cells.

This section briefly considers mathematical modelling in the overlapping areas of molecular biology, biochemistry and genetics. These fields comprise an enormous variety of different applications and models, the complete review of which would be far beyond the scope of this book. Hence, this section exemplarily addresses one specific branch of the above research areas: That is, applications which utilise the framework of chemical reactions for the modelling of selected key processes. This section hence starts with historical background information and a mathematical review on that subject in Sects. 2.3.1 and 2.3.2, followed by an outline of cross connections to other disciplines in Sect. 2.3.3.

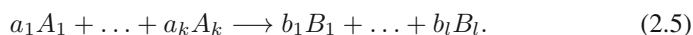
2.3.1 History of Chemical Reaction Modelling

The first landmark in the development of chemical reaction modelling was set in 1850 by Ludwig Wilhelmy, who empirically derived a mathematical expression for the progress of the inversion of cane sugar in the presence of acids (McQuarrie 1967; Arnaut et al. 2007). In several articles published between 1864 and 1879, Cato Maximilian Goldberg and Peter Waage proposed the *law of mass action*, which says that the hazard of an elementary reaction is proportional to the product of the concentrations of all reactants; cf. Sect. 2.3.2 for details. Important further contributions to the understanding of the order and temperature dependence of chemical reactions were made between 1865 and 1889 by Augustus Harcourt, William Esson, Jacobus Henricus van't Hoff, Wilhelm Ostwald and Svante Arrhenius (Laidler 1993). Until 1940, many mathematical models were formulated which described the mechanism of a chemical reaction in a deterministic way. According to McQuarrie (1967), Kramers (1940) was the first author who applied the theory of stochastic processes to chemical reactions models.

Nowadays, detailed knowledge about molecular structures and mechanisms is available, in addition to sophisticated mathematical and statistical modelling tools. This enables the description and analysis of complex chemical networks. A detailed historical review on chemical kinetics modelling is provided by Arnaut et al. (2007). McQuarrie (1967) considers this subject from a statistician's point of view.

2.3.2 Chemical Reaction Kinetics

Chemical reactions are typically specified by reaction equations of the form



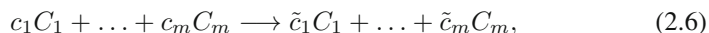
This equation describes a reaction in which k different *reactants* A_1, \dots, A_k are transformed into l distinct *products* B_1, \dots, B_l . The numbers a_i , $i = 1, \dots, k$, and b_j , $j = 1, \dots, l$, are the *stoichiometries* of the reaction and denote the numbers of

reactants A_i and products B_j involved. They are assumed to be natural numbers with greatest common divisor equal to one. In this chapter, equations like (2.5) are declared to represent *elementary reactions*, i.e. reactions that do not consist of several intermediate steps. Equation (2.1) on p. 13 was of type (2.5) as well.

As in the context of modelling the spread of infectious diseases in the previous section, there are various approaches to mathematically describe the dynamics of a process in which reactions such as (2.5) occur. In what follows, three possibilities are briefly introduced in the same order as for the SIR model in Sect. 2.2.2. All representations have in common that they assume the underlying system well-stirred and the process to be Markovian and time-homogeneous. In particular, external parameters such as temperature and pressure are presumed to be constant.

Representation as Pure Markov Jump Process

The sets of reactants $\{A_1, \dots, A_k\}$ and products $\{B_1, \dots, B_l\}$ are typically non-disjoint subsets of a collection $\{C_1, \dots, C_m\}$ of particles that are present in the considered system. The reaction equation (2.5) can hence be rewritten as



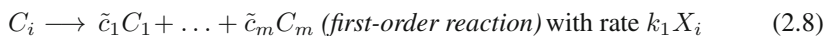
where

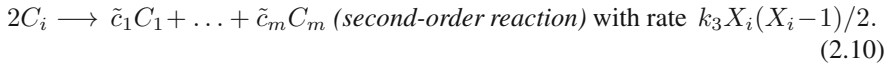
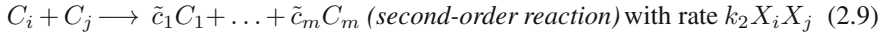
$$c_i = \begin{cases} a_j & \text{if } C_i = A_j \\ 0 & \text{if } C_i \notin \{A_1, \dots, A_k\} \end{cases} \quad \text{and} \quad \tilde{c}_i = \begin{cases} b_j & \text{if } C_i = B_j \\ 0 & \text{if } C_i \notin \{B_1, \dots, B_l\}. \end{cases}$$

For $i \in \{1, \dots, m\}$, let X_i denote the number of particles C_i in the system and define $(X_1, \dots, X_m)'$ as the state variable of a stochastic process describing the system dynamics. The chemical reaction (2.6) then causes a state change

$$\begin{pmatrix} X_1 \\ \vdots \\ X_m \end{pmatrix} \longrightarrow \begin{pmatrix} X_1 - (c_1 - \tilde{c}_1) \\ \vdots \\ X_m - (c_m - \tilde{c}_m) \end{pmatrix}. \quad (2.7)$$

In real applications, one typically has several chemical reactions such as (2.6), each causing a transition like (2.7). Every reaction is associated with a *reaction rate* indicating the hazard with which the specific reaction is going to occur within the next infinitesimal time interval. These rates are assumed to depend on the left hand side of (2.6) only. [Wilkinson \(2006\)](#) exemplarily states the following reactions and associated reaction rates, where the current state of the process is $(X_1, \dots, X_m)'$:





In the second equation, one requires $i \neq j$. The variables $k_1, k_2, k_3 \in \mathbb{R}_+$ are called *rate constants*. They are usually unknown and hence the subject of statistical inference based on available experimental data. The remaining parts of the reaction rates result from combinatorial considerations, counting the number of possible collisions between the reactants, and the fact that the hazard of two specific particles colliding is constant (Gillespie 1992).

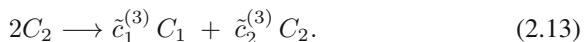
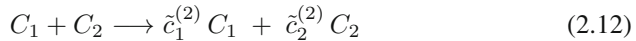
As a consequence of the above specified reaction rates, the probability that, for example, reaction (2.8) will occur within a time interval of length Δt , provided that the current number of particles C_i is X_i , equals $k_1 X_i \Delta t + o(\Delta t)$, where $o(\Delta t)/\Delta t \rightarrow 0$ as $\Delta t \rightarrow 0$. Without any other reactions taking place, the expected time until the occurrence of this reaction is exponentially distributed with mean $k_1^{-1} X_i$.

Representation Through a System of Ordinary Differential Equations

A different possibility to describe the state of a system which is subject to elementary chemical reactions of type (2.6) is via the rates of change of the concentrations of all reaction participants. To that end, consider the concentrations x_1, \dots, x_m of the particles X_1, \dots, X_m . These concentrations are considered continuous rather than discrete quantities. The chemical reaction (2.6) induces a change of the current state $(x_1, \dots, x_m)'$ which is typically described by a set of ordinary differential equations (ODEs): For all $i = 1, \dots, m$, one has

$$dx_i/dt = \bar{k} (\tilde{c}_i - c_i) x_1^{c_1} \cdot \dots \cdot x_m^{c_m}$$

for some positive (*stochastic*) rate constant \bar{k} . This equation results from the law of mass action, which was already mentioned in Sect. 2.3.1. The sum of exponents $c_1 + \dots + c_m$ is called the *order* of the reaction (McQuarrie 1967). The right hand side of the ODE is positive if $c_i < \tilde{c}_i$, i.e. if the chemical reaction described by (2.6) increases the amount of particles X_i in the system. It is negative or equal to zero if the reaction decreases the number X_i or leaves it unaltered, respectively. If there is more than one possible reaction, each reaction is assigned a different rate constant, and the ODEs resulting from each reaction equation are added in order to arrive at a description for the whole reaction dynamics. For example, consider the following set of coupled reactions for $m = 2$, which is a special case of Eqs. (2.8)–(2.10):



For these reactions, one obtains the ODEs

$$dx_1/dt = \bar{k}_1 \left(\tilde{c}_1^{(1)} - 1 \right) x_1 + \bar{k}_2 \left(\tilde{c}_1^{(2)} - 1 \right) x_1 x_2 + \bar{k}_3 \tilde{c}_1^{(3)} x_2^2 \quad (2.14)$$

$$dx_2/dt = \bar{k}_1 \tilde{c}_2^{(1)} x_1 + \bar{k}_2 \left(\tilde{c}_2^{(2)} - 1 \right) x_1 x_2 + \bar{k}_3 \left(\tilde{c}_2^{(3)} - 2 \right) x_2^2 \quad (2.15)$$

for appropriate rate constants $\bar{k}_1, \bar{k}_2, \bar{k}_3 > 0$. Additionally, a suitable initial state of the process needs to be specified. The constants k_1, k_2, k_3 in Eqs. (2.8)–(2.10) and the constants $\bar{k}_1, \bar{k}_2, \bar{k}_3$ in (2.14)–(2.15) depend on the units of X_1, X_2 and x_1, x_2 , respectively, and are not necessarily the same. See [Wilkinson \(2006, Chap. 6.6\)](#) for the conversion from k_i to \bar{k}_i in case the concentrations are measured in moles per litre.

Representation Through a System of Stochastic Differential Equations

Finally, a third way to represent the evolution of a system which is subject to chemical reactions utilises stochastic differential equations (SDEs). In case of the reactions (2.11)–(2.13), the multi-dimensional SDE reads

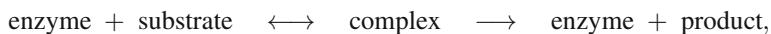
$$\begin{pmatrix} dx_1 \\ dx_2 \end{pmatrix} = \begin{pmatrix} \bar{k}_1 \left(\tilde{c}_1^{(1)} - 1 \right) x_1 + \bar{k}_2 \left(\tilde{c}_1^{(2)} - 1 \right) x_1 x_2 + \bar{k}_3 \tilde{c}_1^{(3)} x_2^2 \\ \bar{k}_1 \tilde{c}_2^{(1)} x_1 + \bar{k}_2 \left(\tilde{c}_2^{(2)} - 1 \right) x_1 x_2 + \bar{k}_3 \left(\tilde{c}_2^{(3)} - 2 \right) x_2^2 \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \begin{pmatrix} dB_1 \\ dB_2 \end{pmatrix},$$

where $\sigma_{11}, \sigma_{12}, \sigma_{21}$ and σ_{22} are functions of the state variables, rate constants and stoichiometries not explicitly given here. The first summand on the right hand side represents the deterministic component of the process and agrees with Eqs. (2.14) and (2.15). The second summand stands for the probabilistic component with B_1 and B_2 being two independent Brownian motion processes. SDEs and Brownian motion will formally be defined in Chap. 3.

2.3.3 Reaction Kinetics in the Biological Sciences

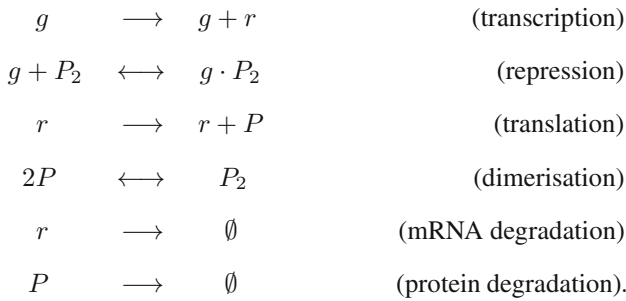
Reaction equations and their associated mathematical theory are convenient tools also in the biological sciences. They are particularly used to describe the natural laws which underlie the functioning of cells. This section gives some examples.

Chemical work can be performed by cells only if there is enough energy available. Such energy is gained through cellular catabolism, which is a mechanism consisting of a series of enzymatic reactions like



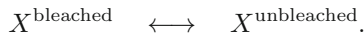
where the enzyme acts as a catalyst (Keener and Sneyd 1989). Double-sided arrows mean that the reaction can take place in both directions. Kinetic models for metabolic systems are, for example, developed by Demin et al. (2005).

Within each cell, there are several thousand types of interacting proteins. Depending on its environment, a cell determines the required amount of each protein by means of *transcription networks* (Alon 2007). Transcription is one out of several regulatory mechanisms in genetic networks and can be described by a set of coupled elementary reactions (Wilkinson 2006). At a less detailed level, transcription and other key processes can be assembled to construct genetic networks. For example, the following components of a prokaryotic auto-regulatory network are summarised by Wilkinson (2006):



In these equations, P stands for a protein, P_2 for the compound of two of these proteins, g for a gene and r for a *transcript* of g . The empty set \emptyset indicates that the product of a reaction is not part of the model, and a dot represents the compound of two components.

The close connection between models for chemical reactions and genetic mechanisms is hardly surprising as genetics is based on the chemistry of nucleic acids. There are, however, also cases of compartmental systems in cellular biology where reaction equations represent transitions other than chemical reactions. In the application in Chap. 9, for example, the location of a diffusing protein between a bleached and an unbleached part of the cell nucleus is observed. This can be written as



A molecule that undergoes this transition does not change any of its chemical or kinetic properties but only its location, so the compartments reflect the spatial dimension of the problem here.

Plenty of further applications are, for example, presented in Jacquez (1972) and McQuarrie (1967). Ehrenberg et al. (2003) give a brief overview about current research questions in systems biology. For general reviews on mathematical models in biology, see Goel and Richter-Dyn (1974), Renshaw (1991), Allen (2003) or Lande et al. (2003).

Though representing entirely different natural phenomena, the above mentioned applications have in common that they are intrinsically stochastic. A number of

papers is devoted to the importance of the utilisation of probabilistic instead of deterministic models in systems biology, biochemistry and genetics, see for example Kimura (1964), Zheng and Ross (1991), Arkin et al. (1998), Svecizer et al. (2001), Rao et al. (2002), Bahcall (2005), Tian et al. (2007) and Boys et al. (2008). In agreement with this point of view, the present book motivates, constructs and statistically infers on stochastic models from life sciences.

2.4 Algorithms for Simulation

In Sects. 2.2 and 2.3, different kinds of processes were considered to represent the dynamics of different phenomena in life sciences. For the simulation of these processes, one requires algorithms for the exact or approximate generation of according sample paths. Such algorithms have already been applied for the generation of Figs. 2.3–2.5.

2.4.1 Simulation of Continuous-Time Markov Jump Processes

Continuous-time pure Markov jump processes can always exactly be simulated. An according algorithm is presented in what follows.

Consider a system consisting of n different types of objects such as molecules in a fluid, predator and prey in a specified region or susceptibles and infectives in a population. Assume that the time-continuous evolution of these objects can be described by a time-homogeneous stochastic Markov process with state variable $\mathbf{X}(t) = (X_1(t), \dots, X_n(t))' \in \mathbb{Z}^n$, where $X_i(t)$ is the number of type i objects at time $t \in \mathbb{R}_+$. Suppose that there are m possible events $k \in \{1, \dots, m\}$ like chemical reactions or interactions within a population, each causing a change $\Delta_k \in \mathbb{Z}^n \setminus \{0\}$ in the state variable. Let $\lambda_k = f_k(\mathbf{X})$ denote the hazard for event k , where f_k is an appropriate function depending on the state \mathbf{X} . That means, the probability that a type k event will occur within the next time interval of length Δt conditioned on the current state \mathbf{X} is $\lambda_k \Delta t + o(\Delta t)$, where $o(\Delta t)/\Delta t \rightarrow 0$ as $\Delta t \rightarrow 0$. The objective is to exactly simulate realisations of the considered process, that means to successively draw pairs $(\tau, k) \in \mathbb{R}_+ \times \{1, \dots, m\}$, where τ is the waiting time until the occurrence of the next event, and k is the type of event happening at that time.

Denote by $p(\tau, k)$ the joint probability density function of τ and k . Under the assumption that only one event can happen at the same time, Gillespie (1976) shows that

$$p(\tau, k) = \lambda_k \exp \left(-\tau \sum_{j=1}^m \lambda_j \right) = \lambda_k \exp(-\lambda \tau) \text{ for } \tau \in \mathbb{R}_+ \text{ and } k \in \{1, \dots, m\},$$

where $\lambda = \sum_{j=1}^m \lambda_j$. This joint density can be expressed as $p(\tau, k) = p(\tau)p(k|\tau)$, where

$$p(\tau) = \sum_{k=1}^m p(\tau, k) = \lambda \exp(-\lambda\tau), \quad \text{i.e.} \quad \tau \sim \text{Exp}(\lambda),$$

and

$$p(k|\tau) = \frac{p(\tau, k)}{p(\tau)} = \frac{\lambda_k}{\lambda} \quad (2.16)$$

are the density of τ and the conditional probability function of k , respectively.

This leads to an exact and efficient method to obtain sample trajectories of the considered process on a time interval $[t_{\min}, t_{\max}]$. The procedure has been called *stochastic simulation algorithm (SSA)* by its originator, but is usually known as *Gillespie's algorithm*:

Algorithm 2.1 (Gillespie's Algorithm, Gillespie 1976).

1. Set $t = t_{\min}$ and initialise $\mathbf{X}(t)$.
2. While $t < t_{\max}$:
 - i. Calculate λ_k for all k and their sum λ . Terminate if the system has reached an absorbing state, i.e. $\lambda = 0$.
 - ii. Draw $\tau \sim \text{Exp}(\lambda)$. Set $\tau^* = \min\{\tau, t_{\max} - t\}$.
 - iii. Draw k from (2.16).
 - iv. Set $\mathbf{X}(s) = \mathbf{X}(t)$ for all $s \in (t, t + \tau^*)$ and $\mathbf{X}(t + \tau^*) = \mathbf{X}(t) + \Delta_k \mathbb{1}(\tau^* = \tau)$.
 - v. Set $t = t + \tau$.

Estimates of the average or the variation of the sample paths can be obtained by respective Monte Carlo statistics. For further details and experimental results, see Gillespie (1976, 1977). Extensions, later elaborations and improvements with respect to computing time are contained in Gillespie (2007). Manninen et al. (2006) provide ample references for different implementations of the Gillespie algorithm, such as the *next reaction method* by Gibson and Bruck (2000), and alternative approaches, for example the *StochSim algorithm* by Le Novère and Shimizu (2001). Another good review is Wilkinson (2006, Chap. 8).

2.4.2 Simulation of Solutions of ODEs and SDEs

When a system consists of a large number of objects, the just described simulation of a pure Markov jump process becomes expensive in terms of computing time. In contrast, the most convenient process with respect to its simulation is the deterministic process described by a set of ODEs, because this process has no random component. If there is an analytically explicit solution of the ODEs available, one can simply calculate the according multivariate sample path without

any approximation error. Otherwise, numerical schemes such as the Euler scheme can be applied to obtain approximate trajectories. Such algorithms can be found in any standard textbook on numerical mathematics.

Similarly, a stochastic process described by a set of SDEs can exactly be simulated if an explicit solution for the differential equations is known. Otherwise, numerical approximation schemes are utilised. The consideration of respective procedures is postponed to Sect. 3.3 in the next chapter, because this subject requires a preliminary introduction to stochastic calculus. The numerical approximation of a solution of an ODE arises as a special case of the algorithm for an SDE.

2.5 Conclusion

Assessment of key mechanisms in life sciences cannot be imagined without the application of mathematical models. Moreover, real situations can particularly be rendered by the consideration of random events. This chapter provided an introduction to established models in life sciences, starting with the general class of compartment models in Sect. 2.1 and then proceeding to applications in mathematical epidemiology and biology in Sects. 2.2 and 2.3. To that end, three types of processes were considered, namely stochastic jump processes, deterministic continuous processes and stochastic diffusion processes, the simulation of which is the subject of Sect. 2.4. The latter type of process emerges as a convenient compromise between the former two, and hence this book focuses on diffusion processes.

However, diffusions have not been defined formally in this book yet. For that reason, Chap. 3 introduces the theory of stochastic calculus to an extent which is oriented towards the needs of subsequent chapters. Chapter 4 discusses the application of the three above process classes and considers the derivation of diffusion processes from the compartmental description of some phenomenon. This methodology is applied in Chap. 5, where a multitype SIR model for heterogeneous contact patterns is developed.

Until that point, this book is mainly concerned with the construction of models, which enables the simulation of a considered mechanism for given sets of model parameters. In practice, however, such parameters are unknown and hence to be estimated statistically based on available observations. Therefore, Chaps. 6 and 7 consider the important subject of statistical inference for diffusion processes.

The methodology of all preceding parts is applied in Chaps. 8 and 9 on the example of modelling the spread of influenza and the binding behaviour of molecules, respectively. These chapters also point out challenges arising from typical data situations such as partial observations or measurement errors.

References

- Abbey H (1952) An examination of the Reed-Frost theory of epidemics. *Hum Biol* 24:201–233
- Allen L (2003) An introduction to stochastic processes with applications to biology. Pearson Prentice Hall, Upper Saddle River
- Alon U (2007) An Introduction to systems biology. Design principles of biological circuits. Chapman and Hall, Boca Raton
- Anderson R (1982) The population dynamics of infectious diseases. Chapman and Hall, London
- Andersson H, Britton T (2000) Stochastic epidemic models and their statistical analysis. Lecture Notes in Statistics, vol 151. Springer, New York
- Anderson R, May R (1985) Vaccination and herd immunity to infectious diseases. *Nature* 318: 323–329
- Anderson R, May R (1991) Infectious diseases of humans. Oxford University Press, Oxford
- Arkin A, Ross J, McAdams H (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage λ -infected *Escherichia coli* cells. *Genetics* 149:1633–1648
- Arnaut L, Formosinho S, Burrows H (2007) Chemical kinetics: from molecular structure to chemical reactivity. Elsevier, Amsterdam/Oxford
- Bahcall O (2005) Single cell resolution in regulation of gene expression. *Mol Syst Biol* 1 (article number 2005.0015)
- Bailey N (1975) The mathematical theory of infectious diseases, 2nd edn. Charles Griffin, London
- Ball F (1983) The threshold behaviour of epidemic models. *J Appl Probab* 20:227–241
- Ball F (1986) A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models. *Adv in Appl Probab* 18:289–310
- Ball F, Mollison D, Scalia-Tomba G (1997) Epidemics with two levels of mixing. *Ann Appl Probab* 7:46–89
- Bartlett M (1949) Some evolutionary stochastic processes. *J R Stat Soc Ser B* 11:211–229
- Becker N (1989) Analysis of infectious disease data. Monographs on statistics and applied probability. Chapman and Hall, London
- Boys R, Wilkinson D, Kirkwood T (2008) Bayesian inference for a discretely observed stochastic kinetic model. *Stat Comput* 18:125–135
- Brauer F (2009) Mathematical epidemiology is not an oxymoron. *BMC Public Health* 9:S2
- Costa Maia J (1952) Some mathematical developments on the epidemic theory formulated by Reed and Frost. *Hum Biol* 24:167–200
- Cunha B (2004) Historical aspects of infectious diseases, part I. *Infect Dis Clin N Am* 18(1):xi–xv
- Daley D, Gani J (1999) Epidemic modelling: an introduction. Cambridge studies in mathematical biology, vol 15. Cambridge University Press, Cambridge
- Demin O, Plyusnina T, Lebedeva G, Zobova E, Metelkin E, Kolupaev A, Goryanin I, Tobin F (2005) Kinetic modelling of the *E. coli* metabolism. In: Alberghina L, Westerhoff H (eds) Systems biology. Definitions and perspectives. Springer, Berlin/Heidelberg, pp 31–67
- Diekmann O, Heesterbeek J (2000) Mathematical epidemiology of infectious diseases: model building, analysis and interpretation. Wiley, Chichester
- Dietz K (1967) Epidemics and rumours: a survey. *J R Stat Soc Ser A* 130:505–528
- Dobson A, Carper E (1996) Infectious diseases and human population history. *Bioscience* 46: 115–126
- Ehrenberg M, Elf J, Aurell E, Sandberg R, Tegnér J (2003) Systems biology is taking off. *Genome Res* 13:2377–2380
- En'ko P (1889) On the course of epidemics of some infectious diseases. *Vrach St Petersburg* 10:1008–1010, 1039–1042, 1061–1063
- Fine P (1993) Herd immunity: history, theory, practice. *Epidemiol Rev* 15:265–302
- Galvani A, May R (2005) Dimensions of superspreading. *Nature* 438:293–295
- Gibson M, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. *J Phys Chem A* 104:1876–1889

- Gillespie D (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comput Phys* 22:403–434
- Gillespie D (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81:2340–2361
- Gillespie D (1992) A rigorous derivation of the chemical master equation. *Phys A* 188:404–425
- Gillespie D (2007) Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem* 58:35–55
- Goel N, Richter-Dyn N (1974) Stochastic models in biology. Academic, New York
- Greenwood M (1931) On the statistical measure of infectiousness. *J Hyg* 31:336–351
- Hamer W (1906) The Milroy lectures on epidemic disease in England – the evidence of variability and of persistency of type (Lecture I). *Lancet* 167:569–574
- Hethcote H (1994) A thousand and one epidemic models. In: Levin S (ed) *Frontiers in mathematical biology, Lecture notes in biomathematics*. Springer, Berlin, pp 504–515
- Hethcote H (2000) The mathematics of infectious diseases. *SIAM Rev* 42:599–653
- Ireland J, Mestel B, Norman R (2007) The effect of seasonal host birth rates on disease persistence. *Math Biosci* 206:31–45
- Isham V (2004) Stochastic models for epidemics. Research Report No 263, Department of Statistical Science, University College London
- Jacquez J (1972) *Compartmental analysis in biology and medicine*. Elsevier, Amsterdam
- Keeling M, Rohani P (2008) *Modeling infectious disease in humans and animals*. Princeton University Press, Princeton
- Keener J, Sneyd J (1989) *Mathematical physiology*. Springer, New York
- Kelly H, Peck H, Laurie K, Wu P, Nishiura H, Cowling B (2011) The age-specific cumulative incidence of infection with pandemic influenza H1N1 2009 was similar in various countries prior to vaccination. *PLoS ONE* 6:e21 828
- Kermack W, McKendrick A (1927) A contribution to the mathematical theory of epidemics. *Proc R Soc London Ser A* 115:700–721
- Kimura M (1964) Diffusion models in population genetics. *J Appl Probab* 1:177–232
- Kramers H (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 7:284–304
- Laidler K (1993) *The world of physical chemistry*. Oxford University Press, New York
- Lande R, Engen S, Sæther B (2003) *Stochastic population dynamics in ecology and conservation*. Oxford University Press, New York
- Le Novère N, Shimizu T (2001) StochSim: modelling of stochastic biomolecular processes. *Bioinformatics* 17:575–576
- Lloyd-Smith J, Schreiber S, Kopp P, Getz W (2005) Superspreading and the effect of individual variation on disease emergence. *Nature* 438:355–359
- Manninen T, Linne ML, Ruohonen K (2006) Developing Itô stochastic differential equation models for neuronal signal transduction pathways. *Comput Biol Chem* 30:280–291
- McKendrick A (1926) Application of mathematics to medical problems. *Proc Edinb Math Soc* 44:98–130
- McNeill W (1976) *Plagues and people*. Anchor, New York
- McQuarrie D (1967) Stochastic approach to chemical kinetics. *J Appl Probab* 4:413–478
- Morton R, Wickwire K (1974) On the optimal control of a deterministic epidemic. *Adv Appl Probab* 6:622–635
- Neal P (2007) Coupling of two SIR epidemic models with variable susceptibilities and infectivities. *J Appl Probab* 44:41–57
- Oldstone M (2010) *Viruses, plagues, and history: past, present and future*. Oxford University Press, Oxford/New York
- Rao C, Wolf D, Arkin A (2002) Control, exploitation and tolerance of intracellular noise. *Nature* 420:231–237
- Renshaw E (1991) *Modelling biological populations in space and time*. Cambridge University Press, Cambridge
- Riley S (2007) Large-scale spatial-transmission models of infectious disease. *Science* 316:1298–1301

- Ross R (1915) Some a priori pathometric equations. *Br Med J* 1:546–547
- Rushton S, Mautner A (1955) The deterministic model of a simple epidemic for more than one community. *Biometrika* 42:126–132
- Sattenspiel L (1987) Population structure and the spread of disease. *Hum Biol* 59:411–438
- Sattenspiel L, Dietz K (1995) A structured epidemic model incorporating geographic mobility among regions. *Math Biosci* 128:71–91
- Sherman I (2006) *The power of plagues*. ASM Press, Washington, DC
- Smallman-Raynor M, Cliff A (2004) Impact of infectious diseases on war. *Infect Dis Clin N Am* 18:341–368
- Sveiczzer A, Tyson J, Novak B (2001) A stochastic, molecular model of the fission yeast cell cycle: role of the nucleocytoplasmic ratio in cycle time regulation. *Biophys Chem* 92:1–15
- Tian T, Xu S, Gao J, Burrage K (2007) Simulated maximum likelihood method for estimating kinetic rates in gene expression. *Bioinformatics* 23:84–91
- UNAIDS (2009) AIDS epidemic update: November 2009. WHO library cataloguing-in-publication data. Available at <http://www.unaids.org>
- Vasold M (2008) *Grippe, Pest und Cholera: eine Geschichte der Seuchen in Europa*. Franz Steiner Verlag, Stuttgart
- Whittle P (1955) The outcome of a stochastic epidemic – a note on Bailey’s paper. *Biometrika* 42:116–122
- WHO (2010) Pandemic (H1N1) 2009 – update 112 (from 6 August 2010). Available at http://www.who.int/csr/don/2010_08_06/en/index.html
- Wilkinson D (2006) *Stochastic modelling for systems biology*. Chapman and Hall, Boca Raton
- Williams T (1971) An algebraic proof of the threshold theorem for the general stochastic epidemic. *Adv Appl Probab* 3:223
- Zheng Q, Ross J (1991) Comparison of deterministic and stochastic kinetics for nonlinear systems. *J Chem Phys* 94:3644–3648



<http://www.springer.com/978-3-642-25968-5>

Inference for Diffusion Processes
With Applications in Life Sciences

Fuchs, C.

2013, XIX, 430 p., Hardcover

ISBN: 978-3-642-25968-5