Compartmental Models in Epidemiology

Chapter *in* Lecture Notes in Mathematics -Springer-verlag- · April 2008

DOI: 10.1007/978-3-540-78911-6_2

CITATIONS
121

READS
4,503

1 author:

Fred Brauer
University of British Columbia - Vancouver
122 PUBLICATIONS **3,405** CITATIONS

SEE PROFILE

Chapter 2
Compartmental Models in Epidemiology

Fred Brauer

Abstract We describe and analyze compartmental models for disease transmission. We begin with models for epidemics, showing how to calculate the basic reproduction number and the final size of the epidemic. We also study models with multiple compartments, including treatment or isolation of infectives. We then consider models including births and deaths in which there may be an endemic equilibrium and study the asymptotic stability of equilibria. We conclude by studying age of infection models which give a unifying framework for more complicated compartmental models.

2.1 Introduction

Communicable diseases such as measles, influenza, or tuberculosis, are a fact of modern life. The mechanism of transmission of infections is now known for most diseases. Generally, diseases transmitted by viral agents, such as influenza, measles, rubella (German measles), and chicken pox, confer immunity against reinfection, while diseases transmitted by bacteria, such as tuberculosis, meningitis, and gonorrhea, confer no immunity against reinfection. Other diseases, such as malaria, are transmitted not directly from human to human but by vectors, which are agents (usually insects) who are infected by humans and who then transmit the disease to humans. The West Nile virus involves two vectors, mosquitoes and birds. For sexually transmitted diseases with heterosexual transmission each sex acts as a vector and disease is transmitted back and forth between the sexes.

Department of Mathematics, University of British Columbia, 1984, Mathematics Road, Vancouver BC, Canada V6T 1Z2 brauer@math.ubc.ca

We will be concerned both with epidemics which are sudden outbreaks of a disease, and endemic situations, in which a disease is always present. Epidemics such as the 2002 outbreak of SARS, the Ebola virus and avian flu outbreaks are events of concern and interest to many people. The 1918 Spanish flu epidemic caused millions of deaths, and a recurrence of a major influenza epidemic is a dangerous possibility. An introduction of smallpox is of considerable concern to government officials dealing with terrorism threats.

An endemic situation is one in which a disease is always present. The prevalence and effects of many diseases in less developed countries are probably less well-known but may be of even more importance. Every year millions of people die of measles, respiratory infections, diarrhea and other diseases that are easily treated and not considered dangerous in the Western world. Diseases such as malaria, typhus, cholera, schistosomiasis, and sleeping sickness are endemic in many parts of the world. The effects of high disease mortality on mean life span and of disease debilitation and mortality on the economy in afflicted countries are considerable.

Our goal is to provide an introduction to mathematical epidemiology, including the development of mathematical models for the spread of disease as well as tools for their analysis. Scientific experiments usually are designed to obtain information and to test hypotheses. Experiments in epidemiology with controls are often difficult or impossible to design and even if it is possible to arrange an experiment there are serious ethical questions involved in withholding treatment from a control group. Sometimes data may be obtained after the fact from reports of epidemics or of endemic disease levels, but the data may be incomplete or inaccurate. In addition, data may contain enough irregularities to raise serious questions of interpretation, such as whether there is evidence of chaotic behaviour [12]. Hence, parameter estimation and model fitting are very difficult. These issues raise the question of whether mathematical modeling in epidemiology is of value.

Our response is that mathematical modeling in epidemiology provides understanding of the underlying mechanisms that influence the spread of disease and, in the process, it suggests control strategies. In fact, models often identify behaviours that are unclear in experimental data – often because data are non-reproducible and the number of data points is limited and subject to errors in measurement. For example, one of the fundamental results in mathematical epidemiology is that *most* mathematical epidemic models, including those that include a high degree of heterogeneity, usually exhibit "threshold" behaviour. In epidemiological terms this can be stated as follows: *If the average number of secondary infections caused by an average infective, called the basic reproduction number, is less than one a disease will die out, while if it exceeds one there will be an epidemic.* This broad principle, consistent with observations and quantified via epidemiological models, has been consistently used to estimate the effectiveness of vaccination policies and the likelihood that a disease may be eliminated or eradicated. Hence, even if it is not possible to verify hypotheses accurately, agreement with hypotheses of

a qualitative nature is often valuable. Expressions for the basic reproduction number for HIV in various populations have been used to test the possible effectiveness of vaccines that may provide temporary protection by reducing either HIV-infectiousness or susceptibility to HIV. Models are used to estimate how widespread a vaccination plan must be to prevent or reduce the spread of HIV.

In the mathematical modeling of disease transmission, as in most other areas of mathematical modeling, there is always a trade-off between simple models, which omit most details and are designed only to highlight general qualitative behaviour, and detailed models, usually designed for specific situations including short-term quantitative predictions. Detailed models are generally difficult or impossible to solve analytically and hence their usefulness for theoretical purposes is limited, although their strategic value may be high. In these notes we describe simple models in order to establish broad principles. Furthermore, these simple models have additional value as they are the building blocks of models that include more detailed structure.

Many of the early developments in the mathematical modeling of communicable diseases are due to public health physicians. The first known result in mathematical epidemiology is a defense of the practice of inoculation against smallpox in 1760 by Daniel Bernoulli, a member of a famous family of mathematicians (eight spread over three generations) who had been trained as a physician. The first contributions to modern mathematical epidemiology are due to P.D. En'ko between 1873 and 1894 [11], and the foundations of the entire approach to epidemiology based on compartmental models were laid by public health physicians such as Sir Ross, R.A., W.H. Hamer, A.G. McKendrick and W.O. Kermack between 1900 and 1935, along with important contributions from a statistical perspective by J. Brownlee. A particularly instructive example is the work of Ross on malaria. Dr. Ross was awarded the second Nobel Prize in Medicine for his demonstration of the dynamics of the transmission of malaria between mosquitoes and humans. Although his work received immediate acceptance in the medical community, his conclusion that malaria could be controlled by controlling mosquitoes was dismissed on the grounds that it would be impossible to rid a region of mosquitoes completely and that in any case mosquitoes would soon reinvade the region. After Ross formulated a mathematical model that predicted that malaria outbreaks could be avoided if the mosquito population could be reduced below a critical threshold level, field trials supported his conclusions and led to sometimes brilliant successes in malaria control. However, the Garki project provides a dramatic counterexample. This project worked to eradicate malaria from a region temporarily. However, people who have recovered from an attack of malaria have a temporary immunity against reinfection. Thus elimination of malaria from a region leaves the inhabitants of this region without immunity when the campaign ends, and the result can be a serious outbreak of malaria.

We will begin with an introduction to epidemic models. Next, we will incorporate demographic effects into the models to explore endemic states,

and finally we will describe models with infectivity depending on the age of infection. Our approach will be *qualitative*. By this we mean that rather than attempting to find explicit solutions of the systems of differential equations which will form our models we will be concerned with the asymptotic behaviour, that is, the behaviour as $t \to \infty$ of solutions.

This material is meant to be an introduction to the study of compartmental models in mathematical epidemiology. More advanced material may be found in many other sources, including Chaps. 5–9 of this volume, the case studies in Chaps. 11–14, and [2, 4–6, 9, 17, 29, 35].

2.1.1 Simple Epidemic Models

An epidemic, which acts on a short temporal scale, may be described as a sudden outbreak of a disease that infects a substantial portion of the population in a region before it disappears. Epidemics usually leave many members untouched. Often these attacks recur with intervals of several years between outbreaks, possibly diminishing in severity as populations develop some immunity. This is an important aspect of the connection between epidemics and disease evolution.

The Book of Exodus describes the plagues that Moses brought down upon Egypt, and there are several other biblical descriptions of epidemic outbreaks. Descriptions of epidemics in ancient and medieval times frequently used the term "plague" because of a general belief that epidemics represented divine retribution for sinful living. More recently some have described AIDS as punishment for sinful activities. Such views have often hampered or delayed attempts to control this modern epidemic .

There are many biblical references to diseases as historical influences, such as the decision of Sennacherib, the king of Assyria, to abandon his attempt to capture Jerusalem about 700 BC because of the illness of his soldiers (Isaiah 37, 36–38). The fall of empires has been attributed directly or indirectly to epidemic diseases. In the second century AD the so-called Antonine plagues (possibly measles and smallpox) invaded the Roman Empire, causing drastic population reductions and economic hardships. These led to disintegration of the empire because of disorganization, which facilitated invasions of barbarians. The Han empire in China collapsed in the third century AD after a very similar sequence of events. The defeat of a population of millions of Aztecs by Cortez and his 600 followers can be explained in part by a smallpox epidemic that devastated the Aztecs but had almost no effect on the invading Spaniards thanks to their built-in immunities. The Aztecs were not only weakened by disease but also confounded by what they interpreted as a divine force favoring the invaders. Smallpox then spread southward to the Incas in Peru and was an important factor in the success of Pizarro's invasion a few years later. Smallpox was followed by other diseases such as measles

and diphtheria imported from Europe to North America. In some regions, the indigenous populations were reduced to one tenth of their previous levels by these diseases. Between 1519 and 1530 the Indian population of Mexico was reduced from 30 million to 3 million.

The Black Death spread from Asia throughout Europe in several waves during the fourteenth century, beginning in 1346, and is estimated to have caused the death of as much as one third of the population of Europe between 1346 and 1350. The disease recurred regularly in various parts of Europe for more than 300 years, notably as the Great Plague of London of 1665–1666. It then gradually withdrew from Europe. As the plague struck some regions harshly while avoiding others, it had a profound effect on political and economic developments in medieval times. In the last bubonic plague epidemic in France (1720–1722), half the population of Marseilles, 60% of the population in nearby Toulon, 44% of the population of Arles and 30% of the population of Aix and Avignon died, but the epidemic did not spread beyond Provence.

The historian W.H. McNeill argues, especially in his book [26], that the spread of communicable diseases has frequently been an important influence in history. For example, there was a sharp population increase throughout the world in the eighteenth century; the population of China increased from 150 million in 1760 to 313 million in 1794 and the population of Europe increased from 118 million in 1700 to 187 million in 1800. There were many factors involved in this increase, including changes in marriage age and technological improvements leading to increased food supplies, but these factors are not sufficient to explain the increase. Demographic studies indicate that a satisfactory explanation requires recognition of a decrease in the mortality caused by periodic epidemic infections. This decrease came about partly through improvements in medicine, but a more important influence was probably the fact that more people developed immunities against infection as increased travel intensified the circulation and co-circulation of diseases.

Perhaps the first epidemic to be examined from a modeling point of view was the Great Plague in London (1665–1666). The plague was one of a sequence of attacks beginning in the year 1346 of what came to be known as the Black Death. It is now identified as the bubonic plague, which had actually invaded Europe as early as the sixth century during the reign of the Emperor Justinian of the Roman Empire and continued for more than three centuries after the Black Death. The Great Plague killed about one sixth of the population of London. One of the few "benefits" of the plague was that it caused Cambridge University to be closed for two years. Isaac Newton, who was a student at Cambridge at the time, was sent to his home and while "in exile" he had one of the most productive scientific periods of any human in history. He discovered his law of gravitation, among other things, during this period.

The characteristic features of the Great Plague were that it appeared quite suddenly, grew in intensity, and then disappeared, leaving part of the

population untouched. The same features have been observed in many other epidemics, both of fatal diseases and of diseases whose victims recovered with immunity against reinfection.

In the nineteenth century recurrent invasions of cholera killed millions in India. The influenza epidemic of 1918–1919 killed more than 20 million people overall, more than half a million in the United States. One of the questions that first attracted the attention of scientists interested in the study of the spread of communicable diseases was why diseases would suddenly develop in a community and then disappear just as suddenly without infecting the entire community. One of the early triumphs of mathematical epidemiology [21] was the formulation of a simple model that predicted behaviour very similar to the behaviour observed in countless epidemics. The Kermack–McKendrick model is a compartmental model based on relatively simple assumptions on the rates of flow between different classes of members of the population.

There are many questions of interest to public health physicians confronted with a possible epidemic. For example, how severe will an epidemic be? This question may be interpreted in a variety of ways. For example, how many individuals will be affected altogether and thus require treatment? What is the maximum number of people needing care at any particular time? How long will the epidemic last? How much good would quarantine or isolation of victims do in reducing the severity of the epidemic? These are some of the questions we would like to study with the aid of models.

2.1.2 The Kermack–McKendrick Model

We formulate our descriptions as *compartmental models*, with the population under study being divided into compartments and with assumptions about the nature and time rate of transfer from one compartment to another. Diseases that confer immunity have a different compartmental structure from diseases without immunity. We will use the terminology SIR to describe a disease which confers immunity against re-infection, to indicate that the passage of individuals is from the susceptible class S to the infective class I to the removed class R. On the other hand, we will use the terminology SIS to describe a disease with no immunity against re-infection, to indicate that the passage of individuals is from the susceptible class to the infective class and then back to the susceptible class. Other possibilities include $SEIR$ and $SEIS$ models, with an exposed period between being infected and becoming infective, and $SIRS$ models, with temporary immunity on recovery from infection.

The independent variable in our compartmental models is the time t and the rates of transfer between compartments are expressed mathematically as derivatives with respect to time of the sizes of the compartments, and as a result our models are formulated initially as *differential equations*. Possible

generalizations, which we shall not explore in these notes, include models in which the rates of transfer depend on the sizes of compartments over the past as well as at the instant of transfer, leading to more general types of functional equations, such as differential-difference equations, integral equations, or integro-differential equations.

In order to model such an epidemic we divide the population being studied into three classes labeled S, I, and R. We let $S(t)$ denote the number of individuals who are susceptible to the disease, that is, who are not (yet) infected at time t. $I(t)$ denotes the number of infected individuals, assumed infectious and able to spread the disease by contact with susceptibles. $R(t)$ denotes the number of individuals who have been infected and then removed from the possibility of being infected again or of spreading infection. Removal is carried out either through isolation from the rest of the population or through immunization against infection or through recovery from the disease with full immunity against reinfection or through death caused by the disease. These characterizations of removed members are different from an epidemiological perspective but are often equivalent from a modeling point of view which takes into account only the state of an individual with respect to the disease.

In formulating models in terms of the derivatives of the sizes of each compartment we are assuming that the number of members in a compartment is a differentiable function of time. This may be a reasonable approximation if there are many members in a compartment, but it is certainly suspect otherwise. In formulating models as differential equations, we are assuming that the epidemic process is *deterministic*, that is, that the behaviour of a population is determined completely by its history and by the rules which describe the model. In other chapters of this volume Linda Allen and Ping Yan describe the study of *stochastic* models in which probabilistic concepts are used and in which there is a distribution of possible behaviours. The developing study of network science, introduced in Chap. 4 of this volume and described in [28, 30, 33], is another approach.

The basic compartmental models to describe the transmission of communicable diseases are contained in a sequence of three papers by W.O. Kermack and A.G. McKendrick in 1927, 1932, and 1933 [21–23]. The first of these papers described epidemic models. What is often called the Kermack–McKendrick epidemic model is actually a special case of the general model introduced in this paper. The general model included dependence on age of infection, that is, the time since becoming infected. Curiously, Kermack and McKendrick did not explore this situation further in their later models which included demographic effects. Age of infection models have become important in the study of HIV/AIDS, and we will return to them in the last section of this chapter.

The special case of the model proposed by Kermack and McKendrick in 1927 which is the starting point for our study of epidemic models is

$$S' = -\beta SI$$
$$I' = \beta SI - \alpha I$$
$$R' = \alpha I \ .$$

A flow chart is shown in Fig. 2.1. It is based on the following assumptions:

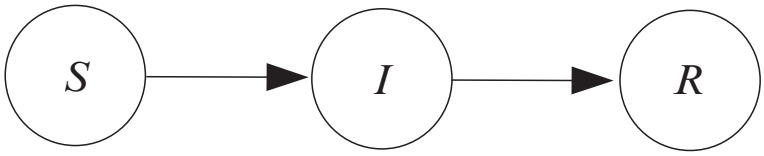

Fig. 2.1 Flow chart for the SIR model

(1) An average member of the population makes contact sufficient to transmit infection with βN others per unit time, where N represents total population size (mass action incidence) .
(2) Infectives leave the infective class at rate αI per unit time.
(3) There is no entry into or departure from the population, except possibly through death from the disease.

According to (1), since the probability that a random contact by an infective is with a susceptible, who can then transmit infection, is S/N, the number of new infections in unit time per infective is $(\beta N)(S/N)$, giving a rate of new infections $(\beta N)(S/N)I = \beta SI$. Alternately, we may argue that for a contact by a susceptible the probability that this contact is with an infective is I/N and thus the rate of new infections per susceptible is $(\beta N)(I/N)$, giving a rate of new infections $(\beta N)(I/N)S = \beta SI$. Note that both approaches give the same rate of new infections; there are situations which we shall encounter where one is more appropriate than the other. We need not give an algebraic expression for N since it cancels out of the final model, but we should note that for a disease that is fatal to all who are infected $N = S+I$; while, for a disease from which all infected members recover with immunity, $N = S+I+R$. Later, we will allow the possibility that some infectives recover while others die of the disease. The hypothesis (3) really says that the time scale of the disease is much faster than the time scale of births and deaths so that demographic effects on the population may be ignored. An alternative view is that we are only interested in studying the dynamics of a single epidemic outbreak. In later sections we shall consider models that are the same as those considered in this first section except for the incorporation of demographic effects (births and deaths) along with the corresponding epidemiological assumptions.

The assumption (2) requires a fuller mathematical explanation, since the assumption of a recovery rate proportional to the number of infectives has no clear epidemiological meaning. We consider the "cohort" of members who were all infected at one time and let $u(s)$ denote the number of these who are still infective s time units after having been infected. If a fraction α of these leave the infective class in unit time then

$$u' = -\alpha u \ ,$$

and the solution of this elementary differential equation is

$$u(s) = u(0)\, e^{-\alpha s} \ .$$

Thus, the fraction of infectives remaining infective s time units after having become infective is $e^{-\alpha s}$, so that the length of the infective period is distributed exponentially with mean $\int_0^\infty e^{-\alpha s} ds = 1/\alpha$, and this is what (2) really assumes.

The assumptions of a rate of contacts proportional to population size N with constant of proportionality β, and of an exponentially distributed recovery rate are unrealistically simple. More general models can be constructed and analyzed, but our goal here is to show what may be deduced from extremely simple models. It will turn out that many more realistic models exhibit very similar qualitative behaviours.

In our model R is determined once S and I are known, and we can drop the R equation from our model, leaving the system of two equations

$$\begin{aligned} S' &= -\beta SI \\ I' &= (\beta S - \alpha) I \ . \end{aligned} \tag{2.1}$$

We are unable to solve this system analytically but we learn a great deal about the behaviour of its solutions by the following qualitative approach. To begin, we remark that the model makes sense only so long as $S(t)$ and $I(t)$ remain non-negative. Thus if either $S(t)$ or $I(t)$ reaches zero we consider the system to have terminated. We observe that $S' < 0$ for all t and $I' > 0$ if and only if $S > \alpha/\beta$. Thus I increases so long as $S > \alpha/\beta$ but since S decreases for all t, I ultimately decreases and approaches zero. If $S(0) < \alpha/\beta, I$ decreases to zero (no epidemic), while if $S(0) > \alpha/\beta, I$ first increases to a maximum attained when $S = \alpha/\beta$ and then decreases to zero (epidemic). We think of introducing a small number of infectives into a population of susceptibles and ask whether there will be an epidemic. The quantity $\beta S(0)/\alpha$ is a threshold quantity, called the *basic reproduction number* and denoted by $\mathcal{R}_0$, which determines whether there is an epidemic or not. If $\mathcal{R}_0 < 1$ the infection dies out, while if $\mathcal{R}_0 > 1$ there is an epidemic.

The definition of the basic reproduction number $\mathcal{R}_0$ is that the basic reproduction number is the number of secondary infections caused by a single infective introduced into a wholly susceptible population of size $K \approx S(0)$

over the course of the infection of this single infective. In this situation, an infective makes βK contacts in unit time, all of which are with susceptibles and thus produce new infections, and the mean infective period is $1/\alpha$; thus the basic reproduction number is actually $\beta K/\alpha$ rather than $\beta S(0)/\alpha$.

Instead of trying to solve for S and I as functions of t, we divide the two equations of the model to give

$$\frac{I'}{S'} = \frac{dI}{dS} = \frac{(\beta S - \alpha)I}{-\beta SI} = -1 + \frac{\alpha}{\beta S} ,$$

and integrate to find the orbits (curves in the (S, I)-plane, or phase plane)

$$I = -S + \frac{\alpha}{\beta} \log S + c , \tag{2.2}$$

with c an arbitrary constant of integration. Here, we are using log to denote the natural logarithm. Another way to describe the orbits is to define the function

$$V(S, I) = S + I - \frac{\alpha}{\beta} \log S$$

and note that each orbit is a curve given implicitly by the equation $V(S, I) = c$ for some choice of the constant c. The constant c is determined by the initial values $S(0)$, $I(0)$ of S and I, respectively, because $c = V(S(0), I(0)) = S(0)+I(0)-\alpha \log S(0)/\beta$. Note that the maximum value of I on each of these orbits is attained when $S = \alpha/\beta$. Note also that since none of these orbits reaches the I - axis, $S > 0$ for all times. In particular, $S_\infty = \lim_{t\to\infty} S(t) > 0$, which implies that part of the population escapes infection.

Let us think of a population of size K into which a small number of infectives is introduced, so that $S_0 \approx K$, $I_0 \approx 0$, and $\mathcal{R}_0 = \beta K/\alpha$. If we use the fact that $\lim_{t\to\infty} I(t) = 0$, and let $S_\infty = \lim_{t\to\infty} S(t)$, then the relation $V(S_0, I_0) = V(S_\infty, 0)$ gives

$$K - \frac{\alpha}{\beta} \log S_0 = S_\infty - \frac{\alpha}{\beta} \log S_\infty ,$$

from which we obtain an expression for β/α in terms of the measurable quantities S_0 and S_∞, namely

$$\frac{\beta}{\alpha} = \frac{(\log S_0 - \log S_\infty)}{K - S_\infty} .$$

We may rewrite this in terms of $\mathcal{R}_0$ as the *final size relation*

$$\log S_0 - \log S_\infty = \mathcal{R}_0 \left[1 - \frac{S_\infty}{K}\right] . \tag{2.3}$$

In particular, since the right side of (2.3) is finite, the left side is also finite, and this shows that $S_\infty > 0$.

It is generally difficult to estimate the contact rate β which depends on the particular disease being studied but may also depend on social and behavioural factors. The quantities S_0 and S_∞ may be estimated by serological studies (measurements of immune responses in blood samples) before and after an epidemic, and from these data the basic reproduction number $\mathcal{R}_0$ may be estimated by using (2.3). This estimate, however, is a retrospective one which can be determined only after the epidemic has run its course.

Initially, the number of infectives grows exponentially because the equation for I may be approximated by

$$I' = (\beta K - \alpha)I$$

and the initial growth rate is

$$r = \beta K - \alpha = \alpha(\mathcal{R}_0 - 1) \ .$$

This initial growth rate r may be determined experimentally when an epidemic begins. Then since K and α may be measured β may be calculated as

$$\beta = \frac{r + \alpha}{K} \ .$$

However, because of incomplete data and under-reporting of cases this estimate may not be very accurate. This inaccuracy is even more pronounced for an outbreak of a previously unknown disease, where early cases are likely to be misdiagnosed.

The maximum number of infectives at any time is the number of infectives when the derivative of I is zero, that is, when $S = \alpha/\beta$. This maximum is given by

$$I_{max} = S_0 + I_0 - \frac{\alpha}{\beta} \log S_0 - \frac{\alpha}{\beta} + \frac{\alpha}{\beta} \log \frac{\alpha}{\beta} \ , \tag{2.4}$$

obtained by substituting $S = \alpha/\beta$, $I = I_{max}$ into (2.2).

Example. (The Great Plague in Eyam) The village of Eyam near Sheffield, England suffered an outbreak of bubonic plague in 1665–1666 the source of which is generally believed to be the Great Plague of London. The Eyam plague was survived by only 83 of an initial population of 350 persons. As detailed records were preserved and as the community was persuaded to quarantine itself to try to prevent the spread of disease to other communities, the disease in Eyam has been used as a case study for modeling [31]. Detailed examination of the data indicates that there were actually two outbreaks of which the first was relatively mild. Thus we shall try to fit the model (2.1) over the period from mid-May to mid-October 1666, measuring time in months with an initial population of seven infectives and 254 susceptibles, and a final population of 83. Values of susceptibles and infectives in Eyam are given in [31] for various dates, beginning with $S(0) = 254, I(0) = 7$, shown in Table 2.1.

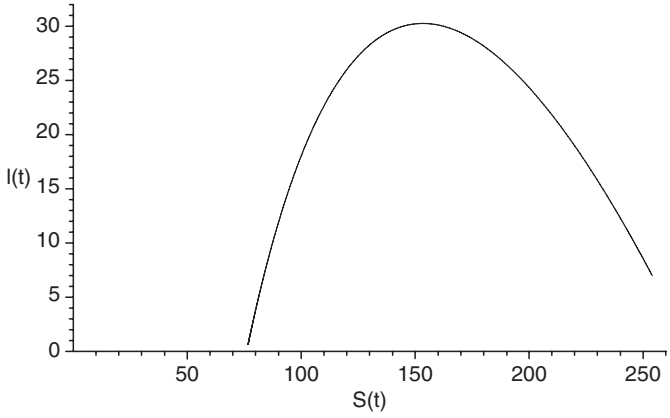

Fig. 2.2 The S–I plane

Table 2.1 Eyam Plague data

Date (1666)	Susceptibles	Infectives
July 3/4	235	14.5
July 19	201	22
August 3/4	153.5	29
August 19	121	21
September 3/4	108	8
September 19	97	8
October 4/5	Unknown	Unknown
October 20	83	0

The relation (2.3) with $S_0 = 254$, $I_0 = 7$, $S_\infty = 83$ gives $\beta/\alpha = 6.54 \times 10^{-3}$, $\alpha/\beta = 153$. The infective period was 11 days, or 0.3667 month, so that $\alpha = 2.73$. Then $\beta = 0.0178$. The relation (2.4) gives an estimate of 30.4 for the maximum number of infectives. We use the values obtained here for the parameters β and α in the model (2.1) for simulations of both the phase plane, the (S, I)-plane, and for graphs of S and I as functions of t (Figs. 2.2, 2.3, and 2.4). Figure 2.5 plots these data points together with the phase portrait given in Fig. 2.2 for the model (2.1).

The actual data for the Eyam epidemic are remarkably close to the predictions of this very simple model. However, the model is really too good to be true. Our model assumes that infection is transmitted directly between people. While this is possible, bubonic plague is transmitted mainly by rat fleas. When an infected rat is bitten by a flea, the flea becomes extremely hungry and bites the host rat repeatedly, spreading the infection in the rat. When the host rat dies its fleas move on to other rats, spreading the disease further. As the number of available rats decreases the fleas move to human hosts, and this is how plague starts in a human population (although the second phase of the epidemic may have been the pneumonic form of bubonic plague, which

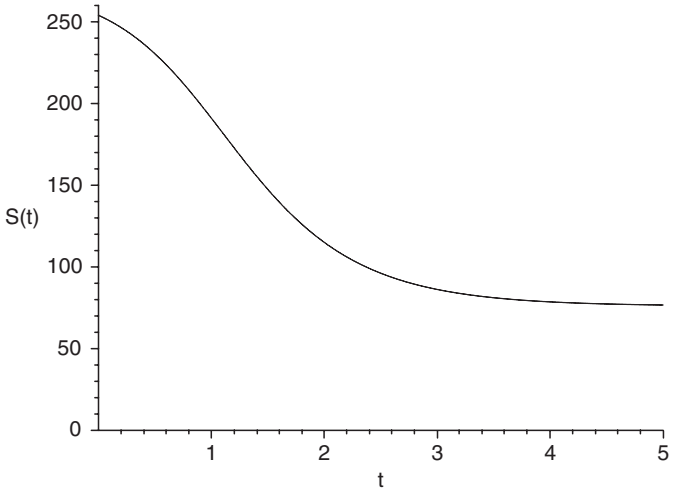

Fig. 2.3 S as a function of t

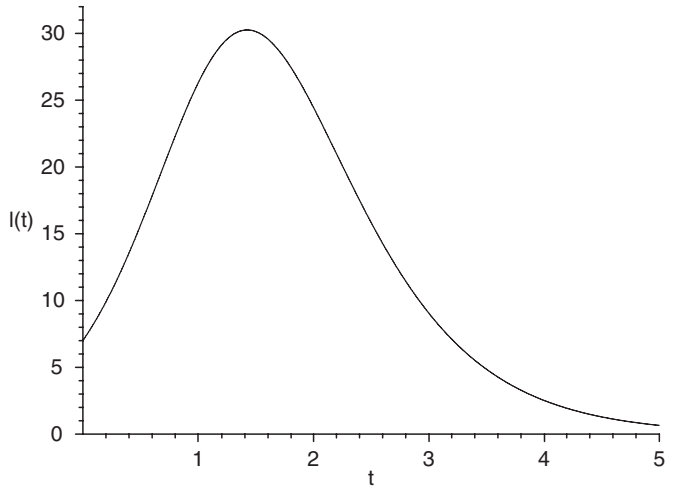

Fig. 2.4 I as a function of t

can be spread from person to person). One of the main reasons for the spread of plague from Asia into Europe was the passage of many trading ships; in medieval times ships were invariably infested with rats. An accurate model of plague transmission would have to include flea and rat populations, as well as movement in space. Such a model would be extremely complicated and its predictions might well not be any closer to observations than our simple unrealistic model. In [31] a stochastic model was also used to fit the data, but the fit was rather poorer than the fit for the simple deterministic model (2.1).

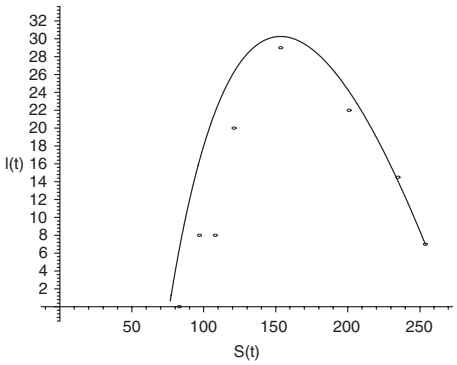

Fig. 2.5 The S–I plane, model and data

In the village of Eyam the rector persuaded the entire community to quarantine itself to prevent the spread of disease to other communities. This policy actually increased the infection rate in the village by keeping fleas, rats, and people in close contact with one another, and the mortality rate from bubonic plague was much higher in Eyam than in London. Further, the quarantine could do nothing to prevent the travel of rats and thus did little to prevent the spread of disease to other communities. One message this suggests to mathematical modelers is that control strategies based on false models may be harmful, and it is essential to distinguish between assumptions that simplify but do not alter the predicted effects substantially, and wrong assumptions which make an important difference.

2.1.3 Kermack–McKendrick Models with General Contact Rates

The assumption in the model (2.1) of a rate of contacts per infective which is proportional to population size N, called *mass action incidence* or bilinear incidence, was used in all the early epidemic models. However, it is quite unrealistic, except possibly in the early stages of an epidemic in a population of moderate size. It is more realistic to assume a contact rate which is a non-increasing function of total population size. For example, a situation in which the number of contacts per infective in unit time is constant, called *standard incidence*, is a more accurate description for sexually transmitted diseases.

We generalize the model (2.1) by replacing the assumption (1) by the assumption that an average member of the population makes $C(N)$ contacts in unit time with $C'(N) \geq 0$ [7, 10], and we define

$$\beta(N) = \frac{C(N)}{N} .$$

It is reasonable to assume $\beta'(N) \leq 0$ to express the idea of saturation in the number of contacts. Then mass action incidence corresponds to the choice $C(N) = \beta N$, and standard incidence corresponds to the choice $C(N) = \lambda$. The assumptions $C(N) = N\beta(N), C'(N) \geq 0$ imply that

$$\beta(N) + N\beta'(N) \geq 0 . \tag{2.5}$$

Some epidemic models [10] have used a Michaelis–Menten type of interaction of the form

$$C(N) = \frac{aN}{1 + bN} .$$

Another form based on a mechanistic derivation for pair formation [14] leads to an expression of the form

$$C(N) = \frac{aN}{1 + bN + \sqrt{1 + 2bN}} .$$

Data for diseases transmitted by contact in cities of moderate size [25] suggests that data fits the assumption of a form

$$C(N) = \lambda N^a$$

with $a = 0.05$ quite well. All of these forms satisfy the conditions $C'(N) \geq 0, \beta'(N) \leq 0$.

Because the total population size is now present in the model we must include an equation for total population size in the model. This forces us to make a distinction between members of the population who die of the disease and members of the population who recover with immunity against reinfection. We assume that a fraction f of the αI members leaving the infective class at time t recover and the remaining fraction $(1 - f)$ die of disease. We use S, I, and N as variables, with $N = S + I + R$. We now obtain a three-dimensional model

$$\begin{aligned} S' &= -\beta(N)SI \\ I' &= \beta(N)SI - \alpha I \\ N' &= -(1 - f)\alpha I . \end{aligned} \tag{2.6}$$

We also have the equation $R' = -f\alpha I$, but we need not include it in the model since R is determined when S, I, and N are known. We should note

that if $f = 1$ the total population size remains equal to the constant K, and the model (2.6) reduces to the simpler model (2.1) with β replaced by the constant $\beta(K)$.

We wish to show that the model (2.6) has the same qualitative behaviour as the model (2.1), namely that there is a basic reproduction number which distinguishes between disappearance of the disease and an epidemic outbreak, and that some members of the population are left untouched when the epidemic passes. These two properties are the central features of all epidemic models.

For the model (2.6) the basic reproduction number is given by

$$\mathcal{R}_0 = \frac{K\beta(K)}{\alpha}$$

because a single infective introduced into a wholly susceptible population makes $C(K) = K\beta(K)$ contacts in unit time, all of which are with susceptibles and thus produce new infections, and the mean infective period is $1/\alpha$. In addition to the basic reproduction number $\mathcal{R}_0$ there is also a time-dependent running reproduction number which we call $\mathcal{R}^*$, representing the number of secondary infections caused by a single individual in the population who becomes infective at time t. In this situation, an infective makes $C(N) = N\beta(N)$ contacts in unit time and a fraction S/N are with susceptibles and thus produce new infections. Thus it is easy to see that for the model (2.6) the running reproduction number is given by

$$\mathcal{R}^* = \frac{S\beta(N)}{\alpha}.$$

If $\mathcal{R}^* < 1$ for all large t, the epidemic will pass. We may calculate the rate of change of the running reproduction number with respect to time, using (2.6) and (2.5) to find that

$$\begin{aligned}\frac{d}{dt}\mathcal{R}^* &= \frac{S'(t)\beta(N) + S(t)\beta'(N)N'(t)}{\alpha} = \frac{(-\beta(N))^2 SI - S\alpha(1-f)\beta'(N)}{\alpha}. \\ &\leq \frac{\beta(N)SI}{\alpha} \cdot \left[\beta(N) - \frac{(1-f)\alpha}{N}\right].\end{aligned}$$

Thus $\frac{d}{dt}\mathcal{R}^* < 0$ if $N\beta(N) > \alpha(1-f)$, or $\mathcal{R}^* > (1-f)S/N$. This means that $\mathcal{R}^*$ decreases whenever $\mathcal{R}^* > 1$. Thus if $\mathcal{R}^* < 1$ for $t = T$ then $\mathcal{R}^* < 1$ for $t > T$. If $\mathcal{R}_0 > 1$ then $I'(0) = \alpha(\mathcal{R}_0 - 1)I(0) > 0$, and an epidemic begins. However, $\mathcal{R}^*$ decreases until it is less than 1 and then remains less than 1. Thus the epidemic will pass. If $\mathcal{R}_0 < 1$ then $I'(0) = \alpha(\mathcal{R}_0 - 1)I(0) < 0, \mathcal{R}^* < 1$ for all t, and there is no epidemic.

From (2.6) we obtain

$$S' + I' = -\alpha I$$
$$N' = -\alpha(1-f)I\ .$$

Integration of these equations from 0 to t gives

$$S(t) + I(t) - S(0) - I(0) = -\alpha \int_0^t I(s)ds \tag{2.7}$$
$$N(t) - N(0) = -\alpha(1-f)\int_0^t I(s)ds\ .$$

When we combine these two equations, eliminating the integral expression, and use $N(0) = S(0) + I(0) = K$, we obtain

$$K - N(t) = (1-f)[K - S(t) - I(t)]\ .$$

If we let $t \to \infty$, $S(t)$ and $N(t)$ decrease monotonically to limits S_∞ and N_∞ respectively and $I(t) \to 0$. This gives the relation

$$K - N_\infty = (1-f)[K - S_\infty]\ . \tag{2.8}$$

In this equation, $K - N_\infty$ is the change in population size, which is the number of disease deaths over the course of the epidemic, while $K - S_\infty$ is the change in the number of susceptibles, which is the number of disease cases over the course of the epidemic. In this model, (2.8) is obvious, but we shall see in a more general setting how to derive an analogous equation from which we can calculate an average disease mortality. Equation (2.8) generalizes to the infection age epidemic model of Kermack and McKendrick.

If we use the same approach as was used for (2.1) to show that $S_\infty > 0$, we obtain

$$\frac{dI}{dS} = -1 + \frac{\alpha}{S\beta(N)}$$

and we are unable to proceed because of the dependence on N. However, we may use a different approach to obtain the desired result. We assume that $\beta(0)$ is finite, thus ruling out standard incidence. If we let $t \to \infty$ in the second equation of (2.7) we obtain

$$\alpha \int_0^\infty I(s)ds = S(0) + I(0) - S_\infty = K - S_\infty.$$

The first equation of (2.6) may be written as

$$-\frac{S'(t)}{S(t)} = \beta(N(t))I(t).$$

Since

$$\beta(N) \le \beta(0),$$

integration from 0 to ∞ gives

$$\begin{aligned}\log\frac{S(0)}{S_\infty} &= \int_0^\infty \beta(N(t))I(t)dt \\ &\le \beta(0)\int_0^\infty I(t)dt \\ &= \frac{\beta(0)(K-S_\infty)}{\alpha K}.\end{aligned}$$

Since the right side of this inequality is finite, the left side is also finite and this establishes that $S_\infty > 0$.

In addition, if we use the same integration together with the inequality

$$\beta(N) \ge \beta(K),$$

we obtain a final size inequality

$$\begin{aligned}\log\frac{S(0)}{S_\infty} &= \int_0^\infty \beta(N(t))I(t)dt \\ &\ge \beta(K)\int_0^\infty I(t)dt = \mathcal{R}_0\left[1-\frac{S_\infty}{K}\right].\end{aligned}$$

If $\beta(N) \to \infty$ as $N \to 0$ we must use a different approach to analyze the limiting behaviour. It is possible to show that $S_\infty = 0$ is possible only if $N \to 0$ and $\int_0^K \beta(N)dN$ diverges, and this is possible only if $f = 0$, that is, only if all infectives ide of disease. The assumption that $\beta(N)$ is unbounded as $N \to 0$ is biologically unreasonable. In particular, standard incidence is not realistic for small population sizes. A more realistic assumption would be that the number of contacts per infective in unit time is linear for small population size and saturates for larger population sizes, which rules out the possibility that the epidemic sweeps through the entire population.

2.1.4 Exposed Periods

In many infectious diseases there is an exposed period after the transmission of infection from susceptibles to potentially infective members but before these potential infectives can transmit infection. If the exposed period is short it is often neglected in modeling. A longer exposed period could perhaps lead to significantly different model predictions, and we need to show that this is not the case. To incorporate an exponentially distributed exposed period with mean exposed period $1/\kappa$ we add an exposed class E and use compartments S, E, I, R and total population size $N = S+E+I+R$ to give a generalization of the epidemic model (2.6).

$$\begin{aligned} S' &= -\beta(N)SI \\ E' &= \beta(N)SI - \kappa E \\ I' &= \kappa E - \alpha I \\ N' &= -(1-f)\alpha I \ . \end{aligned} \tag{2.9}$$

We also have the equation $R' = -f\alpha I$, but we need not include it in the model since R is determined when S, I, and N are known. A flow chart is shown in Fig. 2.6.

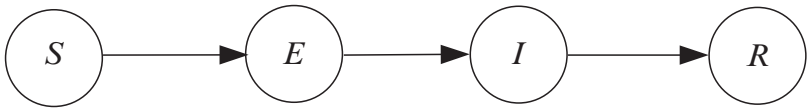

Fig. 2.6 Flow chart for the $SEIR$ model

The analysis of this model is the same as the analysis of (2.6), but with I replaced by $E + I$. That is, instead of using the number of infectives as one of the variables we use the total number of infected members, whether or not they are capable of transmitting infection.

Some diseases have an asymptomatic stage in which there is some infectivity rather than an exposed period. This may be modeled by assuming infectivity reduced by a factor ε_E during an exposed period. A calculation of the rate of new infections per susceptible leads to a model

$$\begin{aligned} S' &= -\beta(N)S(I + \varepsilon_E E) \\ E' &= \beta(N)S(I + \varepsilon_E E) - \kappa E \\ I' &= \kappa E - \alpha I \ . \end{aligned} \tag{2.10}$$

For this model

$$\mathcal{R}_0 = \frac{K\beta(K)}{\alpha} + \varepsilon_E \frac{K\beta(K)}{\kappa} \ .$$

There is a final size relation like (2.3) for the model (2.9). Integration of the sum of the first two equations of (2.9) from 0 to ∞ gives

$$K - S_\infty = \kappa \int_0^\infty E(s)ds$$

and division of the first equation of (2.9) by S followed by integration from 0 to ∞ gives

$$\begin{aligned}\log S_0 - \log S_\infty &= \int_0^\infty \beta(N(s))[I(s) + \epsilon_E E(s)ds \\ &\geq \beta(K) \int_0^\infty [I(s) + \epsilon_E E(s)ds \\ &= \beta(K)[\epsilon_E + \frac{\kappa}{\alpha}] \int_0^\infty E(s)ds \\ &= \mathcal{R}_0 \left[1 - \frac{S_\infty}{K}\right].\end{aligned}$$

The same integration using $\beta(N) \leq \beta(0) < \infty$ shows as in the previous section that $S_\infty > 0$.

2.1.5 Treatment Models

One form of treatment that is possible for some diseases is vaccination to protect against infection before the beginning of an epidemic. For example, this approach is commonly used for protection against annual influenza outbreaks. A simple way to model this would be to reduce the total population size by the fraction of the population protected against infection. However, in reality such inoculations are only partly effective, decreasing the rate of infection and also decreasing infectivity if a vaccinated person does become infected. To model this, it would be necessary to divide the population into two groups with different model parameters and to make some assumptions about the mixing between the two groups. We will not explore such more complicated models here.

If there is a treatment for infection once a person has been infected, we model this by supposing that a fraction γ per unit time of infectives is selected for treatment, and that treatment reduces infectivity by a fraction δ. Suppose that the rate of removal from the treated class is η. The $SITR$ model, where T is the treatment class, is given by

$$\begin{aligned} S' &= -\beta(N)S[I + \delta T] \\ I' &= \beta(N)S[I + \delta T] - (\alpha + \gamma)I \\ T' &= \gamma I - \eta T \\ N' &= -(1 - f)\alpha I - (1 - f_T)\eta T. \end{aligned} \tag{2.11}$$

A flow chart is shown in Fig. 2.7.

It is not difficult to prove, much as was done for the model (2.1) that

$$S_\infty = \lim_{t\to\infty} S(t) > 0, \quad \lim_{t\to\infty} I(t) = \lim_{t\to\infty} T(t) = 0.$$

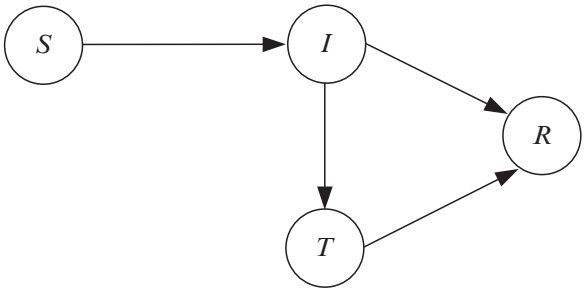

Fig. 2.7 Flow chart for the $SITR$ model

In order to calculate the basic reproduction number, we may argue that an infective in a totally susceptible population causes βK new infections in unit time, and the mean time spent in the infective compartment is $1/(\alpha+\gamma)$. In addition, a fraction $\gamma/(\alpha+\gamma)$ of infectives are treated. While in the treatment stage the number of new infections caused in unit time is $\delta\beta K$, and the mean time in the treatment class is $1/\eta$. Thus $\mathcal{R}_0$ is

$$\mathcal{R}_0 = \frac{\beta K}{\alpha+\gamma} + \frac{\gamma}{\alpha+\gamma}\frac{\delta\beta K}{\eta}. \tag{2.12}$$

It is also possible to establish the final size relation (2.3) by means similar to those used for the simple model (2.1). We integrate the first equation of (2.11) to obtain

$$\begin{aligned}\log\frac{S(0)}{S_\infty} &= \int_0^\infty \beta(N(t))[I(t)+\delta T(t)]dt \\ &\geq \beta(K)\int_0^\infty [I(t)+\delta T(t)]dt.\end{aligned}$$

Integration of the third equation of (2.11) gives

$$\gamma\int_0^\infty I(t)dt = \eta\int_0^\infty T(t)dt.$$

Integration of the sum of the first two equations of (2.11) gives

$$K - S_\infty = (\alpha+\gamma)\int_0^\infty I(t)dt.$$

Combination of these three equations and (2.12) gives

$$\log \frac{S(0)}{S_\infty} \geq \mathcal{R}_0 \left[\frac{K - S_\infty}{K} \right].$$

If β is constant, this relation is an equality, and is the same as (2.3).

2.1.6 An Epidemic Management (Quarantine-Isolation) Model

An actual epidemic differs considerably from the idealized models (2.1) or (2.6), as was shown by the SARS epidemic of 2002–3. Some notable differences are:

1. As we have seen in the preceding section, at the beginning of an epidemic the number of infectives is small and a deterministic model, which presupposes enough infectives to allow homogeneous mixing, is inappropriate.
2. When it is realized that an epidemic has begun, individuals are likely to modify their behaviour by avoiding crowds to reduce their contacts and by being more careful about hygiene to reduce the risk that a contact will produce infection.
3. If a vaccine is available for the disease which has broken out, public health measures will include vaccination of part of the population. Various vaccination strategies are possible, including vaccination of health care workers and other first line responders to the epidemic, vaccination of members of the population who have been in contact with diagnosed infectives, or vaccination of members of the population who live in close proximity to diagnosed infectives.
4. Diagnosed infectives may be hospitalized, both for treatment and to isolate them from the rest of the population.
5. Contact tracing of diagnosed infectives may identify people at risk of becoming infective, who may be quarantined (instructed to remain at home and avoid contacts) and monitored so that they may be isolated immediately if and when they become infective.
6. In some diseases, exposed members who have not yet developed symptoms may already be infective, and this would require inclusion in the model of new infections caused by contacts between susceptibles and asymptomatic infectives from the exposed class.
7. Isolation may be imperfect; in-hospital transmission of infection was a major problem in the SARS epidemic.

In the SARS epidemic of 2002–2003 in-hospital transmission of disease from patients to health care workers or visitors because of imperfect isolation accounted for many of the cases. This points to an essential heterogeneity in

disease transmission which must be included whenever there is any risk of such transmission.

All these generalizations have been considered in studies of the SARS epidemic of 2002–3. While the ideas were suggested in SARS modelling, they are in fact relevant to any epidemic. One beneficial effect of the SARS epidemic has been to draw attention to epidemic modelling which may be of great value in coping with future epidemics.

If a vaccine is available for a disease which threatens an epidemic outbreak, a vaccinated class which is protected at least partially against infection should be included in a model. While this is not relevant for an outbreak of a new disease, it would be an important aspect to be considered in modelling an influenza epidemic or a bioterrorist outbreak of smallpox.

For an outbreak of a new disease, where no vaccine is available, isolation and quarantine are the only control measures available. Let us formulate a model for an epidemic once control measures have been started. Thus, we assume that an epidemic has started, but that the number of infectives is small and almost all members of the population are still susceptible.

We formulate a model to describe the course of an epidemic when control measures are begun under the assumptions:

1. Exposed members may be infective with infectivity reduced by a factor $\varepsilon_E, 0 \leq \varepsilon_E < 1$.
2. Exposed members who are not isolated become infective at rate κ_1.
3. We introduce a class Q of quarantined members and a class J of isolated members.
4. Exposed members are quarantined at a proportional rate γ_1 in unit time (in practice, a quarantine will also be applied to many susceptibles, but we ignore this in the model). Quarantine is not perfect, but reduces the contact rate by a factor ε_Q. The effect of this assumption is that some susceptibles make fewer contacts than the model assumes.
5. There may be transmission of disease by isolated members, with an infectivity factor of ε_J.
6. Infectives are diagnosed at a proportional rate γ_2 per unit time and isolated. In addition, quarantined members are monitored and when they develop symptoms at rate κ_2 they are isolated immediately.
7. Infectives leave the infective class at rate α_1 and a fraction f_1 of these recover, and isolated members leave the isolated class at rate α_2 with a fraction f_2 recovering.

These assumptions lead to the $SEQIJR$ model [13]

$$\begin{aligned} S' &= -\beta(N)S[\varepsilon_E E + \varepsilon_E\varepsilon_Q Q + I + \varepsilon_J J] \\ E' &= \beta(N)S[\varepsilon_E E + \varepsilon_E\varepsilon_Q Q + I + \varepsilon_J J] - (\kappa_1 + \gamma_1)E \\ Q' &= \gamma_1 E - \kappa_2 Q \\ I' &= \kappa_1 E - (\alpha_1 + \gamma_2)I \\ J' &= \kappa_2 Q + \gamma_2 I - \alpha_2 J \\ N' &= -(1 - f_1)\alpha_1 I - (1 - f_2)\alpha_2 J \, . \end{aligned} \tag{2.13}$$

Here, we have used an equation for N to replace the equation

$$R' = f_1\alpha_1 I + f_2\alpha_2 J \, .$$

The model before control measures are begun is the special case

$$\gamma_1 = \gamma_2 = \kappa_2 = \alpha_2 = f_2 = 0, \;\; Q = J = 0$$

of (2.13). It is the same as (2.10).

We define the *control reproduction number* $\mathcal{R}_c$ to be the number of secondary infections caused by a single infective in a population consisting essentially only of susceptibles with the control measures in place. It is analogous to the basic reproduction number but instead of describing the very beginning of the disease outbreak it describes the beginning of the recognition of the epidemic. The basic reproduction number is the value of the control reproduction number with

$$\gamma_1 = \gamma_2 = \kappa_2 = \alpha_2 = f_2 = 0 \, .$$

In addition, there is a time-dependent *effective reproduction number* $\mathcal{R}^*$ which continues to track the number of secondary infections caused by a single infective as the epidemic continues with control measures (quarantine of asymptomatics and isolation of symptomatics) in place. It is not difficult to show that if the inflow into the population from travellers and new births is small (i.e., if the epidemiological time scale is much faster than the demographic time scale), our model implies that $\mathcal{R}^*$ will become and remain less than unity, so that the epidemic will always pass. Even if $\mathcal{R}_c > 1$, the epidemic will abate eventually when the effective reproduction number becomes less than unity. The effective reproduction number $\mathcal{R}^*$ is essentially $\mathcal{R}_c$ multiplied by a factor S/N, but allows time-dependent parameter values as well.

However, it should be remembered that if the epidemic takes so long to pass that there are enough new births and travellers to keep $\mathcal{R}^* > 1$, there will be an endemic equilibrium meaning that the disease will establish itself and remain in the population.

We have already calculated $\mathcal{R}_0$ for (2.10) and we may calculate $\mathcal{R}_c$ in the same way but using the full model with quarantined and isolated classes. We obtain

$$\mathcal{R}_c = \frac{\varepsilon_E K\beta(K)}{D_1} + \frac{K\beta(K)\kappa_1}{D_1 D_2} + \frac{\varepsilon_Q \varepsilon_E K\beta(K)\gamma_1}{D_1\kappa_2} + \frac{\varepsilon_J K\beta(K)\kappa_1\gamma_2}{\alpha_2 D_1 D_2} + \frac{\varepsilon_J K\beta(K)\gamma_1}{\alpha_2 D_1}$$
$$\mathcal{R}^* = \mathcal{R}_c \frac{S}{N} ,$$

where $D_1 = \gamma_1 + \kappa_1, \quad D_2 = \gamma_2 + \alpha_1$.

Each term of $\mathcal{R}_c$ has an epidemiological interpretation. The mean duration in E is $1/D_1$ with contact rate $\varepsilon_E\beta$, giving a contribution to $\mathcal{R}_c$ of $\varepsilon_E K\beta(K)/D_1$. A fraction κ_1/D_1 goes from E to I, with contact rate β and mean duration $1/D_2$, giving a contribution of $K\beta(K)\kappa_1/D_1D_2$. A fraction γ_1/D_1 goes from E to Q, with contact rate $\varepsilon_E\varepsilon_Q\beta$ and mean duration $1/\kappa_2$, giving a contribution of $\varepsilon_E\varepsilon_Q K\beta(K)\gamma_1/D_1\kappa_2$. A fraction $\kappa_1\gamma_2/D_1D_2$ goes from E to I to J, with a contact rate of $\varepsilon_J\beta$ and a mean duration of $1/\alpha_2$, giving a contribution of $\varepsilon_J K\beta(K)\kappa_1\gamma_2/\alpha_2 D_1 D_2$. Finally, a fraction γ_1/D_1 goes from E to Q to J with a contact rate of $\varepsilon_J\beta$ and a mean duration of $1/\alpha_2$ giving a contribution of $\varepsilon_J K\beta(K)\gamma_1/D_1\alpha_2$. The sum of these individual contributions gives $\mathcal{R}_c$.

In the model (2.13) the parameters γ_1 and γ_2 are *control* parameters which may be varied in the attempt to manage the epidemic. The parameters ϵ_Q and ϵ_J depend on the strictness of the quarantine and isolation processes and are thus also control measures in a sense. The other parameters of the model are specific to the disease being studied. While they are not variable, their measurements are subject to experimental error.

The linearization of (2.13) at the disease-free equilibrium $(K, 0, 0, 0, 0, K)$ has matrix

$$\begin{bmatrix} \epsilon_E K\beta(K) - (\kappa_1 + \gamma_1) & \varepsilon_E\varepsilon_Q\beta(K) & K\beta(K) & \varepsilon_J K\beta(K) \\ \gamma_1 & -\kappa_2 & 0 & 0 \\ \kappa_1 & 0 & -(\alpha_1 + \gamma_2) & 0 \\ 0 & \kappa_2 & \gamma_2 & -\alpha_2 \end{bmatrix} .$$

The corresponding characteristic equation is a fourth degree polynomial equation whose leading coefficient is 1 and whose constant term is a positive constant multiple of $1 - \mathcal{R}_c$, thus positive if $\mathcal{R}_c < 1$ and negative if $\mathcal{R}_c > 1$. If $\mathcal{R}_c > 1$ there is a positive eigenvalue, corresponding to an initial exponential growth rate of solutions of (2.13). If $\mathcal{R}_c < 1$ it is possible to show that all eigenvalues of the coefficient matrix have negative real part, and thus solutions of (2.13) die out exponentially [38].

Next, we wish to show that analogues of the relation (2.8) and $S_\infty > 0$ derived for the model (2.6) are valid for the management model (2.13). We begin by integrating the equations for $S + E, Q, I, J$, and N of (2.13) with respect to t from $t = 0$ to $t = \infty$, using the initial conditions

$$S(0) + E(0) = N(0) = K, \qquad Q(0) = I(0) = J(0) = 0 .$$

We obtain, since E, Q, I, and J all approach zero at $t \to \infty$,

$$K - S_\infty = (\kappa_1 + \gamma_1) \int_0^\infty E(s)ds$$
$$\gamma_1 \int_0^\infty E(s)ds = \kappa_2 \int_0^\infty Q(s)ds$$
$$\kappa_1 \int_0^\infty E(s)ds = (\alpha_1 + \gamma_2) \int_0^\infty I(s)ds$$
$$\kappa_2 \int_0^\infty Q(s)ds = \alpha_2 \int_0^\infty J(s)ds - \gamma_2 \int_0^\infty I(s)ds$$
$$K - N_\infty = (1 - f_1)\alpha_1 \int_0^\infty I(s)ds + (1 - f_2)\alpha_2 \int_0^\infty J(s)ds \, .$$

In order to relate $(K - S_\infty)$ to $(K - N_\infty)$, we need to express $\int_0^\infty I(s)ds$ and $\int_0^\infty J(s)ds$ in terms of $\int_0^\infty E(s)ds$.

From the three above relations for integrals we obtain

$$(\alpha_1 + \gamma_2) \int_0^\infty I(s)ds = \kappa_1 \int_0^\infty E(s)ds$$
$$\alpha_2 \int_0^\infty J(s)ds = \frac{\gamma_1\alpha_1 + \gamma_1\gamma_2 + \kappa_1\gamma_2}{\alpha_1 + \gamma_2} \int_0^\infty E(s)ds \, .$$

Thus we have

$$K - N_\infty = \frac{(1 - f_1)\alpha_1\kappa_1 + (1 - f_2)(\gamma_1\alpha_1 + \gamma_1\gamma_2 + \kappa_1\gamma_2)}{\alpha_1 + \gamma_2} \int_0^\infty E(s)ds$$
$$= \frac{(1 - f_1)\alpha_1\kappa_1 + (1 - f_2)(\gamma_1\alpha_1 + \gamma_1\gamma_2 + \kappa_1\gamma_2)}{(\kappa_1 + \gamma_1)(\alpha_1 + \gamma_2)}[K - S_\infty] \, .$$

This has the form, analogous to (2.8),

$$K - N_\infty = c[K - S_\infty] \tag{2.14}$$

with c, the disease death rate, given by

$$c = \frac{(1 - f_1)\alpha_1\kappa_1 + (1 - f_2)(\gamma_1\alpha_1 + \gamma_1\gamma_2 + \kappa_1\gamma_2)}{(\kappa_1 + \gamma_1)(\alpha_1 + \gamma_2)} \, .$$

The mean disease death rate may be measured and this expression gives information about some of the parameters in the model which can not be measured directly. It is easy to see that $0 \le c \le 1$ with $c = 0$ if and only if $f_1 = f_2 = 1$, that is, if and only if there are no disease deaths, and $c = 1$ if and only if $f_1 = f_2 = 0$, that is, if and only if the disease is universally fatal.

An argument similar to the one used for (2.6) but technically more complicated may be used to show that $S_\infty > 0$ for the treatment model (2.13). Thus the asymptotic behaviour of the management model (2.13) is the same as that of the simpler model (2.6). If the control reproduction number $\mathcal{R}_c$ is

less than 1 the disease dies out and if $\mathcal{R}_c > 1$ there is an epidemic which will pass leaving some members of the population untouched.

2.1.7 Stochastic Models for Disease Outbreaks

The underlying assumptions of the models of Kermack–McKendrick type studied in this chapter are that the sizes of the compartments are large enough that the mixing of members is homogeneous. While these assumptions are probably reasonable once an epidemic is well underway, at the beginning of a disease outbreak the situation may be quite different. At the beginning of an epidemic most members of the population are susceptible, that is, not (yet) infected, and the number of infectives (members of the population who are infected and may transmit infection) is small. The transmission of infection depends strongly on the pattern of contacts between members of the population, and a description should involve this pattern. Since the number of infectives is small a description involving an assumption of mass action should be replaced by a model which incorporates stochastic effects.

One approach would be a complete description of stochastic epidemic models, for which we refer the reader to the chapter on stochastic models in this volume by Linda Allen. Another approach would be to consider a stochastic model for an outbreak of a communicable disease to be applied so long as the number of infectives remains small, distinguishing a (minor) disease outbreak confined to this initial stage from a (major) epidemic which occurs if the number of infectives begins to grow at an exponential rate. Once an epidemic has started we may switch to a deterministic compartmental model. This approach is described in Chap. 4 on network models in this volume. There is an important difference between the behaviour of network models and the behaviour of models of Kermack–McKendrick type, namely that for a stochastic disease outbreak model if $\mathcal{R}_0 < 1$ the probability that the infection will die out is 1, while if $\mathcal{R}_0 > 1$ there is a positive probability that the infection will persist, and will lead to an epidemic and a positive probability that the infection will increase initially but will produce only a minor outbreak and will die out before triggering a major epidemic.

2.2 Models with Demographic Effects

2.2.1 The SIR Model

Epidemics which sweep through a population attract much attention and arouse a great deal of concern. As we have mentioned in the introduction,

the prevalence and effects of many diseases in less developed countries are probably less well-known but may be of even more importance. There are diseases which are endemic in many parts of the world and which cause millions of deaths each year. We have omitted births and deaths in our description of models because the time scale of an epidemic is generally much shorter than the demographic time scale. In effect, we have used a time scale on which the number of births and deaths in unit time is negligible. To model a disease which may be endemic we need to think on a longer time scale and include births and deaths.

For diseases that are endemic in some region public health physicians need to be able to estimate the number of infectives at a given time as well as the rate at which new infections arise. The effects of quarantine or vaccine in reducing the number of victims are of importance, just as in the treatment of epidemics. In addition, the possibility of defeating the endemic nature of the disease and thus controlling or even eradicating the disease in a population is worthy of study.

Measles is a disease for which endemic equilibria have been observed in many places, frequently with sustained oscillations about the equilibrium. The epidemic model of the first section assumes that the epidemic time scale is so short relative to the demographic time scale that demographic effects may be ignored. For measles, however, the reason for the endemic nature of the disease is that there is a flow of new susceptible members into the population, and in order to try to model this we must include births and deaths in the model. The simplest way to incorporate births and deaths in an infectious disease model is to assume a constant number of births and an equal number of deaths per unit time so that the total population size remains constant. This is, of course, feasible only if there are no deaths due to the disease. In developed countries such an assumption is plausible because there are few deaths from measles. In less developed countries there is often a very high mortality rate for measles and therefore other assumptions are necessary.

The first attempt to formulate an SIR model with births and deaths to describe measles was given in 1929 by H.E. Soper [32], who assumed a constant birth rate μK in the susceptible class and a constant death rate μK in the removed class. His model is

$$\begin{aligned} S' &= -\beta SI + \mu K \\ I' &= \beta SI - \gamma I \\ R' &= \gamma I - \mu K \ . \end{aligned}$$

This model is unsatisfactory biologically because the linkage of births of susceptibles to deaths of removed members is unreasonable. It is also an improper model mathematically because if $R(0)$ and $I(0)$ are sufficiently small then $R(t)$ will become negative. For any disease model to be plausible it is essential that the problem be properly posed in the sense that the number of

members in each class must remain non-negative. A model that does not satisfy this requirement cannot be a proper description of a disease model and therefore must contain some assumption that is biologically unreasonable. A full analysis of a model should include verification of this property.

A model of Kermack and McKendrick [22] includes births in the susceptible class proportional to total population size and a death rate in each class proportional to the number of members in the class. This model allows the total population size to grow exponentially or die out exponentially if the birth and death rates are unequal. It is applicable to such questions as whether a disease will control the size of a population that would otherwise grow exponentially. We shall return to this topic, which is important in the study of many diseases in less developed countries with high birth rates. To formulate a model in which total population size remains bounded we could follow the approach suggested by [15] in which the total population size is held constant by making birth and death rates equal. Such a model is

$$\begin{aligned} S' &= -\beta SI + \mu(K - S) \\ I' &= \beta SI - \gamma I - \mu I \\ R' &= \gamma I - \mu R \, . \end{aligned}$$

Because $S + I + R = K$, we can view R as determined when S and I are known and consider the two-dimensional system

$$\begin{aligned} S' &= -\beta SI + \mu(K - S) \\ I' &= \beta SI - \gamma I - \mu I \, . \end{aligned}$$

We shall examine a slightly more general SIR model with births and deaths for a disease that may be fatal to some infectives. For such a disease the class R of removed members should contain only recovered members, not members removed by death from the disease. It is not possible to assume that the total population size remain constant if there are deaths due to disease; a plausible model for a disease that may be fatal to some infectives must allow the total population to vary in time. The simplest assumption to allow this is a constant birth rate Λ, but in fact the analysis is quite similar if the birth rate is a function $\Lambda(N)$ of total population size N.

Let us analyze the model

$$\begin{aligned} S' &= \Lambda - \beta SI - \mu S \\ I' &= \beta SI - \mu I - \alpha I \\ N' &= \Lambda - (1 - f)\alpha I - \mu N \, , \end{aligned} \tag{2.15}$$

where $N = S + I + R$, with a mass action contact rate, a constant number of births Λ per unit time, a proportional natural death rate μ in each class, and a rate of recovery or disease death α of infectives with a fraction f of infectives recovering with immunity against reinfection. In this model if $f = 1$

the total population size approaches a limit $K = \Lambda/\mu$. Then K is the carrying capacity of the population. If $f < 1$ the total population size is not constant and K represents a carrying capacity or maximum possible population size, rather than a population size. We view the first two equations as determining S and I, and then consider the third equation as determining N once S and I are known. This is possible because N does not enter into the first two equations. Instead of using N as the third variable in this model we could have used R, and the same reduction would have been possible.

If the birth or recruitment rate $\Lambda(N)$ is a function of total population size then in the absence of disease the total population size N satisfies the differential equation

$$N' = \Lambda(N) - \mu N \ .$$

The *carrying capacity* of population size is the limiting population size K, satisfying

$$\Lambda(K) = \mu K, \qquad \Lambda'(K) < \mu \ .$$

The condition $\Lambda'(K) < \mu$ assures the asymptotic stability of the equilibrium population size K. It is reasonable to assume that K is the only positive equilibrium, so that

$$\Lambda(N) > \mu N$$

for $0 \le N \le K$. For most population models,

$$\Lambda(0) = 0, \qquad \Lambda''(N) \le 0 \ .$$

However, if $\Lambda(N)$ represents recruitment into a behavioural class, as would be natural for models of sexually transmitted diseases, it would be plausible to have $\Lambda(0) > 0$, or even to consider $\Lambda(N)$ to be a constant function. If $\Lambda(0) = 0$, we require $\Lambda'(0) > \mu$ because if this requirement is not satisfied there is no positive equilibrium and the population would die out even in the absence of disease.

We have used a mass action contact rate for simplicity, even though a more general contact rate would give a more accurate model, just as in the epidemics considered in the preceding section. With a general contact rate and a density-dependent birth rate we would have a model

$$\begin{aligned} S' &= \Lambda(N) - \beta(N)SI - \mu S \\ I' &= \beta(N)SI - \mu I - \alpha I \\ N' &= \Lambda(N) - (1-f)\alpha I - \mu N. \end{aligned} \tag{2.16}$$

If $f = 1$, so that there are no disease deaths, the equation for N is

$$N' = \Lambda(N) - \mu N \ ,$$

so that $N(t)$ approaches a limiting population size K. The theory of *asymptotically autonomous systems* [8, 24, 34, 37] implies that if N has a constant

limit then the system is equivalent to the system in which N is replaced by this limit. Then the system (2.16) is the same as the system (2.15) with β replaced by the constant $\beta(K)$ and N by K, and $\Lambda(N)$ replaced by the constant $\Lambda(K) = \mu K$.

We shall analyze the model (2.15) qualitatively. In view of the remark above, our analysis will also apply to the more general model (2.16) if there are no disease deaths. Analysis of the system (2.16) with $f < 1$ is much more difficult. We will confine our study of (2.16) to a description without details.

The first stage of the analysis is to note that the model (2.15) is a properly posed problem. That is, since $S' \geq 0$ if $S = 0$ and $I' \geq 0$ if $I = 0$, we have $S \geq 0, I \geq 0$ for $t \geq 0$ and since $N' \leq 0$ if $N = K$ we have $N \leq K$ for $t \geq 0$. Thus the solution always remains in the biologically realistic region $S \geq 0, I \geq 0, 0 \leq N \leq K$ if it starts in this region. By rights, we should verify such conditions whenever we analyze a mathematical model, but in practice this step is frequently overlooked.

Our approach will be to identify equilibria (constant solutions) and then to determine the asymptotic stability of each equilibrium. Asymptotic stability of an equilibrium means that a solution starting sufficiently close to the equilibrium remains close to the equilibrium and approaches the equilibrium as $t \to \infty >$, while instability of the equilibrium means that there are solutions starting arbitrarily close to the equilibrium which do not approach it. To find equilibria (S_∞, I_∞) we set the right side of each of the two equations equal to zero. The second of the resulting algebraic equations factors, giving two alternatives. The first alternative is $I_\infty = 0$, which will give a disease-free equilibrium, and the second alternative is $\beta S_\infty = \mu + \alpha$, which will give an endemic equilibrium, provided $\beta S_\infty = \mu + \alpha < \beta K$. If $I_\infty = 0$ the other equation gives $S_\infty = K = \Lambda/\mu$. For the endemic equilibrium the first equation gives

$$I_\infty = \frac{\Lambda}{\mu + \alpha} - \frac{\mu}{\beta} \,. \tag{2.17}$$

We linearize about an equilibrium (S_∞, I_∞) by letting $y = S - S_\infty$, $z = I - I_\infty$, writing the system in terms of the new variables y and z and retaining only the linear terms in a Taylor expansion. We obtain a system of two linear differential equations,

$$\begin{aligned} y' &= -(\beta I_\infty + \mu)y - \beta S_\infty z \\ z' &= \beta I_\infty y + (\beta S_\infty - \mu - \alpha)z \,. \end{aligned}$$

The coefficient matrix of this linear system is

$$\begin{bmatrix} -\beta I_\infty - \mu & -\beta S_\infty \\ \beta I_\infty & \beta S_\infty - \mu - \alpha \end{bmatrix}.$$

We then look for solutions whose components are constant multiples of $e^{\lambda t}$; this means that λ must be an eigenvalue of the coefficient matrix. The

condition that all solutions of the linearization at an equilibrium tend to zero as $t \to \infty$ is that the real part of every eigenvalue of this coefficient matrix is negative. At the disease-free equilibrium the matrix is

$$\begin{bmatrix} -\mu & -\beta K \\ 0 & \beta K - \mu - \alpha \end{bmatrix},$$

which has eigenvalues $-\mu$ and $\beta K - \mu - \alpha$. Thus, the disease-free equilibrium is asymptotically stable if $\beta K < \mu + \alpha$ and unstable if $\beta K > \mu + \alpha$. Note that this condition for instability of the disease-free equilibrium is the same as the condition for the existence of an endemic equilibrium.

In general, the condition that the eigenvalues of a 2×2 matrix have negative real part is that the determinant be positive and the trace (the sum of the diagonal elements) be negative. Since $\beta S_\infty = \mu + \alpha$ at an endemic equilibrium, the matrix of the linearization at an endemic equilibrium is

$$\begin{bmatrix} -\beta I_\infty - \mu & -\beta S_\infty \\ \beta I_\infty & 0 \end{bmatrix} \tag{2.18}$$

and this matrix has positive determinant and negative trace. Thus, the endemic equilibrium, if there is one, is always asymptotically stable. If the quantity

$$R_0 = \frac{\beta K}{\mu + \alpha} = \frac{K}{S_\infty} \tag{2.19}$$

is less than one, the system has only the disease-free equilibrium and this equilibrium is asymptotically stable. In fact, it is not difficult to prove that this asymptotic stability is *global*, that is, that every solution approaches the disease-free equilibrium. If the quantity $\mathcal{R}_0$ is greater than one then the disease-free equilibrium is unstable, but there is an endemic equilibrium that is asymptotically stable. Again, the quantity $\mathcal{R}_0$ is the basic reproduction number. It depends on the particular disease (determining the parameter α) and on the rate of contacts, which may depend on the population density in the community being studied. The disease model exhibits a *threshold* behaviour: If the basic reproduction number is less than one the disease will die out, but if the basic reproduction number is greater than one the disease will be endemic. Just as for the epidemic models of the preceding section, the basic reproduction number is the number of secondary infections caused by a single infective introduced into a wholly susceptible population because the number of contacts per infective in unit time is βK, and the mean infective period (corrected for natural mortality) is $1/(\mu + \alpha)$.

There are two aspects of the analysis of the model (2.16) which are more complicated than the analysis of (2.15). The first is in the study of equilibria. Because of the dependence of $\Lambda(N)$ and $\beta(N)$ on N, it is necessary to use two of the equilibrium conditions to solve for S and I in terms of N and then substitute into the third condition to obtain an equation for N. Then by

comparing the two sides of this equation for $N = 0$ and $N = K$ it is possible to show that there must be an endemic equilibrium value of N between 0 and K.

The second complication is in the stability analysis. Since (2.16) is a three-dimensional system which can not be reduced to a two-dimensional system, the coefficient matrix of its linearization at an equilibrium is a 3×3 matrix and the resulting characteristic equation is a cubic polynomial equation of the form

$$\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0 \, .$$

The *Routh–Hurwitz* conditions

$$a_1 > 0, \qquad a_1 a_2 > a_3 > 0$$

are necessary and sufficient conditions for all roots of the characteristic equation to have negative real part. A technically complicated calculation is needed to verify that these conditions are satisfied at an endemic equilibrium for the model (2.16).

The asymptotic stability of the endemic equilibrium means that the compartment sizes approach a steady state. If the equilibrium had been unstable, there would have been a possibility of sustained oscillations. Oscillations in a disease model mean fluctuations in the number of cases to be expected, and if the oscillations have long period could also mean that experimental data for a short period would be quite unreliable as a predictor of the future. Epidemiological models which incorporate additional factors may exhibit oscillations. A variety of such situations is described in [18, 19].

The epidemic models of the first section also exhibited a threshold behaviour but of a slightly different kind. For these models, which were SIR models without births or natural deaths, the threshold distinguished between a dying out of the disease and an epidemic, or short term spread of disease.

From the third equation of (2.15) we obtain

$$N' = \Lambda - \mu N - (1 - f)\alpha I \, ,$$

where $N = S + I + R$. From this we see that at the endemic equilibrium $N = K - (1 - f)\alpha I/\mu$, and the reduction in the population size from the carrying capacity K is

$$(1 - f)\frac{\alpha}{\mu} I_\infty = (1 - f)[\frac{\alpha K}{\mu + \alpha} - \frac{\alpha}{\beta}] \, .$$

The parameter α in the SIR model may be considered as describing the pathogenicity of the disease. If α is large it is less likely that $\mathcal{R}_0 > 1$. If α is small then the total population size at the endemic equilibrium is close to the carrying capacity K of the population. Thus, the maximum population decrease caused by disease will be for diseases of intermediate pathogenicity.

2.2.2 The SIS Model

In order to describe a model for a disease from which infectives recover with immunity against reinfection and that includes births and deaths as in the model (2.16), we may modify the model (2.16) by removing the equation for R and moving the term $f\alpha I$ describing the rate of recovery from infection to the equation for S. This gives the model

$$\begin{aligned} S' &= \Lambda(N) - \beta(N)SI - \mu S + f\alpha I \\ I' &= \beta(N)SI - \alpha I - \mu I \end{aligned} \tag{2.20}$$

describing a population with a density-dependent birth rate $\Lambda(N)$ per unit time, a proportional death rate μ in each class, and with a rate α of departure from the infective class through recovery or disease death and with a fraction f of infectives recovering with no immunity against reinfection. In this model, if $f < 1$ the total population size is not constant and K represents a *carrying capacity*, or maximum possible population size, rather than a constant population size.

It is easy to verify that

$$\mathcal{R}_0 = \frac{K\beta(K)}{\mu + \alpha} \ .$$

If we add the two equations of (2.20), and use $N = S + I$ we obtain

$$N' = \Lambda(N) - \mu N - (1 - f)\alpha I \ .$$

For the SIS model we are able to carry out the analysis with a general contact rate. If $f = 1$ the equation for N is

$$N' = \Lambda(N) - \mu N$$

and N approaches the limit K. The system (2.20) is asymptotically autonomous and its asymptotic behaviour is the same as that of the single differential equation

$$I' = \beta(K)I(K - I) - (\alpha + \mu)I \ , \tag{2.21}$$

where S has been replaced by $K - I$. But (2.21) is a logistic equation which is easily solved analytically by separation of variables or qualitatively by an equilibrium analysis. We find that $I \to 0$ if $K\beta(K) < (\mu + \alpha)$, or $\mathcal{R}_0 < 1$ and $I \to I_\infty > 0$ with

$$I_\infty = K - \frac{\mu + \alpha}{\beta(K)} = K(1 - \frac{1}{\mathcal{R}_0})$$

if $K\beta(K) > (\mu + \alpha)$ or $\mathcal{R}_0 > 1$.

To analyze the SIS model if $f < 1$, it is convenient to use I and N as variables instead of S and I, with S replaced by $N - I$. This gives the model

$$\begin{aligned} I' &= \beta(N)I(N-I) - (\mu+\alpha)I \\ N' &= \Lambda(N) - \mu N - (1-f)\alpha I \ . \end{aligned} \tag{2.22}$$

Equilibria are found by setting the right sides of the two differential equations equal to zero. The first of the resulting algebraic equations factors, giving two alternatives. The first alternative is $I = 0$, which will give a disease-free equilibrium $I = 0, N = K$, and the second alternative is $\beta(N)(N - I) = \mu + \alpha$, which may give an endemic equilibrium. For an endemic equilibrium (I_∞, N_∞) the first equation gives

$$I_\infty \beta(N_\infty) = N_\infty \beta(N_\infty) - (\mu + \alpha) \ .$$

Substitution into the other equilibrium condition gives

$$\Lambda(N_\infty) = \mu N_\infty + (1-f)\alpha[N_\infty - \frac{\mu+\alpha}{\beta(N_\infty)}] \ ,$$

which can be simplified to

$$\beta(N_\infty)\Lambda(N_\infty) = \mu N_\infty \beta(N_\infty) + (1-f)\alpha \left[N_\infty \beta(N_\infty) - (\mu+\alpha)\right] \ . \tag{2.23}$$

At $N = 0$ the left side of (2.23) is $\beta(0)\Lambda(0) \geq 0$, while the right side is $-(1-f)\alpha(\mu+\alpha)$, which is negative since $f < 1$. At $N = K$ the left side of (2.23) is

$$\beta(K)\Lambda(K) = \mu K \beta(K)$$

while the right side of (2.23) is

$$\mu K \beta(K) + (1-f)\alpha[K\beta(K) - (\mu+\alpha)] \ .$$

Since

$$\mathcal{R}_0 = \frac{K\beta(K)}{\mu+\alpha} \ ,$$

if $\mathcal{R}_0 > 1$ the left side of (2.23) is less than the right side of (2.23), and this implies that (2.23) has a solution for $N, 0 < N < K$. Thus there is an endemic equilibrium if $\mathcal{R}_0 > 1$. If $\mathcal{R}_0 < 1$ this reasoning may be used to show that there is no endemic equilibrium.

The linearization of (2.22) at an equilibrium (I_∞, N_∞) has coefficient matrix

$$\begin{bmatrix} \beta(N_\infty)(N_\infty - 2I_\infty) - (\mu+\alpha) & \beta(N_\infty)I_\infty + \beta'(N_\infty)I_\infty(N_\infty - I_\infty) \\ -(1-f)\alpha & \Lambda'(N_\infty) - \mu. \end{bmatrix}$$

At the disease-free equilibrium the matrix is

$$\begin{bmatrix} K\beta(K) - (\mu + \alpha) & 0 \\ -(1-f)\alpha & \Lambda'(K) - \mu \end{bmatrix},$$

which has eigenvalues $\Lambda'(K) - \mu$ and $K\beta K - (\mu + \alpha)$. Thus, the disease-free equilibrium is asymptotically stable if $K\beta(K) < \mu + \alpha$, or $\mathcal{R}_0 < 1$, and unstable if $K\beta(K) > \mu + \alpha$, or $\mathcal{R}_0 > 1$. Note that the condition for instability of the disease-free equilibrium is the same as the condition for the existence of an endemic equilibrium.

At an endemic equilibrium, since $\beta(N_\infty)(N_\infty - I_\infty) = \mu + \alpha$, the matrix is

$$\begin{bmatrix} -I\beta(N_\infty) & I_\infty\beta(N_\infty) + I_\infty(N_\infty - I_\infty)\beta'(N_\infty) \\ -(1-f)\alpha & \Lambda'(N_\infty) - \mu \end{bmatrix}.$$

Since $\beta'(N_\infty) \leq 0$

$$\beta(N_\infty) + (N_\infty - I_\infty)\beta'(N_\infty) \geq \beta(N_\infty) + N_\infty\beta'(N_\infty) \geq 0\,.$$

Thus if $\Lambda'(N_\infty) < \mu$ the coefficient matrix has sign structure

$$\begin{bmatrix} - & + \\ - & - \end{bmatrix}.$$

It is clear that the coefficient matrix has negative trace and positive determinant if $\Lambda'(N) < \mu$ and this implies that the endemic equilibrium is asymptotically stable. Thus, the endemic equilibrium, which exists if $\mathcal{R}_0 > 1$, is always asymptotically stable. If $\mathcal{R}_0 < 1$ the system has only the disease-free equilibrium and this equilibrium is asymptotically stable. In the case $f = 1$ the verification of these properties remains valid if there are no births and deaths. This suggests that a requirement for the existence of an endemic equilibrium is a flow of new susceptibles either through births, as in the SIR model or through recovery without immunity against reinfection, as in the SIS model with or without births and deaths.

If the epidemiological and demographic time scales are very different, for the SIR model we observed that the approach to endemic equilibrium is like a rapid and severe epidemic. The same happens in the SIS model, especially if there is a significant number of deaths due to disease. If there are few disease deaths the number of infectives at endemic equilibrium may be substantial, and there may be damped oscillations of large amplitude about the endemic equilibrium.

For both the SIR and SIS models we may write the differential equation for I as

$$I' = I[\beta(N)S - (\mu + \alpha)] = \beta(N)I[S - S_\infty]\,,$$

which implies that whenever S exceeds its endemic equilibrium value S_∞, I is increasing and epidemic-like behaviour is possible. If $\mathcal{R}_0 < 1$ and $S < K$ it follows that $I' < 0$, and thus I is decreasing. Thus, if $\mathcal{R}_0 < 1$, I cannot increase and no epidemic can occur.

Next, we will turn to some applications of SIR and SIS models, taken mainly from [3].

2.3 Some Applications

2.3.1 Herd Immunity

In order to prevent a disease from becoming endemic it is necessary to reduce the basic reproduction number $\mathcal{R}_0$ below one. This may sometimes be achieved by immunization. If a fraction p of the Λ newborn members per unit time of the population is successfully immunized, the effect is to replace K by $K(1-p)$, and thus to reduce the basic reproduction number to $\mathcal{R}_0(1-p)$. The requirement $\mathcal{R}_0(1-p) < 1$ gives $1-p < 1/\mathcal{R}_0$, or

$$p > 1 - \frac{1}{\mathcal{R}_0} \, .$$

A population is said to have *herd immunity* if a large enough fraction has been immunized to assure that the disease cannot become endemic. The only disease for which this has actually been achieved worldwide is smallpox for which $\mathcal{R}_0$ is approximately 5, so that 80% immunization does provide herd immunity.

For measles, epidemiological data in the United States indicate that $\mathcal{R}_0$ for rural populations ranges from 5.4 to 6.3, requiring vaccination of 81.5–84.1% of the population. In urban areas $\mathcal{R}_0$ ranges from 8.3 to 13.0, requiring vaccination of 88.0–92.3% of the population. In Great Britain, $\mathcal{R}_0$ ranges from 12.5 to 16.3, requiring vaccination of 92–94% of the population. The measles vaccine is not always effective, and vaccination campaigns are never able to reach everyone. As a result, herd immunity against measles has not been achieved (and probably never can be). Since smallpox is viewed as more serious and requires a lower percentage of the population be immunized, herd immunity was attainable for smallpox. In fact, smallpox has been eliminated; the last known case was in Somalia in 1977, and the virus is maintained now only in laboratories (although there is currently some concern that it may be reintroduced as a bioterrorism attack). The eradication of smallpox was actually more difficult than expected because high vaccination rates were achieved in some countries but not everywhere, and the disease persisted in some countries. The eradication of smallpox was possible only after an intensive campaign for worldwide vaccination [16].

2.3.2 Age at Infection

In order to calculate the basic reproduction number $\mathcal{R}_0$ for a disease, we need to know the values of the contact rate β and the parameters μ, K, and α. The parameters μ, K, and α can usually be measured experimentally but the contact rate β is difficult to determine directly. There is an indirect means of estimating $\mathcal{R}_0$ in terms of the life expectancy and the mean age at infection which enables us to avoid having to estimate the contact rate. In this calculation, we will assume that β is constant, but we will also indicate the modifications needed when β is a function of total population size N. The calculation assumes exponentially distributed life spans and infective periods. In fact, the result is valid so long as the life span is exponentially distributed.

Consider the "age cohort" of members of a population born at some time t_0 and let a be the age of members of this cohort. If $y(a)$ represents the fraction of members of the cohort who survive to age (at least) a, then the assumption that a fraction μ of the population dies per unit time means that $y'(a) = -\mu y(a)$. Since $y(0) = 1$, we may solve this first order initial value problem to obtain $y(a) = e^{-\mu a}$. The fraction dying at (exactly) age a is $-y'(a) = \mu y(a)$. The mean life span is the average age at death, which is $\int_0^\infty a[-y'(a)]da$, and if we integrate by parts we find that this life expectancy is

$$\int_0^\infty [-ay'(a)]\,da = [-ay(a)]_0^\infty + \int_0^\infty y(a)\,da = \int_0^\infty y(a)\,da\ .$$

Since $y(a) = e^{-\mu a}$, this reduces to $1/\mu$. The life expectancy is often denoted by L, so that we may write

$$L = \frac{1}{\mu}\ .$$

The rate at which surviving susceptible members of the population become infected at age a and time $t_0 + a$, is $\beta I(t_0 + a)$. Thus, if $z(a)$ is the fraction of the age cohort alive and still susceptible at age a, $z'(a) = -[\mu+\beta I(t_0+a)]z(a)$. Solution of this first linear order differential equation gives

$$z(a) = e^{-[\mu a+\int_0^a \beta I(t_0+b)\,db]} = y(a)e^{-\int_0^a \beta I(t_0+b)\,db}\ .$$

The mean length of time in the susceptible class for members who may become infected, as opposed to dying while still susceptible, is

$$\int_0^\infty e^{-\int_0^a \beta I(t_0+b)db}da\ ,$$

and this is the mean age at which members become infected. If the system is at an equilibrium I_∞, this integral may be evaluated, and the mean age at infection, denoted by A, is given by

$$A = \int_0^\infty e^{-\beta I_\infty a}\, da = \frac{1}{\beta I_\infty}.$$

For our model the endemic equilibrium is

$$I_\infty = \frac{\mu K}{\mu + \alpha} - \frac{\mu}{\beta}\,,$$

and this implies

$$\frac{L}{A} = \frac{\beta I_\infty}{\mu} = \mathcal{R}_0 - 1\,. \tag{2.24}$$

This relation is very useful in estimating basic reproduction numbers. For example, in some urban communities in England and Wales between 1956 and 1969 the average age of contracting measles was 4.8 years. If life expectancy is assumed to be 70 years, this indicates $\mathcal{R}_0 = 15.6$.

If β is a function $\beta(N)$ of total population size the relation (2.24) becomes

$$\mathcal{R}_0 = \frac{\beta(K)}{\beta(N)}\left[1 + \frac{L}{A}\right].$$

If disease mortality does not have a large effect on total population size, in particular if there is no disease mortality, this relation is very close to (2.24).

The relation between age at infection and basic reproduction number indicates that measures such as inoculations, which reduce $\mathcal{R}_0$, will increase the average age at infection. For diseases such as rubella (German measles), whose effects may be much more serious in adults than in children, this indicates a danger that must be taken into account: While inoculation of children will decrease the number of cases of illness, it will tend to increase the danger to those who are not inoculated or for whom the inoculation is not successful. Nevertheless, the number of infections in older people will be reduced, although the fraction of cases which are in older people will increase.

2.3.3 The Interepidemic Period

Many common childhood diseases, such as measles, whooping cough, chicken pox, diphtheria, and rubella, exhibit variations from year to year in the number of cases. These fluctuations are frequently regular oscillations, suggesting that the solutions of a model might be periodic. This does not agree with the predictions of the model we have been using here; however, it would not be inconsistent with solutions of the characteristic equation, which are complex conjugate with small negative real part corresponding to lightly damped oscillations approaching the endemic equilibrium. Such behaviour would look like recurring epidemics. If the eigenvalues of the matrix of the linearization at an endemic equilibrium are $-u \pm iv$, where $i^2 = -1$, then the solutions of the

linearization are of the form $Be^{-ut}\cos(vt+c)$, with decreasing "amplitude" Be^{-ut} and "period" $\frac{2\pi}{v}$.

For the model (2.15) we recall from (2.17) that at the endemic equilibrium we have

$$\beta I_\infty + \mu = \mu\mathcal{R}_0, \qquad \beta S_\infty = \mu + \alpha$$

and from (2.18) the matrix of the linearization is

$$\begin{bmatrix} -\mu R_0 & -(\mu+\alpha) \\ \mu(\mathcal{R}_0 - 1) & 0 \end{bmatrix}$$

The eigenvalues are the roots of the quadratic equation

$$\lambda^2 + \mu\mathcal{R}_0\lambda + \mu(\mathcal{R}_0 - 1)(\mu+\alpha) = 0,$$

which are

$$\lambda = \frac{-\mu\mathcal{R}_0 \pm \sqrt{\mu^2{\mathcal{R}_0}^2 - 4\mu(\mathcal{R}_0-1)(\mu+\alpha)}}{2}\,.$$

If the mean infective period $1/\alpha$ is much shorter than the mean life span $1/\mu$, we may neglect the terms that are quadratic in μ. Thus, the eigenvalues are approximately

$$\frac{-\mu\mathcal{R}_0 \pm \sqrt{-4\mu(\mathcal{R}_0-1)\alpha}}{2}\,,$$

and these are complex with imaginary part $\sqrt{\mu(\mathcal{R}_0-1)\alpha}$. This indicates oscillations with period approximately

$$\frac{2\pi}{\sqrt{\mu(\mathcal{R}_0-1)\alpha}}\,.$$

We use the relation $\mu(\mathcal{R}_0-1) = \mu L/A$ and the mean infective period $\tau = 1/\alpha$ to see that the interepidemic period T is approximately $2\pi\sqrt{A\tau}$. Thus, for example, for recurring outbreaks of measles with an infective period of 2 weeks or 1/26 year in a population with a life expectancy of 70 years with $\mathcal{R}_0$ estimated as 15, we would expect outbreaks spaced 2.76 years apart. Also, as the "amplitude" at time t is $e^{-\mu\mathcal{R}_0 t/2}$, the maximum displacement from equilibrium is multiplied by a factor $e^{-(15)(2.76)/140} = 0.744$ over each cycle. In fact, many observations of measles outbreaks indicate less damping of the oscillations, suggesting that there may be additional influences that are not included in our simple model. To explain oscillations about the endemic equilibrium a more complicated model is needed. One possible generalization would be to assume seasonal variations in the contact rate. This is a reasonable supposition for a childhood disease most commonly transmitted through school contacts, especially in winter in cold climates. Note, however, that data from observations are never as smooth as model predictions and models are inevitably gross simplifications of reality which cannot account for random

variations in the variables. It may be difficult to judge from experimental data whether an oscillation is damped or persistent.

2.3.4 "Epidemic" Approach to the Endemic Equilibrium

In the model (2.15) the demographic time scale described by the birth and natural death rates Λ and μ and the epidemiological time scale described by the rate α of departure from the infective class may differ substantially. Think, for example, of a natural death rate $\mu = 1/75$, corresponding to a human life expectancy of 75 years, and epidemiological parameters $\alpha = 25, f = 1$, describing a disease from which all infectives recover after a mean infective period of 1/25 year, or two weeks. Suppose we consider a carrying capacity $K = 1,000$ and take $\beta = 0.1$, indicating that an average infective makes $(0.1)(1,000) = 100$ contacts per year. Then $\mathcal{R}_0 = 4.00$, and at the endemic equilibrium we have $S_\infty = 250.13$, $I_\infty = 0.40$, $R_\infty = 749.47$. This equilibrium is globally asymptotically stable and is approached from every initial state.

However, if we take $S(0) = 999$, $I(0) = 1$, $R(0) = 0$, simulating the introduction of a single infective into a susceptible population and solve the system numerically we find that the number of infectives rises sharply to a maximum of 400 and then decreases to almost zero in a period of 0.4 year, or about 5 months. In this time interval the susceptible population decreases to 22 and then begins to increase, while the removed (recovered and immune against reinfection) population increases to almost 1,000 and then begins a gradual decrease. The size of this initial "epidemic" could not have been predicted from our qualitative analysis of the system (2.15). On the other hand, since μ is so small compared to the other parameters of the model, we might consider neglecting μ, replacing it by zero in the model. If we do this, the model reduces to the simple Kermack–McKendrick epidemic model (without births and deaths) of the first section.

If we follow the model (2.15) over a longer time interval we find that the susceptible population grows to 450 after 46 years, then drops to 120 during a small epidemic with a maximum of 18 infectives, and exhibits widely spaced epidemics decreasing in size. It takes a very long time before the system comes close to the endemic equilibrium and remains close to it. The large initial epidemic conforms to what has often been observed in practice when an infection is introduced into a population with no immunity, such as the smallpox inflicted on the Aztecs by the invasion of Cortez.

If we use the model (2.15) with the same values of β, K and μ, but take $\alpha = 25, f = 0$ to describe a disease fatal to all infectives, we obtain very similar results. Now the total population is $S + I$, which decreases from an initial size of 1,000 to a minimum of 22 and then gradually increases and

eventually approaches its equilibrium size of 250.53. Thus, the disease reduces the total population size to one-fourth of its original value, suggesting that infectious diseases may have large effects on population size. This is true even for populations which would grow rapidly in the absence of infection, as we shall see later.

2.3.5 Disease as Population Control

Many parts of the world experienced very rapid population growth in the eighteenth century. The population of Europe increased from 118 million in 1700 to 187 million in 1800. In the same time period the population of Great Britain increased from 5.8 million to 9.15 million, and the population of China increased from 150 million to 313 million [27]. The population of English colonies in North America grew much more rapidly than this, aided by substantial immigration from England, but the native population, which had been reduced to one tenth of their previous size by disease following the early encounters with Europeans and European diseases, grew even more rapidly. While some of these population increases may be explained by improvements in agriculture and food production, it appears that an even more important factor was the decrease in the death rate due to diseases. Disease death rates dropped sharply in the eighteenth century, partly from better understanding of the links between illness and sanitation and partly because the recurring invasions of bubonic plague subsided, perhaps due to reduced susceptibility. One plausible explanation for these population increases is that the bubonic plague invasions served to control the population size, and when this control was removed the population size increased rapidly.

In developing countries it is quite common to have high birth rates and high disease death rates. In fact, when disease death rates are reduced by improvements in health care and sanitation it is common for birth rates to decline as well, as families no longer need to have as many children to ensure that enough children survive to take care of the older generations. Again, it is plausible to assume that population size would grow exponentially in the absence of disease but is controlled by disease mortality.

The SIR model with births and deaths of Kermack and McKendrick [22] includes births in the susceptible class proportional to population size and a natural death rate in each class proportional to the size of the class. Let us analyze a model of this type with birth rate r and a natural death rate $\mu < r$. For simplicity we assume the disease is fatal to all infectives with disease death rate α, so that there is no removed class and the total population size is $N = S + I$. Our model is

$$\begin{aligned} S' &= r(S+I) - \beta SI - \mu S \\ I' &= \beta SI - (\mu + \alpha)I \ . \end{aligned} \tag{2.25}$$

From the second equation we see that equilibria are given by either $I = 0$ or $\beta S = \mu + \alpha$. If $I = 0$ the first equilibrium equation is $rS = \mu S$, which implies $S = 0$ since $r > \mu$. It is easy to see that the equilibrium (0,0) is unstable. What actually would happen if $I = 0$ is that the susceptible population would grow exponentially with exponent $r - \mu > 0$. If $\beta S = \mu + \alpha$ the first equilibrium condition gives

$$r\frac{\mu+\alpha}{\beta} + rI - (\mu+\alpha)I - \frac{\mu(\mu+\alpha)}{\beta} = 0\,,$$

which leads to

$$(\alpha + \mu - r)I = \frac{(r-\mu)(\mu+\alpha)}{\beta}\,.$$

Thus, there is an endemic equilibrium provided $r < \alpha + \mu$, and it is possible to show by linearizing about this equilibrium that it is asymptotically stable. On the other hand, if $r > \alpha + \mu$ there is no positive equilibrium value for I. In this case we may add the two differential equations of the model to give

$$N' = (r-\mu)N - \alpha I \geq (r-\mu)N - \alpha N = (r - \mu - \alpha)N$$

and from this we may deduce that N grows exponentially. For this model either we have an asymptotically stable endemic equilibrium or population size grows exponentially. In the case of exponential population growth we may have either vanishing of the infection or an exponentially growing number of infectives.

If only susceptibles contribute to the birth rate, as may be expected if the disease is sufficiently debilitating, the behaviour of the model is quite different. Let us consider the model

$$\begin{aligned} S' &= rS - \beta SI - \mu S = S(r - \mu - \beta I) \\ I' &= \beta SI - (\mu+\alpha)I = I(\beta S - \mu - \alpha) \end{aligned} \tag{2.26}$$

which has the same form as the celebrated Lotka–Volterra predator–prey model of population dynamics. This system has two equilibria, obtained by setting the right sides of each of the equations equal to zero, namely (0,0) and an endemic equilibrium $((\mu+\alpha)/\beta, (r-\mu)/\beta)$. It turns out that the qualitative analysis approach we have been using is not helpful as the equilibrium (0,0) is unstable and the eigenvalues of the coefficient matrix at the endemic equilibrium have real part zero. In this case the behaviour of the linearization does not necessarily carry over to the full system. However, we can obtain information about the behaviour of the system by a method that begins with the elementary approach of separation of variables for first order differential equations. We begin by taking the quotient of the two differential equations and using the relation

$$\frac{I'}{S'} = \frac{dI}{dS}$$

to obtain the separable first order differential equation

$$\frac{dI}{dS} = \frac{I(\beta S - \mu - \alpha)}{S(r - \beta I)} .$$

Separation of variables gives

$$\int \Big(\frac{r}{I} - \beta\Big) dI = \int \Big(\beta - \frac{\mu + \alpha}{S}\Big) dS .$$

Integration gives the relation

$$\beta(S + I) - r \log I - (\mu + \alpha) \log S = c$$

where c is a constant of integration. This relation shows that the quantity

$$V(S, I) = \beta(S + I) - r \log I - (\mu + \alpha) \log S$$

is constant on each orbit (path of a solution in the $(S, I-$ plane). Each of these orbits is a closed curve corresponding to a periodic solution.

This model is the same as the simple epidemic model of the first section except for the birth and death terms, and in many examples the time scale of the disease is much faster than the time scale of the demographic process. We may view the model as describing an epidemic initially, leaving a susceptible population small enough that infection cannot establish itself. Then there is a steady population growth until the number of susceptibles is large enough for an epidemic to recur. During this growth stage the infective population is very small and random effects may wipe out the infection, but the immigration of a small number of infectives will eventually restart the process. As a result, we would expect recurrent epidemics. In fact, bubonic plague epidemics did recur in Europe for several hundred years. If we modify the demographic part of the model to assume limited population growth rather than exponential growth in the absence of disease, the effect would be to give behaviour like that of the model studied in the previous section, with an endemic equilibrium that is approached slowly in an oscillatory manner if $\mathcal{R}_0 > 1$.

Example. (Fox rabies) Rabies is a viral infection to which many animals, especially foxes, coyotes, wolves, and rats, are highly susceptible. While dogs are only moderately susceptible, they are the main source of rabies in humans. Although deaths of humans from rabies are few, the disease is still of concern because it is invariably fatal. However, the disease is endemic in animals in many parts of the world. A European epidemic of fox rabies thought to have begun in Poland in 1939 and spread through much of Europe has been modeled. We present here a simplified version of a model due to R.M. Anderson and coworkers [1].

We begin with the demographic assumptions that foxes have a birth rate proportional to population size but that infected foxes do not produce

offspring (because the disease is highly debilitating), and that there is a natural death rate proportional to population size. Experimental data indicate a birth rate of approximately 1 per capita per year and a death rate of approximately 0.5 per capita per year, corresponding to a life expectancy of 2 years. The fox population is divided into susceptibles and infectives, and the epidemiological assumptions are that the rate of acquisition of infection is proportional to the number of encounters between susceptibles and infectives. We will assume a contact parameter $\beta = 80$, in rough agreement with observations of frequency of contact in regions where the fox density is approximately 1 fox/km^2, and we assume that all infected foxes die with a mean infective period of approximately 5 days or 1/73 year. These assumptions lead to the model

$$S' = -\beta SI + rS - \mu S$$
$$I' = \beta SI - (\mu + \alpha)I$$

with $\beta = 80$, $r = 1.0$, $\mu = 0.5$, $\alpha = 73$. As this is of the form (2.26), we know that the orbits are closed curves in the (S, I) plane, and that both S and I are periodic functions of t. We illustrate with some simulations obtained using Maple (Figs. 2.8, 2.9, and 2.10). It should be noted from the graphs of I in terms of t that the period of the oscillation depends on the amplitude, and thus on the initial conditions, with larger amplitudes corresponding to longer periods.

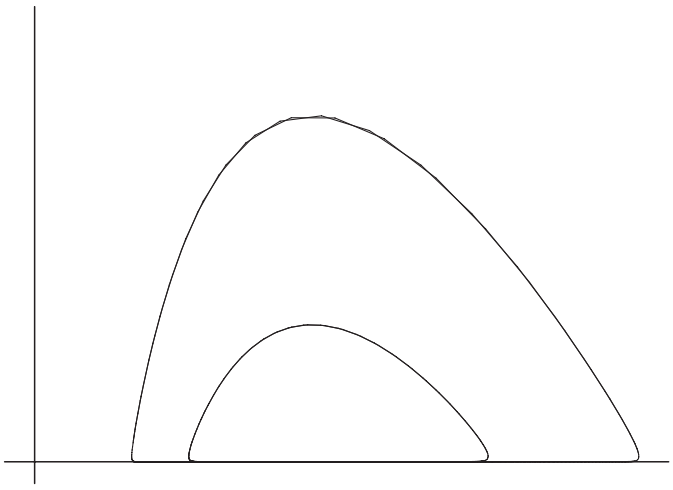

Fig. 2.8 The (S, I) plane

A warning is in order here. The model predicts that for long time intervals the number of infected foxes is extremely small. With such small numbers, the continuous deterministic models we have been using (which assume that

population sizes are differentiable functions) are quite inappropriate. If the density of foxes is extremely small an encounter between foxes is a random event, and the number of contacts cannot be described properly by a function of population densities. To describe disease transmission properly when population sizes are very small we would need to use a stochastic model.

Now let us modify the demographic assumptions by assuming that the birth rate decreases as population size increases. We replace the birth rate of r per susceptible per year by a birth rate of re^{-aS} per susceptible per year, with a a positive constant. Then, in the absence of infection, the fox population is given by the first order differential equation

$$N' = N\left(re^{-aN} - \mu\right)$$

and equilibria of this equation are given by $N = 0$ and $re^{-aN} = \mu$, which reduces to $e^{aN} = r/\mu$, or

$$N = \frac{1}{a}\log\frac{r}{\mu}\ .$$

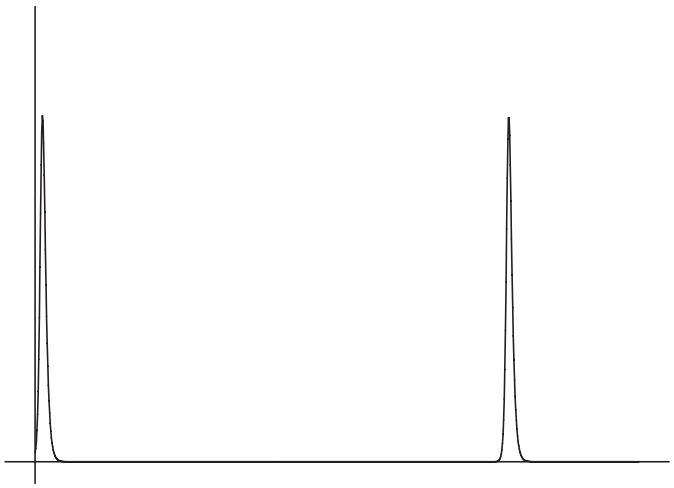

Fig. 2.9 I as a function of t (larger amplitude)

We omit the verification that the equilibrium $N = 0$ is unstable while the positive equilibrium $N = (1/a)\log(r/\mu)$ is asymptotically stable. Thus, the population has a carrying capacity given by

$$K = \frac{1}{a}\log\frac{r}{\mu}\ .$$

The model now becomes

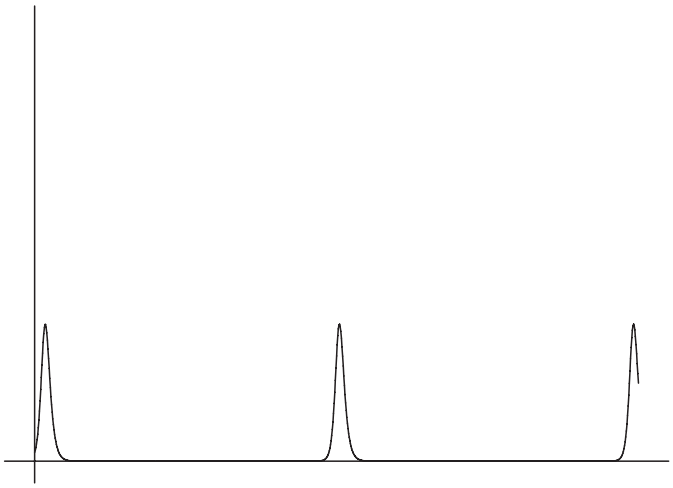

Fig. 2.10 I as a function of t (smaller amplitude)

$$S' = rSe^{-aS} - \beta SI - \mu S$$
$$I' = \beta SI - (\mu + \alpha)I \ .$$

We examine this by looking for equilibria and analyzing their stability. From the second equation, equilibria satisfy either $I = 0$ or $\beta S = \mu + \alpha$. If $I = 0$ the first equilibrium condition reduces to the same equation that determined the carrying capacity, and we have a disease-free equilibrium $S = K,\ I = 0$. If $\beta S = \mu + \alpha$ there is an endemic equilibrium with $\beta I + \mu = re^{-aS}$. A straightforward computation, which we shall not carry out here shows, that the disease-free equilibrium is asymptotically stable if $\mathcal{R}_0 = \beta K/(\mu + \alpha) < 1$ and unstable if $\mathcal{R}_0 > 1$, while the endemic equilibrium, which exists if and only if $\mathcal{R}_0 > 1$, is always asymptotically stable. Another way to express the condition for an endemic equilibrium is to say that the fox population density must exceed a threshold level K_T given by

$$K_T = \frac{\mu + \alpha}{\beta} \ .$$

With the parameter values we have been using, this gives a threshold fox density of $0.92\,\mathrm{fox/km^2}$. If the fox density is below this threshold value, the fox population will approach its carrying capacity and the disease will die out. Above the threshold density, rabies will persist and will regulate the fox population to a level below its carrying capacity. This level may be approached in an oscillatory manner for large $\mathcal{R}_0$.

2.4 Age of Infection Models

*2.4.1 The Basic SI*R Model*

The 1927 epidemic model of Kermack and McKendrick is considerably more general than what is usually called the Kermack–McKendrick model, which was analyzed in the first section. The general model described by Kermack and McKendrick included a dependence of infectivity on the time since becoming infected (age of infection). The 1932 and 1933 models of Kermack and McKendrick, which incorporated births and deaths, did not include this dependence. While age of infection models have not played a large role in studies of epidemics, they are very important in studies of HIV/AIDS. HIV/AIDS acts on a very long time scale and it is essential to include demographic effects (recruitment into and departure from a population of sexually active individuals). Also, the infectivity of HIV-positive people is high for a relatively short time after becoming infected, then very low for a long period, possibly several years, and then high shortly before developing into full-blown AIDS. Thus, the age of infection models described by Kermack and McKendrick for epidemics but not for endemic situations, have become important in endemic situations.

We will describe a general age of infection model and carry out a partial analysis; there are many unsolved problems in the analysis. We continue to let $S(t)$ denote the number of susceptibles at time t and $R(t)$ the number of members recovered with immunity, but now we let $I^*(t)$ denote the number of infected (but not necessarily infective) members.

We make the following assumptions:

1. The population has a birth rate $\Lambda(N)$, and a natural death rate μ giving a carrying capacity K such that $\Lambda(K) = \mu K, \Lambda'(K) < \mu$.
2. An average infected member makes $C(N)$ contacts in unit time of which S/N are with susceptibles. We define $\beta(N) = C(N)/N$ and it is reasonable to assume that $\beta'(N) \leq 0, C'(N) \geq 0$.
3. $B(\tau)$ is the fraction of infecteds remaining infective if alive when infection age is τ and $B_\mu(\tau) = e^{-\mu\tau}B(\tau)$ is the fraction of infecteds remaining alive and infected when infection age is τ. Let $\hat{B_\mu}(0) = \int_0^\infty B_\mu(\tau)d\tau$.
4. A fraction f of infected members recovers with immunity and a fraction $(1-f)$ dies of disease.
5. $\pi(\tau)$ with $0 \leq \pi(\tau) \leq 1$ is the infectivity at infection age τ; let $A(\tau) = \pi(\tau)B(\tau), A_\mu(\tau) = \pi(\tau)B_\mu(\tau), \hat{A_\mu}(0) = \int_0^\infty A_\mu(\tau)d\tau$.

In previous sections we have used $B(\tau) = e^{-\alpha\tau}$, which would give $B_\mu(\tau) = e^{-(\mu+\alpha)\tau}$. We let $i_0(t)$ be the number of new infecteds at time $t, i(t,\tau)$ be the number of infecteds at time t with infection age τ, and let $\phi(t)$ be the total infectivity at time t. Then

$$
\begin{aligned}
i(t,\tau) &= i_0(t-\tau)B_\mu(\tau), \quad 0 \le \tau \le t \\
i_0(t) &= S\beta(N)\phi(t)
\end{aligned}
$$

and

$$
\begin{aligned}
S' &= \Lambda(N) - \mu S - \beta(N)S\phi \\
I^*(t) &= \int_0^\infty i(t,\tau)d\tau \\
&= \int_0^\infty i_0(t-\tau)B_\mu(\tau)d\tau \\
&= \int_0^\infty \beta(N(t-\tau))S(t-\tau)\phi(t-\tau)B_\mu(\tau)d\tau \\
\phi(t) &= \int_0^\infty i_0(t-\tau)A_\mu(\tau)d\tau \\
&= \int_0^\infty \beta(N(t-\tau))S(t-\tau)\phi(t-\tau)A_\mu(\tau)d\tau \ .
\end{aligned}
$$

Differentiation of the equation for I^* gives three terms, including the rate of new infections and the rate of natural deaths. The third term gives the rate of recovery plus the rate of disease death as

$$
-\int_0^\infty \beta(N(t-\tau))S(t-\tau)\phi(t-\tau)e^{-\mu\tau}B'(\tau)d\tau \ .
$$

Thus the SI^*R model is

$$
\begin{aligned}
S' &= \Lambda(N) - \mu S - \beta(N)S\phi \\
\phi(t) &= \int_0^\infty \beta(N(t-\tau))S(t-\tau)\phi(t-\tau)A_\mu(\tau)d\tau \\
N'(t) &= \Lambda(N) - \mu N \\
&\quad + (1-f)\int_0^\infty \beta(N(t-\tau))S(t-\tau)\phi(t-\tau)e^{-\mu\tau}B'(\tau)d\tau \ .
\end{aligned}
\tag{2.27}
$$

Since I^* is determined when S, ϕ, N are known we have dropped the equation for I^* from the model, but it will be convenient to recall

$$
I^*(t) = \int_0^\infty \beta(N(t-\tau))S(t-\tau)\phi(t-\tau)B_\mu(\tau)d\tau \ .
$$

If $f = 1$ then $N(t)$ approaches the limit K, the model is asymptotically autonomous and its dimension may be reduced to two, replacing N by the constant K. We note, for future use, that

$$\hat{B_\mu}(0) = \int_0^\infty e^{-\mu\tau} B(\tau) d\tau \leq \int_0^\infty e^{-\mu\tau} d\tau = 1/\mu \ ,$$

so that

$$0 \leq 1 - \mu \hat{B_\mu}(0) \leq 1 \ .$$

We define $M = (1-f)(1-\mu\hat{B_\mu}(0)$, and $0 \leq M \leq 1$. We note, however, that if $f = 1$ then $M = 0$. We also have, using integration by parts,

$$-\int_0^\infty e^{-\mu\tau} B'(\tau) d\tau = 1 - \mu \hat{B_\mu}(0) \geq 0 \ .$$

If a single infective is introduced into a wholly susceptible population, making $K\beta(K)$ contacts in unit time, the fraction still infective at infection age τ is $B_\mu(\tau)$ and the infectivity at infection age τ is $A_\mu(\tau)$. Thus $\mathcal{R}_0$, the total number of secondary infections caused, is

$$\int_0^\infty K\beta(K) A_\mu(\tau) d\tau = K\beta(K)\hat{A_\mu}(0) \ .$$

Example. (Exposed periods) One common example of an age of infection model is a model with an exposed period, during which individuals have been infected but are not yet infected. Thus we may think of infected susceptibles going into an exposed class (E), proceeding from the exposed class to the infective class (I) at rate κE and out of the infective class at rate αI. Exposed members have infectivity 0 and infective members have infectivity 1. Thus $I^* = E + I$ and $\phi = I$.

We let $u(\tau)$ be the fraction of infected members with infection age τ who are not yet infective if alive and $v(\tau)$ the fraction of infected members who are infective if alive. Then the fraction becoming infective at infection age τ if alive is $\kappa u(\tau)$, and we have

$$\begin{aligned} u'(\tau) &= -\kappa u(\tau), \qquad u(0) = 1 \\ v'(\tau) &= \kappa u(\tau) - \alpha v(\tau) \quad v(0) = 0 \ . \end{aligned} \tag{2.28}$$

The solution of the first of the equations of (2.28) is

$$u(\tau) = e^{-\kappa\tau}$$

and substitution of this into the second equation gives

$$v'(\tau) = \kappa e^{-\kappa\tau} - \alpha v(\tau) \ .$$

When we multiply this equation by the integrating factor $e^{\alpha\tau}$ and integrate, we obtain the solution

$$v(\tau) = \frac{\kappa}{\kappa - \alpha}[e^{-\alpha\tau} - e^{-\kappa\tau}],$$

and this is the term $A_\mu(\tau)$ in the general model. The term $B(\tau)$ is $u(\tau)+v(\tau)$. Thus we have

$$\begin{aligned} A(\tau) &= \frac{\kappa}{\kappa-\alpha}[e^{-\alpha\tau} - e^{-\kappa\tau}] \\ B(\tau) &= \frac{\kappa}{\kappa-\alpha}e^{-\alpha\tau} - \frac{\alpha}{\kappa-\alpha}e^{-\kappa\tau} \\ e^{-\mu\tau}B'(\tau) &= -\frac{\alpha\kappa}{\kappa-\alpha}[e^{-(\mu+\alpha)\tau} - e^{-(\mu+\kappa)\tau}] \ . \end{aligned}$$

With these choices and the identifications

$$I = \phi, \qquad E = I^* - \phi$$

we may verify that the system (2.27) reduces to

$$\begin{aligned} S' &= \Lambda(N) - \beta(N)SI - \mu S \\ E' &= \beta(N)SI - \kappa E \\ I' &= \kappa E - (\mu+\alpha)I \\ N' &= \Lambda(N) - (1-f)\alpha I - \mu N \ , \end{aligned}$$

which is a standard $SEIR$ model.

For some diseases there is an asymptomatic period during which individuals have some infectivity rather than an exposed period. If the infectivity during this period is reduced by a factor ε, then the model can be described by the system

$$\begin{aligned} S' &= \Lambda(N) - \beta(N)S(I+\varepsilon E) - \mu S \\ E' &= \beta(N)S(I+\varepsilon E) - \kappa E \\ I' &= \kappa E - (\mu+\alpha)I \\ N' &= \Lambda(N) - (1-f)\alpha I - \mu N \ . \end{aligned}$$

This may be considered as an age of infection model with the same identifications of the variables and the same choice of $u(\tau), v(\tau)$ but with $A(\tau) = \varepsilon u(\tau) + v(\tau)$.

2.4.2 Equilibria

There is a disease-free equilibrium $S = N = K, \phi = 0$ of (2.27). Endemic equilibria (S, N, ϕ) are given by

$$\begin{aligned} \Lambda(N) &= \mu S + S\phi\beta(N) \\ S\beta(N)\hat{A}_\mu(0) &= 1 \\ \Lambda(N) &= \mu N + (1-f)(1-\mu\hat{B}_\mu(0))S\beta(N)\phi \ . \end{aligned}$$

If $f = 1$ the third condition gives $\Lambda(N) = \mu N$, which implies $N = K$. Then the second condition may be solved for S, after which the first condition may be solved for ϕ. Thus, there is always an endemic equilibrium.

If $f < 1$ the second of the equilibrium conditions gives

$$\phi = \frac{\hat{A}_\mu(0)}{M}[\Lambda(N) - \mu N].$$

Now substitution of the first two equilibrium conditions into the third gives an equilibrium condition for N, namely

$$\begin{aligned} (1-M)\Lambda(N) &= \mu N - \frac{\mu M}{\beta(N)\hat{A}_\mu(0)} \\ &= \mu N\left[1 - \frac{M}{C(N)\hat{A}_\mu(0)}\right]. \end{aligned} \tag{2.29}$$

If $\mathcal{R}_0 < 1$,

$$C(N)\hat{A}_\mu(0) \le C(K)\hat{A}_\mu(0) = \mathcal{R}_0 < 1$$

so that

$$1 - \frac{M}{C(N)\hat{A}_\mu(0)} < 1 - M \ .$$

Then we must have $\Lambda(N) < \mu N$. However, this would contradict the demographic condition $\Lambda(N) > \mu N, 0 < N < K$ imposed earlier. This shows that if $\mathcal{R}_0 < 1$ there is no endemic equilibrium.

If $\mathcal{R}_0 > 1$ for $N = 0$ the left side of (2.29) is non-negative while the right side is negative. For $N = K$ the left side of (2.29) is $\mu K(1-M)$ while the right side is

$$\mu K - \frac{M\mu K}{\mathcal{R}_0} > \mu K(1-M) \ .$$

This shows that there is an endemic equilibrium solution for N.

2.4.3 The Characteristic Equation

The linearization of (2.27) at an equilibrium (S, N, ϕ) is

$$
\begin{aligned}
x' &= -[\mu + \phi\beta(N)]x + [\Lambda'(N) - S\phi\beta'(N)]y - S\beta(N)z \\
y' &= [\Lambda'(N) - \mu]y \\
&\quad +(1-f)\int_0^\infty e^{-\mu\tau}B'(\tau)[\phi\beta(N)x(t-\tau) + S\phi\beta'(N)y(t-\tau) + S\beta(N)z(t-\tau)]d\tau \\
z(t) &= \int_0^\infty A_\mu(\tau)[\phi\beta(N)x(t-\tau) + S\phi\beta'(N)y(t-\tau) + S\beta(N)z(t-\tau)]d\tau \ .
\end{aligned}
$$

The condition that this linearization has solutions which are constant multiples of $e^{-\lambda\tau}$ is that λ satisfies a characteristic equation. The characteristic equation at an equilibrium (S, N, ϕ) is

$$
\det\begin{bmatrix} -[\lambda + \mu + \phi\beta(N)] & [\Lambda'(N) - S\phi\beta'(N)] & -S\beta(N) \\ -\phi\beta(N)Q(\lambda) & -[\lambda - \Lambda'(N) + \mu] - S\phi\beta'(N)Q(\lambda) & -S\phi\beta(N)Q(\lambda) \\ \phi\beta(N)\hat{A}_\mu(\lambda) & S\phi\beta'(N)\hat{A}_\mu(\lambda) & S\beta(N)\hat{A}_\mu(\lambda) - 1 \end{bmatrix} = 0
$$

with

$$
\begin{aligned}
\hat{A}_\mu(\lambda) &= \int_0^\infty e^{-\lambda\tau}A_\mu(\tau)d\tau \\
\hat{B}_\mu(\lambda) &= \int_0^\infty e^{-\lambda\tau}B_\mu(\tau)d\tau \\
Q(\lambda) &= (1-f)[1 - (\lambda+\mu)\hat{B}_\mu(\lambda)] \ .
\end{aligned}
$$

Here, the choice of $Q(\lambda)$ is motivated by the integration by parts formula

$$
\int_0^\infty e^{-(\lambda+\mu)\tau}B'(\tau)d\tau = -1 + \hat{B}_\mu(\lambda) \ .
$$

The characteristic equation then reduces to

$$
\begin{aligned}
&S\beta(N)\hat{A}_\mu(\lambda) + (1-f)\phi S\beta'(N)\hat{B}_\mu(\lambda) \\
&= 1 + \frac{f\phi\beta(N)}{\lambda+\mu} + \frac{(1-f)\phi P}{\lambda+\mu-\Lambda'(N)}[1 - \Lambda'(N)\hat{B}_\mu(\lambda)] \ , \qquad (2.30)
\end{aligned}
$$

where $P = \beta(N) + S\beta'(N) \geq 0$.

The characteristic equation for a model consisting of a system of ordinary differential equations is a polynomial equation. Now we have a transcendental characteristic equation, but there is a basic theorem that if all roots of the characteristic equation at an equilibrium have negative real part then the equilibrium is asymptotically stable [39, Chap. 4].

At the disease-free equilibrium $S = N = K, \phi = 0$ the characteristic equation is

$$
K\beta(K)\hat{A}_\mu(\lambda) = 1 \ .
$$

Since the absolute value of the left side of this equation is no greater than $K\beta(K)\hat{A}_\mu(0)$ if $\Re\lambda \geq 0$ the disease-free equilibrium is asymptotically stable

if and only if

$$\mathcal{R}_0 = K\beta(K)\hat{A}_\mu(0) < 1 \, .$$

2.4.4 The Endemic Equilibrium

In the analysis of the characteristic equation (2.30) it is helpful to make use of the following elementary result:

If $|P(\lambda)| \leq 1, \Re g(\lambda) > 0$ for $\Re\lambda \geq 0$, then all roots of the characteristic equation

$$P(\lambda) = 1 + g(\lambda)$$

satisfy $\Re\lambda < 0$.

To prove this result, we observe that if $\Re\lambda \geq 0$ the left side of the characteristic equation has absolute value at most 1 while the right side has absolute value greater than 1.

If $f = 1$, the characteristic equation reduces to

$$S\beta(N)\hat{A}_\mu(\lambda) = 1 + \frac{\phi\beta(N)}{\lambda + \mu} \, .$$

We have

$$|S\beta(N)\hat{A}_\mu(\lambda)| \leq S_\beta(N)\hat{A}_\mu(0) = 1$$

The term

$$\frac{\phi\beta(N)}{\lambda + \mu}$$

in (2.30) has positive real part if $\Re\lambda \geq 0$. It follows from the above elementary result that all roots satisfy $\Re\lambda < 0$, so that the endemic equilibrium is asymptotically stable. Thus all roots of the characteristic equation (2.30) have negative real part if $f = 1$.

The analysis if $f < 1$ is more difficult. The roots of the characteristic equation depend continuously on the parameters of the equation. In order to have a root with $\Re\lambda \geq 0$ there must be parameter values for which either there is a root at "infinity", or there is a root $\lambda = 0$ or there is a pair of pure imaginary roots $\lambda = \pm iy, y > 0$. Since the left side of (2.30) approaches 0 while the right side approaches 1 as $\lambda \to \infty, \Re\lambda \geq 0$, it is not possible for a root to appear at "infinity". For $\lambda = 0$, since $S\beta(N)\hat{A}_\mu(0) = 1$ and $\beta'(N) \leq 0$ the left side of (2.30) is less than 1 at $\lambda = 0$, while the right side is greater than 1 since

$$1 - \Lambda'(N)\hat{B}_\mu(0) > 1 - \Lambda'(N)/\mu > 0$$

if $\Lambda'(N) < \mu$. This shows that $\lambda = 0$ is not a root of (2.30), and therefore that all roots satisfy $\Re\lambda < 0$ unless there is a pair of roots $\lambda = \pm iy, y > 0$. According to the Hopf bifurcation theorem [20] a pair of roots $\lambda = \pm iy, y > 0$

indicates that the system (2.27) has an asymptotically stable periodic solution and there are sustained oscillations of the system.

A somewhat complicated calculation using the fact that since $B_\mu(\tau)$ is monotone non-increasing,

$$\int_0^\infty B_\mu(\tau)\sin y\tau dy \geq 0$$

for $0 \leq y < \infty$ shows that the term

$$\frac{(1-f)\phi P}{\lambda+\mu-\Lambda'(N)}\cdot[1-\Lambda'(N)\hat{B_\mu}(\lambda)]$$

in (2.30) has positive real part at least if

$$-\mu \leq \Lambda'(N) \leq \mu\ .$$

Thus if $-\mu \leq \Lambda'(N) \leq \mu$, instability of the endemic equilibrium is possible only if the term

$$(1-f)\phi S\beta'(N)\hat{B_\mu}(iy)$$

in (2.30) has negative real part for some $y > 0$. This is not possible with mass action incidence, since $\beta'(N) = 0$; thus with mass action incidence the endemic equilibrium of (2.27) is always asymptotically stable. Since $\beta'(N) \leq 0$, instability requires

$$\Re\hat{B_\mu}(iy) = \int_0^\infty B_\mu(\tau)\cos y\tau d\tau < 0$$

for some $y > 0$. If the function $B(\tau)$ is non-increasing and convex, that is, if $B'(\tau) \leq 0, B''(\tau) \geq 0$, then it is easy to show using integration by parts that

$$\int_0^\infty B_\mu(\tau)\cos y\tau d\tau \geq 0$$

for $0 < y < \infty$. Thus if $B(\tau)$ is convex, which is satisfied, for example, by the choice

$$B(\tau) = e^{-\alpha\tau}$$

the endemic equilibrium of (2.22) is asymptotically stable if $-\mu \leq \Lambda'(N) \leq \mu$.

There are certainly less restrictive conditions which guarantee asymptotic stability. However, examples have been given [36,37] of instability, even with $f = 0, \Lambda'(N) = 0$, where constant infectivity would have produced asymptotic stability. Their results indicate that concentration of infectivity early in the infected period is conducive to such instability. In these examples, the instability arises because a root of the characteristic equation crosses the imaginary axis as parameters of the model change, giving a pure imaginary root of the characteristic equation. This translates into oscillatory solutions

of the model. Thus infectivity which depends on infection age can cause instability and sustained oscillations.

2.4.5 An SI^*S Model

In order to formulate an SI^*S age of infection model we need only take the SI^*R age of infection model (2.22) and move the recovery term from the equation for R (which was not listed explicitly in the model) to the equation for S. We obtain the model

$$\begin{aligned} S'(t) &= \Lambda(N) - \mu S - \beta(N)S\phi \\ &\quad - f\int_0^\infty \beta(N(t-\tau))S(t-\tau)\phi(t-\tau)e^{-\mu\tau}B'(\tau)d\tau \qquad (2.31) \\ \phi(t) &= \int_0^\infty \beta(N(t-\tau))S(t-\tau)\phi(t-\tau)A_\mu(\tau)d\tau \qquad (2.32) \\ N'(t) &= \Lambda(N) - \mu N \\ &\quad + (1-f)\int_0^\infty \beta(N(t-\tau))S(t-\tau)\phi(t-\tau)e^{-\mu\tau}B'(\tau)d\tau\ . \end{aligned}$$

Although we will not carry out any analysis of this model, it may be attacked using the same approach as that used for (2.27). It may be shown that if $\mathcal{R}_0 = K\beta(K)\hat{A}_\mu(0) < 1$ the disease-free equilibrium is asymptotically stable. If $\mathcal{R}_0 > 1$ there is an endemic equilibrium and the characteristic equation at this equilibrium is

$$\begin{aligned} &S\beta(N)\hat{A}_\mu(\lambda) + (1-f)\phi S\beta'(N)\hat{B}_\mu(\lambda) \\ &= 1 + f\phi\beta(N)\hat{B}_\mu(\lambda) + \frac{(1-f)\phi P}{\lambda+\mu-\Lambda'(N)}\cdot[1-\Lambda'(N)\hat{B}_\mu(\lambda)] \qquad (2.33) \end{aligned}$$

where $P = \beta(N) + S\beta'(N) \geq 0$.

Many diseases, including most strains of influenza, impart only temporary immunity against reinfection on recovery. Such disease may be described by SIS age of infection models, thinking of the infected class I^* as comprised of the infective class I together with the recovered and immune class R. In this way, members of R neither spread or acquire infection. We assume that immunity is lost at a proportional rate κ.

We let $u(\tau)$ be the fraction of infected members with infection age τ who are infective if alive and $v(\tau)$ the fraction of infected members who are not recovered and still immune if alive. Then the fraction becoming immune at infection age τ if alive is $\alpha u(\tau)$, and we have

$$u'(\tau) = -\alpha u(\tau), \qquad u(0) = 1 \tag{2.34}$$
$$v'(\tau) = \alpha u(\tau) - \kappa v(\tau) \quad v(0) = 0 \, .$$

These equations are the same as (2.28) obtained in formulating the $SEIR$ model with α and κ interchanged. Thus we may solve to obtain

$$u(\tau) = e^{-\alpha\tau}$$
$$v(\tau) = \frac{\alpha}{\kappa - \alpha}[e^{-\alpha\tau} - e^{-\kappa\tau}] \, .$$

We take $B(\tau) = u(\tau)+v(\tau), A(\tau) = u(\tau)$. Then if we define $I = \phi, R = I^* - \phi$, the model (2.31) is equivalent to the system

$$S' = \Lambda(N) - \beta(N)SI - \mu S + \kappa R$$
$$I' = \beta(N)SI - (\mu + \alpha)I$$
$$R' = f\alpha E - (\mu + \kappa)R$$
$$N' = \Lambda(N) - (1 - f)\alpha I - \mu N,$$

which is a standard $SIRS$ model.

If we assume that, instead of an exponentially distributed immune period, that there is an immune period of fixed length ω we would again obtain $u(\tau) = e^{-\alpha\tau}$, but now we may calculate that

$$v(\tau) = 1 - e^{-\alpha\tau}, (\tau \leq \omega), \quad v(\tau) = e^{-\alpha\tau}(e^{\alpha\omega} - 1), (\tau > \omega) \, .$$

To obtain this, we note that

$$v'(\tau) = \alpha u(\tau), (\tau \leq \omega), \quad v'(\tau) = \alpha u(\tau) - \alpha u(\tau - \omega), (\tau > \omega) \, .$$

From these we may calculate $A(\tau), B(\tau)$ for an SI^*S model. Since it is known that the endemic equilibrium for an $SIRS$ model with a fixed removed period can be unstable [19], this shows that (2.33) may have roots with non-negative real part and the endemic equilibrium of an SI^*S age of infection model is not necessarily asymptotically stable.

The SI^*R age of infection model is actually a special case of the SI^*S age of infection model. We could view the class R as still infected but having no infectivity, so that $v(\tau) = 0$. The underlying idea is that in infection age models we divide the population into members who may become infected and members who can not become infected, either because they are already infected or because they are immune.

2.4.6 An Age of Infection Epidemic Model

We conclude by returning to the beginning, namely an infection age epidemic model closely related to the original Kermack–McKendrick epidemic model [21]. We simply remove the birth and natural death terms from the SI^*R model (2.27). The result is

$$\begin{aligned} S' &= -\beta(N)S\phi \\ \phi(t) &= \int_0^\infty \beta(N(t-\tau))S(t-\tau)\phi(t-\tau)A(\tau)d\tau \\ N'(t) &= (1-f)\int_0^\infty \beta(N(t-\tau))S(t-\tau)\phi(t-\tau)B'(\tau)d\tau \end{aligned}$$

which we may rewrite as

$$\begin{aligned} S' &= -\beta(N)S\phi \\ \phi(t) &= \int_0^\infty [-S'(t-\tau)]A(\tau)d\tau \\ N'(t) &= (1-f)\int_0^\infty [-S'(t-\tau]B'(\tau)d\tau \ . \end{aligned} \tag{2.35}$$

Then integration of the equation for N with respect to t from 0 to ∞ gives

$$\begin{aligned} K - N_\infty &= (1-f)\int_0^\infty [\int_0^\infty [-S'(t-\tau)]B'(\tau)d\tau dt \\ &= (1-f)\int_0^\infty [\int_0^\infty [-S'(t-\tau)]dtB'(\tau)d\tau \\ &= (1-f)\int_0^\infty [S(-\tau) - S_\infty]B'(\tau)d\tau \\ &= (1-f)(K - S_\infty) \ , \end{aligned}$$

which is the same relation (2.8) obtained for the model (2.6). In this calculation we use the initial data to give $S(-\tau) = K$ and

$$\int_0^\infty B'(\tau)d\tau = B(\infty) - B(0) = -1 \ .$$

The argument that $S_\infty > 0$ for the model (2.35) is analogous to the argument for (2.10). From (2.35) we have

$$-\frac{S'(t)}{S(t)} = \beta(N(t))\int_0^\infty [-S'(t-\tau)]A(\tau)d\tau$$

and integration with respect to t from 0 to ∞ gives

$$\begin{aligned}\log\frac{S(0)}{S_\infty} &= \int_0^\infty \beta(N(t))\int_0^\infty [-S'(t-\tau)]A(\tau)d\tau dt\\ &= \int_0^\infty A(\tau)\int_0^\infty \beta(N(t))[-S'(t-\tau)]dtd\tau\\ &\le \beta(0)\int_0^\infty A(\tau)\int_0^\infty [-S'(t-\tau)]dtd\tau\\ &= \beta(0)\int_0^\infty A(\tau)[S(-\tau)-S_\infty]dsd\tau\\ &= \beta(0)(K-S_\infty)\int_0^\infty A(\tau)d\tau\end{aligned}$$

and this shows that $S_\infty > 0$. We recall that we are assuming here that $\beta(0)$ is finite; in other words we are ruling out standard incidence. It is possible to show that S_∞ can be zero only if $N \to 0$ and $\int_0^K \beta(N)dN$ diverges. However, from (2.8) we see that this is possible only if $f = 0$. If there are no disease deaths, so that the total population size N is constant, or if β is constant (mass action incidence), the above integration gives the final size relation

$$\log\frac{S(0)}{S_\infty} = \mathcal{R}_0\left[1-\frac{S_\infty}{K}\right].$$

We may view the epidemic management model (2.13) as an age of infection model. We define $I^* = E+Q+I+J$, and we need only calculate the kernels $A(\tau), B(\tau)$. We let $u(\tau)$ denote the number of members of infection age τ in E, $v(\tau)$ the number of members of infection age τ in Q, $w(\tau)$ the number of members of infection age τ in I, and $z(\tau)$ the number of members of infection age τ in J. Then (u, v, w, z) satisfies the linear homogeneous system with constant coefficient

$$\begin{aligned}u'(\tau) &= -(\kappa_1+\gamma_1)u(\tau)\\ v'(\tau) &= \gamma_1 u(\tau)-\kappa_2 v(\tau)\\ w'(\tau) &= \kappa_1 u(\tau)-\alpha_1 w(\tau)-\gamma_2 w(\tau)\\ z'(\tau) &= \gamma_2 w(\tau)+\kappa_2 v(\tau)-\alpha_2 z(\tau)\end{aligned}$$

with initial conditions $u(0) = 1, v(0) = 0, w(0) = 0, z(0) = 0$. This system is easily solved recursively, and then the system (2.13) is an age of infection epidemic model with

$$A(\tau) = \varepsilon_E u(\tau)+\varepsilon_E\varepsilon_Q v(\tau)+w(\tau)+\varepsilon_J z(\tau), B(\tau) = u(\tau)+v(\tau)+w(\tau)+z(\tau)\ .$$

In particular, it now follows from the argument carried out just above that $S_\infty > 0$ for the model (2.13). The proof is less complicated technically than the proof obtained for the specific model (2.13). The generalization to age

of infection models both unifies the theory and makes some calculations less complicated.

References

1. R.M. Anderson, H.C. Jackson, R.M. May, A.M. Smith: Population dynamics of fox rabies in Europe. Nature, **289**, 765–771 (1981)
2. R.M. Anderson, R.M. May: Infectious Diseases of Humans. Oxford Science Publications, Oxford (1991)
3. F. Brauer, C. Castillo-Chavez: Mathematical Models in Population Biology and Epidemiology. Springer, New York (2001)
4. S. Busenberg, K.L. Cooke: Vertically Transmitted Diseases: Models and Dynamics. Biomathematics, Vol. 23. Springer, Berlin Heidelberg New York (1993)
5. C. Castillo-Chavez with S. Blower, P. van den Driessche, D. Kirschner, A.A. Yakubu (eds.): Mathematical Approaches for Emerging and Reemerging Infectious Diseases: An Introduction. Springer, Berlin Heidelberg New York (2001)
6. C. Castillo-Chavez, with S.Blower, P. van den Driessche, D. Kirschner, A.A. Yakubu (eds.): Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods and Theory. Springer, Berlin Heidelberg New York (2001)
7. C. Castillo-Chavez, K.L. Cooke, W. Huang, S.A. Levin: The role of long incubation periods in the dynamics of HIV/AIDS. Part 1: Single populations models. J. Math. Biol., **27**, 373–98 (1989)
8. C. Castillo-Chavez, H.R. Thieme: Asymptotically autonomous epidemic models. In: O. Arino, D. Axelrod, M. Kimmel, M. Langlais (eds.) Mathematical Population Dynamics: Analysis of Heterogeneity, Vol. 1: Theory of Epidemics. Wuerz, Winnipeg, pp. 33–50 (1993)
9. D.J. Daley, J. Gani: Epidemic Modelling: An Introduction. Cambridge Studies in Mathematical Biology, Vol. 16. Cambridge University Press, Cambridge (1999)
10. K. Dietz: Overall patterns in the transmission cycle of infectious disease agents. In: R.M. Anderson, R.M. May (eds.) Population Biology of Infectious Diseases. Life Sciences Research Report, Vol. 25. Springer, Berlin Heidelberg New York, pp. 87–102 (1982)
11. K. Dietz: The first epidemic model: a historical note on P.D. En'ko. Aust. J. Stat., **30**, 56–65 (1988)
12. S. Ellner, R. Gallant, J. Theiler: Detecting nonlinearity and chaos in epidemic data. In: D. Mollison (ed.) Epidemic Models: Their Structure and Relation to Data. Cambridge University Press, Cambridge, pp. 229–247 (1995)
13. A. Gumel, S. Ruan, T. Day, J. Watmough, P. van den Driessche, F. Brauer, D. Gabrielson, C. Bowman, M.E. Alexander, S. Ardal, J. Wu, B.M. Sahai: Modeling strategies for controlling SARS outbreaks based on Toronto, Hong Kong, Singapore and Beijing experience. Proc. R. Soc. Lond. B Biol. Sci., **271**, 2223–2232 (2004)
14. J.A.P. Heesterbeek, J.A.J. Metz: The saturating contact rate in marriage and epidemic models. J. Math. Biol., **31**: 529–539 (1993)
15. H.W. Hethcote: Qualitative analysis for communicable disease models. Math. Biosci., **28**, 335–356 (1976)
16. H.W. Hethcote: An immunization model for a hetereogeneous population. Theor. Popul. Biol., **14**, 338–349 (1978)
17. H.W. Hethcote: The mathematics of infectious diseases. SIAM Rev., **42**, 599–653 (2000)
18. H.W. Hethcote, S.A. Levin: Periodicity in epidemic models. In: S.A. Levin, T.G. Hallam, L.J. Gross (eds.) Applied Mathematical Ecology. Biomathematics, Vol. 18. Springer, Berlin Heidelberg New York, pp. 193–211 (1989)

19. H.W. Hethcote, H.W. Stech, P. van den Driessche: Periodicity and stability in epidemic models: a survey. In: S. Busenberg, K.L. Cooke (eds.) Differential Equations and Applications in Ecology, Epidemics and Population Problems. Academic, New York, pp. 65–82 (1981)
20. E. Hopf: Abzweigung einer periodischen Lösungen von einer stationaren Lösung eines Differentialsystems. Berlin Math-Phys. Sachsiche Akademie der Wissenschaften, Leipzig, **94**, 1–22 (1942)
21. W.O. Kermack, A.G. McKendrick: A contribution to the mathematical theory of epidemics. Proc. R. Soc. Lond. B Biol. Sci., **115**, 700–721 (1927)
22. W.O. Kermack, A.G. McKendrick: Contributions to the mathematical theory of epidemics, part. II. Proc. R. Soc. Lond. B Biol. Sci., **138**, 55–83 (1932)
23. W.O. Kermack, A.G. McKendrick: Contributions to the mathematical theory of epidemics, part. III. Proc. R. Soc. Lond. B Biol. Sci., **141**, 94–112 (1932)
24. L. Markus: Asymptotically autonomous differential systems. In: S. Lefschetz (ed.) Contributions to the Theory of Nonlinear Oscillations III. Annals of Mathematics Studies, Vol. 36. Princeton University Press, Princeton, NJ, pp. 17–29 (1956)
25. J. Mena-Lorca, H.W. Hethcote: Dynamic models of infectious diseases as regulators of population size. J. Math. Biol., **30**, 693–716 (1992)
26. W.H. McNeill: Plagues and Peoples. Doubleday, New York (1976)
27. W.H. McNeill: The Global Condition. Princeton University Press, Princeton, NJ (1992)
28. L.A. Meyers, B. Pourbohloul, M.E.J. Newman, D.M. Skowronski, R.C. Brunham: Network theory and SARS: predicting outbreak diversity. J. Theor. Biol., **232**, 71–81 (2005)
29. D. Mollison (ed.): Epidemic Models: Their Structure and Relation to Data. Cambridge University Press, Cambridge (1995)
30. M.E.J. Newman: The structure and function of complex networks. SIAM Rev., **45**, 167–256 (2003)
31. G.F. Raggett: Modeling the Eyam plague. IMA J., **18**, 221–226 (1982)
32. H.E. Soper: Interpretation of periodicity in disease prevalence. J. R. Stat. Soc. B, **92**, 34–73 (1929)
33. S.H. Strogatz: Exploring complex networks. Nature, **410**, 268–276 (2001)
34. H.R. Thieme: Asymptotically autonomous differential equations in the plane. Rocky Mt. J. Math., **24**, 351–380 (1994)
35. H.R. Thieme: Mathematics in Population Biology. Princeton University Press, Princeton, NJ (2003)
36. H.R. Thieme, C. Castillo-Chavez: On the role of variable infectivity in the dynamics of the human immunodeficiency virus. In: C. Castillo-Chavez (ed.) Mathematical and Statistical Approaches to AIDS Epidemiology. Lecture Notes in Biomathematics, Vol. 83. Springer, Berlin Heidelberg New York, pp. 200–217 (1989)
37. H.R. Thieme, C. Castillo-Chavez: How may infection-age dependent infectivity affect the dynamics of HIV/AIDS? SIAM J. Appl. Math., **53**, 1447–1479 (1989)
38. P. van den Driessche, J. Watmough: Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci., **180**, 29–48 (2002)
39. G.F. Webb: Theory of Nonlinear Age-Dependent Population Dynamics. Marcel Dekker, New York (1985)