Texts in Applied Mathematics 61

Maia Martcheva

An Introduction to Mathematical Epidemiology

Springer

Texts in Applied Mathematics

Volume 61

More information about this series at http://www.springer.com/series/1214

Maia Martcheva

An Introduction to Mathematical Epidemiology

Springer

Maia Martcheva
Department of Mathematics
University of Florida
Gainesville, FL

ISSN 0939-2475 ISSN 2196-9949 (electronic)
Texts in Applied Mathematics
ISBN 978-1-4899-7611-6 ISBN 978-1-4899-7612-3 (eBook)
DOI 10.1007/978-1-4899-7612-3

Library of Congress Control Number: 2015947259

Mathematics Subject Classification (2010): 92D30, 34D20, 35F50

Springer New York Heidelberg Dordrecht London

Printed on acid-free paper

Springer Science+Business Media LLC New York is part of Springer Science+Business Media (www.springer.com)

Preface

This book is an introductory text to the methods and tools that are nowadays widely used and accepted in the mathematical epidemiology literature. It is intended to start from a beginner level and accelerate to research level. It targets upper undergraduate mathematics students, and mathematics, physics, and engineering graduate students. The book will also be suitable for mathematics researchers who wish to build a background and advance to research level in mathematical epidemiology. The book is expected to be useful to mathematical epidemiologists as a reference text.

Chapter 1 discusses some historical aspects of modeling infectious diseases. It also introduces a number of epidemiological concepts with their definitions. Finally, Chap. 1 includes general ideas about the modeling process, how we go from a biological question to a mathematical model, and how we answer biological questions based on conclusions from the model. Chapter 2 introduces the basic epidemic models without demography as well as a comparison to data. Simple single-equation epidemic models are analyzed. Chapter 3 deals with how and when demographic variables are included in epidemic models. It also includes complete analysis of the basic SIR model with demography. Analysis of planar systems is presented. Hopf bifurcation is introduced and applied to a model with saturating incidence. Chapter 4 brings in vector-borne diseases and treats their modeling in the context of ordinary differential equations (ODEs) and delay-differential equations. Chapter 5 is devoted to building more complex models with various components. Most of the widespread techniques for the computation of the reproduction number are introduced and illustrated on examples. Chapter 6 discusses statistical techniques for fitting models to data and for selecting the model that best represents the data. The first six chapters use standard and relatively elementary mathematical techniques and may be appropriate for a more general audience.

Chapters 7–11 treat more advanced epidemiological models and build appropriate mathematical techniques that are more involved. Chapter 7 is devoted to the mathematical analysis of complex epidemiological models. Global stability is considered via Lyapunov functions. The possibility of multiple equilibria via backward bifurcation is also introduced. Chapter 8 is devoted to multistrain interactions. The chapter begins by establishing the competitive exclusion principle. Furthermore,

it introduces mechanisms for coexistence. Analysis of multistrain models is illustrated. Chapter 9 is devoted to modeling control strategies in the context of single-strain and multistrain diseases. Herd immunity and proportion vaccinated are introduced. The chapter discusses the phenomenon of strain replacement. Optimal control techniques are presented and applied to specific examples. Chapter 10 introduces the most basic ecological models such as predation and competition and includes the spread of a disease in animal populations subject to predation or competition. Complex dynamical behavior such as chaos is shown in a three-dimensional system of ODEs. Chapter 11 focuses on zoonotic diseases in general and avian influenza in particular. Basic models of avian influenza are introduced and compared to data. Control strategies are evaluated. Nonautonomous modeling is treated.

Chapters 16–14 are devoted to partial differential equation (PDE) age-structured epidemic modeling. Chapter 16 introduces host–age structured models. After a brief introduction of age-structured population models, the chapter proceeds by discussing age-structured SIS (susceptible–infected–susceptible) and SIR (susceptible–infected–removed) models. Chapter 13 introduces a basic SI model with age since infection. Pease's influenza model is also discussed, and oscillations are obtained. Chapter 14 is devoted to immuno-epidemiological modeling. It includes immunological modeling, linking within-host models to epidemiological models and computation of relevant reproduction numbers. Chapter 15 is devoted to spatial aspects of epidemiology. Multipatch models with Lagrangian and Eulerian movement are considered. Furthermore, simple diffusion models are introduced, and a rabies epidemic model with diffusion is discussed. Chapter 13 introduces the basic epidemic models in a discrete setting. It also includes tools for local analysis of discrete dynamical systems. More complex discrete models are built, and the next-generation approach for the computation of the reproduction number in discrete settings is discussed.

This book is intended as a comprehensive text on mathematical epidemiology. However, it does not include some important topics that are an integral part of the subject. Two important topics that are missing are stochastic epidemic modeling and network disease modeling. These were not included to keep the length of the book within limits and because of the limited expertise of the author in those specific topics. There are several books that focus on these topic separately and involve epidemic modeling.

Chapters from this book can be used as an upper undergraduate or graduate text, as well as for summer courses. Chapters 1–6 are appropriate for an undergraduate course on mathematical epidemiology or as a supplemental text to an undergraduate mathematical biology class. Chapters 2–11 and 16 are appropriate for an introductory graduate class on mathematical epidemiology. Chapters 16–15 are more advanced and are appropriate for a PDE-focused graduate mathematical biology class. Exercises are provided after each chapter to help the reader understand and retain the material as well as to develop necessary skills to advance to research.

Acknowledgments

This book could not have been written without the help and support of my family and colleagues. First and foremost I am in debt to my family for their continuous support. My husband, Tzvetan, and my son, Theodore, were very supportive and patiently endured my preoccupation while this book was being written.

Next, I am very much in debt to my former postdoctoral associate Necibe Tuncer, who helped immensely with this book by writing necessary MATLAB code and proofreading the book. Without her moral and practical support, this book would not have reached its current state and would not have contained useful code. Chapters of this book were proofread and commented on by my students, particularly Hayriye Gulbudak.

Next, I am much in debt to several anonymous reviewers among my colleagues, who worked tirelessly on the book and provided many useful comments that improved the book. Some of these reviewers also edited the book for correctness and language. I am deeply thankful for the enormous support that I received.

I am also in debt to my editor, Donna Chernyk, who checked on my progress regularly and helped me stay on track, and the National Science Foundation for its continuous support.

Finally, I would like to thank everyone who feels that they contributed to this book being written and has not been explicitly mentioned. Without the generous support I received, I would have not been able to bring this book to completion.

Edgewater, NJ, USA Maia Martcheva
January 2015

Contents

Chapter 1
Introduction

1.1 Epidemiology

Epidemiology is the subject that studies the patterns of health and illness and associated factors at the population level. The word "epidemiology" is derived from the Greek terms *epi*, which means "upon," *demos*, which means " people," and *logos*, which means "study." This etymology implies that the subject of epidemiology applies only to human populations. The role of father of epidemiology is often assigned to the Greek physician Hippocrates (460–377 B.C.E.), who described the connection between disease and environment [83]. The term "epidemiology" appears to have first been used to describe the study of epidemics in 1802 by the Spanish physician de Villalba in *Epidemiologia Espanola* [30]. Until the twentieth century, epidemiological studies were mostly concerned with infectious diseases. Nowadays, the leading causes of deaths worldwide are diseases such as stroke and coronary heart disease [132], positioning diseases that do not transmit from one person to another as a central concern of epidemiology. Among infectious diseases, those that dominate worldwide as a cause of death include lower respiratory infections (such as pneumonia) and HIV. In this book, we will be concerned with mathematical modeling of infectious diseases.

1.2 Classification of Infectious Diseases

An *infectious disease* is a clinically evident illness resulting from the presence of a pathogenic microbial agent. The microbial agent causing the disease can be bacterial, viral, fungal, parasitic, or it can be toxic proteins, called prions. Infectious diseases caused by bacteria include tuberculosis and pneumonia; viral diseases include HIV and influenza; the most widespread fungal diseases are dermatomycoses; parasitic infections are caused by macroparasites such as protozoa, helminths, trematodes, and cestodes. An example of a prion-caused disease is

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_1

Creutzfeldt–Jakob disease. *Communicable diseases* are infectious diseases that can be transmitted from one infectious person to another, directly or indirectly. Often, we do not make a distinction between infectious diseases and communicable diseases, since many of the infectious diseases are in fact communicable diseases. However, there are diseases that are infectious but not communicable. Tetanus is an example of such a disease. *Transmittable diseases* are infectious diseases that can be transmitted from one person to another through unnatural routes. For instance, Creutzfeldt–Jakob disease can be passed from one patient to another through surgical instruments or transplants. Nonetheless, the distinction between infectious diseases, communicable diseases, and transmittable diseases is subtle, and infectious diseases are often called communicable diseases or transmittable diseases because of their potential to be transmitted from one person to another.

Transmission of infectious diseases may occur through a variety of pathways. According to the means of transmission, infectious diseases are classified as follows:

- **Person-to-person transmitted diseases** are diseases that require **direct** or **indirect contact**. Direct contact includes touching or sexual contact. Diseases that are transmitted through sexual contact are called *sexually transmitted diseases*. Sexually transmitted diseases include HIV, gonorrhea, and syphilis. Indirect contact includes exchange of an infected object, blood, or other body fluids. Influenza can be transmitted through indirect contact.
- **Airborne transmission** occurs on inhalation of infected air. Airborne transmitted diseases include influenza, smallpox, measles, chickenpox, and tuberculosis.
- **Food- and waterborne diseases** are transmitted through ingestion of contaminated food or water. Cholera is a waterborne disease. Foodborne diseases include salmonella and stomach flu.
- **Vector-borne diseases** are transmitted by a vector, most often an arthropod such as a mosquito or tick, or a mollusk such as a snail. Examples of vector-borne diseases are malaria, dengue, and West Nile virus, which are transmitted by mosquitoes.
- **Vertical transmission** occurs when a disease is transmitted through the placenta from a mother to a child before or at birth. Examples of such diseases are HIV, hepatitis B, syphilis, rubella, and herpes simplex virus.

For modeling purposes, we distinguish four types of transmission: *direct*, when the causative pathogen is transmitted from one person to another; *vector-transmitted*, when the causative agent is transmitted from a vector to a human; *environmental transmission*, when a human becomes infected through contact with a pathogen present in the environment; and *vertical*, when the pathogen is transmitted from mother to child at birth. Person-to-person and airborne diseases are usually modeled as directly transmitted when transmission occurs through contact between one person and another. What constitutes a "contact" sufficient for transmission in these diseases depends on the specific disease. In sexually transmitted diseases, sexual contact is necessary, while airborne diseases, which are often modeled as directly transmitted, require a certain degree of physical proximity without the necessity of touching. Modeling vector-borne transmission requires the inclusion of the

dynamics of the vector in addition to the dynamics of the infected individuals. Environmentally transmitted diseases are typically modeled by separately modeling the dynamics of the virus in the environment and the transmission that occurs on contact between an individual and the free pathogen.

A pathogen reservoir is an ecological niche in which a pathogen lives and multiplies. Such a reservoir plays a significant role in the spread of the pathogen. According to their reservoir, the microbial agents are classified as **human**, **animal**, and **environmental**. Human pathogens circulate mostly among humans, and humans play a role in their transmission. Animal pathogens have vertebrate animals as a reservoir, and circulate primarily among animals. Epidemiologically, this is significant, because many such pathogens adapt to infect humans through animal-to-human transmission. Infections that spread from vertebrate animals to humans are called **zoonoses**. Environmental pathogens multiply primarily in the environment (typically water and soil) and spread from there to animal and human populations.

Many infectious diseases have more than one pathway of transmission. For instance, HIV is primarily transmitted through sexual contact, but it can also be transmitted by blood transfusion or needle sharing. Furthermore, HIV can be transmitted vertically at birth from an infected mother to her child. Avian influenza H5N1 is primarily transmitted through direct contact with infected poultry and rarely directly from human to human. However, significant evidence now exists that H5N1 can persist in the environment, and the environmental route of transmission is gaining more importance.

1.3 Basic Definitions in the Epidemiology of Infectious Diseases

There are a number of concepts in epidemiology strictly related to infectious diseases. These concepts play an important role in the construction of mathematical models by adding various features to the model. Some of the most widely used concepts are listed below. Others will be introduced as new models are discussed.

- **Exposed Individuals.** When a healthy individual who is vulnerable to contracting a disease makes a potentially disease-transmitting contact, that individual becomes *exposed*. Exposed individuals may or may not develop the disease. These individuals are typically not infectious. In mathematical models, we often assume that all exposed individuals eventually develop the disease.
- **Infected and Infectious Individuals.** If the pathogen establishes itself in an exposed individual, then that individual becomes *infected*. Infected individuals who can transmit the disease are called *infectious*. Infected individuals may not be infectious during the entire time of being infected.
- **Latent Individuals.** These are individuals that are infected but not yet infectious. The *latent period* is defined as the time from infection to when the host is able to transmit the infectious agent to another individual.
- **Incubation Period.** The incubation period is the period between exposure to an infectious agent and the onset of symptoms of the disease. In infectious diseases, the incubation period is the time required for the infectious agent to multiply

to a threshold necessary to produce symptoms or laboratory evidence of infection. The incubation period does not necessarily coincide with the latent period. For instance, in influenza, individuals become infectious approximately one day *before* they exhibit visible flu symptoms.

- **Incidence.** Incidence is defined as the number of individuals who become ill during a specified interval of time (e.g., one year). Sometimes, incidence is the number of individuals who become ill during a specified interval of time divided by the total population. In most cases, incidence is determined from the number of clinical cases, which underestimates the true incidence, since it ignores the subclinical cases.
- **Prevalence.** The prevalence of a disease is the number of people who have the disease at a specific time. Sometimes, prevalence is defined as the number of people who have the disease at a specific time divided by the total population size.
- **Case Fatality Proportion (CFP).** The case fatality proportion is given as the ratio of people who die of a disease to those who contract it. For instance, as of June 27, 2014, 667 people have been diagnosed with H5N1 avian influenza, and 393 of them have died. The CFP is 0.59.
- **Disease-Induced Mortality.** Disease-induced mortality is the number of people who have died from the disease in one unit of time (e.g., one year) divided by the entire population.

This list is by no means exhaustive. More complete lists of terms used in infectious disease epidemiology can be found in the many excellent books on this subject (e.g., [56]).

1.4 Historical Remarks on Infectious Diseases and Their Modeling

The first significant epidemic described by historians was the plague of Athens, which struck the city of Athens in 430–426 B.C.E. The most precise description of that plague was provided by the scientific historian Thucydides (460–400 B.C.E.) in his *History of the Peloponnesian War*. His description is based on personal experience and includes symptoms, progression of the disease, and numbers of deaths. The causative agent of the plague of Athens is still being debated [131, 134]. Hippocrates (459–337 B.C.E.), in his treatise *Epidemics*, delineates the factors that affected the spread of disease at that time. In 165–180 C.E., the Roman Empire and Egypt were affected by smallpox. Tens of millions of people died [5].

One of the most well documented epidemics that devastated Europe was the Black Death. The Black Death spread throughout the Mediterranean and Europe and is estimated to have killed about 50–100 million people in the years 1348–1350 [5]. Recent DNA evidence from victims in Europe suggests that the pathogen responsible was the *Yersinia pestis* bacterium, which causes several forms of

plague [69]. The Black Death pathogen reappeared in Europe in multiple locations into the nineteenth century. Another disastrous epidemic attacked the Aztec population in the sixteenth century. This smallpox epidemic killed an estimated 35 million people. In the early twentieth century, an influenza pandemic killed an estimated 20 million of the world's population. At present, we still have significant outbreaks of epidemics: The Bombay plague 1905–1906, the 2003 severe acute respiratory syndrome (SARS), and the H1N1 swine flu pandemic of 2009. Threats of epidemics and pandemics exist continually, since viruses mutate very quickly and can jump species barriers, infecting humans, potentially on a mass scale.

Although epidemiology itself has a long history, the mathematical study of diseases and their spread is only about 350 years old. The first statistical study of infectious diseases is attributed to John Graunt (1620–1674), whose 1663 book *Natural and Political Observations Made upon the Bills of Mortality* was concerned with methods of public health statistics. A century later, Daniel Bernoulli used mathematical methods to analyze mortality from smallpox. In 1766, he published what is now considered the first epidemiological model (reviewed in [22]). Bernoulli argued that inoculation with live virus obtained from a mild case of smallpox would reduce the death rate and thereby increase the population, even if the inoculation itself might occasionally be fatal. A contemporary reformulation of Bernoulli's approach in terms of differential equations is given in [55].

In the mid nineteenth century, Louis Pasteur made remarkable breakthroughs in the causes and prevention of disease. He reduced mortality from puerperal fever and created the first vaccines for rabies and anthrax. His medical discoveries provided direct support for the germ theory of disease. Around the same time, the founder of modern bacteriology, Robert Koch, identified the specific causative agents of tuberculosis, cholera, and anthrax, thus giving experimental support to the concept of infectious disease. He was also famous for the development of Koch's postulates. In the late 1800s, science could finally explain the mechanism of how one becomes ill. The concept of passing a bacterial disease through contact between an infected individual and a healthy one became known. This paved the way for the mathematical modeling of infectious diseases.

Mathematical modeling of infectious diseases made significant strides with the work of William Hamer, in the early twentieth century. He was looking for an explanation of the recurrence of measles. It appears that Hamer was the first to use the mass action law in modeling infectious diseases. But it is Sir Ronald Ross who is considered the father of modern mathematical epidemiology. He did pioneering work on malaria and discovered that it is transmitted between humans and mosquitoes. For his work on malaria, Ross received the Nobel Prize in 1902. Sir Ronald Ross was concerned with prevention of malaria. Despite his contributions, he could not convince his contemporaries that malaria could be eradicated simply by reducing the number of mosquitoes. In the second edition of his book *The Prevention of Malaria*, published in 1911, he developed mathematical models of malaria transmission and derived a threshold quantity, nowadays known as the basic reproduction number. In Ross's time, mathematical modeling of infectious diseases was

not well accepted. Nonetheless, Ross was a supporter of the use of mathematical tools in epidemiology. A British Medical Journal quotes him as follows [71]:

> As a matter of fact all epidemiology, concerned as it is with the variation of disease from time to time or from place to place, must be considered mathematically, however many variables are implicated, if it is to be considered scientifically at all. To say that a disease depends upon certain factors is not to say much, until we can also form an estimate as to how largely each factor influences the whole result. And the mathematical method of treatment is really nothing but the application of careful reasoning to the problems at issue.

Mathematical epidemiology was raised to a new level by the model of the spread of infectious diseases, published by Kermack and McKendrick in 1927. In their joint article "A contribution to the mathematical theory of epidemics" [84], Kermack and McKendrick published for the first time a deterministic epidemic model that included susceptible, infected, and removed individuals, much like the one we will discuss in Chap. 2. In fact, their model is an age-since-infection model, whose contemporary version is discussed in Chap. 13. Their model does not include natural birth and death rates and, consequently, models only disease outbreaks. To capture epidemic modeling of diseases that can become established in a population and persist, Kermack and McKendrick published Part II and Part III of their "A contribution to the mathematical theory of epidemics" in 1932 and 1933 respectively. Because of their seminal importance to mathematical epidemiology, the Kermack–McKendrick fundamental trilogy of papers was reprinted in 1991 [85, 86, 87].

Mathematical modeling of infectious diseases gained importance in the 1980s with the advent of the HIV epidemics. Since then, a very large number of models have been created, analyzed, and employed to study the spread of infectious disease. Today, mathematical epidemiology has a steady presence in the research literature, and mathematical modeling is making significant contributions to mathematics and public health [74, 75, 162].

1.5 General Approach to Modeling

A mathematical model is a description of a system using mathematical tools and language. The process of developing mathematical models is called mathematical modeling. We will be concerned with modeling infectious diseases and their spread in populations, but in principle, mathematical modeling can be applied to any system, biological or otherwise. Mathematical models are developed to help explain a system, to study the effects of its various components, and to make predictions about their behavior.

The modeling process, schematically depicted in Fig. 1.1, requires translation of a biological scenario into a mathematical problem. The modeling process typically begins with a clear description of the processes based on the scientist's understanding of the system. The translation into mathematical equations should be made with a specific goal or biological question in mind. Then the verbal description of the system is encoded in mathematical equations. The model should incorporate only

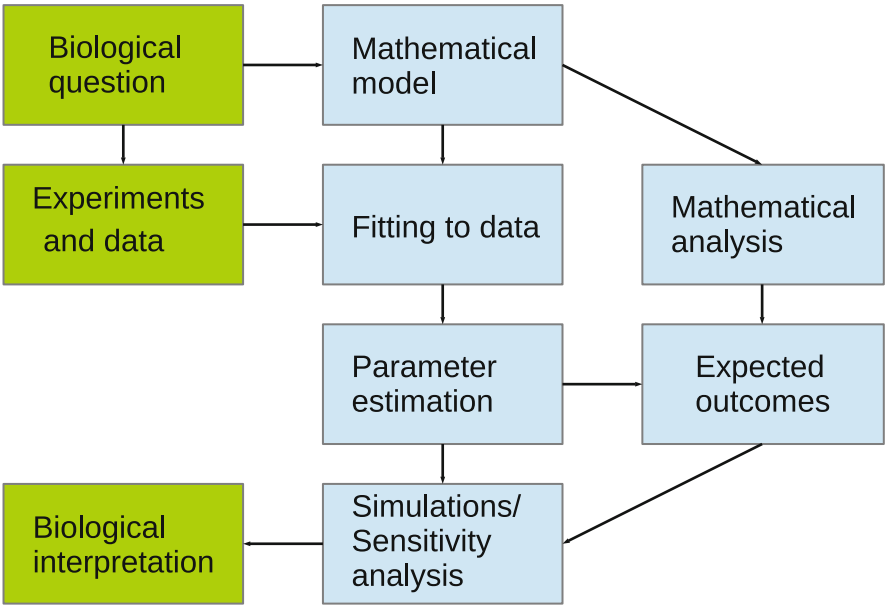

Fig. 1.1 Modeling diagram

those features that are relevant to the specific goal or biological question in mind. Once the model is formulated, it can be investigated with a number of mathematical tools:

- it may be analyzed to produce critical quantities that govern the overall behavior of the solutions;
- it may be fitted to available data or used to stimulate experiments that can produce data;
- parameters of the model may be estimated;
- it may be simulated to understand how important each parameter is to the solution.

After the model has been understood, we must interpret its results in the light of the biological scenario considered and potentially seek the answer of the biological question that was set forth at the beginning. At the very least we must address these questions: What did we learn about the real world from the model? Is our model's message supported by the information about the system?

Mathematical models usually consist of parameters and variables that are connected by relationships. Variables are abstractions of the system's properties that can be quantified or measured. Models can be classified in multiple ways:

- **Linear/nonlinear**. A model is classified as nonlinear if it contains a nonlinear dependence on the variables (e.g., a product of variables). Otherwise, it is classified as linear. The models we will construct and use in this book will be nonlinear.

- **Static/dynamic**. A dynamic model accounts for time-dependent changes in the state of the system, while a static model calculates system quantities assuming that it does not change in time and thus is time-invariant. Dynamic models typically employ differential equations or difference equations. The models that we will consider in this book will be dynamic models.
- **Discrete/continuous**. Discrete models treat time or system states as discrete. Continuous models incorporate time and system states as continuous.
- **Deterministic/stochastic**. A deterministic model is one in which every set of variable states is uniquely determined by the parameters in the model and the initial state of the variables. Stochastic models are characterized by randomness, and variable states are described by probability distributions. The models that we will consider in this book will be deterministic models, although stochastic epidemic models have also been developed and used in the literature.

In this book we will primarily use differential equation models to model the distribution of infectious diseases in a population. The main modeling tool will be ordinary differential equations, but we will introduce epidemic models of delay-differential equations, age-since-infection structured partial differential equations, age-structured partial differential equations, and diffusion partial differential equations. We will also discuss discrete epidemic models. Several types of models used for epidemic modeling will be left out, primarily stochastic epidemic models and network models. For these, you may consult some excellent books and book chapters on the subject (e.g., [7, 124]).

Mathematical models are of great importance in the natural sciences, including biology and epidemiology. They help us to gain new understanding about a system, organize and make sense of biological data, obtain the response behavior of the system, seek optimal performance and intervention strategies, and make predictions about the system. Mathematical models of infectious diseases are the focal point of this book.

Chapter 2
Introduction to Epidemic Modeling

2.1 Kermack–McKendrick SIR Epidemic Model

Introduction to epidemic modeling is usually made through one of the first epidemic models proposed by Kermack and McKendrick in 1927, a model known as the SIR epidemic model [84].

2.1.1 Deriving the Kermack–McKendrick Epidemic Model

When a disease spreads in a population, it splits the population into nonintersecting classes. In one of the simplest scenarios, there are three such classes:

- The class of individuals who are healthy but can contract the disease. These are called *susceptible individuals* or *susceptibles*. The size of this class is usually denoted by S.
- The class of individuals who have contracted the disease and are now sick with it, called *infected individuals*. In this model, it is assumed that infected individuals are also *infectious* (see Chap. 1 for distinction between infected and infectious individuals). The size of the class of infectious/infected individuals is denoted by I.
- The class of individuals who have recovered and cannot contract the disease again are called *removed/recovered individuals*. The class of recovered individuals is usually denoted by R.

The number of individuals in each of these classes changes with time, that is, $S(t)$, $I(t)$, and $R(t)$ are functions of time t. The total population size N is the sum of the sizes of these three classes:

$$N = S(t) + I(t) + R(t).$$

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_2

To formulate a model, we have to make assumptions to simplify reality. The first assumption for the Kermack–McKendrick model is that infected individuals are also infectious. The second assumption of the model is that the total population size remains constant.

Epidemiological models consist of systems of ODEs that describe the dynamics in each class. One of the simplest models involves the dynamics of susceptible, infectious, and recovered individuals. The model was first proposed by Kermack and McKendrick in 1927 [84].

To derive the differential equations, we consider how the classes change over time. When a susceptible individual enters into contact with an infectious individual, that susceptible individual becomes infected with a certain probability and moves from the susceptible class into the infected class. The susceptible population decreases in a unit of time by all individuals who become infected in that time. At the same time, the class of infectives increases by the same number of newly infected individuals. The number of individuals who become infected per unit of time in epidemiology is called *incidence*, and the rate of change of the susceptible class is given by

$$S'(t) = -\text{incidence}.$$

How can we represent the incidence? Consider one infectious individual. Assume:

- cN is the number of contacts per unit of time this infectious individual makes. Here we assume that the number of contacts made by one infectious individual is proportional to the total population size with per capita contact rate c.
- $\frac{S}{N}$ is the probability that a contact is with a susceptible individual. Thus,
- $cN\frac{S}{N}$ is number of contacts with susceptible individuals that one infectious individual makes per unit of time. Not every contact with a susceptible individual necessarily leads to transmission of the disease. Suppose p is the probability that a contact with a susceptible individual results in transmission. Then,
- pcS is number of susceptible individuals who become infected per unit of time per infectious individual.
- βSI is the number of individuals who become infected per unit of time (incidence). Here we have set $\beta = pc$.

If we define $\lambda(t) = \beta I$, then the number of individuals who become infected per unit of time is equal to $\lambda(t)S$. The function $\lambda(t)$ is called the *force of infection*. The coefficient β is the constant of proportionality called the *transmission rate constant*. The number of infected individuals in the population $I(t)$ is called the *prevalence* of the disease.

There are different types of incidence depending on the assumption made about the form of the force of infection. One form is called *mass action incidence*. With this form of incidence, we obtain the following differential equation for susceptible individuals:

$$S'(t) = -\beta IS.$$

The susceptible individuals who become infected move to the class I. Those individuals who recover or die leave the infected class at constant per capita probability

per unit of time α, called the *recovery rate*. That is, αI is the number of infected individuals per unit of time who recover. So,

$$I'(t) = \beta IS - \alpha I.$$

Individuals who recover leave the infectious class and move to the recovered class

$$R'(t) = \alpha I.$$

Thus, the whole model is given by the following system of ODEs:

$$\begin{aligned} S'(t) &= -\beta IS, \\ I'(t) &= \beta IS - \alpha I, \\ R'(t) &= \alpha I. \end{aligned} \tag{2.1}$$

To be well defined mathematically, this system is equipped with given initial conditions $S(0)$, $I(0)$, and $R(0)$.

When we formulate a model, we need to be concerned with the units of the quantities involved. Units are also helpful when we estimate parameters from data. The units of both sides of the above equations must be the same. All derivatives have units number of people per unit of time (why?). Hence, each term on the right-hand side should have the same units. From the first equation, we see that since I and S have units number of people, the units of β must be 1/[number of people$\times$unit of time]. Since $\beta = pc$ and p is a probability, which has no units, the units of c must be 1/[number of people$\times$unit of time]. Thus the contact rate cN has units 1/unit of time. Similarly, from the second equation, we see that the units of α are 1/unit of time, so the term αI has units number of people/unit of time.

Loosely speaking, a differential equation model such as the model (2.1) is *well posed* if through every point (initial condition), there exists a unique solution. Differential equation models must be well posed to be mathematically acceptable and biologically significant. Because the dependent variables in the model denote physical quantities, for most models in biology and epidemiology, we also require that solutions that start from positive (nonnegative) initial conditions remain positive (nonnegative) for all time.

We denote by N the total population size at time zero $N = S(0) + I(0) + R(0)$. Adding all three equations in system (2.1), we obtain $N'(t) = S'(t) + I'(t) + R'(t) = 0$. Hence, $N(t)$ is constant and equal to its initial value, $N(t) = N$. This model is called the SIR model or SIR system. It is a special type of model called a *compartmental model*, because each letter refers to a "compartment" in which an individual can reside. Each individual can reside in exactly one compartment and can move from one compartment to another. Compartmental models are schematically described by a diagram often called a *flowchart*. Each compartment in a flowchart is represented by a box indexed by the name of the class. Arrows indicate the direction of movement of individuals between the classes. The movement arrows are typically labeled by the transition rates (see Fig. 2.1).

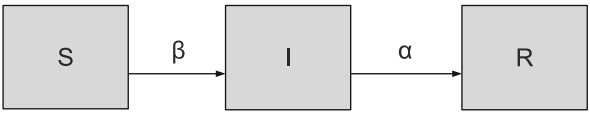

Fig. 2.1 Flowchart of the Kermack–McKendrick SIR epidemic model

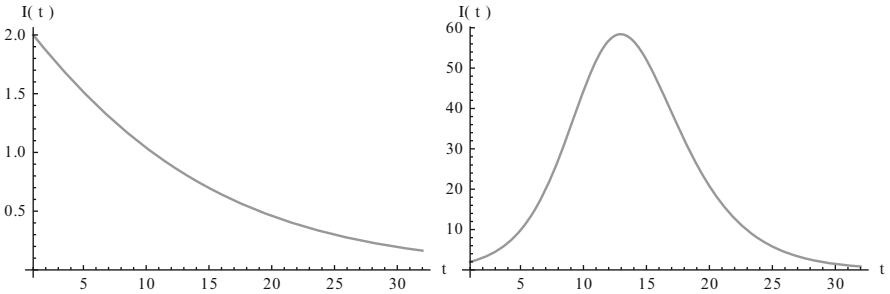

Fig. 2.2 *Left:* shows the prevalence monotonically decreasing. *Right:* shows the prevalence first increasing and then decreasing to zero

2.1.2 Mathematical Properties of the SIR Model

The Kermack–McKendrick epidemic model (2.1) has very distinctive dynamics. Because $S' < 0$ for all t, the number of susceptible individuals is always declining, independently of the initial condition $S(0)$. Since $S(t)$ is monotone and positive, we have

$$\lim_{t\to\infty} S(t) = S_\infty.$$

The number of recovered individuals also has monotone behavior, independently of the initial conditions. Since $R' > 0$ for all t, the number of recovered individuals is always increasing. Since the number of recovered is monotone and bounded by N, we have

$$\lim_{t\to\infty} R(t) = R_\infty.$$

On the other hand, the number of infected individuals may be monotonically decreasing to zero, or may have nonmonotone behavior by first increasing to some maximum level, and then decreasing to zero (see Fig. 2.2). The prevalence first starts increasing if $I'(0) = (\beta S(0) - \alpha)I(0) > 0$. Hence, a necessary and sufficient condition for an initial increase in the number of infecteds is $\beta S(0) - \alpha > 0$, or

$$\frac{\beta S(0)}{\alpha} > 1.$$

This sudden increase in the prevalence and then a decline to zero is a classical model of an **epidemic** or **outbreak**. Threshold conditions for an epidemic to occur are

common in epidemiology, and we will discuss them in detail later on. To determine the limits S_∞ and R_∞, we divide the equation for S and the equation for R. Hence,

$$\frac{dS}{dR} = -\frac{\beta}{\alpha} S.$$

Solving, we have

$$S = S(0)e^{-\frac{\beta}{\alpha}R} \geq S(0)e^{-\frac{\beta}{\alpha}N} > 0.$$

We conclude that $S_\infty > 0$. The quantity S_∞ is called the **final size of the epidemic**. We see that the epidemic does not end, because *all* susceptible individuals have been infected and are now immune. Some individuals always escape a disease—an observation that was made in practice is also confirmed by the SIR model.

Finally, we show that the epidemic dies out. If

$$\lim_{t\to\infty} I(t) = I_\infty,$$

then $I_\infty = 0$. This is evident from the plots in Fig. 2.2, but a mathematical argument can establish the result for all parameters. To see this, we integrate the first equation in (2.1):

$$\begin{aligned}
\int_0^\infty S'(t)dt &= -\beta \int_0^\infty S(t)I(t)dt, \\
S_\infty - S_0 &= -\beta \int_0^\infty S(t)I(t)dt, \\
S_0 - S_\infty &= \beta \int_0^\infty S(t)I(t)dt, \\
S_0 - S_\infty &\geq \beta S_\infty \int_0^\infty I(t)dt. \qquad (2.2)
\end{aligned}$$

The last inequality implies that $I(t)$ is integrable on $[0,\infty)$. Hence, $\lim_{t\to\infty} I(t) = 0$.

The Kermack–McKendrick model is based on several assumptions: (1) There are no births and deaths in the population. (2) The population is **closed**, that is, no one from the outside enters the population, and no one leaves the population, and finally, (3) All recovered individuals have complete immunity and cannot be infected again. These assumptions seem very restrictive, but within limits, they can be satisfied. We will see a specific example in Sect. 2.3. Diseases that lead to permanent immunity and are well modeled by the SIR epidemic model are most diseases typical of childhood years, often called **childhood diseases**. These include chickenpox, smallpox, rubella, and mumps.

To solve the system, we first notice that the variable R does not participate in the first two equations. Thus we can consider only the equations for S and I, which are coupled, and leave out the equation for R. The variable R can then be obtained in this model from the relation $R = N - S - I$:

$$\begin{aligned} S'(t) &= -\beta IS, \\ I'(t) &= \beta IS - \alpha I. \end{aligned} \tag{2.3}$$

Dividing the two equations, we obtain

$$\frac{I'}{S'} = \frac{\beta SI - \alpha I}{-\beta SI} = -1 + \frac{\alpha}{\beta S}.$$

Separating the variables, we have

$$I' = \left(-1 + \frac{\alpha}{\beta S}\right) S'.$$

Integrating leads to

$$I = -S + \frac{\alpha}{\beta} \ln S + C,$$

where C is an arbitrary constant. Thus, the orbits of the solution are given implicitly by the equation

$$I + S - \frac{\alpha}{\beta} \ln S = C. \tag{2.4}$$

The Kermack–McKendrick model is equipped with initial conditions: $S_0 = S(0)$ and $I_0 = I(0)$. Those are given. We also have that $\lim_{t\to\infty} I(t) = 0$, while $S_\infty = \lim_{t\to\infty} S(t)$ gives the final number of susceptible individuals after the epidemic is over. The above equality holds both for (S_0, I_0) and for $(S_\infty, 0)$. Thus,

$$I_0 + S_0 - \frac{\alpha}{\beta} \ln S_0 = C.$$

Consequently,

$$I_0 + S_0 - \frac{\alpha}{\beta} \ln S_0 = S_\infty - \frac{\alpha}{\beta} \ln S_\infty.$$

Rearranging terms, we get

$$I_0 + S_0 - S_\infty = \frac{\alpha}{\beta} (\ln S_0 - \ln S_\infty).$$

Therefore,

$$\frac{\beta}{\alpha} = \frac{\ln \frac{S_0}{S_\infty}}{S_0 + I_0 - S_\infty}. \tag{2.5}$$

We note that since $S(t)$ is a decreasing function, we have $S_\infty < S_0 + I_0$. The implicit solution also allows us to compute the maximum number of infected individuals that is attained. This number occurs when $I' = 0$, that is, when

$$S = \frac{\alpha}{\beta}.$$

From

$$I + S - \frac{\alpha}{\beta} \ln S = I_0 + S_0 - \frac{\alpha}{\beta} \ln S_0,$$

substituting the expression for S and moving all terms but I to the right-hand side leads to

$$I_{\max} = -\frac{\alpha}{\beta} + \frac{\alpha}{\beta} \ln \frac{\alpha}{\beta} + S_0 + I_0 - \frac{\alpha}{\beta} \ln S_0. \tag{2.6}$$

Here $I_{\max}$ is the maximum number of infected individuals reached in the epidemic. It signifies the maximum severity of the epidemic. If we are able to estimate $I_{\max}$ for a newly occurring infectious disease, we will know when the number of infections will begin to decline.

2.2 The Kermack–McKendrick Model: Estimating Parameters from Data

When we are given specific disease and time series data for it, we can estimate the parameters of the SIR model and compare the solution of the model with the data. This section follows the description in [27]. See [27] for a different example.

2.2.1 Estimating the Recovery Rate

For many diseases, information about the mean duration of the exposed period or the infectious period can easily be obtained. For instance, for influenza, the duration of the infectious period is 3–7 days with mean 4–5 days. How can that help us estimate the recovery rate α? To approach that question, let us assume that there is no inflow in the infectious class and a certain number of individuals I_0 have been put in the infectious class at time zero. Then the differential equation that gives the dynamics of this class is given by

$$I'(t) = -\alpha I, \qquad I(0) = I_0.$$

This equation can be easily solved. Therefore, the number of people in the infectious class at time t is given by

$$I(t) = I_0 e^{-\alpha t}.$$

Consequently,

$$\frac{I(t)}{I_0} = e^{-\alpha t}$$

for $t \geq 0$ gives the proportion of people who are still infectious at time t, or in probability language, it gives the probability of being still infectious at time t. We can compute the fraction of individuals who have left the infectious class,

$$1 - e^{-\alpha t},$$

or in probability terms,

$$F(t) = 1 - e^{-\alpha t} \qquad t \geq 0$$

is the probability of recovering/leaving the infectious class in the interval $[0,t)$. Clearly, $F(t)$ is a probability distribution (if defined as zero for $t < 0$). The probability density function is $f(t) = dF/dt$. Consequently,

$$f(t) = \alpha e^{-\alpha t}.$$

Note: $f(t) = 0$ for $t < 0$. Furthermore, the average time spent in the infectious class is given by the mean (expected value of a random variable X, denoting time to exiting the infectious class),

$$E[X] = \int_{-\infty}^{\infty} t f(t) dt.$$

Therefore, computing that integral yields

$$\int_{-\infty}^{\infty} t f(t) dt = \int_{-\infty}^{\infty} t \alpha e^{-\alpha t} dt = \frac{1}{\alpha}.$$

Thus we conclude that

mean time spent in the infectious class = $\frac{1}{\alpha}$.

For influenza, we are sick with it for 3–7 days. Say that the mean time spent as infectious is 5 days. Thus the recovery rate, measured in units of $[\text{days}]^{-1}$, is 1/5.

Estimating the transmission rate β is quite a bit more difficult. Estimating β is possible for the Kermack–McKendrick model, because that model is relatively simple. In particular, we can obtain an implicit solution. An implicit solution is rarely obtainable for epidemic models, and estimating parameters for epidemic models requires techniques different from the one presented below. We will discuss these techniques in Chap. 6.

2.2.2 The SIR Model and Influenza at an English Boarding School 1978

In January and February 1978, an epidemic of influenza occurred in a boarding school in the north of England. The boarding school housed a total of 763 boys, all of whom were at risk during the epidemic. The spring term began on January 10. The boys returned from their Christmas vacation spent at many different locations in the world. A boy returning from Hong Kong exhibited elevated temperature during the period 15–18 January. On January 22, three boys were sick. Table 2.1 gives the number of boys ill on the nth day beginning January 22 ($n = 1$).

Table 2.1 Daily number of influenza-infected boys

Day	No. infected[a]	Day	No. infected
3	25	9	192
4	75	10	126
5	227	11	71
6	296	12	28
7	258	13	11
8	236	14	7

[a]Data taken from "Influenza in a Boarding School," *British Medical Journal*, 4 March 1978

The number of boys who escaped influenza was 19. The average time spent sick was 5–6 days. However, since boys were isolated in the infirmary, they spent perhaps about 2 days as infectious. A swab taken from some of the boys revealed that they were infected with H1N1 influenza A virus. The staff of the boarding school remained healthy, with only one staff member displaying symptoms of illness.

These data give the following values: $S_3 = 738$, $I_3 = 25$, $S_\infty = 19$.

From the computations above, we have

$$\frac{\beta}{\alpha} = \frac{\ln \frac{S_3}{S_\infty}}{S_3 + I_3 - S_\infty} = \frac{\ln \frac{738}{19}}{763 - 19} = 0.00491869. \tag{2.7}$$

We measure time in days. We take $t_0 = 0$ to be January 21. The first datum is given on January 22, which gives $t = 1$. We have that $t_{\text{end}} = 14$ is the February 4, 1978.

We take the infective period to be 2.1 days. This value can be obtained as the best fit as values around 2 days are tried with the procedure below. After we fix the duration of the infectious period, we compute α as the reciprocal of the time spent as an infectious individual (infectious period):

$$\alpha = \frac{1}{2.1} = 0.476.$$

From Eq. (2.7) and using the value for α, we can obtain the value for β:

$$\beta = 0.004918\alpha = 0.004918 * 0.476 = 0.002342.$$

From Eq. (2.6) for the $I_{\max}$, we can estimate the maximum number of infectives during the epidemic. First, notice that $\alpha/\beta = 203.306$. Thus,

$$I_{\max} = -203.306 + 203.306 \ln 203.306 + 738 + 25 - 203.306 \ln 738 = 298.$$

Notice that the data give the maximum number of infective individuals as 296. We illustrate the fit between the model and the data in Fig. 2.3.

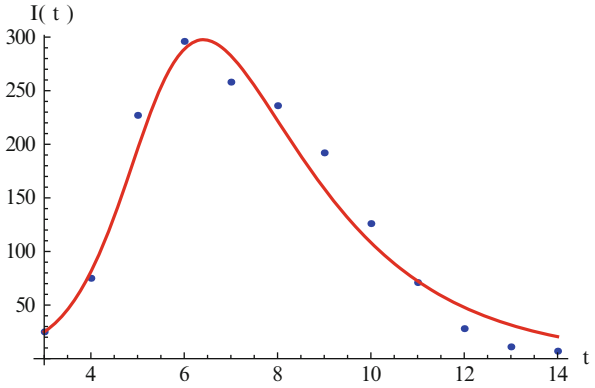

Fig. 2.3 English boarding school influenza epidemic: agreement between Kermack–McKendrick SIR epidemic model and data

2.3 A Simple SIS Epidemic Model

We want to relax the assumption for permanent immunity after recovery to model diseases that can infect us repeatedly, such as influenza. We may assume in the simplest scenario that individuals who recover become immediately susceptible again. Thus, individuals who are susceptible may become infected (and infectious) and then recover into being susceptible again. The model is described with the flowchart in Fig. 2.4.

The model takes the form

$$\begin{aligned} S'(t) &= -\beta IS + \alpha I, \\ I'(t) &= \beta IS - \alpha I. \end{aligned} \tag{2.8}$$

System (2.8) is called an SIS epidemic model and is perhaps the simplest model in mathematical epidemiology. Here, if $N = S + I$ and we add the two equations, we

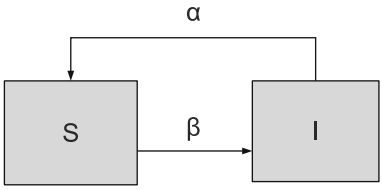

Fig. 2.4 Flowchart of a simple SIS epidemic model

again obtain $N' = 0$. Hence the total population size is N, where N is a constant in time. The system is equipped with initial conditions $S(0)$ and $I(0)$, so that $N = S(0) + I(0)$.

2.3.1 Reducing the SIS Model to a Logistic Equation

Because the total population size is constant and known, the system (2.8) can be reduced to a single equation. This technique is commonly used for the reduction of the dimension of an epidemiological model. We express S as $S = N - I$ and substitute it in the second equation. The resulting equation is a variant of the **logistic equation**:

$$I'(t) = \beta I(N - I) - \alpha I. \tag{2.9}$$

We rewrite this equation in the form of a logistic equation,

$$I'(t) = rI\left(1 - \frac{I}{K}\right),$$

where $r = \beta N - \alpha$ and $K = r/\beta$. To see this, first factor out I and then $r = \beta N - \alpha$. The logistic equation is one of the classical models in population dynamics. It typically models the total population size of a population of individuals. We will use it later on for models in which the total population size does not remain constant. The parameter r is often referred to as the *growth rate*. We can see that r can be positive or negative, so we consider two cases.

$r < 0$ If the growth rate is negative, $r < 0$, then the number of infected individuals $I(t)$ tends to 0 as $t \to \infty$. To see this, notice that if $r < 0$, then $K < 0$. Hence,

$$I'(t) \leq rI(t).$$

The solutions of this simple differential inequality are $I(t) = I(0)e^{rt}$, and they approach zero for $r < 0$. This implies that if $r < 0$, the disease gradually disappears from the population on its own.

$r > 0$ The logistic equation can be solved, and in this case, we need to solve it to have an explicit expression for $I(t)$. The logistic equation is a differential equation of separable type. It is solved by a method called *separation of*

variables. To separate the variables I and t, we move all terms that contain I to the left-hand side of the equation, and all terms that contain t, namely dt, to the right-hand side:

$$\frac{1}{I\left(1-\frac{I}{K}\right)}dt = rdt.$$

Notice that while dividing by $\left(1-\frac{I}{K}\right)$, we have assumed that $I(t) \neq K$. But $I(t) = K$ is a solution of the original logistic equation. On the other hand, $I = K$ is not a solution of the derived equation above, so it will have to be artificially added to the solution set.
Using partial fraction decomposition, we can integrate both sides of that equation:

$$\int \left(\frac{1}{I} + \frac{1}{K-I}\right) dI = r \int 1 dt.$$

Hence,

$$\ln \frac{I}{|K-I|} = rt + C,$$

where C is an arbitrary constant of integration, and the absolute value in the logarithm is necessary, since we can compute logarithms only of positive values, but we do not know whether $K-I$ is positive. To determine C, we use the initial conditions. Assuming that the initial conditions are given at 0, we have

$$\ln \frac{I(0)}{|K-I(0)|} = C.$$

Replacing C with the above expression, we obtain

$$\ln \frac{I}{|K-I|} - \ln \frac{I(0)}{|K-I(0)|} = rt.$$

Hence,

$$\ln \frac{I|K-I(0)|}{I(0)|K-I|} = rt.$$

The absolute values above can be disregarded, since $K-I(0)$ and $K-I$ have the same sign: they are both positive or both negative. Taking an exponent, we obtain

$$\frac{I}{K-I} = \frac{I(0)}{K-I(0)} e^{rt}.$$

Finally, we solve for I to obtain an explicit solution for $I(t)$ in terms of the initial conditions, r and K:

$$I(t) = \frac{KBe^{rt}}{1+Be^{rt}},$$

where $B = I(0)/(K - I(0))$. We see from this that

$$\lim_{t\to\infty} I(t) = K,$$

and the disease remains in the population indefinitely.

The threshold condition $r > 0$ can be rewritten as $\mathscr{R}_0 > 1$, where

$$\mathscr{R}_0 = \frac{\beta N}{\alpha}$$

is called *basic reproduction number* of the disease. Mathematically, the reproduction number plays the role of a threshold value for the dynamics of the system and the disease. If $\mathscr{R}_0 > 1$, the disease remains in the population, and the number of infecteds stabilizes around K. In this case, we say that the disease has become **endemic** in the population. This implies that the simple SIS model is a model of endemic disease. If $\mathscr{R}_0 < 1$, the number of infecteds gradually declines to zero, and the disease disappears from the population.

Epidemiologically, the reproduction number gives the number of secondary cases one infectious individual will produce in a population consisting only of susceptible individuals.

To see this interpretation in the formula for $\mathscr{R}_0$, notice that the number of new cases per unit of time produced by all infectious individuals is given by the incidence βSI. If there is only one infectious individual, we have $I = 1$, and the number of secondary cases produced by one infectious individual will be βS. If the entire population consists of susceptible individuals, we have $S = N$. Hence, the number of secondary cases one infectious individuals will produce in a unit of time is βN. Since one infectious individual remains infectious for $1/\alpha$ time units, the number of secondary cases it will produce during its lifespan is $\mathscr{R}_0 = \beta N/\alpha$.

2.3.2 *Qualitative Analysis of the Logistic Equation*

The information we derived about the behavior of the solutions was obtained from the explicit solution. Many single-equation models in biology cannot be solved explicitly. We need tools to deduce the properties of the solutions directly from the differential equation. These tools can readily be extended to systems of equations.

From the explicit solution of the logistic equation, we saw that in the long run, the disease will become endemic and persist in the population if $\mathscr{R}_0 > 1$. We also learned that in the long run, the number of infected individuals in the population will be approximately $K = (\beta N - \alpha)/\beta$. Furthermore, if $\mathscr{R}_0 < 1$, the disease will die out. Ideally, we would like to be able to obtain such results without having to solve the equation explicitly.

A nonlinear differential equation model with constant coefficients typically has time-independent solutions, that is, solutions that are constant in time. Such solutions are called **equilibrium points**. Equilibrium points play an important role in the long-term behavior of the solutions. They are easy to find from the differential equation even if we don't know the explicit solution, since their derivative with respect to time is zero. Thus, for the equation $\frac{dI}{dt} = f(I)$, the equilibria are the solutions of the equation $f(I) = 0$. We set the right-hand side of Eq. (2.9) equal to zero:

$$\beta I(N - I) - \alpha I = 0.$$

This equation has two solutions, $I_1^* = 0$ and $I_2^* = K$, which give the two equilibrium points. The equilibrium I_1^* always exists. In the mathematical epidemiology literature, the equilibrium I_1^* is referred to as a **disease-free equilibrium**, since the disease is not present in the population, and the entire population is susceptible. The equilibrium I_2^* exists only if $\mathscr{R}_0 > 1$. The equilibrium I_2^* is called an **endemic equilibrium**, since the disease is present in the population.

In the case $\mathscr{R}_0 > 1$, both $I_1(t) = 0$ and $I_2(t) = K$ are solutions to Eq. (2.9). Since the model is well posed, no other solution can cross them. So solutions that start in the interval $(0, K)$ stay in that interval for all time:

$$0 < I(0) < K \Longrightarrow 0 < I(t) < K.$$

Furthermore, solutions that start from a value above K stay above K:

$$I(0) > K \Longrightarrow I(t) > K.$$

If $0 < I(t) < K$, then $f(I) > 0$, which means that $\frac{dI}{dt} > 0$. This means that the solutions in that interval are increasing functions of time. Since $I(t)$ is increasing and bounded, it follows that $I(t)$ converges to a finite limit as $t \to \infty$. To deduce the behavior of the derivative, we use the following corollary.

Corollary 2.1 (Thieme [151]). *Assume that $f(t)$ converges as $t \to \infty$. Assume also that $f'(t)$ is uniformly continuous. Then $f'(t) \to 0$ as $t \to \infty$.*

It can be shown (see (2.10) below) that the second derivative $\frac{d^2I}{dt^2}$ is continuous and bounded. Hence, the corollary above implies that $I'(t) \to 0$ and the limit of $I(t)$, say L, satisfies the equilibrium equation $f(L) = 0$. This implies that $L = 0$ or $L = K$. Since $I(t)$ is positive and increasing, we have $I(t) \to K$ as $t \to \infty$. If $I(0) > K$, then $I(t) > K$ for all t. Thus, $\frac{dI}{dt} < 0$, and $I(t)$ is decreasing and bounded below by K. Similar reasoning as above implies that $I(t) \to K$.

We can further investigate the concavity of the solutions by looking at the second derivative:

$$\frac{d^2I}{dt^2} = r\left(1 - \frac{2I}{K}\right)\frac{dI}{dt} = r^2\left(1 - \frac{2I}{K}\right)I\left(1 - \frac{I}{K}\right). \tag{2.10}$$

For solutions in the interval $0 < I(t) < K$, the second derivative changes sign when $I(t)$ crosses the horizontal line $y = \frac{K}{2}$. Thus, for values of t such that $I(t) < \frac{K}{2}$, the

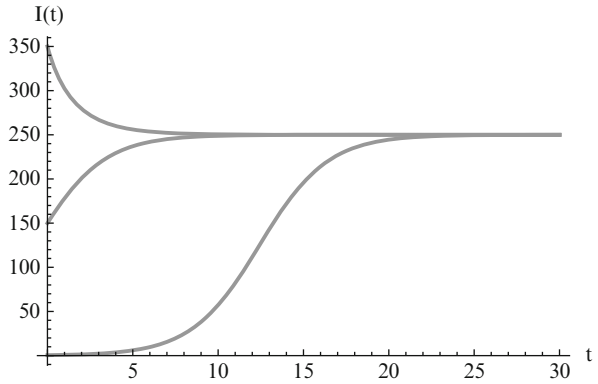

Fig. 2.5 Solutions to the logistic equation (2.9) converge to the endemic equilibrium

second derivative of I is positive, and $I(t)$ is concave up. For values of t for which $I(t) > \frac{K}{2}$, the second derivative of I is negative, and $I(t)$ is concave down. This is illustrated in Fig. 2.5. For solutions for which $I(t) > K$, the second derivative $\frac{d^2I}{dt^2}$ is positive. Consequently, $I(t)$ is decreasing and concave up.

2.3.3 General Techniques for Local Analysis of Single-Equation Models

We saw that if $\mathscr{R}_0 < 1$, then all solutions of Eq. (2.9) approach the unique equilibrium $I^* = 0$. That is, *all* solutions converge to zero, $I(t) \to 0$, for *every* initial condition $I(0) > 0$. In this case, we say that the disease-free equilibrium is **globally stable**. In the case $\mathscr{R}_0 > 1$, there are two equilibria: the disease-free $I_1^* = 0$ and the endemic equilibrium $I_2^* = K$. We see that all solutions that start from $I(0) > 0$ move away from the disease-free equilibrium. Hence, the disease-free equilibrium in this case is **unstable**. At the same time, all solutions that start from $I(0) > 0$ approach the endemic equilibrium $I_2^* = K$. In this case, we call the endemic equilibrium **globally stable**.

For many models, even models given by a single equation, we may not be able to solve the equation(s) explicitly or perform detailed analysis of the behavior of the solutions. In addition, if there are multiple endemic equilibria, there may not be a globally stable equilibrium. In these cases, the concept of a locally stable equilibrium is an applicable and useful tool. Loosely speaking, an equilibrium is locally asymptotically stable if solutions that *start close* to the equilibrium approach that equilibrium as $t \to \infty$. Stability of a nonlinear system can often be inferred from the stability of a corresponding linear system obtained through the process of **linearization**. For a general differential equation

$$x'(t) = f(x), \tag{2.11}$$

x^* is an equilibrium if and only if $f(x^*) = 0$. The idea of the linearization is to shift the equilibrium to zero. Thus, we denote by $u(t) = x(t) - x^*$ the perturbation that gives the deviation of a solution of (2.11) from an equilibrium. Solutions of (2.11) starting from a neighborhood of x^* approach x^* if $u(t)$ approaches zero. The perturbation $u(t)$ is assumed small. Notice that $u(t)$ can be positive or negative, even if $x(t) > 0$. We have $x(t) = u(t) + x^*$. We replace $x(t)$ with its equal in the differential equation and expand f around x^* in a Taylor series, assuming that f is sufficiently differentiable:

$$u'(t) = f(x^*) + f'(x^*)u(t) + \frac{f''(\xi)}{2!}(u(t))^2,$$

where ξ is between x^* and $x^* + u(t)$. Assuming that f has two continuous derivatives, the second derivative f'' is bounded, and the last term in the expansion with $(u(t))^2$ is small and can be neglected. Since x^* is an equilibrium, we also have $f(x^*) = 0$. Thus, the equation for the perturbations becomes

$$u'(t) = f'(x^*)u(t). \tag{2.12}$$

This is the linearized equation of the nonlinear equation (2.11). This equation is linear in the dependent variable $u(t)$. The quantity $f'(x^*)$ is a given known constant. If we define $\lambda = f'(x^*)$ then the linearized equation becomes

$$u'(t) = \lambda u(t),$$

whose solution is $u(t) = u(0)e^{\lambda t}$. These solutions approach ∞ or $-\infty$ exponentially, depending on $u(0)$, if $\lambda > 0$ and approach zero if $\lambda < 0$. Thus, if $\lambda < 0$, then $u(t) \to 0$. Hence, $x(t) - x^* \to 0$ or $x(t) \to x^*$ as $t \to \infty$. We conclude that solutions of (2.11) that start from an initial condition that is sufficiently close to the equilibrium converge to this equilibrium if $\lambda < 0$. In this case, the equilibrium x^* is called **locally asymptotically stable**. If $\lambda > 0$, then $|u(t)| \to \infty$, and $x(t)$ moves away from the equilibrium x^*. In this case, the equilibrium x^* is called **unstable**. We summarize this result in the following theorem.

Theorem 2.1. *An equilibrium x^* of the differential equation $x'(t) = f(x)$ is locally asymptotically stable if $f'(x^*) < 0$ and is unstable if $f'(x^*) > 0$.*

This theorem does not tell us anything about the stability of the equilibrium x^* if $f'(x^*) = 0$. An equilibrium for which $f'(x^*) \neq 0$ is called **hyperbolic**. If $f'(x^*) = 0$, the equilibrium is called **nonhyperbolic**.

We apply Theorem 2.1 to the logistic version of Eq. (2.9). If $\mathscr{R}_0 < 1$, we found only one equilibrium $I_1^* = 0$. If $\mathscr{R}_0 > 1$, we found two equilibria: $I_1^* = 0$ and $I_2^* = K$. We compute the derivative of $f(I)$,

$$f'(I^*) = r\left(1 - \frac{I^*}{K}\right) - \frac{r}{K}I^*,$$

and its value of each equilibrium,

$$f'(0) = r \qquad f'(K) = -r.$$

We conclude that if $\mathscr{R}_0 < 1$, the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, the disease-free equilibrium is unstable, while the endemic equilibrium is locally asymptotically stable.

The problem of determining equilibria and their stability has a very elegant graphical solution. For the equation $x' = f(x)$, if we plot the function $f(x)$ as a function of x, then the places where $f(x)$ intersects the x-axis give the equilibria. The stability of each equilibrium can then be read off the graph from the slope of the graph as it passes through the equilibrium. If the slope of the tangent line to the graph at the point of the equilibrium is positive, then that equilibrium is unstable; if the slope of the tangent is negative, then that equilibrium is locally stable. If the slope of the tangent to the graph at the equilibrium is zero, then the stability of that equilibrium cannot be inferred from the graph. To illustrate this concept, consider the equation $x' = f(x)$, where $f(x)$ is plotted in Fig. 2.6. The equilibria and their stability are explained in the figure caption.

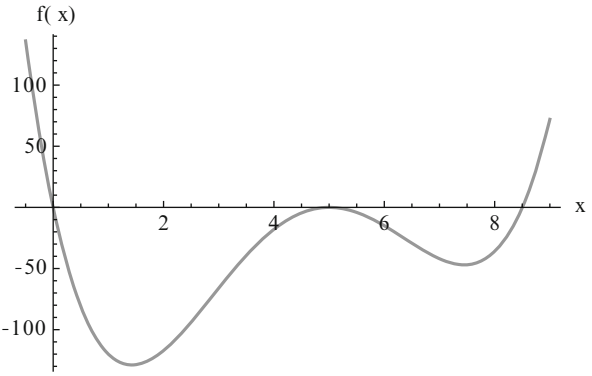

Fig. 2.6 Graph of the function $f(x)$. Figure shows that equilibria are $x_1^* = 0$, $x_2^* = 5$, and $x_3^* = 8.5$. The equilibrium x_1^* is locally stable because the slope of the tangent to x_1^* is negative. The slope of the tangent to x_2^* is zero, so its stability cannot be determined from the graph. Equilibrium x_3^* is unstable, since the slope of the tangent to x_3^* is positive

2.4 An SIS Epidemic Model with Saturating Treatment

We illustrate the concepts of the previous section on an SIS model with saturating treatment/recovery rate. Suppose that in model (2.8), the per capita recovery rate α depends on treatment. In this case, we may assume that treatment resources are

limited and the per capita treatment rate α is not constant, but is decreasing with the number of infected individuals. A reasonably simple form of such a function would be

$$\alpha(I) = \frac{\alpha}{1+I},$$

where the constant α is the treatment/recovery rate when there are few infectives. We use this function in model (2.8) to obtain the following SIS model with saturating treatment:

$$\begin{aligned} S'(t) &= -\beta IS + \frac{\alpha I}{1+I}, \\ I'(t) &= \beta IS - \frac{\alpha I}{1+I}. \end{aligned} \tag{2.13}$$

System (2.13) is called an SIS epidemic model with saturating treatment. Here, if $N = S+I$ and we add the two equations, we again obtain $N' = 0$. Hence the total population size is N, where N is a constant in time. The system is equipped with initial conditions $S(0)$ and $I(0)$, so that $N = S(0)+I(0)$.

2.4.1 *Reducing the SIS Model with Saturating Treatment to a Single Equation*

Since the total population size in model (2.13) is a given constant, we may write $S(t) = N - I(t)$ and substitute it in the second equation of system (2.13). Therefore, we obtain a single equation in the number of infected individuals:

$$I'(t) = \beta I(N-I) - \frac{\alpha I}{1+I}. \tag{2.14}$$

In principle, Eq. (2.14) is a separable equation and can be solved. However, to illustrate common methodologies, we will try to investigate the properties of this equation without solving it. First, we look for the equilibria. We denote by $f(I)$ the right-hand side:

$$f(I) = \beta I(N-I) - \frac{\alpha I}{1+I}.$$

To find the equilibria, we set $f(I) = 0$. Clearly, $I_1^* = 0$ is an equilibrium. This gives the disease-free equilibrium of the equation. To look for endemic equilibria, we cancel one I and we rewrite the equation $f(I) = 0$ as an equality of two functions:

$$\beta(N-I) = \frac{\alpha}{1+I}.$$

This equation can be rewritten as a quadratic equation, which can have zero, one, or two positive roots. We will investigate graphically the options and the conditions

for each to occur. We rewrite the above equation as

$$(N-I)(1+I)=\frac{\alpha}{\beta}. \tag{2.15}$$

Let $g(I)=(N-I)(1+I)$. Then $g(I)$ is a parabola that opens downward. Clearly, $g(0)=N$. The right-hand side of the above equation is $y=\dfrac{\alpha}{\beta}$ and can be graphed as a horizontal line.

- If $g(0)=N>\dfrac{\alpha}{\beta}$, then Eq. (2.15) always has a unique positive solution I_2^*. Then the system (2.13) has one endemic equilibrium. We define the reproduction number of the system as

$$\mathscr{R}_0=\frac{\beta N}{\alpha}.$$

Hence, if $\mathscr{R}_0>1$, there is a unique endemic equilibrium. We illustrate this situation in Fig. 2.7.

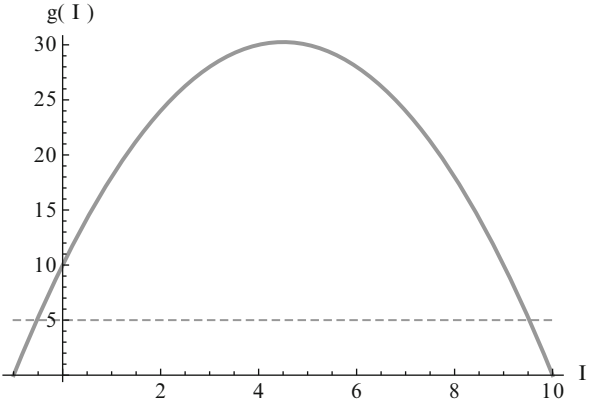

Fig. 2.7 Graph of the function $g(I)$ and the horizontal line $y=\dfrac{\alpha}{\beta}$. The figure shows the existence of a unique intersection for positive I, giving a unique positive equilibrium

- If $g(0)=N<\dfrac{\alpha}{\beta}$, then Eq. (2.15) has either two or zero solutions. In this case,

$$\mathscr{R}_0<1.$$

To specify additional conditions so that Eq. (2.15) has two positive solutions, we must notice that we need two things to happen:
(1) The maximum of the parabola must be to the right of the y-axis. The parabola intersects the x-axis at the points N and -1. Hence, its maximum occurs at their average,

$$I_m=\frac{N-1}{2}>0.$$

This poses the requirement that $N > 1$. (2) The line $y = \dfrac{\alpha}{\beta}$ must lie below the maximum of the parabola. That is, we must have

$$(N - I_m)(1 + I_m) > \frac{\alpha}{\beta}. \tag{2.16}$$

Therefore, if $\mathscr{R}_0 < 1$, $N > 1$, and condition (2.16) are satisfied, then the system (2.13) has two endemic equilibria I_{11}^* and I_{12}^*; otherwise, it has no endemic equilibria. We illustrate these two situations in Fig. 2.8.

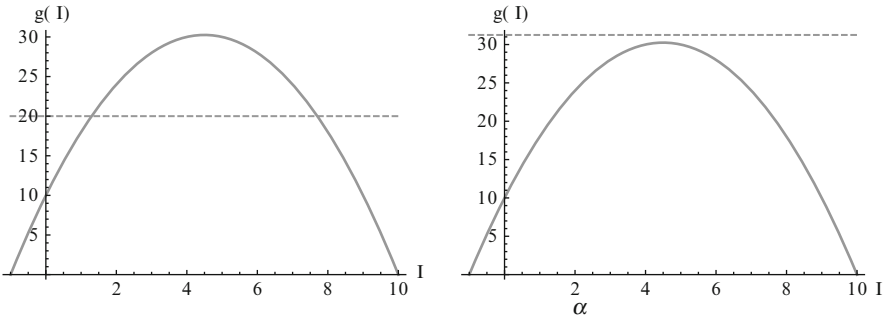

Fig. 2.8 Graph of the function $g(I)$ and the horizontal line $y = \dfrac{\alpha}{\beta}$. *Left:* the existence of two intersections for positive I, giving two positive equilibria. *Right:* no intersections of the function $g(I)$ and the horizontal line $y = \dfrac{\alpha}{\beta}$. Thus, there are no positive equilibria

2.4.2 Bistability

To decide the stability of equilibria, we have to derive the sign of $f'(I^*)$ for each equilibrium I^*. That may not be an easy task to do analytically. Fortunately, the stability of the equilibria can be read off the graph of the function $f(I)$ for each of the three cases above. If $\mathscr{R}_0 < 1$ and there are no nontrivial equilibria, then all solutions of Eq. (2.15) are attracted by the disease-free equilibrium. So the disease-free equilibrium is globally stable in this case. For each of the other two cases, we graph the function $f(I)$ in Fig. 2.9. Looking at Fig. 2.9, we see that in the case $\mathscr{R}_0 > 1$ (left figure), we have $f'(0) > 0$. Hence, the disease-free equilibrium is unstable. Furthermore, $f'(I_2^*) < 0$. Hence, the endemic equilibrium is locally stable. We can argue, as we did in the case of the logistic equation, that the equilibrium is globally stable. In the case $\mathscr{R}_0 < 1$, there are three equilibria: $I_1^* = 0$, $I_{11}^* < I_{12}^*$. For solutions $I(t)$ that start from $I(0) = I_0$ satisfying $0 < I_0 < I_{11}^*$, we have $0 < I(t) < I_{11}^*$ for all t. Furthermore, $f(I) < 0$ for such solutions (the graph of $f(I)$ is below the x-axis), so that $\frac{dI}{dt} < 0$. Hence, $I(t)$ is decreasing and $\lim_{t\to\infty} I(t) = 0$. For solutions $I(t)$ that start from $I(0) = I_0$ satisfying $I_{11}^* < I_0 < I_{12}^*$, we have $I_{11}^* < I(t) < I_{12}^*$ for all t.

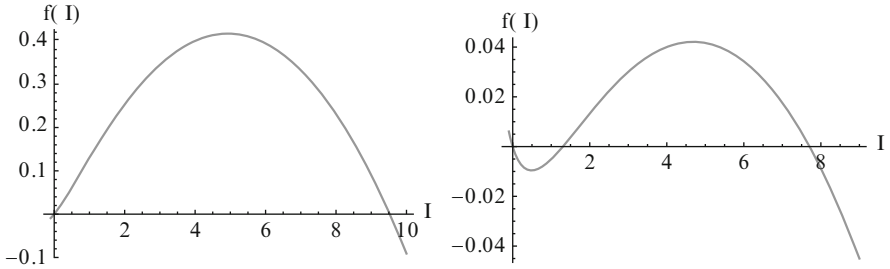

Fig. 2.9 Graph of the function $f(I)$. *Left:* the case $\mathscr{R}_0 > 1$ and the existence of two intersections for nonnegative I, giving two nonnegative equilibria. *Right:* the case $\mathscr{R}_0 < 1$ and three intersections of the function $f(I)$ and the x-axis, giving three nonnegative equilibria. Stabilities explained in text

Furthermore, $f(I) > 0$ for such solutions (the graph of $f(I)$ is above the x-axis), so that $\frac{dI}{dt} > 0$. Hence, $I(t)$ is increasing and $\lim_{t\to\infty} I(t) = I^*_{12}$. For solutions $I(t)$ that start from $I(0) = I_0$ satisfying $I^*_{12} < I_0$, we have $I^*_{12} < I(t)$ for all t. Furthermore, $f(I) < 0$ for such solutions (the graph of $f(I)$ is below the x-axis), so that $\frac{dI}{dt} < 0$. Hence, $I(t)$ is decreasing and $\lim_{t\to\infty} I(t) = I^*_{12}$. We notice that depending on the initial conditions, we have solutions that converge to the disease-free equilibrium and solutions that converge to the endemic equilibrium. Such a situation is called **bistability**. In this case, there is no globally stable equilibrium. The region $0 < I_0 < I^*_{11}$ is called a *domain of attraction* of the disease-free equilibrium. The region $I^*_{11} < I_0$ is called a *domain of attraction* of the endemic equilibrium.

Problems

2.1. Show that the model (2.1) is well posed.

2.2. Use a computer algebra system to graph the solutions (2.4).

2.3. The simplest model of malaria assumes that the mosquito population is at equilibrium and models the proportion of the infected humans I with the following equation:

$$I' = \frac{\alpha\beta I}{\alpha I + r}(1 - I) - \mu I,$$

where r is the natural death rate of mosquitoes, μ is the death rate of humans, β is the transmission rate from infected mosquitoes to susceptible humans, and α is the transmission rate from humans to mosquitoes.

(a) Compute the reproduction number of malaria.
(b) Find the equilibria of the model and their stabilities.
(c) Use a computer algebra system to graph several solutions.

2.4. Consider the model of malaria in Problem 2.3 and assume that saturating treatment is applied:

$$I' = \frac{\alpha\beta I}{\alpha I + r}(1 - I) - \mu I - \frac{\gamma I}{A + I},$$

where r is the natural death rate of mosquitoes, μ is the death rate of humans, β is the transmission rate from infected mosquitoes to susceptible humans, α is the transmission rate from humans to mosquitoes, γ is the treatment rate, and A is the half-saturation constant.

(a) Compute the reproduction number of malaria with saturating treatment.
(b) Find the equilibria of the model and the conditions for their existence.
(c) Find the stabilities of the equilibria.
(d) Use a computer algebra system to graph several solutions.

2.5. Consider the SIS model with constant population size N and saturating incidence in the size of the susceptibles:

$$\begin{aligned} S'(t) &= -\frac{\beta IS}{1 + \sigma S} + \alpha I, \\ I'(t) &= \frac{\beta IS}{1 + \sigma S} - \alpha I. \end{aligned} \tag{2.17}$$

(a) Reduce the SIS model to a single equation.
(b) Determine the threshold condition for the existence of endemic equilibria.
(c) Use a computer algebra system to plot the solutions of (2.17) for $N = 100$, $\beta = 0.5$, $\sigma = 0.01$, $\alpha = 0.05$.

2.6. Consider the SIS model with constant population size N:

$$\begin{aligned} S'(t) &= -\frac{\beta I^p S}{1 + \sigma I^q} + \alpha I, \\ I'(t) &= \frac{\beta I^p S}{1 + \sigma I^q} - \alpha I. \end{aligned} \tag{2.18}$$

(a) Reduce the SIS model to a single equation.
(b) For the case $p < 1$, $q = p - 1$, determine the threshold condition for the existence of endemic equilibria.
(c) For the case $p > 1$, $p = q$, determine the threshold condition for the existence of endemic equilibria.

2.7. Plague in Eyam [27]
The Derbyshire village of Eyam, England, suffered an outbreak of bubonic plague in 1665–1666. The source of that plague was believed to be the Great Plague of London. The village is best known for being the "plague village" that chose to isolate itself when the plague was discovered there in August 1665 rather than let the

infection spread. Detailed records were preserved. The initial population of Eyam was 350. In mid-May 1666, nine months after the beginning of the epidemic, there were 254 susceptibles and 7 infectives. The data about the epidemic in the remaining months are given in Table 2.2. The infective period of the bubonic plague is 11 days.

(a) Estimate α
(b) Use the implicit solution of the SIR model to estimate β.
(c) Plot S and I alongside the data. Do they fit?

2.8. A first-order differential equation is given by $x'(t) = f(x)$, where $f(x)$ is defined by Fig. 2.10.

(a) Determine the equilibria of the model $x' = f(x)$.
(b) Determine the local stabilities of the equilibria of the model $x' = f(x)$.
(c) Graph the solutions $x(t)$ of the model $x' = f(x)$ as a function of time.
(d) What is the limit

$$\lim_{t\to\infty} x(t)$$

if $x(0) = 15$? What about if $x(0) = 1$?

Table 2.2 Number of susceptible and infected individuals during the Great Plague of Eyam

Date 1666	No. susceptible	No. infected
Mid-May	254	7
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 3/4	Unknown	Unknown
October 20	83	0

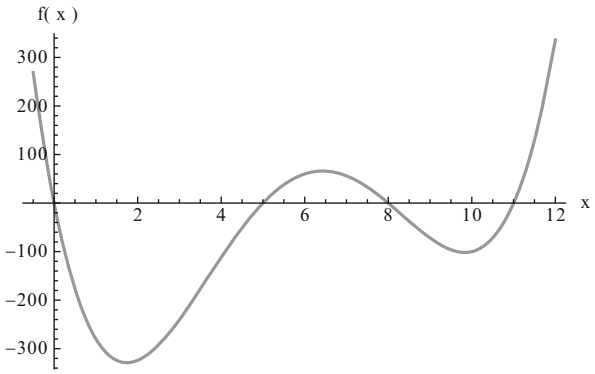

Fig. 2.10 Graph of the function $f(x)$

Chapter 3
The SIR Model with Demography: General Properties of Planar Systems

3.1 Modeling Changing Populations

Models that do not explicitly include births and deaths occurring in the population are called *epidemic models without explicit demography*. They are useful for epidemic modeling on a short time scale, particularly for modeling epidemic outbreaks such as influenza. Omitting population change requires that the disease develop on a much shorter time scale than the period in which significant change in the population size can occur (such as births and deaths). This is valid for *fast* diseases like the childhood diseases and influenza. On the other hand, there are *slow* diseases, such as HIV, tuberculosis, and hepatitis C, that develop for a long period of time even on an individual level. In this case, the total population does not remain constant for long periods of time, and the demography of the population cannot be ignored.

To incorporate the population change in epidemic models, we need population models of the growth of the human population. There are several classical population models that are typically considered in the literature.

Population growth is the rate of change in a population over time, and it can be approximated as the change in the number of individuals of any species in a population per unit time. The study of growth and change of human populations is called *demography*. Modeling and projecting the growth of human populations is in general not a simple matter, but for the purposes of epidemic modeling, we will use several simple population models.

3.1.1 The Malthusian Model

The **Malthusian model**, sometimes called the exponential model, is essentially an exponential growth model based on the assumption that the rate of change of a population is proportional to the total population size. The model is named after the Reverend Thomas Malthus (1766–1834), who authored *An Essay on*

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_3

the Principle of Population, one of the earliest and most influential books on populations. The Malthusian model is based on the following assumptions: (1) All individuals are identical, that is, they are not classified by age, sex, or other characteristics. (2) The environment is constant in space and time, in particular, resources are unlimited. With these assumptions, if $N(t)$ is the total population size, and b is the per capita birth rate, while μ is the per capita death rate, then the Malthusian model becomes

$$N'(t) = bN(t) - \mu N(t) = rN(t), \tag{3.1}$$

where $r = b - \mu$ is the population growth rate. The solution to this equation is an exponential $N(t) = N(0)e^{rt}$. The population is growing exponentially if $r > 0$, decreasing exponentially if $r < 0$, and constant if $r = 0$.

We compare the performance of population models with world population data. Table 3.1 gives the world's human population since 1950.

Table 3.1 World population size 1950–2010[a]

Year	Population	Year	Population
1950	2,556,505,579	1980	4,452,686,744
1952	2,635,724,824	1982	4,615,366,900
1954	2,729,267,486	1984	4,776,577,665
1956	2,834,435,383	1986	4,941,825,082
1958	2,947,380,005	1988	5,114,949,044
1960	3,042,389,609	1990	5,288,828,246
1962	3,139,645,212	1992	5,456,405,468
1964	3,280,890,090	1994	5,619,031,095
1966	3,420,438,740	1996	5,779,990,768
1968	3,562,227,755	1998	5,935,741,324
1970	3,712,813,618	2000	6,088,683,554
1972	3,867,163,052	2002	6,241,717,680
1974	4,017,615,739	2004	6,393,120,940
1976	4,161,423,905	2006	6,545,884,439
1978	4,305,496,751	2008	6,700,765,879
–	–	2010	6,853,019,414

[a] Data taken from http://www.census.gov/ipc/www/idb/worldpop.php

With the simple population models in this section, many methods for estimating the parameters can work. One of the most powerful methods, however, is **calibration** or **curve fitting**. Curve fitting is the process of identifying the parameters of a curve, or mathematical function, that has the best fit to a series of data points. We discuss more thoroughly fitting epidemic models to data in Chap. 6. Here we only compare the population models with the available data.

Calibration is greatly expedited through the use of software such as Mathematica, Matlab, or R to fit the model to the data. For the Malthusian model, we have an explicit solution, and we can fit the solution function to the data. Since the initial condition for the data is not at zero, the solution to the Malthus model becomes $N(t) = Ae^{r(t-1950)}$. We fit both A and r. Fitting in Mathematica can be done with the

command `NonlinearModelFit`. The result of the fit of the world population data to the Malthusian model is given in Fig. 3.1, where the population is taken in millions.

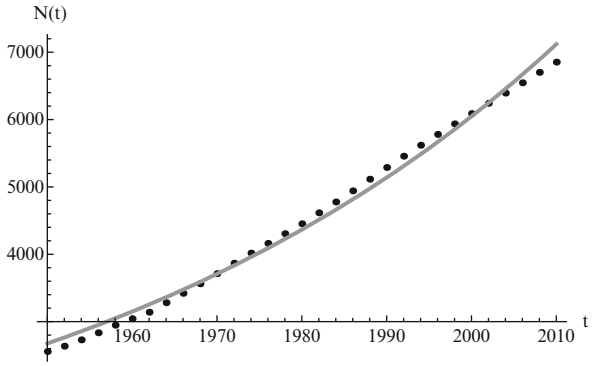

Fig. 3.1 World population data alongside Malthusian model predictions. The estimated values of the parameters are $A = 2676.29$ and $r = 0.0163$. The least-squares error of the fit is $E = 402{,}533$

3.1.2 The Logistic Model as a Model of Population Growth

The Malthus model assumes that the population's per capita growth rate is constant and that the population has unlimited resources by which to grow. In most cases, however, populations live in an environment that has a finite capacity to support only a certain population size. When the population size approaches this capacity, the per capita growth rate declines or becomes negative. This property of the environment to limit population growth is captured by the logistic model. The logistic model was developed by the Belgian mathematician Pierre Verhulst (1838), who suggested that the per capita growth rate of the population may be a decreasing function of population density:

$$\frac{1}{N(t)}N'(t) = r\left(1 - \frac{N}{K}\right),$$

which gives the classical logistic model that we studied in Chap. 2. At low densities $N(t) \approx 0$, the population growth rate is maximal and equals r. The parameter r can be interpreted as the population growth rate in the absence of intraspecific competition. The population growth rate declines with population number N and reaches 0 when $N = K$. The parameter K is the upper limit of population growth, and it is called the **carrying capacity** of the environment. It is usually interpreted as the quantity of resources expressed in the number of organisms that can be supported by those resources. If population number exceeds K, then population growth rate becomes negative, and population declines. The logistic model has been used unsuccessfully for the projection of human populations. The main difficulty appears to be determining the carrying capacity of a human population. It is believed that human populations do not have a carrying capacity, and even if they do, that the

carrying capacity is not constant. For those reasons, the logistic model is rarely used to model human populations. However, when compared to population data, the logistic equation usually performs admirably in modeling the data for short periods of time. We use the logistic model to model the world population data. The results are given in Fig. 3.2.

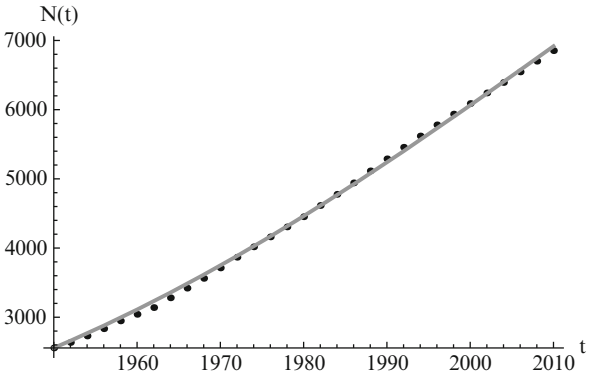

Fig. 3.2 World population data alongside logistic model predictions. The estimated values of the parameters are $K = 13{,}863.9$ and $r = 0.0247$. The least-squares error of the fit is $E = 56{,}659.3$

3.1.3 A Simplified Logistic Model

The third model of population growth is a simplified version of the logistic model. It assumes constant birth rate, independent of population size. It also assumes constant per capita death rate. The model becomes

$$N'(t) = \Lambda - \mu N.$$

Here Λ is the total birth rate, and μ is the per capita natural death rate. Then μN is the total death rate. This model can be solved. The solution is

$$N(t) = N_0 e^{-\mu t} + \frac{\Lambda}{\mu}(1 - e^{-\mu t}).$$

It is not hard to see that if $t \to \infty$, then $N(t) \to \frac{\Lambda}{\mu}$. This limit quantity is called the limit population size. The simplified logistic model is the one most often used to model population dynamics in epidemic models. However, its performance with data is modest. We illustrate how the simplified logistic model fits the world population data in Fig. 3.3. We saw in Chap. 2 that if T is the time spent is a class (or a compartment), then the per capita rate at which the individuals leave that class (compartment) is given by $\frac{1}{T}$. So if the per capita recovery rate was α, then

$$\alpha = \frac{1}{T},$$

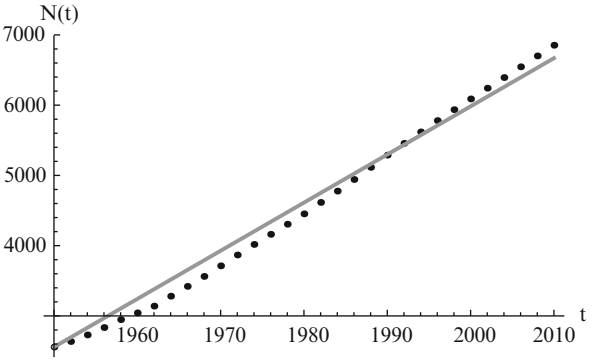

Fig. 3.3 World population data alongside the simplified logistic model predictions. The estimated values of the parameters are $\mu = 5.54 \times 10^{-12}$ and $\Lambda = 68.5$, and we have preset $N(1950) = 2556.5$. The least-squares error of the fit is $E = 703{,}482$

or equivalently, $\frac{1}{\alpha}$ is the time spent in the infectious compartment. Similar reasoning can be applied to the compartment "life." If μ is the natural death rate, then $1/\mu$ should be the average lifespan of an individual human being. From fitting the simplified logistic model to world data, we estimated $\mu = 5.54 \times 10^{-12}$, which gives a lifespan of 1.8×10^{11} years—quite unrealistic. If the lifespan is limited to biologically realistic values, such as a lifespan of 65 years, then the fit becomes worse.

3.2 The SIR Model with Demography

To incorporate the demographics into the SIR epidemic model, we assume that all individuals are born susceptible. Individuals from each class die at a per capita death rate μ, so the total death rate in the susceptible class is μS, while in the infective class, it is μI, and in the removed class, it is μR. The epidemic model with demography becomes

$$\begin{aligned} S'(t) &= \Lambda - \beta IS - \mu S, \\ I'(t) &= \beta IS - \alpha I - \mu I, \\ R'(t) &= \alpha I - \mu R. \end{aligned} \tag{3.2}$$

We add the three equations to obtain the total population. The model of the total population is $N'(t) = \Lambda - \mu N$, where $N = S + I + R$. The population size is not constant, but it is asymptotically constant, since $N(t) \to \frac{\Lambda}{\mu}$ as $t \to \infty$.

When the population is nonconstant and the incidence is proportional to the product of I and S, we say that the incidence is given by the **law of mass action**, analogously to terms from chemical kinetic models, whereby chemicals react by bumping randomly into each other. For this reason, this incidence is called the **mass action incidence**:

$$\text{mass action incidence} = \beta SI.$$

Another type of incidence that is very commonly used in epidemic models is the **standard incidence**. It is similar to the mass action incidence, but it is normalized by the total population size. In particular,

$$\text{standard incidence} = \frac{\beta SI}{N}.$$

The mass action incidence and the standard incidence agree when the total population size is a constant, but they differ if the total population size is variable. Mass action incidence is used in diseases for which disease-relevant contact increases with an increase in the population size. For instance, in influenza and SARS, contacts increase as the population size (and density) increase. Standard incidence is used for diseases for which the contact rate cannot increase indefinitely and is limited even if the population size increases. This is the case in sexually transmitted diseases, where the number of contacts cannot increase indefinitely.

We notice as before that the first two equations in (3.2) are independent of the third, and we consider the two-dimensional system

$$\begin{aligned} S'(t) &= \Lambda - \beta IS - \mu S, \\ I'(t) &= \beta IS - \alpha I - \mu I, \end{aligned} \tag{3.3}$$

where $R = N - S - I$. Mathematically, the SIR system can be written in the general form

$$\begin{aligned} S'(t) &= f(S,I), \\ I'(t) &= g(S,I). \end{aligned} \tag{3.4}$$

This is a system of differential equations with two equations and the two unknowns S and I. The incidence term makes both f and g nonlinear functions. So system (3.4) is a nonlinear system of differential equations. System (3.4) is also **autonomous**, since f and g do not depend explicitly on the time variable; that is, the coefficients of system (3.3) are constants and not functions of time.

What are the units of the quantities in this model? Since S is measured in number of people, it follows that S' is measured in number of people per unit of time. The total birth rate Λ is measured in number of people born per unit of time. The per capita death rate μ is measured in [unit of time]$^{-1}$. Thus, μS is measured again in number of people per unit of time. The most difficult term is βIS. Since the force of infection βI is a per capita rate, it has units [time]$^{-1}$. Consequently, the transmission coefficient β must have units of [number of people$\times$ time]$^{-1}$.

A customary transformation of the system (3.3) that simplifies the system and reduces the number of parameters is often performed. There is a simplification that consists in a change of variables that transforms both the independent variable and the dependent variables into nondimensional quantities. Hence, we say that we have transformed the system into a **nondimensional form**.

Two parameters have units [unit of time]$^{-1}$: α and μ. Since t is in [unit of time], we have to multiply t by one of the rates to obtain a unitless quantity. It is best to define $\tau = (\alpha + \mu)t$. Observe that τ is a dimensionless quantity. Because of the nature of the change, this change will remove the parameter multiplying I. Let $N(t) = N(\frac{\tau}{\alpha+\mu}) = \hat{N}(\tau)$. Similarly, $I(t) = \hat{I}(\tau)$. By the chain rule, we have

$$\begin{aligned} \frac{d\hat{S}}{d\tau} &= \frac{1}{\alpha+\mu}\frac{dS}{dt}, \\ \frac{d\hat{I}}{d\tau} &= \frac{1}{\alpha+\mu}\frac{dI}{dt}. \end{aligned} \tag{3.5}$$

We rescale the $\hat{S}$ and $\hat{I}$ variables with the total limiting population size. Hence $x(t) = \frac{\mu\hat{S}}{\Lambda}$ and $y(t) = \frac{\mu\hat{I}}{\Lambda}$. The new dependent variables $x(\tau)$ and $y(\tau)$ are also dimensionless quantities. The system for them becomes

$$\begin{aligned} x' &= \rho(1-x) - \mathscr{R}_0 xy, \\ y' &= (\mathscr{R}_0 x - 1)y, \end{aligned} \tag{3.6}$$

where

$$\rho = \mu/(\alpha+\mu) \qquad\qquad \mathscr{R}_0 = \frac{\Lambda\beta}{\mu(\alpha+\mu)}$$

are both dimensionless parameters. The notation $\mathscr{R}_0$ is not random. As we will see later, this dimensionless quantity is indeed the reproduction number. Notice that we have reduced the number of parameters from five to two. The dimensionless form of the SIR model with demography is equivalent to the original one, since the solutions of both systems have the same long-term behavior.

3.3 Analysis of Two-Dimensional Systems

We cannot solve the SIR model with demography analytically, but we can obtain some information about the behavior of the solutions. The long-term behavior of the solutions is particularly important from an epidemiological perspective, since we would like to know what will happen to the disease in the long run: will it die out, or will it establish itself in the population and become *endemic*?

3.3.1 Phase-Plane Analysis

We write the system (3.6) in general form

$$\begin{aligned} x' &= f(x,y), \\ y' &= g(x,y), \end{aligned} \tag{3.7}$$

where $f(x,y) = \rho(1-x) - \mathscr{R}_0 xy$ and $g(x,y) = (\mathscr{R}_0 x - 1)y$. To answer the question above, we have to investigate the long-term behavior of the solutions. Instead of considering $x(\tau)$ and $y(\tau)$ as functions of τ, or equivalently, $S(t)$ and $I(t)$ as functions of t, we treat τ as a parameter and consider the curves in the (x,y)-plane, obtained from the points $(x(\tau), y(\tau))$ as τ varies as a parameter. By considering the solution curves in the (x,y)-plane, we say that we are considering the **phase plane**.

Definition 3.1. Curves in the phase plane representing the functional relation between x and y, with τ as a parameter, are called *orbits* or *trajectories*.

The long-term behavior of the trajectories depends largely on the **equilibrium points**, that is, on time-independent solutions of the system. Equilibrium points are solutions for which $x' = 0$ and $y' = 0$.

Definition 3.2. All points (x^*, y^*), where x^* and y^* are constants that satisfy the system

$$\begin{aligned} f(x^*, y^*) &= 0, \\ g(x^*, y^*) &= 0, \end{aligned} \tag{3.8}$$

are called *equilibria* or *singular points*.

For the dimensionless SIR model with demography, we have

$$\begin{aligned} \rho(1-x) - \mathscr{R}_0 xy &= 0, \\ (\mathscr{R}_0 x - 1)y &= 0. \end{aligned} \tag{3.9}$$

We have that if $y = 0$, that is, there are no infectives, then $x = 1$; that is, everyone is susceptible. This gives the first equilibrium in the (x,y)-plane, $(1,0)$. This is the disease-free equilibrium. The disease-free equilibrium is also a **boundary equilibrium**, since it lies on the boundary of the feasible region $x \geq 0$, $y \geq 0$. If $y \neq 0$, then from the second equation, we have $x = 1/\mathscr{R}_0$. From the first equation, we have $y = \rho(1 - 1/\mathscr{R}_0)$. Thus the second equilibrium is the point

$$\mathscr{E} = \left(\frac{1}{\mathscr{R}_0}, \rho\left(1 - \frac{1}{\mathscr{R}_0}\right) \right).$$

This is the endemic equilibrium. The endemic equilibrium exists only in the case $\mathscr{R}_0 > 1$. This equilibrium is also called an **interior equilibrium**.

System (3.6) allows us to compute the slope at each point of a trajectory in the (x,y)-plane. The parameter τ can be eliminated by dividing the equations in system (3.6):

$$\frac{dy}{dx} = \frac{g(x,y)}{f(x,y)}.$$

This quotient is defined for all points in the (x,y)-plane except the equilibria. For any nonequilibrium point (x_0,y_0) in the phase plane, we can compute the expression

$$\frac{dy}{dx}|_{(x_0,y_0)} = \frac{g(x,y)}{f(x,y)}$$

which gives the slope of the trajectory in the (x,y)-plane, with tangent vector

$$(f(x_0,y_0),g(x_0,y_0))^T.$$

This vector also gives the direction of the trajectory. The tangent vector is not defined at the equilibria, since the flow stops at those points and they are fixed points.

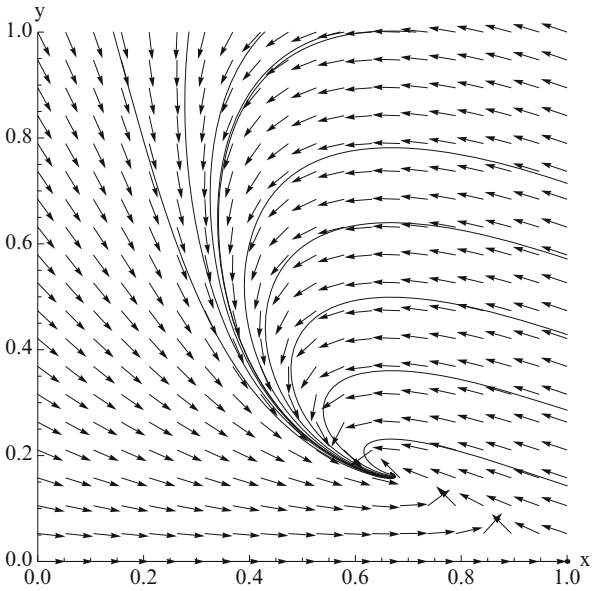

Fig. 3.4 The vector field of the dimensionless SIR model alongside solutions of the model for several initial conditions

The collection of tangent vectors defines a **direction field**. The direction field can be used as a visual aid in sketching a family of solutions called a **phase-plane portrait** or a **phase-plane diagram**. A phase-plane portrait of the dimensionless SIR model is given in Fig 3.4. The creation of the whole phase portrait is a tedious

job and is done only by computer. An easier method to obtain information about the direction of the flow is to analyze the direction of the flow along the x-zero and y-zero isoclines, or nullclines.

Definition 3.3. The *x-zero isocline* or *x-nullcline* for the system (3.7) is the set of all points in the (x,y)-plane satisfying

$$f(x,y) = 0.$$

The *y-zero isocline* or *y-nullcline* for the system (3.7) is the set of all points in the (x,y)-plane satisfying

$$g(x,y) = 0.$$

We can determine the nullclines for the dimensionless SIR model. Setting $\rho(1-x) - \mathscr{R}_0 xy = 0$ gives the x-nullcline

$$y = \frac{\rho}{\mathscr{R}_0}\frac{1-x}{x}.$$

Setting $(\mathscr{R}_0 x - 1)y = 0$ gives two y-nullclines: $y = 0$, which is the x-axis, and the vertical line $x = \frac{1}{\mathscr{R}_0}$. The points where an x-nullcline intersects a y-nullcline give the equilibrium points of the system. There are two scenarios for the SIR dimensionless system.

$\mathscr{R}_0 < 1$ In this case, there is only one intersection of an x-nullcline and a y-nullcline. The x-nullcline intersects the y-nullcline $y = 0$ at the point $(1,0)$, the disease-free equilibrium. Since $1/\mathscr{R}_0 > 1$, the y-nullcline $x = 1/\mathscr{R}_0$ does not intersect the x-nullcline in the positive quadrant.

$\mathscr{R}_0 > 1$ In this case, there are two intersections of an x-nullcline and a y-nullcline. In the first intersection, the x-nullcline intersects the y-nullcline $y = 0$ at the point $(1,0)$, which represents the disease-free equilibrium. In the second intersection, since $1/\mathscr{R}_0 < 1$, the y-nullcline $x = 1/\mathscr{R}_0$ intersects the x-nullcline $1/\mathscr{R}_0 < 1$ at the point $\mathscr{E}$, which gives the endemic equilibrium.

To determine the direction of the vector field along the nullclines, we can use the following general rules:

1. On the x-nullclines, the tangent vector is

$$(0, g(x_0,y_0))^T$$

and is parallel to the y-axis. The direction of the tangent is given by the sign of $g(x_0,y_0)$. If $g(x_0,y_0) > 0$, the direction vector points upward. If $g(x_0,y_0) < 0$, the directional vector points downward.
2. On the y-nullclines, the tangent vector is

$$(f(x_0,y_0), 0)^T$$

and is parallel to the x-axis. The direction of the tangent vector is determined by the sign of $f(x_0, y_0)$. If $f(x_0, y_0) > 0$, the direction vector points to the right. If $f(x_0, y_0) < 0$, the direction vector points to the left.

We determine the direction field along nullclines for the dimensionless SIR model. We consider the case $\mathscr{R}_0 > 1$. The case $\mathscr{R}_0 < 1$ is similar. The results are illustrated in Fig. 3.5.

1. On the x-nullcline, the tangent vector is $(0, g(x_0, y_0))^T$, where (x_0, y_0) is a point on the nullcline. The tangent vector is parallel to the y-axis. Since $g(x_0, y_0) = (\mathscr{R}_0 x_0 - 1) y_0$ and $y_0 > 0$, the sign of $g(x_0, y_0)$ is determined by the first term in the product. Thus, if $x_0 < 1/\mathscr{R}_0$, then $g(x_0, y_0) < 0$, and the vector points downward. If $x_0 > 1/\mathscr{R}_0$, then $g(x_0, y_0) > 0$, and the vector points upward.
2. On the y-nullclines, the tangent vector is $(f(x_0, y_0), 0)^T$, where (x_0, y_0) is a point on a y-nullcline. The tangent vector is parallel to the x-axis. Since there are two y-nullclines, we consider two cases:

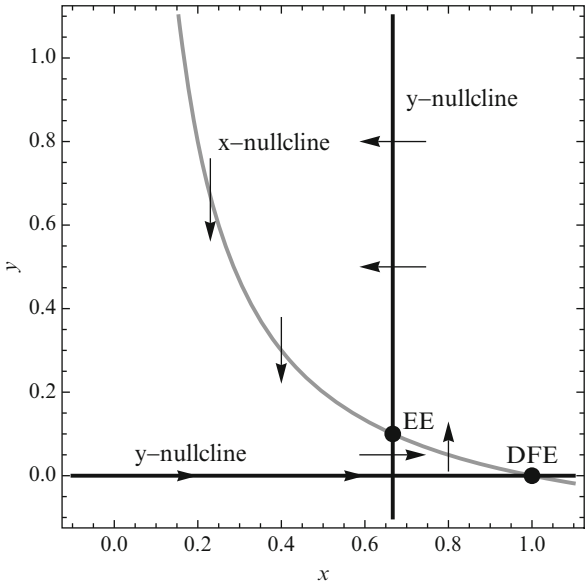

Fig. 3.5 Phase-plane analysis of the dimensionless SIR model. Nullclines and the direction of the vector field along them

$y = 0$ On the nullcline $y = 0$, $f(x_0, y_0) = \rho(1 - x_0)$. We have $f(x_0, y_0) > 0$ if $x_0 < 1$, and the tangent vector points to the right. Furthermore, we have $f(x_0, y_0) < 0$ if $x_0 > 1$, and the tangent vector points to the left.

$x = \frac{1}{\mathscr{R}_0}$ On the y-nullcline $x = \frac{1}{\mathscr{R}_0}$, we have $f(x_0, y_0) = \rho\left(1 - \frac{1}{\mathscr{R}_0}\right) - y_0$. We have that $y_0 > \rho\left(1 - \frac{1}{\mathscr{R}_0}\right)$ if the point (x_0, y_0) is on the y-nullcline

above the intersection of the y-nullcline with the x-nullcline. Hence, $f(x_0, y_0) < 0$, and the tangent vector points to the left. Furthermore, we have that $y_0 < \rho\left(1 - \frac{1}{\mathscr{R}_0}\right)$ if the point (x_0, y_0) is on the y-nullcline *below* the intersection of the y-nullcline with the x-nullcline. Hence, $f(x_0, y_0) > 0$, and the tangent vector points to the right.

3.3.2 Linearization

Just as with first-order nonlinear equations, we can obtain information about the behavior of the solutions near an equilibrium through *linearization*. If (x^*, y^*) is an equilibrium, we consider the perturbation of a solution starting from an initial condition close to the equilibrium:

$$u(\tau) = x(\tau) - x^* \qquad v(\tau) = y(\tau) - y^*.$$

We note again that $u(\tau)$ and $v(\tau)$ are functions of τ but are not necessarily nonnegative. Writing $x(\tau) = u(\tau) + x^*$, $y(\tau) = v(\tau) + y^*$, and substituting in the original system, we obtain

$$\begin{aligned} u' &= f(u + x^*, v + y^*), \\ v' &= g(u + x^*, v + y^*). \end{aligned} \tag{3.10}$$

Assuming that f and g have at least second-order continuous partial derivatives, we expand in a Taylor series using the theorem for functions of two variables. We show that expansion for f; the process for g is the same.

$$\begin{aligned} f(u + x^*, v + y^*) = {} & f(x^*, y^*) + f_x(x^*, y^*)u(\tau) + f_y(x^*, y^*)v(\tau) \\ & + f_{xx}(x^*, y^*)u^2(\tau)/2 + f_{xy}(x^*, y^*)u(\tau)v(\tau) \\ & + f_{yy}(x^*, y^*)v^2(\tau)/2 + \cdots. \end{aligned} \tag{3.11}$$

The terms with the second partial derivatives are multiplied by u^2, uv, and v^2, all second-order terms in the perturbations. If the perturbations are small, $u \approx 0$ and $v \approx 0$, then the second-order terms are even smaller, so we may ignore them. Thus,

$$\begin{aligned} u' &\approx f(x^*, y^*) + f_x(x^*, y^*)u(\tau) + f_y(x^*, y^*)v(\tau), \\ v' &\approx g(x^*, y^*) + g_x(x^*, y^*)u(\tau) + g_y(x^*, y^*)v(\tau). \end{aligned} \tag{3.12}$$

Since (x^*, y^*) is an equilibrium, $f(x^*, y^*) = 0$ and $g(x^*, y^*) = 0$. We obtain the **linearized system**

$$\begin{aligned} u' &= f_x(x^*, y^*)u(\tau) + f_y(x^*, y^*)v(\tau), \\ v' &= g_x(x^*, y^*)u(\tau) + g_y(x^*, y^*)v(\tau). \end{aligned} \tag{3.13}$$

The matrix of the partial derivatives of the functions $f(x,y)$ and $g(x,y)$ is called the Jacobian matrix. The matrix of the system above is the Jacobian matrix evaluated at an equilibrium (x^*,y^*). All entries of this matrix are given constants:

$$J = \begin{pmatrix} f_x(x,y) & f_y(x,y) \\ g_x(x,y) & g_y(x,y) \end{pmatrix} |_{x=x^*,y=y^*} \tag{3.14}$$

An important result, called the **Hartman–Grobman theorem** justifies drawing conclusions about a nonlinear system from studying the linearized system. The Hartman–Grobman theorem says roughly that the solutions of an $n \times n$ autonomous system of ordinary differential equations in a neighborhood of a steady state look "qualitatively" just like the solutions of the linearized system (3.13) near the point $(0,0)$. This result holds only when the equilibrium is a **hyperbolic equilibrium**, that is, when none of the eigenvalues of J have zero real part.

3.3.3 Two-Dimensional Linear Systems

The linearized system (3.13) can be written in the form

$$\begin{aligned} u' &= au(\tau) + bv(\tau), \\ v' &= cu(\tau) + dv(\tau), \end{aligned} \tag{3.15}$$

where a,b,c,d are given constants. The system (3.15) is a two-dimensional linear homogeneous system. The behavior of solutions of such systems has been completely studied. In this subsection, we review what is known about two-dimensional linear systems. The equilibria of linear two-dimensional systems are solutions to the linear system of equations

$$\begin{aligned} au(\tau) + bv(\tau) &= 0, \\ cu(\tau) + dv(\tau) &= 0. \end{aligned} \tag{3.16}$$

Such systems always have $(0,0)$ as a solution. The equilibrium $(0,0)$ is the only equilibrium if the matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{3.17}$$

of the system is invertible, that is, Det$A \neq 0$. We will assume that this condition holds, because if it doesn't, there is a continuum of equilibria. Thus, we assume that $ad - bc \neq 0$. If the matrix A is obtained from the linearization and is the Jacobian evaluated at an equilibrium (x^*,y^*), the condition Det$J \neq 0$ means that the equilibrium is **isolated**; that is, there is a disk around it that does not contain other equilibria. Looking for exponential solutions of the linearized system (3.15), we set

$$u(\tau) = \bar{u}e^{\lambda\tau} \qquad v(\tau) = \bar{v}e^{\lambda\tau},$$

where $\bar{u}$ and $\bar{v}$ are nonzero constants. Substituting in the system and canceling $e^{\lambda\tau}$, we obtain the following system for $\bar{u}$ and $\bar{v}$:

$$\begin{aligned} a\bar{u}+b\bar{v} &= \lambda\bar{u}, \\ c\bar{u}+d\bar{v} &= \lambda\bar{v}. \end{aligned} \tag{3.18}$$

This is a linear homogeneous system for $\bar{u}$ and $\bar{v}$. We want this system to have a nonzero solution, since our perturbations should be nonzero. This can happen only if the determinant of the system is zero, so we have

$$\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = 0. \tag{3.19}$$

By expanding the determinant $(a-\lambda)(d-\lambda)-bc=0$, we obtain the **characteristic equation** of the linearized system:

$$\lambda^2 - p\lambda + q = 0, \tag{3.20}$$

where $p = a+d = \mathrm{Tr}J$, and $q = ad-bc = \mathrm{Det}J$. Thus p is the trace and q is the determinant of the Jacobian matrix. The solutions of the characteristic equation are called the **eigenvalues** of the Jacobian matrix. The main question that we address is when the perturbations u and v approach zero, in which case the equilibrium (x^*, y^*) will be locally asymptotically stable. Given the eigenvalues, we have three cases for the solution of the system of perturbations (3.15).

Case 1 The eigenvalues of the Jacobian are real and distinct, say λ_1 and λ_2. In this case, the solution of the system (3.15) is given by

$$\begin{aligned} u(\tau) &= C_1 e^{\lambda_1\tau} + C_2 e^{\lambda_2\tau}, \\ v(\tau) &= C_3 e^{\lambda_1\tau} + C_4 e^{\lambda_2\tau}, \end{aligned} \tag{3.21}$$

where $C_1, \ldots, C_4$ are arbitrary constants. Clearly, in this case, $u \to 0$ and $v \to 0$ if and only if $\lambda_1 < 0$ and $\lambda_2 < 0$.

Case 2 The eigenvalues of the Jacobian are real and equal, say λ. In this case, the solution of the system (3.15) is given by

$$\begin{aligned} u(\tau) &= C_1 e^{\lambda\tau} + C_2 \tau e^{\lambda\tau}, \\ v(\tau) &= C_3 e^{\lambda\tau} + C_4 \tau e^{\lambda\tau}, \end{aligned} \tag{3.22}$$

where $C_1, \ldots, C_4$ are arbitrary constants. Clearly, in this case, $u \to 0$ and $v \to 0$ if and only if $\lambda < 0$.

Case 3 The eigenvalues of the Jacobian are complex conjugates, say $\lambda_1 = \xi + \eta i$ and $\lambda_2 = \xi - \eta i$. In this case, the real solution of the system (3.15) is given by

$$\begin{aligned} u(\tau) &= C_1 e^{\xi\tau} \sin\eta\tau + C_2 e^{\xi\tau} \cos\eta\tau, \\ v(\tau) &= C_3 e^{\xi\tau} \sin\eta\tau + C_4 e^{\xi\tau} \cos\eta\tau, \end{aligned} \tag{3.23}$$

where $C_1, \ldots, C_4$ are arbitrary constants. Clearly, in this case, $u \to 0$ and $v \to 0$ if and only if $\xi < 0$; that is, the eigenvalues have negative real part.

We summarize this result in the following widely used theorem.

Theorem 3.1. *A necessary and sufficient condition for an equilibrium to be locally asymptotically stable is that all eigenvalues of the Jacobian have negative real part.*

For two-dimensional systems, there is a simple necessary and sufficient condition that all eigenvalues of a matrix have negative real part.

Theorem 3.2. *Assume that J is a 2×2 matrix with constant entries and* $\mathrm{Det}J \neq 0$. *Assume that J has been obtained as a linearization around the equilibrium (x^*, y^*). Then the equilibrium (x^*, y^*) is locally asymptotically stable if and only if*

$$\mathrm{Tr}J < 0 \qquad \text{and} \qquad \mathrm{Det}J > 0.$$

The equilibrium (x^*, y^*) is unstable if and only if

$$\mathrm{Tr}J > 0 \qquad \text{or} \qquad \mathrm{Det}J < 0.$$

Remark 3.1. The asymptotic stability of the equilibrium (x^*, y^*) of the nonlinear system is equivalent to the asymptotic stability of the $(0,0)$ equilibrium of the linear system obtained from the linearization around the equilibrium (x^*, y^*). The only exception occurs when $\mathrm{Tr}\, J = 0$ and $\mathrm{Det}\, J > 0$. In this case, the characteristic equation has eigenvalues with zero real part. Consequently, the $(0,0)$ equilibrium of the linear system may be stable, but there are no implications for the stability of the (x^*, y^*) equilibrium of the nonlinear system.

The origin of a two-dimensional linear system can be classified as one of four types: **node**, **spiral**, **saddle**, or **center**. In addition, the origin can be classified as stable or unstable. This classification depends on whether the eigenvalues are real or complex, positive or negative when real, or with positive or negative real part when complex. We have the following cases:

Node The origin is said to be a *node* if the eigenvalues are real and of the same sign. If the two eigenvalues are negative, the node is a **stable node**. If both eigenvalues are positive, the node is an **unstable node**. If the eigenvalues are real and equal, the node that corresponds to them is called **degenerate**. If two eigenvectors correspond to the double eigenvalue, the degenerate node is called **proper**. If only one eigenvector corresponds to the double eigenvalue, the degenerate node is called **improper**.

Saddle The origin is a saddle if the eigenvalues are real and of opposite sign. A saddle is always unstable.

Spiral The origin is a spiral (or focus) if the eigenvalues are complex with nonzero real part. If the real part is negative, the focus is a **stable focus**; if the real part is positive, the focus is an **unstable focus**.

Center The origin is a center if the eigenvalues are complex with zero real part (purely imaginary). In this case, every orbit is periodic. The center is stable but not asymptotically.

The type of the equilibrium can be inferred from the coefficients of the characteristic equation (3.20). See Table 3.2.

Table 3.2 Relations between the coefficients of the characteristic equation and the type of the equilibrium

Coefficients	Trace and determinant	Type
$q < 0$	Det$J < 0$	Saddle (unstable)
$q > 0$, $p < 0$, $\Delta = p^2 - 4q \geq 0$	Det$J > 0$, Tr $J < 0$	Stable node
$q > 0$, $p < 0$, $\Delta = p^2 - 4q < 0$	Det$J > 0$, Tr $J < 0$	Stable focus
$q > 0$, $p > 0$, $\Delta = p^2 - 4q \geq 0$	Det$J > 0$, Tr $J > 0$	Unstable node
$q > 0$, $p > 0$, $\Delta = p^2 - 4q < 0$	Det$J > 0$, Tr $J > 0$	Unstable focus
$q > 0$, $p = 0$	Det$J > 0$, Tr $J = 0$	Center

3.4 Analysis of the Dimensionless SIR Model

We saw that the dimensionless SIR model (3.6) has two equilibria. The disease-free equilibrium $(1,0)$ always exists, while the endemic equilibrium $\mathscr{E}$ exists only if $\mathscr{R}_0 > 1$.

3.4.1 Local Stability of the Equilibria of the SIR Model

The local stability of equilibria is determined by the eigenvalues of the Jacobian computed at that equilibrium. The Jacobian of the dimensionless SIR model at an equilibrium (x^*, y^*) is

$$J = \begin{pmatrix} -\rho - \mathscr{R}_0 y^* & -\mathscr{R}_0 x^* \\ \mathscr{R}_0 y^* & \mathscr{R}_0 x^* - 1 \end{pmatrix}. \tag{3.24}$$

To obtain the stability of the disease-free equilibrium, we evaluate J at $(1,0)$:

$$J = \begin{pmatrix} -\rho & -\mathscr{R}_0 \\ 0 & \mathscr{R}_0 - 1 \end{pmatrix}. \tag{3.25}$$

The two eigenvalues are $\lambda_1 = -\rho$ and $\lambda_2 = \mathscr{R}_0 - 1$. Since the matrix is upper triangular, the eigenvalues are the diagonal entries of the matrix. The first eigenvalue is clearly negative. The second eigenvalue is negative if $\mathscr{R}_0 < 1$. In this case, the disease-free equilibrium is a stable node. The second eigenvalue λ_2 is positive if $\mathscr{R}_0 > 1$. In this case, the disease-free equilibrium is unstable. It is a saddle.

The next step is to investigate the local stability of the endemic equilibrium. We consider the Jacobian at the endemic equilibrium:

$$J = \begin{pmatrix} -\rho - \mathscr{R}_0 y^* & -\mathscr{R}_0 x^* \\ \mathscr{R}_0 y^* & \mathscr{R}_0 x^* - 1 \end{pmatrix}. \tag{3.26}$$

We notice that from the equilibrium equations we have $\mathscr{R}_0 x^* - 1 = 0$. The Jacobian becomes

$$J = \begin{pmatrix} -\rho - \mathscr{R}_0 y^* & -\mathscr{R}_0 x^* \\ \mathscr{R}_0 y^* & 0 \end{pmatrix}. \tag{3.27}$$

By inspection, the trace of this matrix is negative, $\mathrm{Tr}\, J = -\rho - \mathscr{R}_0 y^* < 0$. The determinant is given by $\mathrm{Det} J = \mathscr{R}_0^2 x^* y^* > 0$. By Theorem 3.2, the endemic equilibrium is locally asymptotically stable.

To determine the type of the endemic equilibrium, we consider the characteristic equation

$$\begin{vmatrix} -\rho - \mathscr{R}_0 y^* - \lambda & -\mathscr{R}_0 x^* \\ \mathscr{R}_0 y^* & -\lambda \end{vmatrix} = 0. \tag{3.28}$$

Expanding the determinant, we obtain the characteristic equation of the endemic equilibrium:

$$\lambda^2 + (\rho + \mathscr{R}_0 y^*)\lambda + \mathscr{R}_0^2 x^* y^* = 0.$$

Since the endemic equilibrium is explicitly known, we can express the coefficients of the characteristic equation in terms of the parameters of the system:

$$\begin{aligned} \rho + \mathscr{R}_0 y^* &= \rho + \mathscr{R}_0 \rho \left(1 - \frac{1}{\mathscr{R}_0}\right) = \rho \mathscr{R}_0, \\ \mathscr{R}_0^2 x^* y^* &= \mathscr{R}_0^2 \frac{1}{\mathscr{R}_0} \rho \left(1 - \frac{1}{\mathscr{R}_0}\right) = \rho(\mathscr{R}_0 - 1). \end{aligned} \tag{3.29}$$

The characteristic equation becomes

$$\lambda^2 + \rho \mathscr{R}_0 \lambda + \rho(\mathscr{R}_0 - 1) = 0.$$

Hence the roots of the characteristic equation are $\lambda_{1,2} = (-\rho \mathscr{R}_0 \pm \sqrt{\Delta})/2$, where $\Delta = (\rho \mathscr{R}_0)^2 - 4\rho(\mathscr{R}_0 - 1)$. Hence if $\Delta > 0$, the characteristic equation has two negative real roots, and the endemic equilibrium is a stable node. If $\Delta < 0$, then the characteristic equation has two complex conjugate roots with negative real part. The endemic equilibrium in this case is a stable focus. The dependent variables $x(\tau)$ and $y(\tau)$ tend to the endemic equilibrium through damped oscillations (see Fig. 3.6). We can compute an approximate period of the oscillation by noting that the mean infectious period $\frac{1}{\alpha}$ is much shorter than the mean lifespan $\frac{1}{\mu}$. That implies that $\alpha \gg \mu$ and $\rho \approx 0$. Hence ρ^2 is very small and can be neglected. Neglecting the quadratic roots for ρ from the expression for the roots of the characteristic equation, we obtain $\lambda_{1,2} = -\rho \mathscr{R}_0/2 \pm i\sqrt{\rho(\mathscr{R}_0 - 1)} = \xi \pm \eta i$. Then the solutions of the linearized problem are of the form $Ce^{\xi\tau}\cos\eta\tau$, that is, functions that oscillate with

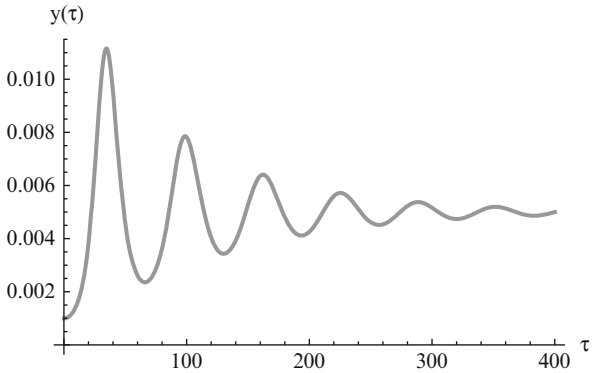

Fig. 3.6 Damped oscillations in the proportion of infectives $x(\tau)$ of the dimensionless SIR model; $\rho = 0.01$, $\mathscr{R}_0 = 2$

decreasing amplitude and approximate period equal to $2\pi/\eta$. Thus, the solution exhibits damped oscillations with period T given by

$$T = \frac{2\pi}{\sqrt{\rho(\mathscr{R}_0 - 1)}}.$$

The following theorem summarizes the results on existence and stability of equilibria of the SIR model with demography.

Theorem 3.3. *Assume $\mathscr{R}_0 < 1$. Then there exists a unique equilibrium, the disease-free equilibrium $(1,0)$, which is locally stable. If $\mathscr{R}_0 > 1$, there are two equilibria: the disease-free equilibrium $(1,0)$, which is unstable, and the endemic equilibrium $\mathscr{E}$, which is locally asymptotically stable.*

3.4.2 The Reproduction Number of the Disease $\mathscr{R}_0$

The expression for $\mathscr{R}_0$ in terms of the original parameters of the system is

$$\mathscr{R}_0 = \frac{\beta\Lambda}{\mu(\alpha+\mu)}.$$

The parameter $\mathscr{R}_0$ is the *reproduction number of the disease*.

Epidemiologically, the reproductive number of the disease tells us how many secondary cases one infected individual will produce in an entirely susceptible population during its period as an infective. Can we see this in the expression that gives $\mathscr{R}_0$?

1. Notice that a population that consists of only susceptible individuals has $\frac{\Lambda}{\mu}$ individuals in the long run.
2. Notice that $\alpha+\mu$ is the rate at which individuals leave the infective class. This means that the average time spent as an infective individual is $\frac{1}{\alpha+\mu}$ time units.

3. The number of transmissions per unit of time is given by the incidence rate βIS. If there is only one infective, $I = 1$, and everybody else is susceptible, $S = \frac{\Lambda}{\mu}$, then the number of transmissions by one infective per unit of time is $\frac{\beta\Lambda}{\mu}$.
4. Thus, the number of transmissions that one infective individual can make during the entire time he/she remains infective if everybody else is susceptible is

$$\frac{\beta\Lambda}{\mu(\alpha+\mu)}.$$

And this is exactly $\mathscr{R}_0$.

The reproduction number of the disease has the following threshold role:

1. If $\mathscr{R}_0 < 1$, then there exists only the disease-free equilibrium. It can be shown that it is attractive, so that every solution of the ODE system approaches this equilibrium, and the disease disappears from the population.
2. If $\mathscr{R}_0 > 1$, then there are two equilibria: the disease-free equilibrium and the endemic equilibrium. The disease-free equilibrium is not attractive in the sense that solutions of the ODE system that start very close to it tend to move away. The endemic equilibrium is attractive, so that solutions of the ODE system approach it as time goes to infinity. Thus, in this case, the disease remains endemic in the population.

3.4.3 Forward Bifurcation

The expression for the endemic equilibrium $\mathscr{E}$ shows that the dimensionless quantity corresponding to infective individuals y^* is a function of the disease reproduction number $\mathscr{R}_0$. It is customary to plot the infective individuals (or y^*) as a function of $\mathscr{R}_0$ in the positive (x,y)-plane, where the x-axis is the reproduction number $\mathscr{R}_0$, and the y-axis is the equilibrium level of the infective individuals y^*. This produces a bifurcation diagram called a **forward bifurcation diagram**, since the endemic equilibrium bifurcates "forward" and exists only for values of the reproduction number greater than one. We have

$$y^* = \begin{cases} 0 & \text{for all } \mathscr{R}_0 < 1 \\ \rho\left(1 - \frac{1}{\mathscr{R}_0}\right) & \mathscr{R}_0 > 1. \end{cases} \tag{3.30}$$

The plot is given in Fig. 3.7.

We plot the locally stable equilibria with solid lines and the unstable equilibria with dashed lines. Hence, since the disease-free equilibrium $y^* = 0$ is locally asymptotically stable for $\mathscr{R}_0 < 1$, it is plotted with a solid line. The endemic equilibrium is also locally asymptotically stable for $\mathscr{R}_0 > 1$ and is also plotted with a solid line. The disease-free equilibrium is unstable for $\mathscr{R}_0 > 1$, and it is plotted with a dashed line.

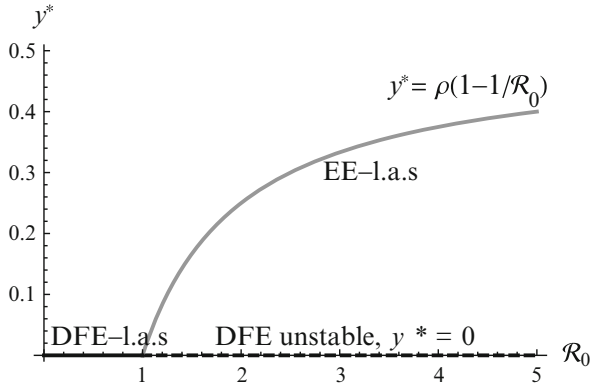

Fig. 3.7 Forward bifurcation diagram with respect to the reproduction number. *Continuous lines* denote stable equilibria. *Dashed lines* denote unstable equilibria

3.5 Global Stability

An equilibrium is called *globally stable* if it is stable for almost all initial conditions, not just those that are close to it. Global stability of an equilibrium cannot always be proved. An equilibrium that is locally stable may be globally stable if there are no other locally stable equilibria coexisting with it. For the SIR model, we have two cases. In the case $\mathscr{R}_0 < 1$, the disease-free equilibrium is the only equilibrium, and it is locally asymptotically stable. It may be expected that it is also globally stable. We establish that in the next subsection. In the case $\mathscr{R}_0 > 1$, the endemic equilibrium is the only locally stable equilibrium, so we may expect that it is also globally stable. We establish that later.

3.5.1 Global Stability of the Disease-Free Equilibrium

Global stability of the disease-free equilibrium can be established for many models, particularly for models for which the disease-free equilibrium is the only equilibrium when $\mathscr{R}_0 < 1$. We note that global stability of the disease-free equilibrium cannot be established for all models. Establishing global stability for the SIR model is perhaps possible through many different techniques. We present one that works well for many models, including partial differential equation models.

Theorem 3.4. *Assume $\mathscr{R}_0 < 1$. Then the disease-free equilibrium is globally stable.*

Proof. Working again with the dimensionless SIR model (3.6), we first notice that if $x(0) > 1$, then $x'(\tau) < 0$, so $x(\tau)$ is a decreasing function if $x > 1$. Assume $\tau_0 > 0$ exits such that $x(\tau_0) = 1$; then $x'(\tau_0) < 1$ and $x(\tau) \leq 1$ for all $\tau \geq \tau_0$. If $x(0) \leq 1$, we may take $\tau_0 = 0$. We consider the equation for $y(\tau)$:

$$y'(\tau) = (\mathscr{R}_0 x - 1)y(\tau). \tag{3.31}$$

For $\tau \geq \tau_0$, we have

$$y'(\tau) \leq (\mathscr{R}_0 - 1)y(\tau).$$

Integrating, we have $y(\tau) = y(\tau_0)e^{(\mathscr{R}_0-1)(\tau-\tau_0)}$. Hence if $\mathscr{R}_0 < 1$, then $\lim_{\tau\to\infty} y(\tau) = 0$. It is somewhat more cumbersome to see that $x \to 1$. First, we notice that $\limsup_{\tau\to\infty} x \leq 1$. We need lim sup, since we do not know that the limit actually exists. From the equation for x, we have

$$x' \leq \rho(1-x),$$

which can be solved in the same way as the corresponding equality would be solved. We have

$$x(\tau) \leq e^{-\rho\tau}x(0) + \rho\int_0^\tau e^{-\rho(\tau-s)}ds.$$

Hence $\limsup_{\tau\to\infty} x \leq 1$. On the other hand, since $\lim_{\tau\to\infty} y = 0$, this implies that for every ε, there exists $\tau_0 > 0$ such that $y \leq \varepsilon$ for $\tau > \tau_0$. For these values of τ, we have

$$x' \geq \rho(1-x) - \varepsilon\mathscr{R}_0 x.$$

Integrating the inequality, we obtain

$$x(\tau) \geq e^{-(\rho+\varepsilon\mathscr{R}_0)\tau}x(0) + \rho\int_0^\tau e^{-(\rho+\varepsilon\mathscr{R}_0)(\tau-s)}ds.$$

This inequality implies that

$$\liminf_{\tau\to\infty} x \geq \frac{\rho}{\rho+\varepsilon\mathscr{R}_0}.$$

Since the inequality holds for every ε, this means that $\liminf_{\tau\to\infty} x \geq 1$. Furthermore, the lim inf and the lim sup are the same, the limit as $\tau \to \infty$ of x exists, and

$$\lim_{\tau\to\infty} x = 1.$$

This completes the proof of the global stability of the disease-free equilibrium. □

3.5.2 Global Stability of the Endemic Equilibrium

We consider again the dimensionless SIR model

$$\begin{aligned} x' &= \rho(1-x) - \mathscr{R}_0 xy, \\ y' &= (\mathscr{R}_0 x - 1)y. \end{aligned} \tag{3.32}$$

This is a planar system. There is theory developed specifically for planar systems that can facilitate understanding of the solution behavior and the proof of global stability. To introduce the main results of that theory, consider a general planar system

$$\begin{aligned} x' &= f(x,y) \qquad x(0) = u_1^0, \\ y' &= g(x,y) \qquad y(0) = u_2^0. \end{aligned} \tag{3.33}$$

Let $u(t) = (x(t), y(t))$ be a solution curve with initial condition $u^0 = (u_1^0, u_2^0)$.

Definition 3.4. The omega limit set of the point u^0, denoted by $\omega(u^0)$, consists of all points $a \in R^2$ for which there is a sequence t_j, $j = 1, 2, \ldots$, such that

$$u(t_j) \to a \qquad t_j \to \infty.$$

Definition 3.5. A *homoclinic orbit* is a trajectory of a flow of a dynamical system that joins a saddle equilibrium point to itself. A *heteroclinic orbit* (sometimes called a heteroclinic connection) is a path in phase space that joins two different equilibrium points.

A *manifold* is a mathematical space that on a small scale resembles Euclidean space of a specific dimension. For instance, a line and a circle are one-dimensional manifolds, while a plane and a sphere are two-dimensional manifolds.

Definition 3.6. A *separatrix* is a phase curve that meets a hyperbolic equilibrium point or connects the stable and unstable manifolds of a pair of equilibrium points. A separatrix marks a boundary between sectors with phase curves with different properties.

Definition 3.7. A *separatrix cycle* consists of the union of a finite number of equilibria p_j for $j = 1, \ldots, m$ and separatrices Γ_j such that the flow on Γ_j is from p_j to p_{j+1} and $p_{m+1} = p_1$.

Definition 3.8. A *compound separatrix cycle* or a *graphic* is the union of a finite number of compatibly oriented separatrix cycles.

The types of omega limit sets for an arbitrary orbit of a planar system is given by the following theorem.

Theorem 3.5 (Poincaré–Bendixson Trichotomy). *Assume that $X \subseteq R^2$, where X is an open set, contains only finitely many equilibria. Let $u(t)$ be a solution in X that is defined and bounded on $[0, \infty)$ with $\omega(u^0) \subseteq X$. Then one of the following holds:*

1. *$\omega(u^0)$ consists of an equilibrium.*
2. *$\omega(u^0)$ is a periodic orbit.*
3. *$\omega(u^0)$ a graphic.*

Assume that X is the open first quadrant. If $\mathscr{R}_0 > 1$, then the dimensionless SIR model has a unique equilibrium in X, the endemic equilibrium. Hence, the omega limit set of every initial point in X is the endemic equilibrium, a potential periodic orbit, or a graphic. To rule out possible periodic orbits and graphics inside X, one can use the Dulac–Bendixson criterion, which applies to planar systems only.

Theorem 3.6 (Dulac–Bendixson Criterion). *Let $Z \subseteq X$ be open and simply connected. Assume the following:*

1. *The functions f and g are continuously differentiable on Z.*
2. *There exists a function $D : Z \to R$, continuously differentiable on Z, such that*

$$\frac{\partial(Df)}{\partial x} + \frac{\partial(Dg)}{\partial y}$$

is either strictly positive almost everywhere on Z or strictly negative almost everywhere on Z.

Then Z contains no periodic orbits or graphics.

Definition 3.9. The function D is called the *Dulac function*.

If $D \equiv 1$, then the Dulac criterion is refereed to as the Bendixson criterion.

Theorem 3.7. *Assume $\mathscr{R}_0 > 1$. The system (3.6) has no periodic orbits or graphics in R^2_+.*

Proof. We will apply the Dulac–Bendixson criterion. Let $Z = X$ be the open first quadrant. Let $f(x,y) = \rho(1-x) - \mathscr{R}_0 xy$ and $g(x,y) = (\mathscr{R}_0 x - 1)y$. Applying the Dulac–Bendixson criterion directly with $D = 1$ gives

$$\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} = -\rho - \mathscr{R}_0 y + \mathscr{R}_0 x - 1.$$

This expression has unspecified sign, which potentially may change. The term that disrupts the definiteness of the sign is $\mathscr{R}_0 x$. Thus, we have to "eliminate" this term. This suggests that we use $D(x,y) = 1/y$. We take Z to be the open first quadrant. Then D is continuously differentiable in Z. Furthermore, we have

$$\frac{\partial Df}{\partial x} + \frac{\partial Dg}{\partial y} = -\frac{\rho}{y} - \mathscr{R}_0 < 0.$$

Thus, the system has no periodic orbits or graphics in the open first quadrant. This implies that choices two and three of the Poincaré–Bendixson theorem are ruled out as an option. □

The next theorem shows the global stability of the endemic equilibrium for system (3.6).

Theorem 3.8. *Assume $\mathscr{R}_0 > 1$. The endemic equilibrium (x^*, y^*) of system (3.6) is globally stable whenever $I(0) > 0$.*

Proof. We apply Poincaré-Bendixson theorem. First, we have to show that all solutions of system (3.6) are bounded. To see this, we add the two equations in (3.6). Set $\widehat{\rho} = \min\{\rho, 1\}$. Then

$$x' + y' \leq \rho - \widehat{\rho}(x+y).$$

Hence,

$$x+y \leq \kappa e^{-\widehat{\rho}t} + \frac{\rho}{\widehat{\rho}}(1-e^{-\widehat{\rho}t}),$$

where κ is the value of the initial condition. We obtain that

$$\limsup_t (x+y) \leq \frac{\rho}{\widehat{\rho}},$$

that is, solutions remain bounded. We conclude that the first quadrant is positively invariant with respect to the solutions of (3.6) and contains the omega limit set of every initial condition. Therefore, we can apply Poincaré–Bendixson theorem. When $\mathscr{R}_0 > 1$, if $y(0) = 0$, then the solutions will stay on the x-axis and converge to the disease-free equilibrium. If $y(0) > 0$, that is, $u^0 = (x(0), y(0)) \in X$, then we claim that the disease-free equilibrium does not belong to the omega limit set of u^0. Suppose the disease-free equilibrium belongs to $\omega(u^0)$. Then, since the disease-free equilibrium is an unstable saddle, it has a stable manifold that is given by the x-axis. That is the case, because each solution that starts from $y(0) = 0$, that is, that starts on the x-axis, stays on the x-axis and converges to the disease-free equilibrium. Hence, the stable manifold of the disease-free equilibrium is not in X. Hence, $\omega(u^0)$ would have to contain another equilibrium, namely the endemic equilibrium. But since the endemic equilibrium is locally asymptotically stable, every solution that gets close to it, stays close to it. Therefore, the disease-free equilibrium does not belong to $\omega(u^0)$. Hence, the omega limit set of u^0 consists of the endemic equilibrium only. All solutions with $I(0) > 0$ converge to the endemic equilibrium. □

3.6 Oscillations in Epidemic Models

In the previous sections, we saw that the most basic SIR epidemic model has a unique endemic equilibrium, which is globally stable if $\mathscr{R}_0 > 1$. This means that every solution converges to a stationary state. On the other hand, many times, the incidence or the prevalence data of various diseases exhibit periodicity. This is particularly true of childhood diseases. For instance, data on measles in New York City for the period 1928–1963 suggests that the disease persisted in the form of periodic outbreaks. That can be clearly seen from the monthly case data on measles for New York City, illustrated in Fig. 3.8.

Can simple epidemic models capture the oscillations exhibited in data? That would be the case if the epidemic model had a stable periodic solution. System (3.7) has a periodic solution (or a cycle) if there is an orbit $(x(t), y(t))$ such that $x(t+T) = x(t)$ and $y(t+T) = y(t)$ for some appropriate value T, called the *period*. The cycle is stable if solutions that start from close initial conditions converge to the cycle. It is well known that ODE models that reduce to a one-dimensional dynamical system do not have cycles, and cannot capture oscillations in data. However, planar ODE systems, including planar epidemic models, can exhibit periodicity.

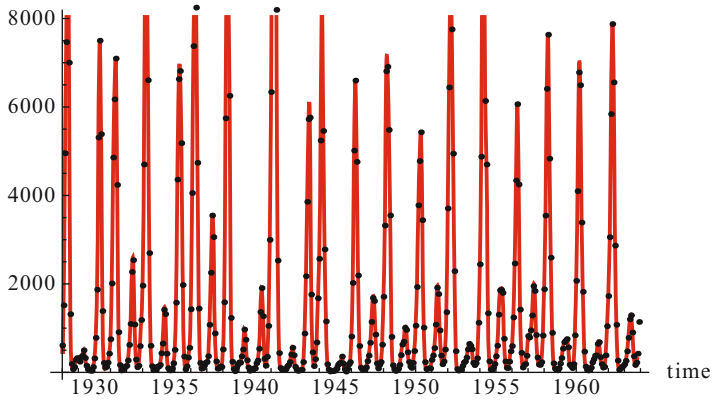

Fig. 3.8 Monthly case data for measles for New York City in the period 1928–1963. The data are given as data points. The *continuous curve* is an interpolation. The figure clearly shows recurrent outbreaks

The periodic solutions typically arise from a single endemic equilibrium that loses stability through a bifurcation called a **Hopf bifurcation**. Hopf bifurcations occur when a pair of complex conjugate eigenvalues of the linearization around a nontrivial fixed point cross the imaginary axis of the complex plane with nonzero speed. In that case, a stable limit cycle may bifurcate from the fixed point, which at the same time loses stability. Hopf bifurcations occur in planar ODE systems as well as in higher-dimensional systems.

The existence of a periodic solution can be deduced from the Hopf bifurcation theorem, which we state below for planar systems. To introduce the theorem, we need to restate the problem (3.7) to include a parameter. We write the system (3.7) in the form

$$\begin{aligned} x' &= f(x,y;\mu), \\ y' &= g(x,y;\mu), \end{aligned} \tag{3.34}$$

where we explicitly acknowledge that f and g depend on the parameter μ. Furthermore, let $(x^*(\mu), y^*(\mu))$ be an equilibrium of the system (3.34) that also depends on the parameter. We linearize the system (3.34) around the equilibrium $(x^*(\mu), y^*(\mu))$. The Jacobian of the linearization is given by

$$J(x^*(\mu), y^*(\mu)) = \begin{pmatrix} f_x(x^*, y^*; \mu) & f_y(x^*, y^*; \mu) \\ g_x(x^*, y^*; \mu) & g_y(x^*, y^*; \mu) \end{pmatrix}. \tag{3.35}$$

Assume that the Jacobian has eigenvalues $\lambda_\pm = \alpha(\mu) \pm i\beta(\mu)$, where $i = \sqrt{-1}$. In terms of the trace of the Jacobian, $\mathrm{Tr}J$, and determinant of the Jacobian, $\mathrm{Det}J$, the eigenvalues are given by

$$\lambda_\pm = \frac{\mathrm{Tr}J \pm \sqrt{(\mathrm{Tr}J)^2 - 4\mathrm{Det}J}}{2}. \tag{3.36}$$

For a Hopf bifurcation to occur, we must have, for some parameter value $\mu = \mu_0$, that the following conditions hold:

$$\begin{aligned} \mathrm{Tr}J(x^*(\mu_0), y^*(\mu_0)) &= 0, \\ \mathrm{Det}J(x^*(\mu_0), y^*(\mu_0)) &> 0. \end{aligned} \tag{3.37}$$

When these conditions are satisfied, the eigenvalues of the Jacobian are purely imaginary. If in addition to the above conditions, the *transversality condition* is satisfied,

$$\frac{d}{d\mu}\alpha(\mu)_{|\mu=\mu_0} = d \neq 0, \tag{3.38}$$

then a Hopf bifurcation occurs at the bifurcation point $(x^*(\mu_0), y^*(\mu_0); \mu_0)$. At such a Hopf bifurcation for some μ near μ_0, small-amplitude oscillations (limit cycles) bifurcate from the equilibrium solution. The amplitude of these oscillations approaches zero as μ approaches μ_0. Hopf theory guarantees the existence of such periodic orbits for $\mu \approx \mu_0$ only; it does not guarantee the existence of the oscillations for μ farther away from μ_0.

To state the Hopf bifurcation theorem, we rewrite system (3.34) in the form

$$\begin{aligned} x' &= j_{11}(\mu)x + j_{12}(\mu)y + f_1(x, y; \mu), \\ y' &= j_{21}(\mu)x + j_{22}(\mu)y + g_1(x, y; \mu), \end{aligned} \tag{3.39}$$

where $j_{11}(\mu) = f_x(x^*, y^*; \mu)$, $j_{12}(\mu) = f_y(x^*, y^*; \mu)$, $j_{21}(\mu) = g_x(x^*, y^*; \mu)$, $j_{22}(\mu) = g_y(x^*, y^*; \mu)$. The complete Hopf bifurcation theorem, which is given below, gives also a third condition that is rarely checked.

Theorem 3.9 (Hopf Bifurcation Theorem). *Let f and g in (3.34) have continuous third-order derivatives in x and y. Assume that $(0,0)$ is an equilibrium of (3.39) and that the Jacobian matrix J defined by (3.35) is valid for all values of $\mu \approx \mu_0$. In addition, assume that the eigenvalues of J are $\alpha(\mu) \pm i\beta(\mu)$. Suppose in addition that for $\mu = \mu_0$, the following conditions hold:*

1. *Nonhyperbolicity condition: $\alpha(\mu_0) = 0$ and $\beta(\mu_0) = \omega \neq 0$.*
2. *Transversality condition: the eigenvalues cross the imaginary axis with nonzero speed*

$$\frac{d}{d\mu}\alpha(\mu)_{|\mu=\mu_0} = d \neq 0. \tag{3.40}$$

3. *Genericity condition: $a \neq 0$, where*

$$\begin{aligned} a = &\frac{1}{16}\left(f_{xxx} + f_{xyy} + g_{xxy} + g_{yyy}\right) \\ &+ \frac{1}{16\omega}\left(f_{xy}(f_{xx} + f_{yy}) - g_{xy}(g_{xx} + g_{yy}) - f_{xx}g_{xx} + f_{yy}g_{yy}\right), \end{aligned} \tag{3.41}$$

where $f_{xy} = \frac{\partial^2 f}{\partial x \partial y}|_{\mu=\mu_0}(x^*, y^*)$, etc.

Then system (3.34) has a periodic solution for $\mu > \mu_0$ if $ad < 0$ and for $\mu < \mu_0$ if $ad > 0$. In the case $ad < 0$, the bifurcation is called supercritical, and the bifurcating

periodic solution is stable. In the case $ad > 0$, the bifurcation is called subcritical, and the bifurcating periodic solution is unstable. An approximate period of the periodic solution is given by

$$T = \frac{2\pi}{\omega}.$$

Before we continue with an example of Hopf bifurcation, we summarize the options for the stability of an equilibrium in a planar system based on the use of the trace and the determinant of the Jacobian. This theorem gives a quick and very efficient way to deduce the stability of an equilibrium.

Theorem 3.10. *Consider the planar system*

$$\begin{aligned} x' &= f(x,y), \\ y' &= g(x,y), \end{aligned} \tag{3.42}$$

and let (x^*,y^*) be an equilibrium of that system. Then the Jacobian of system (3.42) evaluated at that equilibrium is given by

$$J(x^*,y^*) = \begin{pmatrix} f_x(x^*,y^*) & f_y(x^*,y^*) \\ g_x(x^*,y^*) & g_y(x^*,y^*). \end{pmatrix}. \tag{3.43}$$

The following results give the stability of the equilibrium (x^*,y^*):

1. *Equilibrium (x^*,y^*) is locally asymptotically stable if and only if $\mathrm{Tr}J < 0$ and $\mathrm{Det}J > 0$.*
2. *Equilibrium (x^*,y^*) is a saddle if and only if $\mathrm{Det}J < 0$.*
3. *Equilibrium (x^*,y^*) loses stability and undergoes Hopf bifurcation if and only if for some value of the parameter μ, called μ_0, the following hold:*

$$\begin{aligned} \mathrm{Tr}J(x^*(\mu_0),y^*(\mu_0)) &= 0, \\ \mathrm{Det}J(x^*(\mu_0),y^*(\mu_0)) &> 0. \end{aligned} \tag{3.44}$$

In addition, we must also have

$$\frac{d\mathrm{Tr}J}{d\mu}|_{\mu=\mu_0} \neq 0.$$

To illustrate the application of the Hopf bifurcation theorem, we consider a simple modification of the SIR model (3.3). Assume that the transmission coefficient of infection β is not constant but linear in the number of infecteds: $\beta(1+\nu I)$, where $\nu > 0$ is a parameter. This means that either the contact rate increases with the number of infectious individuals or the probability of transmission does so. Thus, new infections occur at a much faster pace compared to the standard mass action incidence. The model becomes [6]

$$\begin{aligned} S'(t) &= \Lambda - \beta(1+\nu I)IS - \mu S, \\ I'(t) &= \beta(1+\nu I)IS - (\alpha+\mu)I, \end{aligned} \tag{3.45}$$

where we have omitted the equation for recovered individuals R. We will investigate this model without nondimensionalizing it. The total population size $N = S+I+R$ satisfies $N' = \Lambda - \mu N$. We assume that the initial total population size is given by $N_0 = S_0+I_0+R_0$. The disease-free equilibrium $\mathscr{E}_0 = \left(\frac{\Lambda}{\mu}, 0\right)$ of model (3.45) always exists. The reproduction number of the model (3.45) is given by

$$\mathscr{R}_0 = \frac{\Lambda\beta}{\mu(\mu+\alpha)}. \tag{3.46}$$

It can be shown (see Problem 3.3) that the disease-free equilibrium is locally stable if $\mathscr{R}_0 < 1$ and unstable if $\mathscr{R}_0 > 1$.

The endemic equilibria of the system are solutions to the following system:

$$\begin{aligned} \Lambda - \beta(1+\nu I)IS - \mu S &= 0, \\ \beta(1+\nu I)S - (\alpha+\mu) &= 0. \end{aligned} \tag{3.47}$$

From the second equation, we see that $\beta(1+\nu I)S = (\alpha+\mu)$. Substituting this expression into the first equation, we can express S in terms of I:

$$S = \frac{\Lambda}{\mu} - \frac{\mu+\alpha}{\mu} I. \tag{3.48}$$

Hence, substituting in $\beta(1+\nu I)S = (\alpha+\mu)$, we obtain the following quadratic equation for I:

$$(1+\nu I)\left[\frac{\Lambda}{\mu} - \frac{\mu+\alpha}{\mu} I\right] = \frac{\mu+\alpha}{\beta}. \tag{3.49}$$

If we denote by $f(I)$ the parabola on the left-hand side of (3.49), then the endemic equilibria of model (3.45) are given by the intersections of the parabola with the horizontal line $y = (\mu+\alpha)/\beta$ (see Fig. 3.9). If $f(0) > (\mu+\alpha)/\beta$, or equivalently, if $\mathscr{R}_0 > 1$, then there is always a unique (positive) equilibrium $\mathscr{E}^* = (S^*, I^*)$. That is the scenario that is shown in Fig. 3.9. If $f(0) < (\mu+\alpha)/\beta$, or equivalently, if $\mathscr{R}_0 < 1$, then there may be two equilibria if the maximum of the parabola occurs to the right of the y-axis and the horizontal line lies below the maximum of the parabola. Since the maximum of the parabola is achieved at

$$I_m = \frac{1}{2\nu}\left[\frac{\Lambda\nu}{\mu+\alpha} - 1\right],$$

the maximum of the parabola is to the right of the y-axis if and only if $I_m > 0$. Hence, there will be two endemic equilibria if the maximum of the parabola is above the horizontal line, that is, if

$$(1+\nu I_m)\left[\frac{\Lambda}{\mu} - \frac{\mu+\alpha}{\mu} I_m\right] > \frac{\mu+\alpha}{\beta}. \tag{3.50}$$

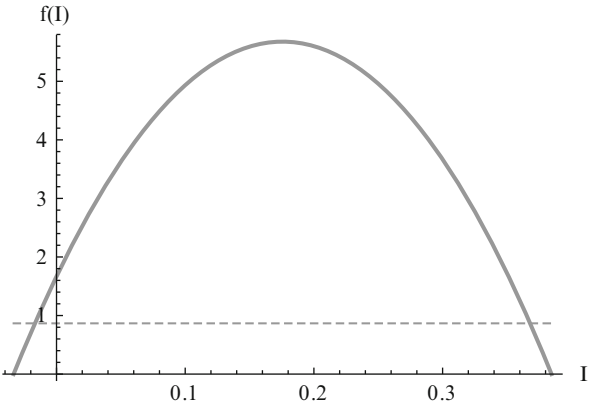

Fig. 3.9 The graph shows intersections of the function $f(I)$ with the *horizontal line* $y = (\mu+\alpha)/\beta$ in Eq. (3.49). Each intersection gives one nontrivial equilibrium of the model (3.45). The figure shows the case $\mathscr{R}_0 > 1$, and there is a unique (positive) equilibrium I^*

If two endemic equilibria exist, we denote them by $\mathscr{E}_1 = (S_1^*, I_1^*)$ and $\mathscr{E}_2 = (S_2^*, I_2^*)$, where $I_1^* < I_2^*$ and the corresponding value of S is computed from (3.48). The stability of the equilibria is given by the Jacobian

$$J = \begin{pmatrix} -\beta(1+\nu I)I - \mu & -\beta\nu IS - (\mu+\alpha) \\ \beta(1+\nu I)I & \beta\nu IS \end{pmatrix}, \tag{3.51}$$

where we have used the equality $\beta(1+\nu I)S = (\alpha+\mu)$ to simplify the Jacobian. The characteristic equation of the Jacobian $|J - \lambda I| = 0$ is a quadratic polynomial in λ given by

$$\lambda^2 + B\lambda + C = 0, \tag{3.52}$$

where B and C are given by

$$\begin{aligned} B &= \mu + \beta(1+\nu I)I - \beta\nu IS \\ C &= \beta(1+\nu I)I(\mu+\alpha) - \mu\beta\nu IS, \end{aligned} \tag{3.53}$$

and I is any equilibrium. We note that the endemic equilibria differ in the slope of the tangent line to the curve of $f(I)$ at each equilibrium. In particular, if $\mathscr{R}_0 > 1$, the slope of the tangent at the equilibrium satisfies $f'(I^*) < 0$. When there are two equilibria, we have $f'(I_1^*) > 0$ while $f'(I_2^*) < 0$. The slope of $f(I)$ is given by

$$f'(I) = \frac{\mu+\alpha}{\mu}\left[\frac{\Lambda\nu}{\mu+\alpha} - 1 - 2\nu I\right].$$

On the other hand, C can be rewritten in the form (where S has been replaced with (3.48))

$$C = (\mu + \alpha)\beta I \left[1 + 2\nu I - \frac{\Lambda \nu}{\mu + \alpha}\right].$$

It is evident from the above two expressions that the sign of C is opposite the sign of $f'(I)$. Hence, if $\mathscr{R}_0 > 1$, the unique endemic equilibrium I^* gives $C > 0$. When $\mathscr{R}_0 < 1$ and there are two equilibria, we have for the lower one $C < 0$ and for the upper one $C > 0$. Hence, when two equilibria are present, the lower one $\mathscr{E}_1$ is unstable, since $C < 0$ means that the characteristic equation (3.52) has one positive and one negative root. The local stability of $\mathscr{E}^*$ and $\mathscr{E}_2$ depends on the sign of B. If $B > 0$, then each of these equilibria is stable. However, for some value of the parameter $\nu = \nu_0$, we may have

$$\mu + \beta(1 + \nu_0 I)I - \beta \nu_0 IS = 0,$$

and then a Hopf bifurcation may occur. To see that this condition may hold, we exhibit a specific numerical example. To decide on parameter values, we first decide on a time unit. We will measure time in years. Since $1/\mu$ gives an average lifespan of individuals, if we take $\mu = 0.2$, that will give a lifespan of 5 years. For the human population, that lifespan will be adequate for some childhood diseases. The lifespan can describe well many animal populations. Furthermore, $1/\alpha$ corresponds to a duration of infectiousness. Hence, if we take $\alpha = 26$, that will correspond to duration of infectiousness of about 2 weeks. The remaining parameters are taken as $\beta = 0.005$ and $\Lambda = 1250$. We think of $B(\nu) = \mu + \beta(1 + \nu I)I - \beta \nu IS$ as a function of the parameter ν. Since these parameters give $\mathscr{R}_0 = 1.19275$, we focus on the stability of the endemic equilibrium $\mathscr{E}^*$. We note that I^* and S^* are also functions of ν. We plot $B(\nu)$ against ν in Fig. 3.10.

We check the transversality condition for $\mathscr{E}^*$ also numerically. We notice that the real part of the eigenvalues of (3.52) is given by

$$\Re(\lambda) = \frac{-(\mu + \beta(1 + \nu I)I - \beta \nu IS)}{2}.$$

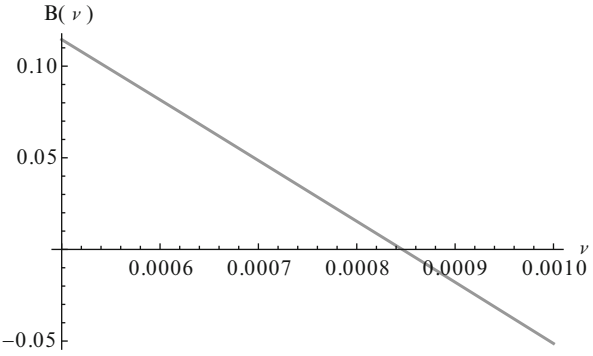

Fig. 3.10 The graph of $B(\nu)$ shows clearly that B changes sign as the parameter ν passes through the critical value $\nu_0 = 0.000846293$

To differentiate $\Re(\lambda)$ with respect to v, we have to differentiate $B(v)$ with respect to v and evaluate the results at $v = v_0$. Figure 3.10 suggests that

$$\frac{dB(v)}{dv}|_{v=v_0} < 0.$$

Differentiating the real part of the roots of the characteristic equation (3.52), we have

$$\frac{\partial \Re(\lambda)}{\partial v} = -\frac{1}{2}\frac{dB(v)}{dv}|_{v=v_0} > 0.$$

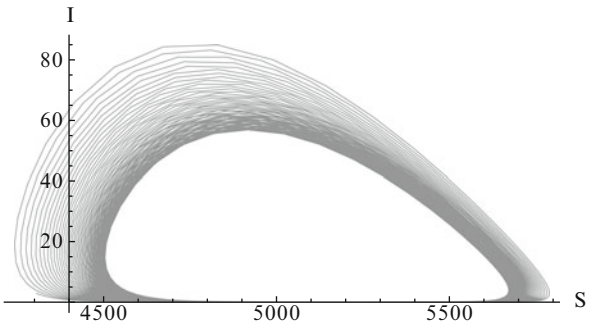

Fig. 3.11 The graph shows oscillations in the (S,I)-plane that converge to a periodic orbit. Initial conditions are $S(0) = 40{,}000$, $I(0) = 15$. The plot is made for $t \geq 200$

Hence, the transversality condition is satisfied. Assuming $a \neq 0$, we may conclude from the Hopf bifurcation theorem that a periodic solution bifurcates from the stable endemic equilibrium $\mathscr{E}^*$. We cannot conclude that the bifurcation is supercritical or subcritical without computing a. Therefore, we do not know whether the bifurcating solution is stable. We checked the stability of the bifurcating oscillatory solution numerically. We chose $v = 0.00117$. The equilibrium is given by $\mathscr{E}^* = (5191, 8)$. The real part of the roots of the characteristic equation is $\Re(\lambda) = 0.054228$. From the simulations we performed, it appears that the bifurcating oscillatory solution in this case is stable (see Fig. 3.11), and the number of susceptible and infected individuals tend to a periodic orbit.

Problems

3.1. Census data for the population of the United States (in millions) are given in Table 3.3. Fit each of the three population models, Malthus model, logistic model, and constrained logistic model, to the data and determine the least-squares error. Which model fits the data best?

Table 3.3 Population of the US (in millions)

Year	Population (million)	Year	Population (million)
1790	3.9	1910	92.0
1800	5.3	1920	105.7
1810	7.2	1930	122.8
1820	9.6	1940	131.7
1830	12.9	1950	150.7
1840	17.1	1960	179.0
1850	23.1	1970	205.0
1860	31.4	1980	226.5
1870	38. 6	1990	248.7
1880	50.2	2000	281.4
1890	62.9	2010	310.0
1900	76.0	–	–

3.2. Consider the following SIS epidemic model with disease-induced mortality γ:

$$\begin{aligned} S' &= \Lambda - \beta IS + \alpha I - \mu S, \\ I' &= \beta IS - (\alpha + \gamma + \mu) I, \end{aligned} \tag{3.54}$$

where S is the number of susceptibles, I is the number of infected, β is the transmission rate, α is the recovery rate, Λ is the birth rate, μ is the natural death rate.

(a) Sketch the nullclines of the system and the direction of the vector field along them. Confirm your results by plotting the vector field with solutions.
(b) Determine the reproduction number and equilibria of the system.
(c) Calculate the Jacobian of each equilibrium and determine the stability.
(d) Use the Dulac criterion to rule out periodic solutions.
(e) Use the Poincaré–Bendixson theorem to show convergence to equilibrium.

3.3. Consider model (3.45):

$$\begin{aligned} S'(t) &= \Lambda - \beta(1 + \nu I) IS - \mu S, \\ I'(t) &= \beta(1 + \nu I) IS - (\alpha + \mu) I, \end{aligned} \tag{3.55}$$

where S is the number of susceptibles, I is the number of infected, β is the transmission rate, α is the recovery rate, Λ is the birth rate, μ is the natural death rate, and ν is a proportionality constant.

(a) Derive the reproduction number $\mathscr{R}_0$ for that model.
(b) Show that the disease-free equilibrium $\left(\frac{\Lambda}{\mu}, 0\right)$ is locally stable if $\mathscr{R}_0 < 1$ and unstable if $\mathscr{R}_0 > 1$.
(c) Argue that if $\mathscr{R}_0 > 1$, there is always a unique endemic equilibrium.

3.4. Non-dimensionalization
Consider the SIS model with saturating incidence in the size of the susceptibles.

$$\begin{aligned} S'(t) &= \Lambda - \frac{\beta IS}{1+\sigma S} + \alpha I - \mu S, \\ I'(t) &= \frac{\beta IS}{1+\sigma S} - (\alpha+\mu)I. \end{aligned} \tag{3.56}$$

(a) What are the units of the parameters?
(b) Rescale the system above into a nondimensional system both in the time variable and in the dependent variables.
(c) Determine conditions for the existence of an endemic equilibrium.

3.5. Multiple Equilibria
Consider the following SIS epidemic model with disease-induced mortality γ:

$$\begin{aligned} S' &= \Lambda - \frac{\beta IS}{A+I^2} + \alpha I - \mu S, \\ I' &= \frac{\beta IS}{A+I^2} - (\alpha+\gamma+\mu)I, \end{aligned} \tag{3.57}$$

where S is the number of susceptibles, I is the number of infected, β is the transmission rate, α is the recovery rate, Λ is the birth rate, μ is the natural death rate.

(a) Determine the reproduction number and equilibria of the system.
(b) Sketch the nullclines of the system and the direction of the vector field along them. Confirm your results by plotting the vector field with solutions.
(c) Calculate the Jacobian of each equilibrium and determine the stability.
(d) Use the Dulac criterion to rule out periodic solutions.

3.6. Fox Rabies
The following model has been proposed to model fox rabies [26]:

$$\begin{aligned} S' &= rSe^{-aS} - \beta IS - \mu S, \\ I' &= \beta IS - (\alpha+\mu)I, \end{aligned} \tag{3.58}$$

where S is the number of susceptibles, I is the number of infected, β is the transmission rate, α is the disease-induced death rate, r and a are constants associated with declining with population size per capita birth rate re^{-aS}, and μ is the natural death rate.

(a) Determine the reproduction number and equilibria of the system.
(b) Sketch the nullclines of the system and the direction of the vector field along them. Confirm your results by plotting the vector field with solutions.
(c) Calculate the Jacobian of each equilibrium and determine the stability.
(d) Use the Dulac criterion to rule out periodic solutions.

3.7. Hopf Bifurcation

The following model has been proposed to model the saturating contact rate:

$$\begin{aligned} S' &= rS\left(1-\frac{S}{K}\right)-\frac{\beta IS}{1+\alpha S}-\mu S, \\ I' &= \frac{\beta IS}{1+\alpha S}-(\gamma+\mu)I, \end{aligned} \tag{3.59}$$

where S is the number of susceptibles, I is the number of infected, β is the transmission rate, α is a parameter that measures the inhibitory effect, γ is the recovery rate, r and K are constants associated with the logistic population growth, and μ is the natural death rate.

(a) Determine the reproduction number and equilibria of the system.
(b) Calculate the Jacobian of each equilibrium and determine the stability.
(c) Use the Hopf bifurcation theorem to show the presence of periodic solutions. Use a computer algebra system to graph the periodic solution in the phase plane together with the vector field.

Chapter 4
Vector-Borne Diseases

4.1 Vector-Borne Diseases: An Introduction

A vector-borne disease is one in which the pathogenic microorganism is transmitted from an infected individual to another individual by an arthropod or other agent, sometimes with vertebrate animals serving as intermediary hosts. The transmission of vector-borne diseases to humans depends on three factors: (1) the pathogenic agent; (2) the arthropod vector; (3) and the human host. Mathematical models of vector-borne diseases typically take into account the dynamics of the vectors as well as the dynamics of the humans, and occasionally the dynamics of the intermediate animal host.

4.1.1 The Vectors

In epidemiology, a *vector* is a living carrier that transmits an infectious agent from one host to another. Vectors do not get ill from the disease, and once infected, they remain infected throughout the remainder of their lives. By common usage, vectors are considered to be invertebrate animals, usually arthropods. Technically, however, vertebrates can also act as vectors, including foxes, raccoons, and skunks, which can all transmit the rabies virus to humans via a bite. Arthropods account for over 85% of all known animal species, and they are the most important disease vectors. Examples of common vectors are mosquitoes, flies, sand flies, lice, fleas, ticks, mites, and cyclops, which transmit a huge number of diseases. Many such vectors feed on blood at some or all stages of their lives. When the arthropod feeds on blood, the parasite enters the blood stream of the host. Table 4.1 gives the most common vectors and the diseases they transmit.

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_4

Table 4.1 Examples of vectors and diseases they transmit

Vector	Diseases	Comment
Flies	Bubonic plague	Flies
Mosquitoes	Malaria	Anopheles genus
Mosquitoes	Avian malaria, dengue, chikungunya	*Aedes* genus
Tsetse flies	African sleeping sickness	Include 34 species of genus *Glossina*
Kissing bugs	Chagas diseases	*Triatominae* subfamily of *Reduviidae*
Ticks	Lyme disease, babesiosis	Genus *Ixodes*
Sand flies	Leishmaniasis, bartonellosis, sandfly fever	Subfamily Phlebotominae
Ticks, lice	Bacterial Rickettsia	

4.1.2 The Pathogen

The pathogen of the vector-borne diseases can fall in one of four categories:

- **Protozoa:** A number of vector-borne diseases are caused by protozoa. The most notable example is malaria, which is caused by the *Plasmodium* parasite. The *Plasmodium* parasite has a complex life cycle that occurs in both the mosquito vector and the human host. Other vector-borne diseases caused by protozoa are trichomoniasis, leishmaniasis, and sleeping sickness.
- **Bacteria:** Bacterial vector-borne diseases include Lyme disease, plague, tick-borne relapsing fever, and tularemia.
- **Virus:** Most of the vector-borne diseases are caused by viruses. Most notable examples include dengue, chikungunya, eastern equine encephalitis, Japanese encephalitis, West Nile encephalitis, caused by the West Nile Virus, Yellow fever, and others. The viruses transmitted by arthropod vectors are known collectively as *arboviruses*.
- **Helminth:** An example of a vector-borne disease cased by helminths (worms) is lymphatic filariasis, which is transmitted by mosquitoes.

4.1.3 Epidemiology of Vector-Borne Diseases

We summarize in Table 4.2 the primary examples of vector-borne diseases, the pathogens that cause them, the vectors that transmit them, and the human population at risk.

The majority of vector-borne diseases persist in nature by utilizing a vertebrate host. For a small number of vector-borne diseases, such as malaria and dengue, humans are the major host, with no significant animal reservoirs. The vector receives the pathogen from an infected host and transmits it either to an intermediary animal host or directly to the human host. The different stages of the pathogen's life cycle

Table 4.2 Primary examples of vector-borne diseases[a]

Disease	Pathogen	Vector	Population at risk (millions)	Prevalence
Malaria	*Plasmodium* spp.	Anopheles mosquito	2100	270 million
Schistosomiasis	*Schistosome* flatworms	Water snails	600	260 million
Dengue	Dengue virus	*Aedes* mosquitoes	2000	50–100 million/year[b]
Leishmaniasis	*Leishmania* spp.	Sand flies	12	350 million
Lymphatic filariasis	Nematode worms	Mosquito	900	65.5 million
River blindness	Onchocerca volvulus	Black fly	90	17.8 million
African trypanosomiasis	*Trypanosoma brucei* spp.	Tsetse fly	50	0.025 million/year
Chagas diseases	*Trypanosoma cruzi* spp.	Kissing bug	100	6.5 million

[a]http://www.ciesin.org/TG/HH/v-bd.html
[b]Data for dengue taken from [24]

occur during this process and are intimately dependent on the availability of suitable vectors and hosts.

Key components that determine the occurrence of vector-borne diseases include:

- the abundance of vectors and intermediate and reservoir hosts;
- the prevalence of disease-causing pathogens suitably adapted to the vectors and the human or animal host;
- the local environmental conditions, especially temperature and humidity;
- the resilience behavior and immune status of the human population.

Vector-borne diseases are prevalent in the tropics and subtropics and are relatively rare in temperate zones. Lyme disease and Rocky Mountain spotted fever persist in temperate regions, including the United States. There are different patterns of vector-borne disease occurrence. Parasitic and bacterial diseases, such as malaria and Lyme disease, tend to produce a high disease incidence but do not cause major outbreaks. An exception to this rule is plague, a bacterial disease that does cause outbreaks. In contrast, many vector viral diseases, such as yellow fever, dengue, and Japanese encephalitis, commonly cause major epidemics.

There has been a worldwide resurgence of vector-borne diseases since the 1970s, including malaria, dengue, yellow fever, plague, leishmaniasis, sleeping sickness, West Nile encephalitis, Lyme disease, Japanese encephalitis, Rift Valley fever, and Crimean-Congo hemorrhagic fever. Reasons for the emergence or resurgence of vector-borne diseases include:

- the development of insecticide and drug resistance;
- decreased resources for surveillance, prevention, and control of vector-borne diseases;
- population growth;
- urbanization;

- changes in agricultural practices;
- deforestation;
- increased travel.

Changes in the distribution and population size of important arthropod disease vectors have also been observed. For instance,

- The yellow fever mosquito, *Aedes aegypti*, has reestablished itself in parts of the Americas where it had been presumed to have been eradicated.
- The Asian tiger mosquito, *Aedes albopictus*, was introduced into the Americas in the 1980s and has spread to Central and South America;
- The black-legged tick, *Ixodes scapularis*, an important transmitter of Lyme disease and other pathogens, has gradually expanded its range in parts of eastern and central North America.

4.2 Simple Models of Vector-Borne Diseases

Historically, the first mathematical models of infectious diseases were derived at the beginning of the twentieth century. These models were developed by Ross and MacDonald to capture the dynamics of malaria and inform public health officials at that time how to combat malaria. Mathematical models of malaria have since become a guiding tool in the process of development and use of models in understanding infectious disease epidemiology and guiding control measures.

4.2.1 Deriving a Model of Vector-Borne Disease

Transmission in malaria, as in other vector-borne diseases, involves at least two species—the vector and the human host. The simplest model for the vector is an SI model, since most vectors once infected do not recover. Let us denote the susceptible vectors by S_v and the infected vectors by I_v. A susceptible vector becomes infected by biting an infected human I_H at a biting rate a and probability of transmission of the disease given by p. The dynamical system that describes the vector is

$$\begin{aligned} S_v' &= \Lambda_v - paS_vI_H - \mu S_v, \\ I_v' &= paS_vI_H - \mu I_v, \end{aligned} \tag{4.1}$$

where Λ_v is the birth rate of the vectors, and μ is the death rate. Since the vectors such as the mosquito usually have a very short life cycle, demography should be included. The dynamics of the total vector population size $N_v = S_v + I_v$ is then given by the simplified logistic equation $N_v' = \Lambda_v - \mu N_v$, whose solution can be obtained in explicit form. Since $N_v(t)$ is essentially a given function of t, we may express the number of susceptible vectors in terms of infected vectors $S_v = N_v - I_v$ and

replace it in the second equation of system (4.1), thus essentially reducing the two-dimensional vector system to one equation:

$$I_v' = pa(N_v(t) - I_v)I_H - \mu I_v. \tag{4.2}$$

We turn now to the system for the humans. Humans generally recover from the disease, but for most vector-borne diseases, recovery is not permanent, and the recovered individual can become infected again. As a starting point, we model the transmission of a vector-borne disease in humans with an SIS model. Some of the vector-borne diseases, such as chikungunya, occur as outbreaks, and in this case, omitting births and deaths for humans is acceptable. Other vector-borne diseases, such as malaria, are endemic, and inclusion of demography in the human portion of the model is necessary. We begin with the simplest human model—an SIS model without demography. Susceptible humans in class S_H become infected when bitten by an infectious vector. If we assume that infected vectors bite at the same rate as susceptible vectors, namely a, and q is the probability of transmission, then the model takes the form

$$\begin{aligned} S_H' &= -qaS_HI_v + \alpha I_H, \\ I_H' &= qaS_HI_v - \alpha I_H, \end{aligned} \tag{4.3}$$

where α is the recovery rate. The total human population size N_H is constant. We can reduce the human system by replacing the susceptible humans S_H with $S_H = N_H - I_H$ in the second equation. The system (4.3) reduces to the following equation:

$$I_H' = qa(N_H - I_H)I_v - \alpha I_H. \tag{4.4}$$

The system for the infected vectors and infected humans becomes

$$\begin{aligned} I_v' &= pa(N_v(t) - I_v)I_H - \mu I_v, \\ I_H' &= qa(N_H - I_H)I_v - \alpha I_H. \end{aligned} \tag{4.5}$$

The right-hand side of this system depends on the unknown dependent variables I_v and I_H, and the known function of time $N_v(t)$. This makes the right-hand side explicitly dependent on time, and the model **nonautonomous**.

Definition 4.1. A differential equation model $y' = f(t,y)$ is called *nonautonomous* if the right-hand-side function $f(t,y)$ depends explicitly on t.

Nonautonomous models are harder to analyze and are often not subjected to the rules of autonomous models that we investigate in this book. However, system (4.5) depends on time only through the function $N_v(t)$, which has a limit as time goes to infinity, namely,

$$N_v(t) \to \frac{\Lambda_v}{\mu} = N_v.$$

Suppose we replace system (4.5) with the following limiting system:

$$\begin{aligned} I_v' &= pa(N_v - I_v)I_H - \mu I_v, \\ I_H' &= qa(N_H - I_H)I_v - \alpha I_H. \end{aligned} \tag{4.6}$$

The limiting system (4.6) is an autonomous system, which is easier to analyze. Nonautonomous models whose limiting system is an autonomous model are called **asymptotically autonomous**.

The main question that remains is whether the global dynamics of system (4.6) are similar to the dynamics of the original nonautonomous system (4.5). If the answer is yes, then we may investigate system (4.6), for which a number of tools exist, and draw conclusions about system (4.5). The following theorem gives a positive answer to this question for planar systems.

Theorem 4.1 (Thieme [150]). *Assume that $x' = f(t,x)$ is a nonautonomous system such that $x \in R^2$ and $y' = g(y)$ is the limiting autonomous system. Let ω be the ω-limit set of a forward bounded solution x of the nonautonomous system. Assume that there exists a neighborhood of ω that contains at most finitely many equilibria of the autonomous system. Then the following trichotomy holds:*

1. *ω consists of an equilibrium of the autonomous system.*
2. *ω is a union of periodic orbits of the autonomous system and possible centers that are surrounded by periodic orbits of the autonomous system lying in ω.*
3. *ω contains equilibria of the autonomous system that are cyclically chained to each other in ω by orbits of the autonomous system.*

First, we show that all solutions of the nonautonomous system (4.5) are bounded. Indeed, suppose for some t^* that $I_v(t^*) > N_v(t^*)$. Then $I_v'(t^*) < 0$, and $I_v(t)$ is decreasing for all t for which it exceeds $N_v(t)$. Hence, $I_v(t) \leq \max\{I_v(0), N_v(t)\}$. Similar reasoning shows that $I_H(t)$ is bounded.

Since all solutions of the nonautonomous system (4.5) are bounded, the above theorem implies that the ω-limit set of the nonautonomous system can be obtained from investigation of the ω-limit set of the autonomous system (4.6). We focus on the investigation of the global behavior of the solutions of system (4.6).

System (4.6) is often recast in terms of proportions. If we set $x = I_H/N_H$ and $z = I_v/N_v$, then the system for the proportions becomes

$$\begin{aligned} z' &= pa(1-z)x - \mu z, \\ x' &= qam(1-x)z - \alpha x, \end{aligned} \tag{4.7}$$

where $m = N_v/N_H$ is the proportion of vectors to humans, and a has absorbed a multiple of the total human population.

4.2.2 Reproduction Numbers, Equilibria, and Their Stability

We will derive reproduction numbers and the equilibria for the original dimensional system (4.6). To compute equilibria, we set the system equal to zero:

$$\begin{aligned} pa(N_v - I_v)I_H - \mu I_v &= 0, \\ qa(N_H - I_H)I_v - \alpha I_H &= 0. \end{aligned} \tag{4.8}$$

The system clearly has the disease-free equilibrium $\mathscr{E} = (0,0)$. To compute the reproduction number and investigate the stability of the disease-free equilibrium, we consider the Jacobian of the system (4.6):

$$J = \begin{pmatrix} -\mu & paN_v \\ qaN_H & -\alpha \end{pmatrix}. \tag{4.9}$$

The disease-free equilibrium is stable if $\mathrm{Tr}J = -\mu - \alpha < 0$, which is clearly satisfied, and also if $\mathrm{Det}J = \mu\alpha - pqa^2N_vN_H > 0$. This last condition, rewritten as an expression whose value is less than 1, that is, $\mathscr{R}_0 < 1$, gives the reproduction number of a vector-borne disease, where

$$\mathscr{R}_0 = \frac{pqa^2N_vN_H}{\mu\alpha}.$$

We see that the local stability/instability of the disease-free equilibrium is connected to the basic disease reproduction number $\mathscr{R}_0$, which identifies the threshold for the local stability of the disease-free equilibrium. The disease-free equilibrium is locally asymptotically stable (the disease dies out) if $\mathscr{R}_0 < 1$, and unstable if $\mathscr{R}_0 > 1$. This is a critical parameter for control of the disease, since if $\mathscr{R}_0 > 1$, introduction of a small amount of disease into the population may cause it to evolve into an endemic prevalence.

The disease reproduction number $\mathscr{R}_0$ is one of the main epidemiological indicators of whether a disease can successfully invade and persist in a population. In vector-borne diseases, just as in directly transmitted diseases, the disease reproduction number gives the number of secondary infections of humans that one infective human individual may produce in a population of susceptible human individuals. The classical approach of Kermack and McKendrick derives $\mathscr{R}_0$ from the condition for local stability of the disease-free equilibrium. Other approaches are possible, and will be discussed later.

Transmission of vector-borne diseases involves two transmission cycles, namely human to vector and vector to human, and each of these transmission processes may be characterized by its own disease reproduction number. These two numbers may be combined to form a single dimensionless number that indicates whether, and to some extent how seriously, the human–vector system is open to invasion by the parasite. The Kermack–McKendrick–MacDonald approach places one infected human in a population of susceptible vectors; there will result $\mathscr{R}_H$ secondary infected vectors. Similarly, placing one infected vector in a population of susceptible humans will produce $\mathscr{R}_M$ infected humans, where

$$\mathscr{R}_H = \frac{paN_v}{\alpha}, \qquad \mathscr{R}_M = \frac{qaN}{\mu}.$$

To understand these definitions, consider the incidence term in the equation for the vectors $pa(N_v - I_v)I_H$, which gives the number of secondary infections of vectors I_H infected humans will produce per unit of time. Then, one infected human will produce paN_v infected vectors in an entirely susceptible vector population per unit of time. One infected human is infectious for $1/\alpha$ time units; hence we obtain $\mathscr{R}_H$. Similar reasoning leads to $\mathscr{R}_M$. To account for the secondary **human** infections that one infected human will produce, we notice that one infected human will produce $\mathscr{R}_H$ infected vectors, each of which will produce $\mathscr{R}_M$ infected humans, giving

$$\mathscr{R}_0 = \mathscr{R}_H \mathscr{R}_M$$

secondary human infections. This expression gives the classical reproduction number of vector-borne diseases.

To compute the endemic equilibrium, we consider again system (4.8). This system is nonlinear in I_v and I_H and not easy to solve directly. To simplify the solution, we adopt the substitution $X = 1/I_H$, $Z = 1/I_v$, which transforms the system into a linear inhomogeneous system of equations for X and Z:

$$\begin{aligned} pa(N_v Z - 1) - \mu X &= 0, \\ qa(N_H X - 1) - \alpha Z &= 0. \end{aligned} \tag{4.10}$$

Solving this linear system, we obtain

$$Z = \frac{qa}{\alpha}(N_H X - 1).$$

Substituting Z in the first equation and solving for X, we obtain X in terms of the parameter values. Further substituting X in the equation for Z, we obtain the following equilibrium:

$$I_H = N_H \frac{\mathscr{R}_0 - 1}{\frac{paN_H}{\mu} + \mathscr{R}_0}, \qquad I_v = N_v \frac{\mathscr{R}_0 - 1}{\frac{qaN_v}{\alpha} + \mathscr{R}_0}. \tag{4.11}$$

From these expressions, it is clear that the endemic equilibrium exists and is positive if and only if $\mathscr{R}_0 > 1$. From the expressions above, we can easily compute the values of the equilibrium for the proportions x and z.

Next, we investigate the local stability of the endemic equilibrium. This can be obtained from the Jacobian matrix at the endemic equilibrium, which is obtained from the linearization of system (4.6) around the endemic equilibrium. The Jacobian is given by the matrix

$$J = \begin{pmatrix} -paI_H - \mu & pa(N_v - I_v) \\ qa(N_H - I_H) & -qaI_v - \alpha \end{pmatrix}. \tag{4.12}$$

The trace of the matrix is negative: $\mathrm{Tr}J = -paI_H - \mu - qaI_v - \alpha < 0$. To determine the sign of the determinant, notice that

$$\begin{aligned} pa(N_v - I_v) &= \mu I_v / I_H, \\ qa(N_H - I_H) &= \alpha I_H / I_v. \end{aligned} \tag{4.13}$$

From these equations, it follows that $paqa(N_v - I_v)(N_H - I_H) = \alpha\mu$. This identity simplifies the determinant of the Jacobian and helps determine its sign: $\mathrm{Det}J = (paI_H+\mu)(qaI_v+\alpha) - paqa(N_v - I_v)(N_H - I_H) = (paI_H+\mu)(qaI_v+\alpha) - \alpha\mu > 0$. Hence, the endemic equilibrium is locally stable. To determine the type of the equilibrium point that gives the endemic equilibrium, we look at the characteristic equation $|J - \lambda I| = 0$, which takes the form

$$\lambda^2 - (paI_H + \mu + qaI_v + \alpha)\lambda + (paI_H + \mu)(qaI_v + \alpha) - \alpha\mu = 0.$$

The discriminant of this quadratic polynomial can be written as $\Delta = (paI_H + \mu - qaI_v - \alpha)^2 + 4\alpha\mu > 0$. We conclude that the endemic equilibrium is a stable node for all parameter values. One can establish the Dulac criterion for the system (4.6) and show global stability of the endemic equilibrium (see Problem 4.3).

4.3 Delay-Differential Equation Models of Vector-Borne Diseases

Delay-differential equations differ from ordinary differential equations in that the derivative at every point in time depends on the solution at prior times. The simplest constant delay equations have the form

$$x'(t) = F(t, x(t), x(t-\tau_1), x(t-\tau_2), \ldots, x(t-\tau_k)),$$

where the *time delays* τ_j are positive constants. Additional information is required to specify a system of delay-differential equations. Because the derivative in the equation above depends on the solution at the previous time $t - \tau_j$, it is necessary to provide an initial history function, or a vector of functions, to specify the value of the solution before time $t = 0$.

Interest in such systems often arises when traditional pointwise modeling assumptions are replaced by more realistic dependence of the rate of change on the prior population numbers, for example when the birth rate of a population is affected by prior levels of the population rather than by only the current levels. In vector-borne diseases, delays naturally occur because of the incubation period of the pathogen in the vector and the humans. Inclusion of the incubation period in the vector, also called the **extrinsic incubation period**, is particularly important, because the length of the incubation period in the vector is often of duration comparable to the mean lifespan of the vector. The fact that vectors may or may not survive the extrinsic incubation period affects significantly the dynamics of the infectious disease. This makes imperative the inclusion of the extrinsic incubation period and the

probability that the vector survives that period as a delay in the Ross–MacDonald model introduced in the previous section. We can further include the incubation period of the pathogen in the human as a second constant delay.

Let the incubation period of the pathogen in a vector have duration τ_1. Thus, of those vectors infected τ_1 units of time ago, only a proportion $pa[N_v - I_v(t - \tau_1)]I_H(t - \tau_1)e^{-\mu\tau_1}$ are infectious at the present time t. The exponent $e^{-\mu\tau_1}$ is the probability that the vector survives the extrinsic incubation period. Similarly, let incubation period of the pathogen in the human have duration τ_2. Thus, of those humans infected τ_2 units of time ago, only a proportion $qa[N_H - I_H(t - \tau_2)]I_v(t - \tau_2)e^{-\alpha\tau_2}$ are infectious at the present time t. The exponent $e^{-\alpha\tau_2}$ is the probability that a human remains infected during the entire incubation period. The delay vector-borne model takes the form [140]

$$\begin{aligned} I_v' &= pae^{-\mu\tau_1}(N_v - I_v(t - \tau_1))I_H(t - \tau_1) - \mu I_v, \\ I_H' &= qae^{-\alpha\tau_2}(N_H - I_H(t - \tau_2))I_v(t - \tau_2) - \alpha I_H. \end{aligned} \tag{4.14}$$

4.3.1 Reducing the Delay Model to a Single Equation

Scientists often use various methods to reduce the dimension of a system. The newly obtained system does not necessarily have the same dynamical behavior as the original one, but it is still useful for obtaining initial insights from a simpler model.

The reduction is typically based on the assumption that the lifespan of the vector is much shorter than that of the humans, that is, we assume that $\mu \gg \alpha$, which leads to much faster equilibration of the dynamics of the vector population. This assumption is common for vector-borne diseases transmitted by mosquitoes, such as malaria. Furthermore, we assume that the incubation period is approximately equal to the extrinsic incubation period, that is, $\tau_1 = \tau_2$. This is certainly the case in malaria, where the incubation period in humans typically lasts between 10 days and 4 weeks. The extrinsic period is often temperature-dependent, but it lasts 10–18 days. With the assumption that the two incubation periods are the same, the model above becomes [111]

$$\begin{aligned} I_v' &= pae^{-\mu\tau}(N_v - I_v(t - \tau))I_H(t - \tau) - \mu I_v, \\ I_H' &= qae^{-\alpha\tau}(N_H - I_H(t - \tau))I_v(t - \tau) - \alpha I_H. \end{aligned} \tag{4.15}$$

Furthermore, since the vector dynamics have reached equilibrium, we have $I_v' = 0$. At equilibrium, the population numbers at time t and $t - \tau$ are approximately the same. Hence, from the first equation, we have

$$I_v(t - \tau) = \frac{pae^{-\mu\tau}N_v I_H(t - \tau)}{pae^{-\mu\tau}I_H(t - \tau) + \mu}.$$

Substituting I_v in the second equation, we obtain the following single delay equation for the dynamics of the humans:

$$I_H' = \frac{pa^2 qe^{-\alpha\tau}e^{-\mu\tau}N_v I_H(t-\tau)}{pae^{-\mu\tau}I_H(t-\tau)+\mu}(N_H - I_H(t-\tau)) - \alpha I_H. \tag{4.16}$$

It is helpful to normalize this equation by setting $x = I_H/N_H$. The equation becomes

$$x' = \frac{pa^2 qme^{-\alpha\tau}e^{-\mu\tau}x(t-\tau)}{pae^{-\mu\tau}x(t-\tau)+\mu}(1 - x(t-\tau)) - \alpha x(t), \tag{4.17}$$

where as before, $m = N_v/N_H$, and aN_H has been replaced again by a.

Delay equations, just like ODEs, have equilibria. The value x^* is an equilibrium of model (4.17) if it satisfies the equation

$$\frac{pa^2 qme^{-\alpha\tau}e^{-\mu\tau}x^*}{pae^{-\mu\tau}x^*+\mu}(1-x^*) - \alpha x^* = 0. \tag{4.18}$$

This equation clearly has the solution $x^* = 0$, which gives the disease-free equilibrium. To investigate the stability of the disease-free equilibrium, we linearize the equation. We look for a solution $x(t) = x^* + y(t)$, where $y(t)$ is the perturbation, and $x^* = 0$. This means that we have to replace x with y and linearize the nonlinear term. Notice that

$$\frac{1}{pae^{-\mu\tau}x(t-\tau)+\mu} = \frac{1}{\mu(\frac{pa}{\mu}e^{-\mu\tau}y(t-\tau)+1)} \approx \frac{1}{\mu}\left[1 - \frac{pa}{\mu}e^{-\mu\tau}y(t-\tau)\right].$$

Hence, the linearization around the disease-free equilibrium is given by

$$y' = \frac{pa^2 qme^{-\alpha\tau}e^{-\mu\tau}y(t-\tau)}{\mu} - \alpha y(t).$$

Looking for a solution of the form $y(t) = \bar{y}e^{\lambda t}$, we obtain the following characteristic equation:

$$\lambda + \alpha = \frac{pa^2 qme^{-\alpha\tau}e^{-\mu\tau}e^{-\lambda\tau}}{\mu}.$$

The above equation is a **transcendental equation**, that is, an equation containing a transcendental function of λ, namely $e^{\lambda\tau}$. If we think of λ as a real variable, the left-hand side of the above equation is an increasing linear function of λ, while the right-hand side is a decreasing function of λ. This equation always has a unique real solution, which is positive if and only if $\mathscr{R}_0 > 1$, where we define the reproduction number $\mathscr{R}_0$ to be

$$\mathscr{R}_0 = \frac{pa^2 qme^{-\alpha\tau}e^{-\mu\tau}}{\mu\alpha}. \tag{4.19}$$

So if $\mathscr{R}_0 > 1$, the disease-free equilibrium is unstable. If $\mathscr{R}_0 < 1$, the unique real eigenvalue is negative. We show that all other eigenvalues, which are complex,

have negative real part. Assume that we have an eigenvalue $\lambda = b + ci$, where i is the imaginary unit, that has nonnegative real part, that is, $b \geq 0$. Then $|\lambda + \alpha| = \sqrt{(b+\alpha)^2 + c^2} \geq b + \alpha \geq \alpha$. At the same time,

$$\begin{aligned} &\left| \frac{pa^2 qme^{-\alpha\tau} e^{-\mu\tau} e^{-\lambda\tau}}{\mu} \right| \\ &= \frac{pa^2 qme^{-\alpha\tau} e^{-\mu\tau} |e^{-\lambda\tau}|}{\mu} \\ &= \frac{pa^2 qme^{-\alpha\tau} e^{-\mu\tau} e^{-b\tau}}{\mu} \\ &\leq \frac{pa^2 qme^{-\alpha\tau} e^{-\mu\tau}}{\mu}, \end{aligned} \tag{4.20}$$

which gives a contradiction to the fact that $\mathscr{R}_0 < 1$, that is, $\alpha > \frac{pa^2 qme^{-\alpha\tau} e^{-\mu\tau}}{\mu}$. Hence, the disease-free equilibrium is locally asymptotically stable if $\mathscr{R}_0 < 1$. We note that if $\mathscr{R}_0 = 1$, then $\lambda = 0$ is an eigenvalue, and we cannot use this argument to draw any conclusions. We consider again the equation for the equilibria. Canceling x^*, we see that the nontrivial endemic equilibria satisfy the equation

$$\frac{pa^2 qme^{-\alpha\tau} e^{-\mu\tau}}{pae^{-\mu\tau} x^* + \mu} (1 - x^*) - \alpha = 0. \tag{4.21}$$

Multiplying by the denominator, we obtain a linear equation in x^*, which can be solved to give the unique endemic equilibrium:

$$x^* = \frac{\mathscr{R}_0 - 1}{\frac{pa}{\mu} e^{-\mu\tau} + \mathscr{R}_0}. \tag{4.22}$$

It is clear from this expression that the endemic equilibrium exists and that it is positive if and only if $\mathscr{R}_0 > 1$.

To investigate the stability of the endemic equilibrium, we linearize around it. Set $x(t) = x^* + y(t)$, where $y(t)$ is the perturbation of the endemic equilibrium. The perturbation y can take positive and negative values. Furthermore, to simplify the notation, we will set $Q = pa^2 qme^{-\alpha\tau} e^{-\mu\tau}$ and $P = pae^{-\mu\tau}$. Substituting in the delay equation (4.17), we obtain the following equation for the perturbation:

$$y'(t) = \frac{Q(x^* + y(t-\tau))}{P(x^* + y(t-\tau)) + \mu} [1 - x^* - y(t-\tau)] - \alpha(x^* + y(t)). \tag{4.23}$$

Taking into account the equation for the equilibrium

$$\frac{Qx^*(1 - x^*)}{Px^* + \mu} = \alpha x^* \tag{4.24}$$

and linearizing as in the case of the disease-free equilibrium, we obtain the following equation for the perturbation y:

$$y'(t) = \frac{Q(1-x^*)y(t-\tau)}{Px^*+\mu} - \frac{Qx^*}{Px^*+\mu}\left[\frac{P(1-x^*)y(t-\tau)}{Px^*+\mu} + y(t-\tau)\right] - \alpha y(t). \tag{4.25}$$

This equation can be simplified as follows:

$$y'(t) = \frac{Q(1-x^*)y(t-\tau)}{Px^*+\mu}\left[1 - \frac{Px^*}{Px^*+\mu}\right] - \frac{Qx^*y(t-\tau)}{Px^*+\mu} - \alpha y(t). \tag{4.26}$$

Using the equation for the equilibrium (4.24) and the fact that $\mathscr{R}_0 = Q/(\alpha\mu)$, we obtain the following simplified linearized equation:

$$y'(t) = \frac{\alpha\mu}{Px^*+\mu}(1-\mathscr{R}_0 x^*)y(t-\tau) - \alpha y(t). \tag{4.27}$$

Looking for an exponential solution $y(t) = \bar{y}e^{\lambda t}$, we obtain a characteristic equation of the form

$$\lambda + \alpha = \frac{\alpha\mu}{Px^*+\mu}(1-\mathscr{R}_0 x^*)e^{-\lambda\tau}. \tag{4.28}$$

If $\mathscr{R}_0 x^* < 1$, the coefficient in front the term $e^{-\lambda\tau}$ is positive and less than α, which corresponds to the case $\mathscr{R}_0 < 1$ in the characteristic equation for the disease-free equilibrium. A similar argument can be used to show that all roots of (4.28) have negative real parts, and the endemic equilibrium is locally asymptotically stable. We summarize these results in the following theorem:

Theorem 4.2. *If $\mathscr{R}_0 < 1$, the delay-differential equation (4.17) has only the disease-free equilibrium $x^* = 0$, which is locally asymptotically stable. If $\mathscr{R}_0 > 1$, the delay-differential equation (4.17) has a disease-free equilibrium and a unique endemic equilibrium x^*. If $\mathscr{R}_0 > 1$, the disease-free equilibrium is unstable. The endemic equilibrium is locally asymptotically stable if in addition, $\mathscr{R}_0 x^* < 1$.*

4.3.2 Oscillations in Delay-Differential Equations

If $\mathscr{R}_0 x^* > 1$, then the coefficient on the right-hand side of the characteristic equation (4.28) is negative, and the equation can have as principal eigenvalues (eigenvalues with the largest real part) a pair of complex conjugate eigenvalues. However, as a parameter changes, this pair of principal eigenvalues may cross the imaginary axis, giving rise to a stable oscillatory solution. At the same time, the principal eigenvalues start having positive real part, and the endemic equilibrium becomes unstable. This process that gives rise to a stable oscillatory solution is called (as in the ODE case) **Hopf bifurcation**. The result, valid for ODEs, is also valid for delay-differential equations. For delay-differential equations, it is given in the Hopf bifurcation Theorem below:

Theorem 4.3. *Consider the delay-differential equation*

$$x'(t) = F(x(t), x(t-\tau_1), \ldots, x(t-\tau_{n-1}), \mu), \tag{4.29}$$

where μ is a parameter. Suppose the following conditions hold:

(a) F is analytic in x and μ in a neighborhood of $(\mathbf{x}^*, \mu_0)$ in $\Re^n \times \Re$.
(b) $F(\mathbf{x}^*, \mu) = 0$ for μ in an open interval containing μ_0, and $x(t) = x^*$ is an isolated stationary solution of (4.29).
(c) The characteristic equation of (4.29) has a pair of complex conjugate eigenvalues λ and $\bar{\lambda}$ such that $\lambda(\mu) = b(\mu) + i\omega(\mu)$, where $\omega(\mu_0) = \omega_0 > 0$, $b(\mu_0) = 0$ and $b'(\mu_0) \neq 0$.
(d) The remaining eigenvalues of the characteristic equation have strictly negative real parts.

Then the delay-differential equation (4.29) has a family of Hopf periodic solutions.

One can apply Theorem 4.3 to show rigorously that a Hopf bifurcation occurs in Eq. (4.17). Instead, we will build a specific numerical example of such an oscillatory solution. To find sustained oscillations in Eq. (4.17), we need to find values of the parameters for which such oscillations occur. We begin from the characteristic equation (4.28), which we simplify further and write as

$$\lambda + \alpha = \rho e^{-\lambda\tau}, \tag{4.30}$$

where $\rho = \frac{\alpha\mu}{Px^*+\mu}(1 - \mathscr{R}_0 x^*)$. We recall that we have assumed that $\rho < 0$. Let $\lambda = b + i\omega$. We separate the real and the imaginary parts:

$$\begin{aligned} b + \alpha &= \rho e^{-b\tau}\cos[\omega\tau] \\ \omega &= -\rho e^{-b\tau}\sin[\omega\tau]. \end{aligned} \tag{4.31}$$

Now we ask whether we can find parameters $\alpha > 0$ and $\rho < 0$ such that the system above has positive solution $b > 0$ and $\omega > 0$. We solve in terms of α and ρ:

$$\begin{aligned} \alpha &= -b - \omega\cot[\omega\tau] \\ \rho &= -\omega e^{b\tau}\csc[\omega\tau]. \end{aligned} \tag{4.32}$$

As we have seen earlier, some of the parameters that have physical meaning can be estimated in advance, or at least reasonable biological ranges can be determined for them. In the equations above, we assume values for b ($b = 0.396053$) and τ ($\tau = 1$) and interpret α and ρ as functions of ω. Using a computer algebra system, we can make a parametric plot of α and ρ in the (α, ρ)-plane. This plot is shown in Fig. 4.1. We pick a value for ω, say $\omega = 2.47471$. From system (4.32), we obtain the values $\alpha = 2.74766$ and $\rho = -5.94512$. The value of α corresponds to an infectious period of $1/2.74766 = 0.3639$ years, which is a reasonable duration for *Plasmodium falciparum* malaria. Now we have to assume values of the remaining parameters, so that the combined value of ρ is as given. We assume the value $\mu = 12$, which gives 1 month as the duration of the vector's lifespan. This duration is realistic for

mosquitoes. Furthermore, we have to find Q and P such that the following system is valid:

$$\begin{aligned} \frac{Q(1-x^*)}{Px^*+\mu} &= \alpha, \\ \frac{\mu\alpha(1-\mathscr{R}_0x^*)}{Px^*+\mu} &= \rho. \end{aligned} \tag{4.33}$$

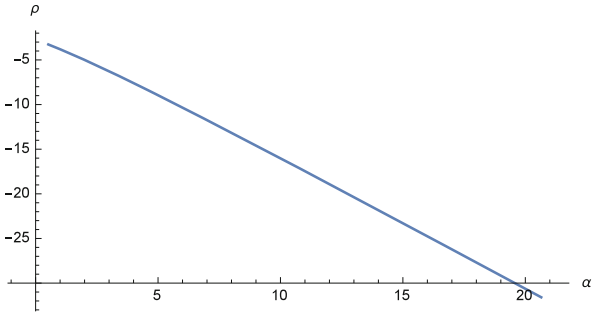

Fig. 4.1 Parametric plot of α and ρ in the (α,ρ)-plane as given by Eq. (4.32). The values of b and τ are taken as $b = 0.396053$, $\tau = 1$. The value of τ, which is equal to 1 year, is rather high for *Plasmodium falciparum* malaria. The plot is made for $2 \le \omega \le 3$

Dividing these two equations, we have

$$\frac{\mathscr{R}_0(1-x^*)}{1-\mathscr{R}_0x^*} = \frac{\alpha}{\rho}.$$

From here, assuming a value of $\mathscr{R}_0x^*$, we can compute $\mathscr{R}_0$:

$$\mathscr{R}_0 = \mathscr{R}_0x^* + \frac{\alpha}{\rho}(1-\mathscr{R}_0x^*).$$

If we take $\mathscr{R}_0x^* = 5$, then $\mathscr{R}_0 = 6.84869$. From here, we can compute $x^* = 0.73$. Finally, $Q = \mathscr{R}_0\alpha\mu = 225.816$. From the second equation in system (4.33), we determine $P = 13.9498$. With these parameters, we plot the solution of Eq. (4.17) in Fig. 4.2. The trajectory in Fig. 4.2 suggests that the endemic equilibrium is indeed unstable. However, the trajectory is not periodic. It is *aperiodic*, suggesting the presence of **chaos** in model (4.17). What is chaos? There are many definitions. Perhaps the most useful in biology is the following [148]:

Definition 4.2. *Chaos* is aperiodic long-term behavior in a deterministic system that exhibits sensitive dependence on initial conditions.

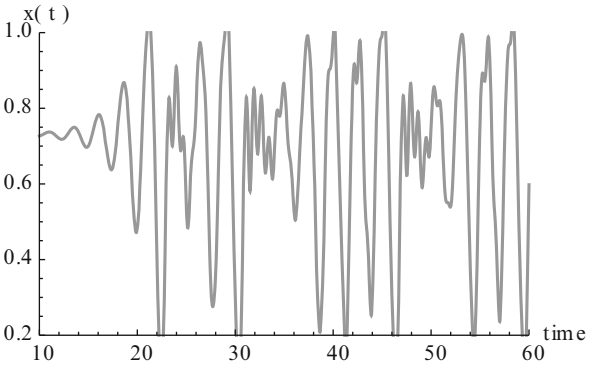

Fig. 4.2 Plot of the solution of Eq. (4.17) with $P = pae^{-\mu\tau} = 13.9498$, $Q = Pqame^{-\alpha\tau} = 225.816$, $\tau = 1$, $\alpha = 2.74768$, $\mu = 12$, and initial condition $x(0) = 0.73$. The resulting trajectory is aperiodic, suggesting the presence of chaotic behavior

This definition has several components:

1. *Aperiodic long-term behavior* means that there are trajectories that do not settle down to fixed points, periodic orbits, or quasiperiodic orbits as $t \to \infty$. For practical purposes, we require that these aperiodic orbits not be too rare.
2. *Deterministic* means that the system has no random or noisy inputs.
3. *Sensitive dependence on initial conditions* means that nearby trajectories separate exponentially fast.

From Fig. 4.2, we see that the delay malaria model (4.17) has solutions that are aperiodic; that is, their trajectory does not repeat even when run for a long time. Furthermore, the trajectories exhibit sensitive dependence on initial data. If we start very close to the trajectory above, the two trajectories "coincide" for a certain amount of time, called the *time horizon*, after which they completely diverge, and one looks nothing like the other. The sensitive dependence is illustrated in Fig. 4.3. The existence of sensitive dependence on initial conditions in simple but chaotic models means that we have lost the ability to make long-term predictions. However, we can still make short-term predictions based on chaotic models. Chaotic behavior emerges from periodic behavior through a process called period doubling. This suggests that if we decrease the bifurcation parameter, which in delay models is usually taken to be the delay τ, we will obtain a regular periodic solution. This is indeed the case. Figure 4.4 shows a periodic trajectory produced with the same parameters as above and $\tau = 0.6$. We see that even first-order deterministic delay models can exhibit chaotic behavior and sustained oscillations. ODE models need three dependent variables to exhibit chaotic behavior, and at least two dependent variables to produce oscillations.

Why is it important to have models that can have oscillatory solutions? Many vector-borne diseases give periodic values for the incidence or the prevalence of the disease in humans. There are two types of periodic behavior exhibited by vector-borne incidence/prevalence data:

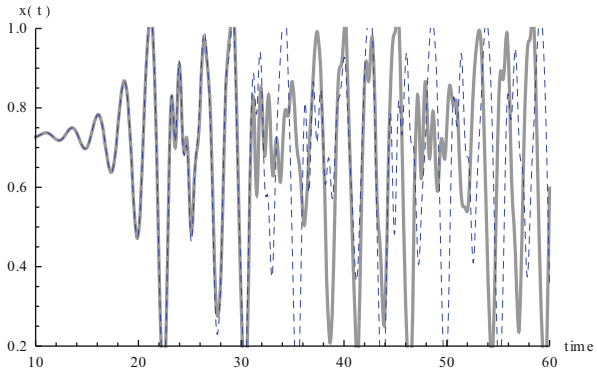

Fig. 4.3 Plot of two solutions of Eq. (4.17) with $P = pae^{-\mu\tau} = 13.9498$, $Q = Pqame^{-\alpha\tau} = 225.816$, $\tau = 1$, $\alpha = 2.74768$, $\mu = 12$, and initial conditions $x_1(0) = 0.73$ and $x_2(0) = 0.730001$. The two close trajectories coincide for a while and then diverge, suggesting sensitive dependence on the initial conditions

1. *Seasonality* is a periodic behavior that corresponds to the seasons in a given region. Since vector demographic and disease characteristics depend on the temperature and humidity, many vector-borne diseases, particularly those transmitted by mosquitoes, are transmitted more effectively and show higher numbers of cases in warm and rainy seasons. This gives periodic behavior that repeats every year. This is the most common periodicity exhibited by vector-borne diseases. It is typically modeled by periodic forcing in the vector demographic characteristics (such as vector birth and death) and results in nonautonomous models with periodic coefficients.
2. *Interannual periodicity* in disease prevalence or incidence occurs when the corresponding disease characteristics repeat with a period of two or more years. Interannual periodicity is much rarer and not well understood. There could be a number of reasons producing this behavior, but some may be internal and should be modeled with autonomous differential equation models such as those considered here.

4.3.3 The Reproduction Number of the Model with Two Delays

In this subsection, we compute the reproduction number of the vector-borne disease model with two delays introduced earlier in this section:

$$\begin{aligned} I_v' &= pae^{-\mu\tau_1}(N_v - I_v(t-\tau_1))I_H(t-\tau_1) - \mu I_v, \\ I_H' &= qae^{-\alpha\tau_2}(N_H - I_H(t-\tau_2))I_v(t-\tau_2) - \alpha I_H. \end{aligned} \tag{4.34}$$

The disease-free equilibrium in the above model is $\mathscr{E}_0 = (0,0)$. Linearizing around the disease-free equilibrium by setting $I_v = x + 0$ and $I_H = y + 0$, we obtain the equations for the perturbations:

$$\begin{aligned} x' &= pae^{-\mu\tau_1}N_v y(t-\tau_1) - \mu x(t), \\ y' &= qae^{-\alpha\tau_2}N_H x(t-\tau_2) - \alpha y(t). \end{aligned} \tag{4.35}$$

Looking for exponential solutions, we obtain the following eigenvalue problem:

$$\begin{aligned} \lambda x &= pae^{-\mu\tau_1}N_v y e^{-\tau_1\lambda} - \mu x, \\ \lambda y &= qae^{-\alpha\tau_2}N_H x e^{-\tau_2\lambda} - \alpha y. \end{aligned} \tag{4.36}$$

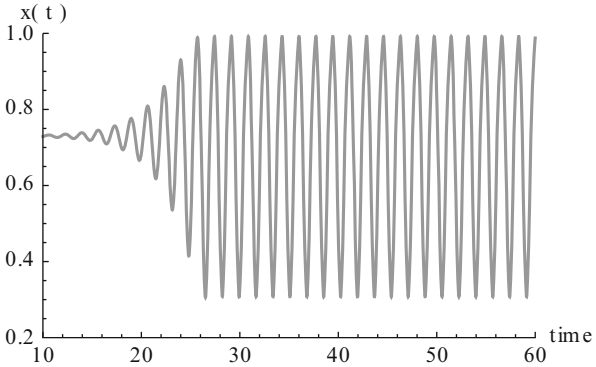

Fig. 4.4 Plot of a periodic solution of Eq. (4.17) with $P = pae^{-\mu\tau} = 13.9498$, $Q = Pqame^{-\alpha\tau} = 225.816$, $\tau = 0.6$, $\alpha = 2.74768$, $\mu = 12$, and initial condition $x(0) = 0.73$

This is a linear system for x and y. The only way this system can have a nonzero solution is if the determinant is zero. We obtain

$$\begin{vmatrix} -(\lambda+\mu) & pae^{-\mu\tau_1}N_v e^{-\tau_1\lambda} \\ qae^{-\alpha\tau_2}N_H e^{-\tau_2\lambda} & -(\lambda+\alpha) \end{vmatrix} = 0. \tag{4.37}$$

This expression gives the following characteristic equation:

$$(\lambda+\mu)(\lambda+\alpha) = pae^{-\mu\tau_1}N_v e^{-\tau_1\lambda} qae^{-\alpha\tau_2}N_H e^{-\tau_2\lambda}. \tag{4.38}$$

We define

$$\mathscr{R}_0 = \frac{pae^{-\mu\tau_1}N_v qae^{-\alpha\tau_2}N_H}{\alpha\mu}.$$

It is clear that Eq. (4.38) has a real positive solution if $\mathscr{R}_0 > 1$. If $\mathscr{R}_0 < 1$, we have $|(\lambda+\mu)(\lambda+\alpha)| \geq \alpha\mu$ for every λ with nonnegative real part. At the same time, $|pae^{-\mu\tau_1}N_v e^{-\tau_1\lambda} qae^{-\alpha\tau_2}N_H e^{-\tau_2\lambda}| \leq pae^{-\mu\tau_1}N_v qae^{-\alpha\tau_2}N_H$ for λ with nonnegative real part. Hence, Eq. (4.38) has no solutions λ with nonnegative real part. We conclude that for $\mathscr{R}_0 > 1$, the disease-free equilibrium is unstable, while if $\mathscr{R}_0 < 1$, all solutions λ of (4.38) have negative real part. Therefore, the disease-free equilibrium is locally asymptotically stable.

4.4 A Vector-Borne Disease Model with Temporary Immunity

Most vector-borne diseases lead to acquired immunity as a result of exposure. This immunity is partial and often temporary. This is in particular the case with malaria and dengue. In this section we build a vector-borne disease model whereby recovered individuals are immune but gradually lose immunity and become susceptible again. The model with permanent immunity was one of the early models of malaria. While malaria does not confer permanent immunity after infection, other vector-borne diseases, such as yellow fever, do. The model with permanent immunity is a **nested** model and is obtained from the model with waning immunity by setting the rate of loss of immunity to zero. To introduce the model, let S_v and I_v be the susceptible and infected vectors, and S, I, and R susceptible, infected, and recovered humans. The vector is modeled with an SI epidemic model. The model for the vector becomes

$$\begin{aligned} S_v' &= \Lambda_v - paS_vI - \mu_v S_v, \\ I_v' &= paS_vI - \mu_v I_v, \end{aligned} \tag{4.39}$$

where p, q, and a have the same meaning as before, and μ_v denotes the vector death rate. The total population size of the vector satisfies

$$N_v' = \Lambda_v - \mu_v N_v,$$

and it is asymptotically constant, that is, $N_v(t) \to \frac{\Lambda_v}{\mu_v}$. The model for the host is an SIRS epidemic model:

$$\begin{aligned} S' &= \Lambda - qaSI_v - \mu S + \gamma R, \\ I' &= qaSI_v - (\mu + \alpha)I, \\ R' &= \alpha I - (\mu + \gamma)R, \end{aligned} \tag{4.40}$$

where μ is the host natural death rate, Λ is the host birth rate, α is the recovery rate, and γ is the rate of loss of immunity. If $\gamma = 0$, then the model is one with permanent immunity. The parameters q and a have the same meaning as before. The total host population size satisfies the equation

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

and it is also asymptotically constant. The model has five equations. Using techniques as before, we may reduce the model to three equations by eliminating the equation for the susceptible vectors and, say, the equation for the recovered hosts. The reduced model has the form

$$\begin{aligned} I_v' &= pa(N_v - I_v)I - \mu_v I_v, \\ S' &= \Lambda - qaSI_v - \mu S + \gamma(N - S - I), \\ I' &= qaSI_v - (\mu + \alpha)I, \end{aligned} \tag{4.41}$$

where $N_v = \frac{\Lambda_v}{\mu_v}$ and $N = \frac{\Lambda}{\mu}$. To compute the reproduction number of the model (4.41), we have to compute the disease-free equilibrium and evaluate the Jacobian there. To compute the disease-free equilibrium, we set the time derivatives equal to zero. We obtain the nonlinear system

$$\begin{aligned} pa(N_v - I_v)I - \mu_v I_v &= 0, \\ \Lambda - qaSI_v - \mu S + \gamma(N - S - I) &= 0, \\ qaSI_v - (\mu + \alpha)I &= 0. \end{aligned} \tag{4.42}$$

Setting $I = 0$, we obtain that $I_v = 0$. From the second equation, we then have $S = N$. Thus, the disease-free equilibrium is given by $\mathscr{E}_0 = (0, N, 0)$, where the variables are ordered as in system (4.41). Stability of the disease-free equilibrium is given by the Jacobian:

$$J = \begin{pmatrix} -\mu_v & 0 & paN_v \\ -qaN & -\mu - \gamma & -\gamma \\ qaN & 0 & -(\mu + \alpha) \end{pmatrix}. \tag{4.43}$$

We look at the characteristic equation given by $|J - \lambda I| = 0$. If we expand by the second column, the second-row, second-column entry gives an eigenvalue $\lambda_1 = -(\mu + \gamma) < 0$. The remaining eigenvalues are the eigenvalues of the matrix

$$J_1 = \begin{pmatrix} -\mu_v & paN_v \\ qaN & -(\mu + \alpha) \end{pmatrix}. \tag{4.44}$$

This matrix has eigenvalues with negative real part if and only if its trace is negative and its determinant is positive. We have $\mathrm{Tr} J_1 = -\mu_v - (\mu + \alpha) < 0$. The condition that the determinant has to be positive gives $\mu_v(\mu + \alpha) - paN_v qaN > 0$. That condition can be rewritten as $\mathscr{R}_0 < 1$, where the reproduction number is given by

$$\mathscr{R}_0 = \frac{paN_v qaN}{\mu_v(\mu + \alpha)}.$$

Acknowledgements A portion of this chapter was previously published in [111]. Martcheva and Prosper [111] contains additional interpretations of delay models of vector-borne diseases as a tool for modeling unstable malaria.

Problems

4.1. Delayed Logistic Model

The delayed logistic equation is given by

$$x'(t) = rx(t)\left(1 - \frac{x(t-\tau)}{K}\right).$$

(a) Find the equilibria of the delayed logistic equation.

(b) Establish the stabilities of the equilibria.
(c) Use the Hopf bifurcation theorem to find an example of oscillations in the delayed logistic equation. Use a computer algebra system to simulate the oscillations.

4.2. Dependence of Equilibrium on $\mathscr{R}_0$
For the model with two delays

$$\begin{aligned} I_v' &= pae^{-\mu\tau_1}(N_v - I_v(t-\tau_1))I_H(t-\tau_1) - \mu I_v, \\ I_H' &= qae^{-\alpha\tau_2}(N_H - I_H(t-\tau_2))I_v(t-\tau_2) - \alpha I_H, \end{aligned} \tag{4.45}$$

do the following:

(a) Compute the endemic equilibrium $\mathscr{E}^*$.
(b) Express the endemic equilibrium in terms of the reproduction number $\mathscr{R}_0$.
(c) Choose parameter values and graph the prevalence I_H^* and I_v^* in the endemic equilibrium as functions of the reproduction number $\mathscr{R}_0$.
(d) Based on the graph in (c), hypothesize how the equilibrial prevalences depend on $\mathscr{R}_0$. Show that your hypothesis is correct for all parameter values.

4.3. Delayed SIR Model
Use the methodology described in this chapter to analyze the stability of equilibria of the delayed SIR model. Do oscillations occur?

$$\begin{aligned} S'(t) &= \Lambda - \mu S - \beta S(t)I(t), \\ I'(t) &= \beta S(t-\tau)I(t-\tau) - (\mu+\alpha)I(t), \\ R'(t) &= \alpha I(t) - \mu R(t). \end{aligned} \tag{4.46}$$

4.4. Global Stability
Consider the model given in (4.6):

$$\begin{aligned} I_v' &= pa(N_v - I_v)I_H - \mu I_v, \\ I_H' &= qa(N_H - I_H)I_v - \alpha I_H. \end{aligned} \tag{4.47}$$

(a) Use Dulac's criterion to rule out periodic orbits in the model.
(b) Use the Poincaré–Bendixson theorem to show global stability of the endemic equilibrium.

4.5. Vector-Borne Disease Model
Consider the following model of vector-borne disease with temporary immunity:

$$\begin{aligned} S'(t) &= \Lambda - \mu S - \frac{\beta N_v S(t)I(t)}{\beta_v I(t) + \mu_v} + \gamma R(t), \\ I'(t) &= \frac{\beta N_v S(t)I(t)}{\beta_v I(t) + \mu_v} - (\mu+\alpha)I(t), \\ R'(t) &= \alpha I(t) - (\mu+\gamma)R(t). \end{aligned} \tag{4.48}$$

(a) Explain how the above model was obtained from the model (4.39) and (4.40).
(b) Reduce the model to a two equation-model by eliminating the $R(t)$ variable.
(c) Find all equilibria of the two-dimensional model and their stabilities.
(d) Draw phase portraits for the cases $\mathscr{R}_0 > 1$ and $\mathscr{R}_0 < 1$.

4.6. Saturating Incidence
Consider the SIS model with saturating incidence:

$$\begin{aligned} S'(t) &= \Lambda - \frac{\beta I(t)S(t)}{1+\alpha_1 S(t)+\alpha_2 I(t)} + \gamma I - \mu S, \\ I'(t) &= \frac{\beta I(t-\tau)S(t-\tau)}{1+\alpha_1 S(t-\tau)+\alpha_2 I(t-\tau)} - (\gamma+\mu)I. \end{aligned} \tag{4.49}$$

(a) Determine the reproduction number $\mathscr{R}_0$ and the disease-free equilibrium of the model above.
(b) Determine biologically sensible parameters that give a reproduction number $\mathscr{R}_0 > 1$. Use a computer algebra system to simulate the model with this parameter set. Vary the delay parameter τ. Do the dynamics of the model change?
(c) Show that if $\mathscr{R}_0 < 1$, the disease-free equilibrium is locally stable. Furthermore, show that if $\mathscr{R}_0 > 1$, the disease-free equilibrium is unstable.
(d) Determine conditions for the existence of an endemic equilibrium.

4.7. Elasticity
The elasticity of quantity Q with respect to a parameter p is defined as

$$\varepsilon_p = \frac{\partial Q}{\partial p}\frac{p}{Q}, \tag{4.50}$$

where Q is any quantity and p is any parameter. Consider the quantity

$$\mathscr{R}_0 = \frac{pae^{-\mu\tau_1}N_v qae^{-\alpha\tau_2}N_H}{\alpha\mu}.$$

The elasticity says that 1% change in the parameter p leads to ε_p% change in Q. If $\varepsilon_p > 0$, then Q is increasing with p; if $\varepsilon_p < 0$, then Q is decreasing with p.

(a) Compute the elasticities of $\mathscr{R}_0$ with respect to the parameters p, q, a, μ, α. To which parameter is the reproduction number most sensitive?
(b) List a set of control measures applied to malaria. Which parameters does each control measure affect?

4.8. Maturation Delays
The following model has been proposed for the transmission of a vector-borne disease transmitted by mosquitoes [59]:

$$\begin{aligned}
\frac{dM_S}{dt} &= rM_s(t-\tau)e^{-d_j\tau}e^{-\alpha N_M(t)} - \beta_m M_S(t)I(t) - dM_S(t),\\
\frac{dM_S}{dt} &= \beta_m M_S(t)I(t) - dM_i(t),\\
\frac{dS}{dt} &= -\beta_h M_i(t)S(t),\\
\frac{dI}{dt} &= \beta_h M_i(t)S(t) - (\mu+\nu)I,
\end{aligned} \tag{4.51}$$

where M_S is the number of susceptible mosquitoes, M_i is the number of infected mosquitoes, S is the number of susceptible hosts, and I is the number of infected hosts. The parameters have the following meanings: r is the maximum daily mosquito-egg production rate, d_j is the death rate of juvenile mosquitoes, d is the death rate of adult mosquitoes, β_m and β_h are the two transmission rates, μ is the recovery rate of the hosts, and ν is the disease-induced death rate.

(a) Explain what biological scenario is modeled by the delay in this model.
(b) Find the differential equation model satisfied by the total mosquito population size. Compute the equilibria of this model.
(c) Determine the stabilities of the equilibria of the differential equation model satisfied by the total mosquito population size.

4.9. Temporary Immunity of Fixed Length
When individuals are subjected to temporary immunity of fixed length, the standard Kermack–McKendrick model with delay becomes

$$\begin{aligned}
S'(t) &= -\beta S(t)I(t) + \gamma I(t-\tau),\\
I'(t) &= \beta S(t)I(t) - \gamma I(t),\\
R'(t) &= \gamma I(t) - \gamma I(t-\tau).
\end{aligned} \tag{4.52}$$

(a) Notice that the total population size is constant. Reduce the system to an SI system.
(b) Compute the reproduction number and the equilibria of the system.
(c) Determine the stability of the disease-free equilibrium.
(d) Determine the linearized equation around the endemic equilibrium. Analyze the stability of the endemic equilibrium. Does Hopf bifurcation occur?

Chapter 5
Techniques for Computing $\mathscr{R}_0$

5.1 Building Complex Epidemiological Models

In the previous chapters, we introduced and studied the simplest epidemic models: the SIR model, the SIS model, and the Ross–MacDonald model of vector-borne disease. The baseline SI, SIS, SIR, and SIRS epidemic models can be extended to incorporate more realistic features of the disease, various control strategies, and heterogeneities of the host and the pathogens. As the models become more complex, their flowcharts also become more complex, with more compartments and transitions. Developing a flowchart of a complex model that is both clear and informative is more of an art than a science. Furthermore, as more realistic features are incorporated in the models and the systems become higher-dimensional, we need better tools for the computation of the reproduction number $\mathscr{R}_0$. In this chapter, we build models with more realism and introduce techniques for the computation of the reproduction number.

The baseline SI, SIS, and SIR models are developed based on essential components of the infectious disease transmission process: the presence of susceptible and infectious individuals. However, for many diseases, the natural progression of the disease may contain various other components that affect the disease transmission, and ultimately the conclusions we will make from the model.

5.1.1 Stages Related to Disease Progression

The most important stages of the disease other than susceptible and infectious that affect transmission are the (1) exposed/latent stage; (2) asymptomatic stage; (3) carrier stage; (4) passive immunity stage.

1. *Exposed/Latent Stage*. For many diseases, the infected individuals do not become immediately infectious, as assumed in the SIS and SIR models. The pathogen needs time to replicate and establish itself in the new host. The time during which

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_5

an individual is infected but not yet infectious is called the *latent* period. As an additional compartment in epidemiological models, the latent period is denoted by $E(t)$ or by $L(t)$. We note that the incubation period of a pathogen is the period between infection and onset of symptoms. The lengths of the latent and incubation periods do not necessarily coincide. For instance, in influenza, the onset of symptoms begins one day after infected individuals have become infectious; that is, the incubation period is one day longer than the latent period. The exposed (latent) period usually follows the susceptible stage. If we introduce a latent period in the SI model, the resulting model is an SEI model; if we introduce a latent period in the SIR model, the resulting model is an SEIR model. Similarly, we can have SEIS and SEIRS models based respectively on SIS and SIRS models. The SEIR model takes the form

$$\begin{aligned} S'(t) &= \Lambda - \beta SI - \mu S, \\ E'(t) &= \beta SI - (\eta + \mu)E, \\ I'(t) &= \eta E - (\alpha + \mu)I, \\ R'(t) &= \alpha I - \mu R, \end{aligned} \tag{5.1}$$

where η is the per capita rate of becoming infectious. We recall that $1/\eta$ is approximately the length of the latent period. The flowchart of the SEIR model is given in Fig. 5.1.

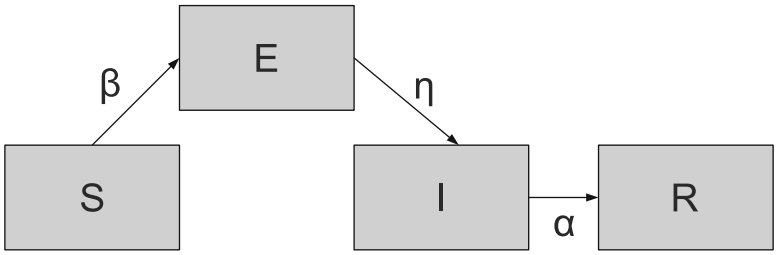

Fig. 5.1 Flowchart of the SEIR model. Demographic rates are not included

2. *Asymptomatic Stage*. The asymptomatic compartment is usually included to incorporate asymptomatic infection. *Asymptomatic infection*, also often called *subclinical infection*, is an infection without symptoms. Individuals with asymptomatic infection are still infectious and contribute to the distribution of the disease. However, because they do not show symptoms, they are much harder to detect. Asymptomatic infection has been shown to exist for many diseases, including HIV, malaria, dengue, measles, polio, and influenza. An asymptomatic compartment is typically included as an alternative to the infectious compartment I. The asymptomatic compartment is denoted by $A(t)$. Exposed individuals progress to the symptomatic infectious compartment with probability

p, and to the asymptomatic infectious compartment with probability $(1-p)$. Asymptomatic individuals are typically assumed to be infectious at a reduced transmission rate $q\beta$:

$$\begin{aligned} S'(t) &= \Lambda - \beta S(I+qA) - \mu S, \\ E'(t) &= \beta S(I+qA) - (\eta+\mu)E, \\ I'(t) &= p\eta E - (\alpha+\mu)I, \\ A'(t) &= (1-p)\eta E - (\gamma+\mu)A, \\ R'(t) &= \alpha I + \gamma A - \mu R. \end{aligned} \tag{5.2}$$

The parameter γ gives the recovery rate of the asymptomatic individuals. Often, the symptomatic infectious period is longer than the asymptomatic: $1/\gamma < 1/\alpha$. A flowchart of the model with asymptomatic infection is shown in Fig. 5.2.

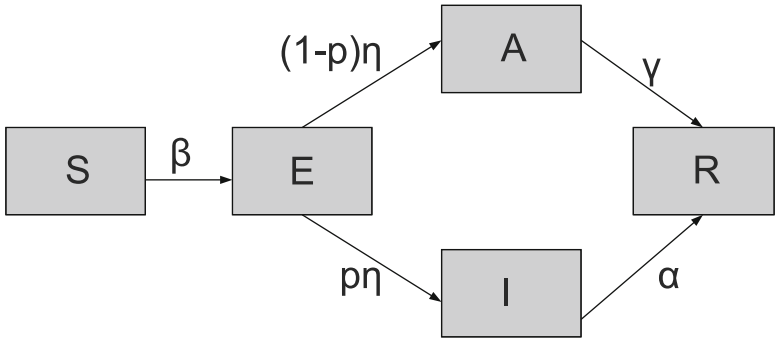

Fig. 5.2 Flowchart of the SEIR model with asymptomatic stage model. Demographic rates are not included

3. *Carrier Stage*. A carrier stage is incorporated to account for individuals who are otherwise healthy but harbor and transmit the pathogen. Carriers show no symptoms or signs of infection, but the disease microorganisms can be recovered from their nose, throat, or feces. Carrier individuals contribute to the distribution of the disease without being sick themselves. Alternatively, a pathogen carrier may be someone who has been infected and was not treated because the infection was asymptomatic, or was incompletely treated. Several viral and bacterial diseases exhibit a carrier stage. Viral diseases, such as viral hepatitis and poliomyelitis, and bacterial diseases, including diphtheria and meningococcal meningitis, often have a carrier stage. The carrier compartment is usually denoted by $C(t)$. An SIRS model with carrier stage is included below:

$$\begin{aligned} S'(t) &= \Lambda - \beta S(I+qC) - \mu S + \rho R, \\ C'(t) &= \beta S(I+qC) - (\eta+\gamma+\mu)C, \\ I'(t) &= \eta C - (\alpha+\mu)I, \\ R'(t) &= \alpha I + \gamma C - (\mu+\rho)R. \end{aligned} \tag{5.3}$$

Individuals are assumed to enter the carrier stage upon infection. Then some progress to being infected and infectious, while others recover from the pathogen without ever being infected. A flowchart of the model is given in Fig. 5.3.

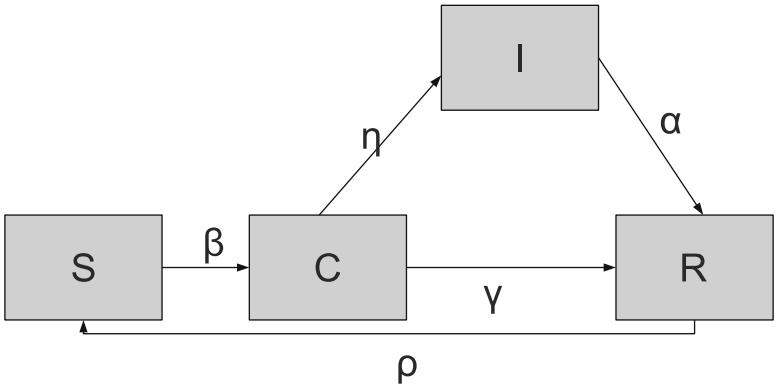

Fig. 5.3 Flowchart of the SCIRS model. Demographic rates are not included

4. *Passive Immunity Stage*. Passive immunity is the transfer of active immunity in the form of antibodies from one individual to another. Passive immunity can occur naturally, when maternal antibodies are transferred to the fetus through the placenta or in the milk during breastfeeding. Passive immunity can also be induced artificially, when high levels of antibodies specific for a pathogen or toxin are transferred to nonimmune individuals. Mathematical models typically involve natural passive immunity. The passive immunity stage is often denoted by $M(t)$, and it is added to the system before individuals become susceptible. An MSIR model with passive immunity has the form

$$\begin{aligned} M'(t) &= \Lambda - \rho M - \mu M, \\ S'(t) &= \rho M - \beta SI - \mu S, \\ I'(t) &= \beta SI - (\alpha + \mu)I, \\ R'(t) &= \alpha I - \mu R, \end{aligned} \tag{5.4}$$

where ρ is the per unit of time rate of loss of maternal antibodies. We assume that babies are completely protected by the maternal antibodies and cannot get infected.

5.1.2 Stages Related to Control Strategies

Besides involving compartments related to the various disease-progression stages, basic baseline SI, SIS, SIR, and SIRS models may involve compartments associated with the various strategies applied for control of the disease. Disease-control strategies involve (1) quarantine/isolation; (2) vaccination; (3) treatment.

1. *Quarantine/isolation*. *Quarantine* is compulsory isolation, typically to contain the spread of a disease. The word comes from the Italian (seventeenth-century Venetian) *quarantena*, meaning a 40-day period. Quarantine is applied to individuals who have come into contact with an infectious individual and may or may not be infected. *Isolation* is confinement of an infectious individual that restricts that individual's contact with healthy susceptible individuals. Isolation is applied to infectious individuals only. The class of quarantines/isolated individuals is typically denoted by $Q(t)$. An SIQR model is given below. Standard, rather than mass-action, incidence is assumed [61]. We note that the "total" population size consists of all individuals that participate in the mixing. We call all individuals that participate in the mixing *active* individuals, and denote the size of the active class by $A(t) = S(t) + I(t) + R(t)$:

$$\begin{aligned} S'(t) &= \Lambda - \beta SI/A - \mu S, \\ I'(t) &= \beta SI/A - (\alpha + \gamma + \mu)I, \\ Q'(t) &= \gamma I - (\eta + \mu)Q, \\ R'(t) &= \alpha I + \eta Q - \mu R. \end{aligned} \tag{5.5}$$

SIQR models are appropriate to model childhood diseases. Quarantine has been found to destabilize the epidemic and lead to sustained oscillations in the dynamics of childhood diseases [61]. A flowchart of the SIQR model is given in Fig. 5.4.

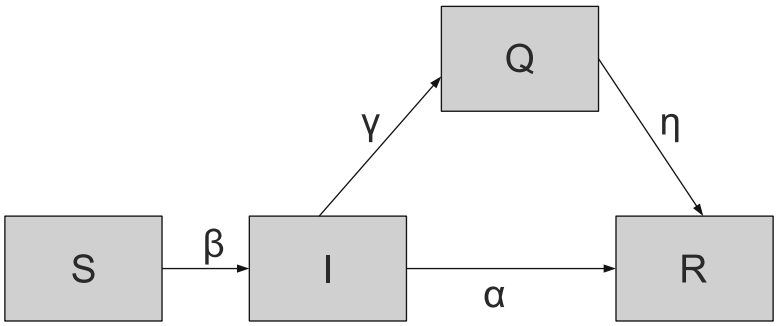

Fig. 5.4 Flowchart of the SIQR model. Demographic rates are not included

2. *Treatment*. Treatment is the care provided to reduce morbidity and mortality of an infectious disease. Treatment usually includes administering medication that mitigates symptoms and helps the immune system fight the pathogen. A treatment compartment $T(t)$ has been included in standard SEI epidemic models as a class that replaces the class of recovered individuals, making the model an SEIT model. Since patients often do not complete the full regimen of treatment, treatment may or may not be successful, and leads to the individual being successfully

treated with probability q or relapsing to the exposed/latent class with probability p, where $p+q=1$. The model with treatment is presented below, and it captures the dynamics of a disease such as tuberculosis [60]:

$$\begin{aligned} S'(t) &= \Lambda - \beta_1 SI/N - \mu S, \\ E'(t) &= \beta_1 SI/N + \beta_2 TI/N - (\mu + \kappa + r_1)E + pr_2 I, \\ I'(t) &= \kappa E - (r_2 + \mu)I, \\ T'(t) &= r_1 E + qr_2 I - \beta_2 TI/N - \mu T, \end{aligned} \tag{5.6}$$

where r_1 is the treatment rate of exposed individuals, r_2 is the treatment rate of infectious individuals, and κ is the progression to the infectious state. The presence of relapsing individuals leads to ambiguity in the computation of the reproduction number, which we will discuss later. A flowchart of the model is given in Fig. 5.5.

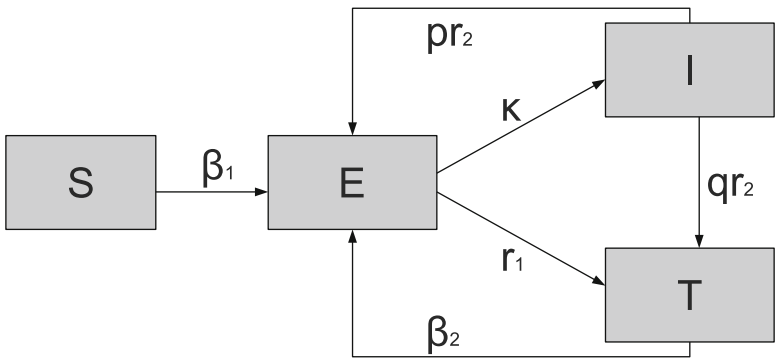

Fig. 5.5 Flowchart of the SEIT model. Demographic rates are not included

3. *Vaccination*. Vaccination is the administration of generally dead or weakened antigenic material (a vaccine) to produce immunity to a disease. There are two main ways to incorporate vaccination in an epidemic model. In the first approach vaccine is administered to individuals who are entering the system. A proportion p of entering individuals enters the susceptible class, while a proportion $1-p$ enters the recovered/immune class. In the second approach, susceptible individuals are continuously being vaccinated and move to the vaccinated class $V(t)$. Vaccination may provide complete or partial immunity to the disease. We will consider vaccination in more detail in Chap. 9.

5.1.3 Stages Related to Pathogen or Host Heterogeneity

The impact of host or pathogen heterogeneities on the dynamics of the disease has been a longstanding question of interest in mathematical epidemiology. We will consider the following heterogeneities:

1. *Pathogen Genetic Heterogeneities*. Many pathogens are represented by multiple genetically distinct variants. The presence of the variants has implication for the disease progression and control as different pathogen variants respond differently to the control measures. To study pathogen heterogeneities, two-strain and multistrain models have been developed. The simplest SIS model with two strains includes the number of infected with strain one I_1 and the number of infected with strain two I_2:

$$\begin{aligned} S'(t) &= \Lambda - \beta_1 SI_1 - \beta_2 SI_2 - \mu S + \alpha_1 I_1 + \alpha_2 I_2, \\ I_1'(t) &= \beta_1 SI_1 - (\mu + \alpha_1) I_1, \\ I_2'(t) &= \beta_2 SI_2 - (\mu + \alpha_2) I_2. \end{aligned} \tag{5.7}$$

   We address the interplay between pathogen heterogeneity and disease transmission dynamics in Chap. 8. A flowchart of the two-strain model above is given in Fig. 5.6.

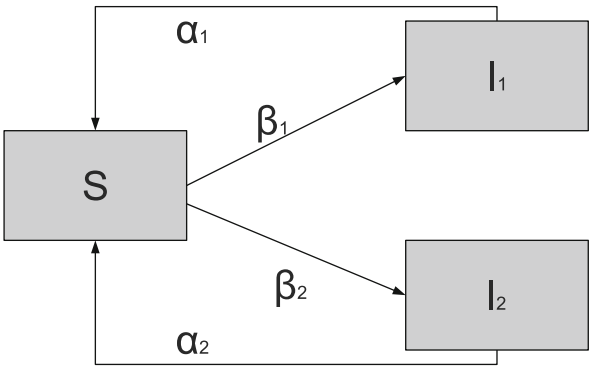

Fig. 5.6 Flowchart of the two-strain model. Demographic rates are not included

2. *Host Heterogeneities*. Host heterogeneities in susceptibility and infectivity associated with host genetic differences can lead to models with multiple host species. For instance, in avian influenza, the pathogen can infect both wild birds and domestic birds. These two categories of birds can have very different demographic characteristics as well as different characteristics with respect to the disease. To account for the differences, a simple SI model should have two susceptible populations S_w and S_d and two infected populations I_w and I_d. The multihost single-pathogen model takes the form

$$
\begin{aligned}
S_w'(t) &= \Lambda_w - \beta_{11} S_w I_w - \beta_{12} S_w I_d - \mu_w S_w, \\
I_w'(t) &= \beta_{11} S_w I_w + \beta_{12} S_w I_d - (\mu_w + \alpha_w) I_w, \\
S_d'(t) &= \Lambda_d - \beta_{21} S_d I_w - \beta_{22} S_d I_d - \mu_d S_d, \\
I_d'(t) &= \beta_{21} S_d I_w + \beta_{22} S_d I_d - (\mu_d + \alpha_d) I_d.
\end{aligned} \tag{5.8}
$$

A flowchart of the model is given in Fig. 5.7.

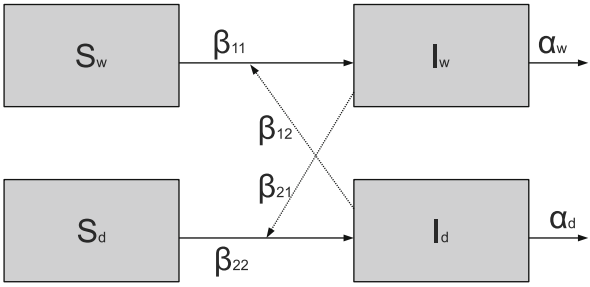

Fig. 5.7 Flowchart of the two-strain model. Demographic rates are not included

3. *Age and Space Heterogeneities*. Age and space heterogeneities are introduced respectively by the age of the host and the presence of a spatial component in the transmission of the disease. Age and space heterogeneities can be incorporated in the model as a continuous component, which results in partial differential equations, or as a discrete component, which results in larger systems of ordinary differential equations. These heterogeneities will be discussed in Chaps. 16 and 15.

5.2 Jacobian Approach for the Computation of $\mathscr{R}_0$

We saw in the previous chapters that mathematically, the reproduction number $\mathscr{R}_0$ gives a threshold condition for the stability of the disease-free equilibrium. Imposing conditions for stability leads to an expression for the reproduction number. This general approach can also be applied in more complex models. In particular, we compute the Jacobian of the system at the disease-free equilibrium, and we pose the condition that all eigenvalues of the corresponding characteristic equation must have negative real parts. In the two-dimensional case, this requirement follows from the conditions that $\mathrm{Tr}J < 0$ and $\mathrm{Det}J > 0$. In the higher-dimensional case, this theorem does not apply, but one can often reduce the characteristic equation to the two-dimensional case using the properties for manipulation of matrices. Even when we have a unique condition for the stability of the disease-free equilibrium, rewriting this condition in the form of a reproduction number can often be done in more than

one way. It is typically expected that the expression for the reproduction number should satisfy the following conditions.

The reproduction number should:

- Be nonnegative for nonnegative parameter values;
- Be zero if there is no transmission;
- Be interpretable as the number of secondary infections.

5.2.1 Examples in Which the Jacobian Reduces to a 2×2 Matrix

As a first higher-dimensional example, we consider the SIRS model with carrier stage (5.3). The disease-free equilibrium for that model is given as $\mathscr{E}_0 = (\frac{\Lambda}{\mu}, 0, 0, 0)$, where the variables are arranged in the same way as the equations in the system. Computing the Jacobian at the disease-free equilibrium gives

$$J = \begin{pmatrix} -\mu & -q\beta S^* & -\beta S^* & \rho \\ 0 & q\beta S^* - (\eta+\gamma+\mu) & \beta S^* & 0 \\ 0 & \eta & -(\alpha+\mu) & 0 \\ 0 & \gamma & \alpha & -(\mu+\rho) \end{pmatrix}, \tag{5.9}$$

where $S^* = \Lambda/\mu$. If we consider $|J - \lambda I| = 0$, we can expand the matrix in terms of the first column. This will give an eigenvalue $\lambda_1 = -\mu$. Then we can expand the remaining matrix around the last column. This will give an eigenvalue $\lambda_2 = -(\mu+\rho)$. The remaining two eigenvalues are the eigenvalues of the following matrix:

$$J_1 = \begin{pmatrix} q\beta S^* - (\eta+\gamma+\mu) & \beta S^* \\ \eta & -(\alpha+\mu) \end{pmatrix}. \tag{5.10}$$

Now we can apply the usual conditions that guarantee that the eigenvalues of J_1 have negative real part. In particular, we want $\mathrm{Tr}J_1 < 0$ and $\mathrm{Det}J_1 > 0$. The second inequality gives $-[q\beta S^* - (\eta+\gamma+\mu)](\alpha+\mu) - \eta\beta S^* > 0$. This condition gives a reproduction number in the form

$$\mathscr{R}_0 = \frac{q\beta S^*}{\eta+\gamma+\mu} + \frac{\eta\beta S^*}{(\eta+\gamma+\mu)(\alpha+\mu)}.$$

We notice that the condition $\mathscr{R}_0 < 1$ implies both $\mathrm{Tr}J_1 < 0$ and $\mathrm{Det}J_1 > 0$. Therefore, if $\mathscr{R}_0 < 1$, the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, the disease-free is unstable. There are multiple ways to rewrite the inequality $\mathrm{Det}J_1 > 0$ as a reproduction number, but the expression above is preferred, because it has a clear interpretation as the number of secondary cases. To see this, first we notice that the reproduction number consists of the sum of two terms. The first

term gives the number of secondary infections produced by one carrier, and the second term gives the number of secondary infections produced by an infectious individual. Since $q\beta SC$ is the incidence of a carrier, the number of secondary infections that one carrier will produce in an entirely susceptible population per unit of time is $q\beta S^*$. The time units spent in the carrier compartment is $1/(\eta+\gamma+\mu)$. Hence, the first term gives the number of secondary infections produced by a carrier in an entirely susceptible population during its lifetime as a carrier. Similarly, βSI is the incidence of an infected individual. Hence, the number of secondary infections that one infectious individual will produce in an entirely susceptible population per unit of time is βS^*. The number of time units spent in the carrier compartment is $1/(\alpha+\mu)$. The newly infected individual becomes a carrier, and the fraction of carriers that survive the carrier stage and become infected is $\eta/(\eta+\gamma+\mu)$. We see that the second term gives the number of secondary infections produced by an infectious individual in an entirely susceptible population during its lifetime as infectious.

5.2.2 Routh–Hurwitz Criteria in Higher Dimensions

For many higher-dimensional models, the Jacobian computed at the disease-free equilibrium cannot be reduced to a 2×2 matrix. The characteristic polynomial then has degree three or higher. In this case, the reproduction number can be obtained from the constant term. Whether the reproduction number is greater or less than 1 determines the sign of the constant term. Nonetheless, we still need tools that give necessary and sufficient conditions for the eigenvalues to have negative real parts. These conditions are given by the **Routh–Hurwitz criterion**, which is stated in the following theorem:

Theorem 5.1 (Routh–Hurwitz Criteria). *Consider the nth-degree polynomial with real constant coefficients*

$$P(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n.$$

Define n Hurwitz matrices using the coefficients a_i of the characteristic polynomial:

$$H_1=(a_1)\qquad H_2=\begin{pmatrix}a_1 & 1\\ a_3 & a_2\end{pmatrix}\qquad H_3=\begin{pmatrix}a_1 & 1 & 0\\ a_3 & a_2 & a_1\\ a_5 & a_4 & a_3\end{pmatrix}$$

and

$$H_n=\begin{pmatrix}a_1 & 1 & 0 & 0 & \cdots & 0\\ a_3 & a_2 & a_1 & 1 & \cdots & 0\\ a_5 & a_4 & a_3 & a_2 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & a_n\end{pmatrix},$$

where $a_j = 0$ if $j > n$. All roots of the polynomial $P(\lambda)$ are negative or have negative real part if and only if the determinants of all Hurwitz matrices are positive:

$$\mathrm{Det}H_j > 0, \qquad j = 1, \ldots, n.$$

For $n = 2$, the Routh–Hurwitz criterion simplifies to $a_1 > 0$ and $a_1 a_2 > 0$. We note that $a_3 = 0$ in H_2. These conditions are equivalent to $a_1 > 0$ and $a_2 > 0$ and are analogous to the conditions we applied before: $\mathrm{Tr}J < 0$ and $\mathrm{Det}J > 0$. The Routh–Hurwitz criteria for polynomials are given in Table 5.1. A necessary but not

Table 5.1 Routh–Hurwitz criteria

n	Coefficient signs	Additional conditions
2	$a_1 > 0, a_2 > 0$	–
3	$a_1 > 0, a_2 > 0, a_3 > 0$	$a_1 a_2 > a_3$
4	$a_1 > 0, a_2 > 0, a_3 > 0, a_4 > 0$	$a_1 a_2 a_3 > a_3^2 + a_1^2 a_4$
5	$a_1 > 0, a_2 > 0, a_3 > 0, a_4 > 0,$ $a_5 > 0$	$a_1 a_2 a_3 > a_3^2 + a_1^2 a_4,$ $(a_1 a_4 - a_5)(a_1 a_2 a_3 - a_3^2 - a_1^2 a_4) > a_5 (a_1 a_2 - a_3)^2 + a_1 a_5^2$

sufficient condition for the roots of the polynomial $P(\lambda)$ to be negative or have negative real part is that all coefficients be strictly positive. This is required in the column labeled *Coefficient signs* in Table 5.1.

We use the Routh–Hurwitz criterion to derive the reproduction number for the SEIR model with asymptomatic stage (5.2). The disease-free equilibrium in that model is given by $\mathscr{E}_0 = (S^*, 0, 0, 0, 0)$, where $S^* = \Lambda/\mu$. The Jacobian at the disease-free equilibrium is

$$J = \begin{pmatrix} -\mu & 0 & -\beta S^* & -q\beta S^* & 0 \\ 0 & -(\eta+\mu) & \beta S^* & q\beta S^* & 0 \\ 0 & p\eta & -(\alpha+\mu) & 0 & 0 \\ 0 & (1-p)\eta & 0 & -(\gamma+\mu) & 0 \\ 0 & 0 & \alpha & \gamma & -\mu \end{pmatrix}. \tag{5.11}$$

Expanding the determinant of the characteristic equation $|J - \lambda I| = 0$ by the first column and then by the last column, we obtain two of the eigenvalues of J: $\lambda_1 = -\mu$ and $\lambda_2 = -\mu$. The remaining three eigenvalues are the eigenvalues of the 3×3 matrix

$$J_1 = \begin{pmatrix} -(\eta+\mu) & \beta S^* & q\beta S^* \\ p\eta & -(\alpha+\mu) & 0 \\ (1-p)\eta & 0 & -(\gamma+\mu) \end{pmatrix}. \tag{5.12}$$

The characteristic equation takes the form

$$\begin{vmatrix} -(\eta+\mu+\lambda) & \beta S^* & q\beta S^* \\ p\eta & -(\alpha+\mu+\lambda) & 0 \\ (1-p)\eta & 0 & -(\gamma+\mu+\lambda) \end{vmatrix} = 0.$$

Expanding the determinant, we have

$$(\eta+\mu+\lambda)(\alpha+\mu+\lambda)(\gamma+\mu+\lambda) \\ -(1-p)\eta q\beta S^*(\alpha+\mu+\lambda)-p\eta\beta S^*(\gamma+\mu+\lambda)=0. \tag{5.13}$$

This leads to the following cubic equation in λ:

$$P(\lambda):=\lambda^3+a_1\lambda^2+a_2\lambda+a_3=0,$$

where the coefficients are the following expressions of the parameters:

$$a_1=\alpha+\mu+\gamma+\mu+\eta+\mu,$$
$$a_2=(\alpha+\mu)(\eta+\mu)+(\alpha+\mu)(\gamma+\mu)+(\gamma+\mu)(\eta+\mu)-(1-p)\eta q\beta S^*-p\eta\beta S^*,$$
$$a_3=(\alpha+\mu)(\eta+\mu)(\gamma+\mu)-(\mu+\alpha)(1-p)\eta q\beta S^*-(\gamma+\mu)p\eta\beta S^*.$$

The condition $\mathscr{R}_0<1$ should be equivalent to the condition $a_3>0$. Hence, we define $\mathscr{R}_0$ as follows:

$$\mathscr{R}_0=\frac{(1-p)\eta q\beta S^*}{(\eta+\mu)(\gamma+\mu)}+\frac{p\eta\beta S^*}{(\alpha+\mu)(\eta+\mu)}.$$

It is easy to see that the condition $a_3>0$ is equivalent to the condition $\mathscr{R}_0<1$. Furthermore, if $\mathscr{R}_0>1$, then $a_3<0$. Since $\lim_{\lambda\to\infty}P(\lambda)=\infty$, the equation $P(\lambda)=0$ has a real positive solution, and the disease-free equilibrium is therefore unstable. It remains to see that if $\mathscr{R}_0<1$, the disease-free equilibrium is locally asymptotically stable. In particular, we have to show that in this case, the equation $P(\lambda)=0$ has only negative roots or roots with negative real part. To see this, we use the Routh–Hurwitz criteria for dimension three. The Routh–Hurwitz conditions for stability for dimension three are listed in Table 5.1. We see that from the condition $\mathscr{R}_0<1$, it follows that $a_3>0$. Clearly $a_1>0$. The condition $\mathscr{R}_0<1$ also gives $(\eta+\mu)(\gamma+\mu)>(1-p)\eta q\beta S^*$ and $(\alpha+\mu)(\eta+\mu)>p\eta\beta S^*$. These two inequalities imply that $a_2>0$. Finally, we need to show that $a_1a_2>a_3$. Since $(\eta+\mu)(\gamma+\mu)>(1-p)\eta q\beta S^*$ and $(\alpha+\mu)(\eta+\mu)>p\eta\beta S^*$, we have that $a_1a_2>(\eta+\mu)(\gamma+\mu)(\alpha+\mu)$. On the other hand, it is clear that $a_3<(\eta+\mu)(\gamma+\mu)(\alpha+\mu)$. Consequently,

$$a_1a_2>(\eta+\mu)(\gamma+\mu)(\alpha+\mu)>a_3.$$

The Routh–Hurwitz criterion then implies that the disease-free equilibrium is locally asymptotically stable if $\mathscr{R}_0<1$. Mathematically, the above definition of $\mathscr{R}_0$ is sound. To be sound epidemiologically, the reproduction number should be interpretable as the number of secondary cases. To see this, first notice that the reproduction number is a sum of two terms. The first term gives the number of secondary infections of asymptomatic individuals produced by one asymptomatic individual. Indeed, $q\beta S^*$ is the number of newly exposed individuals generated by one asymptomatic individual per unit of time in an entirely susceptible population. A fraction $(1-p)\eta/(\eta+\mu)$ survives the exposed stage and progresses to the asymptomatic stage. One asymptomatic individual remains asymptomatic and infects other individuals as asymptomatic for $1/(\gamma+\mu)$ units. Hence,

$$\mathscr{R}_a = \frac{(1-p)\eta q\beta S^*}{(\eta+\mu)(\gamma+\mu)}$$

is the number of secondary infections that one asymptomatic individual will produce in an entirely susceptible population during its lifespan as asymptomatic. Similarly, βS^* is the number of newly exposed individuals generated by one infectious individual per unit of time in an entirely susceptible population. A fraction $p\eta/(\eta+\mu)$ survives the exposed stage and progresses to the infectious stage. One infectious individual remains infectious to susceptible individuals for $1/(\alpha+\mu)$ units. Hence,

$$\mathscr{R}_s = \frac{p\eta q\beta S^*}{(\eta+\mu)(\alpha+\mu)}$$

is the number of secondary infections that one symptomatic infectious individual will produce in an entirely susceptible population during its lifespan. Finally, $\mathscr{R}_0 = \mathscr{R}_a + \mathscr{R}_s$.

5.2.3 Failure of the Jacobian Approach

The Jacobian approach works well for models in which the necessary and sufficient conditions for stability of the Jacobian can be reduced to a single condition. We saw this in the case that the Jacobian can be reduced to a 2×2 matrix. The fact that the trace is negative is either automatically satisfied or follows from the requirement that $\mathscr{R}_0 < 1$, or that the determinant be positive. That, unfortunately, is not always the case. In some cases, the fact that the trace is negative is a completely independent condition, and it is not automatic, nor it does follow from the condition that the determinant should be positive. This situation occurs predominantly in models with host heterogeneities. For instance, problems with defining the reproduction number via the Jacobian approach will occur if the model has male and female susceptible individuals, or the pathogen is spreading in several genetically distinct host populations.

To illustrate the difficulties with defining the reproduction number via the Jacobian approach, we consider the wild bird–domestic bird avian influenza model above. The system is given in Eq. (5.8). The disease-free equilibrium in that system is given by $\mathscr{E}_0 = (S_w^*, 0, S_d^*, 0)$, where $S_w^* = \Lambda_w/\mu_w$ and $S_d^* = \Lambda_d/\mu_d$, and the variables are ordered as in the system. The Jacobian at the disease-free equilibrium is given by

$$J = \begin{pmatrix} -\mu_w & -\beta_{11}S_w^* & 0 & -\beta_{12}S_d^* \\ 0 & \beta_{11}S_w^* - (\mu_w+\alpha_w) & 0 & \beta_{12}S_d^* \\ 0 & -\beta_{21}S_d^* & -\mu_d & \beta_{22}S_d^* \\ 0 & \beta_{21}S_d^* & 0 & \beta_{22}S_d^* - (\mu_d+\alpha_d) \end{pmatrix}. \quad (5.14)$$

This Jacobian has two obvious eigenvalues: $\lambda_1 = -\mu_w$ and $\lambda_2 = -\mu_d$. The remaining eigenvalues are the eigenvalues of the 2×2 matrix

$$J_1 = \begin{pmatrix} \beta_{11} S_w^* - (\mu_w + \alpha_w) & \beta_{12} S_d^* \\ \beta_{21} S_d^* & \beta_{22} S_d^* - (\mu_d + \alpha_d) \end{pmatrix}. \tag{5.15}$$

The conditions that eigenvalues of this matrix are negative or have negative real part are that the trace be negative and the determinant be positive, that is,

$$\begin{aligned} \beta_{11} S_w^* - (\mu_w + \alpha_w) + \beta_{22} S_d^* - (\mu_d + \alpha_d) < 0, \\ (\beta_{22} S_d^* - (\mu_d + \alpha_d))(\beta_{11} S_w^* - (\mu_w + \alpha_w)) - \beta_{12} S_w^* \beta_{21} S_d^* > 0. \end{aligned} \tag{5.16}$$

The second inequality implies that $\beta_{11} S_w^* - (\mu_w + \alpha_w)$ and $\beta_{22} S_d^* - (\mu_d + \alpha_d)$ must have the same sign but does not imply that they both need to be negative. If we assume that they are both negative, then the first inequality holds, but the second may or may not hold. It does not seem obvious how to define $\mathscr{R}_0$ for this Jacobian so that inequalities (5.16) are both valid.

5.3 The Next-Generation Approach

The idea for the next-generation approach rests on the observation that $\mathscr{R}_0$ is characterized by regarding the infection transmission as producing offspring in an epidemiological sense, that is, giving birth to a new infected individual. In that sense, the infection process can corresponded to a demographic process with consecutive generations of infected individuals. If subsequent generations are growing in size, that signifies an epidemic. The growth factor per generation then gives the potential for growth. The mathematical characterization of this factor is $\mathscr{R}_0$. For compartmental models of ordinary differential equations, where the traits are taken into account in discrete categories, one can define a matrix that relates the number of newly infected individuals in the various categories in consecutive generations. This matrix is called the **next-generation matrix** and was introduced by Diekmann and Heesterbeek in 1990 [54]. The reproduction number $\mathscr{R}_0$ is then defined as the spectral radius of the next-generation matrix.

Several techniques have been developed to derive the next-generation matrix from compartmental models. We introduce these techniques below and present a number of examples to illustrate their application.

5.3.1 Van den Driessche and Watmough Approach

This method consists in a technique for the derivation of the next-generation matrix from ordinary differential equation compartmental models for disease transmission.

We will divide the compartments into two broad categories: infected compartments and noninfected (healthy) compartments. A compartment is called an **infected compartment** if the individuals in that compartment are infected. Compartments where individuals are infected but not infectious (such as latent individuals) are also among the infected compartments. The remaining compartments in which the individuals are not infected are the noninfected compartments. Assume that there are n infected compartments and m noninfected compartments, so the entire ordinary differential equation model has $m+n$ dependent variables. Let x be the vector of dependent variables in the infected compartments, and let y be the vector of variables in the noninfected compartments. We have $x \in R^n$ and $y \in R^m$. The method below was introduced in [159]. We then proceed with the following steps:

1. First, we arrange the equations so that the first n components of the ODE system correspond to the infected compartments. Thus, we write the original ODE system as

$$\begin{aligned} x_i' &= f_i(x,y), \qquad i = 1,\dots n, \\ y_j' &= g_j(x,y), \qquad j = 1,\dots,m. \end{aligned} \tag{5.17}$$

2. Second, we split the right-hand side in the infected compartments in the following way:

$$\begin{aligned} x_i' &= \mathscr{F}_i(x,y) - \mathscr{V}_i(x,y), \qquad i = 1,\dots,n, \\ y_j' &= g_j(x,y), \qquad\qquad\qquad\quad j = 1,\dots,m, \end{aligned} \tag{5.18}$$

where

- $\mathscr{F}_i(x,y)$ is the rate of appearance of **new** infections in compartment i;
- $\mathscr{V}_i(x,y)$ incorporates the remaining transitional terms, namely births, deaths, disease progression, and recovery.

We note that this decomposition in infected and noninfected compartments as well as the decomposition into $\mathscr{F}$ and $\mathscr{V}$ may not be unique. Different decompositions may correspond to different interpretations of the disease process and may lead to somewhat different expressions of the reproduction number. The decomposition should satisfy the following *properties*:

- $\mathscr{F}_i(0,y) = 0$ and $\mathscr{V}_i(0,y) = 0$ for $y \geq 0$ and $i = 1,\dots,n$. The first condition says that all new infections are secondary infections arising from infected hosts. The second condition says that there is no immigration of susceptible individuals into the disease compartments.
- $\mathscr{F}_i(x,y) \geq 0$ for all $x,y \geq 0$.
- $\mathscr{V}_i(x,y) \leq 0$ whenever $x_i = 0$, for $i = 1,\dots,n$. Each component $\mathscr{V}_i$ represents the net outflow of a compartment and must give inflow only (that is, be negative) if the compartment is empty.
- $\sum_{i=1}^{n} \mathscr{V}_i(x,y) \geq 0$ for all $x,y \geq 0$. The total outflow of all infected compartments is positive.

3. Assume that the disease-free system

$$y' = g(0, y)$$

has a unique disease-free equilibrium $\mathscr{E}_0 = (0, y_0)$ such that all solutions with initial conditions of the form $(0, y)$ approach $(0, y_0)$ as $t \to \infty$. Determine the disease-free equilibrium $\mathscr{E}_0$.

4. Determine the matrices F and V with components

$$F = \left[\frac{\partial \mathscr{F}_i(0, y_0)}{\partial x_j}\right] \qquad \text{and} \qquad V = \left[\frac{\partial \mathscr{V}_i(0, y_0)}{\partial x_j}\right].$$

These matrices appear from the linearization of the system (5.18) around the disease-free equilibrium. It can be shown that

$$\frac{\partial \mathscr{F}_i(0, y_0)}{\partial y_j} = \frac{\partial \mathscr{V}_i(0, y_0)}{\partial y_j} = 0$$

for every pair (i, j). This implies that the linearized equations for the infected compartments x while computed at the disease-free equilibrium are decoupled from the remaining equations. The linearized system for the infected compartments can be written as

$$x_j' = (F - V)x,$$

where the F and V matrices are defined above.

5. The next-generation matrix is defined as

$$K = FV^{-1}$$

and

$$\mathscr{R}_0 = \rho(FV^{-1}),$$

where $\rho(A)$ denotes the spectral radius of A.

Definition 5.1. The *spectral radius* of a matrix A is defined as the maximum of the absolute values of the eigenvalues of A:

$$\rho(A) = \sup\{|\lambda| : \lambda \in \sigma(A)\},$$

where $\sigma(A)$ denotes the set of eigenvalues of A.

It can be shown that V is a nonsingular M-matrix.

Definition 5.2. A matrix A is called an *M-matrix* if:

- A has the Z-pattern, that is, the off-diagonal elements of A are nonpositive.
- The inverse of A exists and has nonnegative elements: $A^{-1} \geq 0$.

Since V is an M-matrix, $V^{-1} \geq 0$, that is, V^{-1} has only nonnegative entries. Since F also has only nonnegative entries, the next-generation matrix $K = FV^{-1}$ is also

nonnegative. That implies that the next-generation matrix has its spectral radius as an eigenvalue (by the Perron–Frobenius theorem), and there are no other eigenvalues with larger modulus. This largest positive eigenvalue gives $\mathscr{R}_0$. Hence,

> The reproduction number $\mathscr{R}_0$ is computed as the largest positive eigenvalue of the next-generation matrix.

Furthermore, it can be shown that the reproduction number thus defined has the usual mathematical properties expected from a quantity called the reproduction number, that is, if $\mathscr{R}_0 < 1$, the disease-free equilibrium is locally asymptotically stable; otherwise, it is unstable. The disease-free equilibrium is locally asymptotically stable if all eigenvalues of the matrix $F - V$ have negative real part, that is, the **spectral bound** of the matrix $F - V$ is negative.

Definition 5.3. The *spectral bound* of a matrix A is given by the maximum real part of all eigenvalues, that is,

$$m(A) = \sup\{\mathrm{Re}\,\lambda : \lambda \in \sigma(A)\}.$$

The correspondence between the spectral bound of the linearized matrix $F - V$ and the spectral radius of the next-generation matrix FV^{-1} is given in the following theorem:

Theorem 5.2. *We have the following equivalent statements:*

- $m(F-V) < 0$ *if and only if* $\rho(FV^{-1}) < 1$;
- $m(F-V) > 0$ *if and only if* $\rho(FV^{-1}) > 1$.

Depending on the interpretation of the disease processes in the model, the next-generation approach can result in a number of decompositions into matrices F and V, and consequently, the next-generation matrix is usually not unique. Alternative next-generation matrices result in different expressions for the reproduction number $\mathscr{R}_0$. Often, the expressions for the reproduction number obtained from the next-generation approach are different from the expressions obtained from the Jacobian approach. All these different definitions of $\mathscr{R}_0$ lead to different values for the reproduction number, although all values of $\mathscr{R}_0$ are always simultaneously greater than one or smaller than one. The advantage of the next-generation approach is that it can lead to a threshold quantity, a reproduction number, even when the Jacobian approach fails. For instance, the next-generation approach will give an $\mathscr{R}_0$ for model (5.8). The disadvantage of the next-generation approach is that the reproduction number obtained through this technique is often hard to interpret as the number of secondary cases generated by one infectious individual.

5.3.2 Examples

In this subsection, we discuss a number of examples that showcase the strengths and weaknesses of the next-generation approach.

5.3.2.1 A Model of a Vector-Borne Disease

The first example that we consider here is the model of a vector-borne disease with waning immunity that we first introduced at the end of Chap. 4. We include the model here for reference. Let S_v and I_v be respectively the susceptible and infected vectors. Furthermore, if I is the number of infected humans, the equations for the vector dynamics are given by

$$\begin{aligned} S_v' &= \Lambda_v - paS_vI - \mu_vS_v, \\ I_v' &= paS_vI - \mu_vI_v, \end{aligned} \tag{5.19}$$

where p is the probability of transmission given a bite of a susceptible vector on an infectious human, and a is the vector biting rate. Furthermore, μ_v denotes the vector death rate. The total population size of the vector satisfies

$$N_v' = \Lambda_v - \mu_vN_v,$$

and it is asymptotically constant. Let $\lim_t N_v(t) = N_v$. The model for the host is an SIRS epidemic model, where S are the susceptible humans, I the infected humans, and R the recovered humans:

$$\begin{aligned} S' &= \Lambda - qaSI_v - \mu S + \gamma R, \\ I' &= qaSI_v - (\mu + \alpha)I, \\ R' &= \alpha I - (\mu + \gamma)R, \end{aligned} \tag{5.20}$$

where μ is the host natural death rate, Λ is the host birth rate, α is the recovery rate, and γ is the rate of loss of immunity. If $\gamma = 0$, then the model is one with permanent immunity. The parameter q is the probability of transmission given a bite of an infected vector on a susceptible human. The total host population size satisfies the equation

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

and it is asymptotically constant; that is, $\lim_{t\to\infty} N(t) = N$. The reproduction number that we computed for this model using the Jacobian approach is given by

$$\mathscr{R}_0 = \frac{paN_vqaN}{\mu_v(\mu + \alpha)}.$$

This reproduction number has a clear interpretation as the number of secondary cases one infectious host individual will produce in an entirely susceptible host population. To see this, notice that

- paS_v is the number of secondary infections of vectors by one infected host individual per unit of time.
- paN_v is the number of secondary infections of vectors by one infected host individual per unit of time in an entirely susceptible vector population.
- $1/(\mu+\alpha)$ is the lifespan of an infected host individual. Hence,

$$\mathscr{R}_H = \frac{paN_v}{(\alpha+\mu)}$$

  is the number of secondary infections of vectors one infected host individual will produce in an entirely susceptible vector population during its lifespan as infectious.
- Similar reasoning gives that

$$\mathscr{R}_v = \frac{qaN}{\mu_v}$$

  is the number of secondary infections of hosts one infected vector will produce in an entirely susceptible host population during its lifespan as infectious.
- The product $\mathscr{R}_H\mathscr{R}_v$ gives the number of secondary infections one infective host will produce in an entirely susceptible host population during its lifespan as infectious, that is, it gives $\mathscr{R}_0$.

We see that $\mathscr{R}_0$ defined as the product $\mathscr{R}_H\mathscr{R}_v$ has a clear interpretation as the number of secondary infections. Now we apply the next-generation approach to compute the reproduction number. We recall that system (5.19)–(5.20) has a unique disease-free equilibrium: $\mathscr{E}_0 = (N_v, 0, N, 0, 0)$, where N_v and N are the susceptible vector and host populations in the absence of infection. The infected compartments are I_v and I, ordered (I_v, I). The nonlinear terms with new infection $\mathscr{F}$ and the outflow term $\mathscr{V}$ are given by

$$\mathscr{F} = \begin{pmatrix} paS_vI \\ qsSI_v \end{pmatrix} \quad \text{and} \quad \mathscr{V} = \begin{pmatrix} \mu_vI_v \\ (\alpha+\mu)I \end{pmatrix}. \tag{5.21}$$

This gives the following linearized matrices F and V, computed at the disease-free equilibrium:

$$F = \begin{pmatrix} 0 & paN_v \\ qsN & 0 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} \mu_v & 0 \\ 0 & (\alpha+\mu) \end{pmatrix}. \tag{5.22}$$

Since V is diagonal, inverting V is straightforward:

$$V^{-1} = \begin{pmatrix} \dfrac{1}{\mu_v} & 0 \\ 0 & \dfrac{1}{(\alpha+\mu)} \end{pmatrix}. \tag{5.23}$$

Hence, the next-generation matrix is

$$K = FV^{-1} = \begin{pmatrix} 0 & \dfrac{paN_v}{\alpha+\mu} \\ \dfrac{qaN}{\mu_v} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathscr{R}_H \\ \mathscr{R}_v & 0 \end{pmatrix}. \tag{5.24}$$

We see that the entries of the next-generation matrix are $\mathscr{R}_H$, the number of secondary infections that one infected host produces in an entirely susceptible vector population during its lifespan as infective, and $\mathscr{R}_v$, the number of secondary infections that one infected vector produces in an entirely susceptible host population during its lifespan as infective. The reproduction number of the vector–host system is then given by the principal eigenvalue of the next-generation matrix K. Considering $|K - \lambda I| = 0$ gives the following equation for the eigenvalues of K:

$$\lambda^2 - \mathscr{R}_H \mathscr{R}_v = 0.$$

Hence, the next-generation reproduction number is given by

$$\mathscr{R}_0^{NG} = \sqrt{\mathscr{R}_H \mathscr{R}_v}.$$

We see that the reproduction number obtained via the Jacobian approach $\mathscr{R}_0$ is the square of the reproduction number obtained via the next-generation approach:

$$\mathscr{R}_0 = (\mathscr{R}_0^{NG})^2.$$

Only $\mathscr{R}_0$ gives the number of secondary infections that one infective host individual will produce in an entirely susceptible host population during its lifespan as infective. The concept of the next-generation approach is different. It defines the reproduction number as the number of secondary infections generated per stage. Because the vector-borne disease transmission processes involves two stages, host$\rightarrow$ vector transmission and vector$\rightarrow$ host transmission, the reproduction number should reflect the average number of secondary infections for each of the transmission stages, whether the initial infective individual is a human or a vector. Thus, the geometric mean of the two secondary infection numbers is used for $\mathscr{R}_0^{NG}$, and the total reproduction number is used for the two stages $\mathscr{R}_0$. Note that $\mathscr{R}_0$ is the square of the *per stage* reproduction number $\mathscr{R}_0^{NG}$. We note that we have

$$\begin{aligned} \mathscr{R}_0 > 1 \quad &\text{iff} \quad \mathscr{R}_0^{NG} > 1, \\ \mathscr{R}_0 = 1 \quad &\text{iff} \quad \mathscr{R}_0^{NG} = 1, \\ \mathscr{R}_0 < 1 \quad &\text{iff} \quad \mathscr{R}_0^{NG} < 1. \end{aligned}$$

Because the quantity obtained as a threshold from the next-generation approach is hard to interpret as the number of secondary cases, some researchers believe that it should not be called a reproduction number.

5.3.2.2 A Model with Treatment and Relapse

In this section, we consider the SEI model with treatment and relapse that we introduced earlier through system (5.6). The model illustrates how one can obtain erroneous results if the processes that lead to new infections are not properly identified. The ambiguity in that this particular model comes from the fact that individuals in the infectious compartment get treated, and a proportion q move to the treated/recovered class, while a proportion p do not complete treatment and relapse to the latent/exposed class. This causes an inflow in the compartment of latent individuals that may be interpreted as new infections. Thus, we have two options: (1) View the relapse term pr_2I as new infections. (2) View the relapse term as existing infection. We apply the next-generation approach under these two scenarios and obtain two different reproduction numbers. It is customary in the mathematical epidemiology literature to denote the reproduction number in the presence of a control strategy, such as treatment, by $\mathscr{R}_c$ rather than $\mathscr{R}_0$. The infected compartments are E and I. The disease-free equilibrium is given by $(\frac{\Lambda}{\mu},0,0,0)$, where the variables are ordered as (S,E,I,T).

View Relapse Term pr_2I as New Infections

Under this scenario, the right-hand side in infection compartments E and I gives the following $\mathscr{F}$ and $\mathscr{V}$, where the relapse pr_2I is viewed as new infections:

$$\mathscr{F} = \begin{pmatrix} \beta_1 SI/N + \beta_2 TI/N + pr_2 I \\ 0 \end{pmatrix} \qquad \text{and} \qquad \mathscr{V} = \begin{pmatrix} (\mu+\kappa+r_1)E \\ -\kappa E + (r_2+\mu)I \end{pmatrix}.$$

Evaluating the derivatives of $\mathscr{F}$ and $\mathscr{V}$ at the disease-free equilibrium leads to the following matrices F and V:

$$F = \begin{pmatrix} 0 & \beta_1 + pr_2 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad V = \begin{pmatrix} (\mu+\kappa+r_1) & 0 \\ -\kappa & (r_2+\mu) \end{pmatrix}. \tag{5.25}$$

One can use a computer algebra system to invert V:

$$V^{-1} = \begin{pmatrix} \dfrac{1}{\mu+\kappa+r_1} & 0 \\ \dfrac{\kappa}{(\mu+\kappa+r_1)(\mu+r_2)} & \dfrac{1}{\mu+r_2} \end{pmatrix}.$$

We note that the matrix V is an M-matrix. It follows from the general theory that the inverse of V should always have nonnegative elements. The next-generation matrix is then given by

$$K = FV^{-1} = \begin{pmatrix} \dfrac{\kappa\beta_1 + \kappa pr_2}{(\mu+\kappa+r_1)(\mu+r_2)} & \dfrac{\beta_1 + pr_2}{\mu+r_2} \\ 0 & 0 \end{pmatrix}.$$

One of the eigenvalues of this matrix is 0. The other one, which gives the reproduction number, is

$$\mathscr{R}_c^{NG_1} = \frac{\kappa\beta_1 + \kappa p r_2}{(\mu+\kappa+r_1)(\mu+r_2)}.$$

One can interpret this reproduction number as a sum of two quantities. The first one,

$$\frac{\kappa\beta_1}{(\mu+\kappa+r_1)(\mu+r_2)}, \tag{5.26}$$

gives the number of secondary infections one infective individual will produce in an entirely susceptible population during its lifespan. Indeed, $\beta_1 S/N + \beta_2 T/N$ is the number of secondary infections that one infectious individual will produce in a unit of time. If the population is entirely susceptible, then $S/N = 1$, while $T/N = 0$. The lifespan of an infectious individual is $1/(\mu+r_2)$. Hence,

$$\frac{\beta_1}{(\mu+r_2)}$$

is the number of secondary infections produced by one infectious individual in an entirely susceptible population during its lifespan. However, only a fraction $\kappa/(\mu+\kappa+r_1)$ survives the exposed period and moves to the infectious stage. Therefore, the number of secondary *infectious* individuals produced by one infectious individual in an entirely susceptible population during its lifespan is given by (5.26). The fraction

$$\frac{\kappa p r_2}{(\mu+\kappa+r_1)(\mu+r_2)}$$

gives the fraction of infected individuals who relapse, survive the exposed period, and become infectious again. This fraction does *not* give *new infections*. It is somewhat difficult to argue that if we turn off the transmission process, that is, if we assume $\beta_1 = \beta_2 = 0$, then the reproduction number may still be positive, even though in this case, the reproduction number is less than 1, and the disease will eventually die out.

View Relapse Term pr_2I as Existing Infections

Under this scenario, the right-hand side in infection compartments E and I gives the following $\mathscr{F}$ and $\mathscr{V}$, where the relapse pr_2I is viewed as existing infections and becomes a part of $\mathscr{V}$:

$$\mathscr{F} = \begin{pmatrix} \beta_1 SI/N + \beta_2 TI/N \\ 0 \end{pmatrix} \qquad \text{and} \qquad \mathscr{V} = \begin{pmatrix} (\mu+\kappa+r_1)E - pr_2 I \\ -\kappa E + (r_2+\mu)I \end{pmatrix}.$$

Evaluating the derivatives of $\mathscr{F}$ and $\mathscr{V}$ at the disease-free equilibrium leads to the following matrices F and V:

$$F = \begin{pmatrix} 0 & \beta_1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} (\mu+\kappa+r_1) & -pr_2 \\ -\kappa & (r_2+\mu) \end{pmatrix}. \tag{5.27}$$

Using a computer algebra system to invert V yields

$$V^{-1} = \begin{pmatrix} \dfrac{\mu+r_2}{(\mu+\kappa+r_1)(\mu+r_2)-\kappa pr_2} & \dfrac{pr_2}{(\mu+\kappa+r_1)(\mu+r_2)-\kappa pr_2} \\ \dfrac{\kappa}{(\mu+\kappa+r_1)(\mu+r_2)-\kappa pr_2} & \dfrac{\mu+\kappa+r_1}{(\mu+\kappa+r_1)(\mu+r_2)-\kappa pr_2} \end{pmatrix}.$$

We note that again, the matrix V is an M-matrix. The next-generation matrix is given by

$$K = FV^{-1} = \begin{pmatrix} \dfrac{\kappa\beta_1}{(\mu+\kappa+r_1)(\mu+r_2)-\kappa pr_2} & \dfrac{\beta_1(\mu+\kappa+r_1)}{(\mu+\kappa+r_1)(\mu+r_2)-\kappa pr_2} \\ 0 & 0 \end{pmatrix}.$$

One of the eigenvalues of this matrix is 0. The other one gives the reproduction number in this case:

$$\mathscr{R}_c^{NG_2} = \frac{\kappa\beta_1}{(\mu+\kappa+r_1)(\mu+r_2)-\kappa pr_2}.$$

Although the denominator involves a term of negative sign, it is not hard to see that the denominator is positive, since $p < 1$. So the reproduction number this defined is always nonnegative. The challenge here lies in the epidemiological interpretation of $\mathscr{R}_c^{NG_2}$. To interpret this reproduction number, we have to rewrite it in appropriate form. We introduce the following proportions:

$$h_1 = \frac{\kappa}{\mu+\kappa+r_2} \quad \text{and} \quad h_2 = \frac{pr_2}{\mu+r_2}. \tag{5.28}$$

The fraction h_1 gives the fraction of exposed individuals who survive the exposed state and become infectious, and h_2 is the fraction of infected individuals who relapse to the exposed stage. Furthermore, we notice that we can rewrite the current reproduction number as

$$\begin{aligned} \mathscr{R}_c^{NG_2} &= \frac{\beta_1\kappa}{(\mu+\kappa+r_1)(\mu+r_2)}\left[\frac{1}{1-h_1h_2}\right] \\ &= \frac{\beta_1\kappa}{(\mu+\kappa+r_1)(\mu+r_2)}(1+h_1h_2+h_1^2h_2^2+\ldots). \end{aligned} \tag{5.29}$$

Hence, the reproduction number can be written as an infinite sum of reproduction numbers:

$$\mathscr{R}_c^{NG_2} = \sum_{n=0}^{\infty} \mathscr{R}_c^n = \sum_{n=0}^{\infty} \frac{\beta_1 h_1}{\mu + r_2} h_1^n h_2^n,$$

where

$$\mathscr{R}_c^n = \frac{\beta_1 h_1}{\mu + r_2} h_1^n h_2^n$$

is the reproduction number of an individual who relapses *exactly* n times and goes through the infectious stage exactly $n+1$ times. Indeed, for

$$\mathscr{R}_c^0 = \frac{\beta_1 h_1}{\mu + r_2},$$

β_1 is the number of secondary infections one infected individual will produce in an entirely susceptible population per unit of time. An individual in the class I is infectious for $1/(\mu + r_2)$ units of time. Of those who become infected and progress to the exposed stage, only a fraction $\kappa/(\mu + \kappa + r_1)$ survive the exposed class and become infectious. Furthermore, in

$$\mathscr{R}_c^1 = \frac{\beta_1 h_1}{\mu + r_2} h_1 h_2,$$

the number $\frac{\beta_1 h_1}{\mu + r_2}$ gives the number of secondary infectious individuals that one infectious individual will produce. A fraction h_2 of them will survive the infectious period and relapse to the exposed class, and a fraction h_1 of those will survive the exposed class and become infectious again. Thus, $\mathscr{R}_c^n$ gives the number of secondary infections that one infected individual who will relapse exactly n times will produce in an entirely susceptible population.

5.3.2.3 A Multihost Model

In this subsection, we consider the multihost model introduced in (5.8). We saw that the Jacobian approach fails to produce a unique threshold condition that serves as a necessary and sufficient condition for the stability of the disease-free equilibrium. Here, we show that the next-generation approach leads to a reproduction number derived as the principal eigenvalue of the next-generation matrix. The general theory implies that this reproduction number serves as the usual necessary and sufficient threshold condition for the stability of the disease-free equilibrium.

To apply the next-generation approach, we consider system (5.8). The infected classes in this model are (I_w, I_d), ordered as shown. The vector of new infections $\mathscr{F}$ and outflow vector $\mathscr{V}$ are given by

$$\mathscr{F} = \begin{pmatrix} \beta_{11} S_w I_w + \beta_{12} S_w I_d \\ \beta_{21} S_d I_w + \beta_{12} S_d I_d \end{pmatrix} \qquad \text{and} \qquad \mathscr{V} = \begin{pmatrix} (\mu_w + \alpha_w) I_w \\ (\mu_d + \alpha_d) I_d \end{pmatrix}.$$

The disease-free equilibrium in this model is given by $S_w^* = \Lambda_w/\mu_w$, $I_w^* = 0$, $S_d^* = \Lambda_d/\mu_d$, and $I_d^* = 0$. We take the derivatives of $\mathscr{F}$ and $\mathscr{V}$ and evaluate at the disease-free equilibrium to obtain matrices F and V:

$$F = \begin{pmatrix} \beta_{11}S_w^* & \beta_{12}S_w^* \\ \beta_{21}S_d^* & \beta_{22}S_d^* \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} (\mu_w + \alpha_w) & 0 \\ 0 & (\mu_d + \alpha_d) \end{pmatrix}.$$

Inverting V, we obtain the following next-generation matrix:

$$K = \begin{pmatrix} \dfrac{\beta_{11}S_w^*}{\mu_w + \alpha_w} & \dfrac{\beta_{12}S_w^*}{\mu_d + \alpha_d} \\ \dfrac{\beta_{21}S_d^*}{\mu_w + \alpha_w} & \dfrac{\beta_{22}S_d^*}{\mu_d + \alpha_d} \end{pmatrix} = \begin{pmatrix} \mathscr{R}_{11} & \mathscr{R}_{12} \\ \mathscr{R}_{21} & \mathscr{R}_{22} \end{pmatrix}.$$

We see that the next-generation matrix is a full-rank matrix, that is, a matrix of rank two. Consequently, it will have two nonzero eigenvalues. Furthermore, the next-generation matrix is a matrix of reproduction numbers. In particular, $\frac{\beta_{11}S_w^*}{\mu_w+\alpha_w}$ gives the number of secondary infections that one infected wild bird will produce in an entirely susceptible population of wild birds during its lifespan as infectious; $\frac{\beta_{12}S_w^*}{\mu_d+\alpha_d}$ gives the number of secondary infections that one infected domestic bird will produce in an entirely susceptible population of wild birds during its lifespan as infectious; $\frac{\beta_{21}S_d^*}{\mu_w+\alpha_w}$ gives the number of secondary infections that one infected wild bird will produce in an entirely susceptible population of domestic birds during its lifespan as infectious; $\frac{\beta_{22}S_d^*}{\mu_d+\alpha_d}$ gives the number of secondary infections that one infected domestic bird will produce in an entirely susceptible population of domestic birds during its lifespan as infectious. We denote these reproduction numbers by $\mathscr{R}_{11}, \mathscr{R}_{12}, \mathscr{R}_{21}$, and $\mathscr{R}_{22}$. The reproduction number of the whole wild bird–domestic bird system is given by the principal eigenvalue of the next-generation matrix K. The reproduction number $\mathscr{R}_0$ is the larger root of the equation

$$(\mathscr{R}_{11} - \lambda)(\mathscr{R}_{22} - \lambda) - \mathscr{R}_{12}\mathscr{R}_{21} = 0.$$

Hence, the reproduction number is given by

$$\mathscr{R}_0 = \frac{\mathscr{R}_{11} + \mathscr{R}_{22} + \sqrt{(\mathscr{R}_{11} + \mathscr{R}_{22})^2 - 4(\mathscr{R}_{11}\mathscr{R}_{22} - \mathscr{R}_{12}\mathscr{R}_{21})}}{2}.$$

This reproduction number derived by the next-generation approach serves as a threshold condition for the stability of the disease-free equilibrium, but it is difficult to interpret epidemiologically.

A Special Case

The problem with the interpretation of $\mathscr{R}_0$ above may be resolved by assuming that the transmission rates are separable. Suppose that the transmission rates are products of the infectivity rate of infectious individuals and the susceptibility of the susceptible individuals. In particular, if a_w and a_d are the susceptibilities of susceptible wild and domestic birds respectively, and b_w and b_d are the infectivities of the infected wild and domestic birds, then $\beta_{11} = a_w b_w$, $\beta_{12} = a_w b_d$, $\beta_{21} = a_d b_w$, and $\beta_{22} = a_d b_d$. In this case, the determinant of the next-generation matrix is zero, that is,

$$\mathscr{R}_{11}\mathscr{R}_{22} = \mathscr{R}_{12}\mathscr{R}_{21},$$

and the reproduction number is given by the sum of the reproduction number of an infected wild bird in the wild bird population and the reproduction number of an infected domestic bird in the domestic bird population:

$$\mathscr{R}_0 = \mathscr{R}_{11} + \mathscr{R}_{22}.$$

5.3.3 The Castillo-Chavez, Feng, and Huang Approach

Another variation of the next-generation approach, introduced in [36], suggests that we split the epidemic system into three groups of compartments: compartments of noninfected individuals, compartments of infected but not infectious individuals (such as exposed/latent individuals), and compartments of infectious individuals. If the epidemic system has no compartments of infected but not infectious individuals, then this approach is somewhat similar to the previous one, although it does not require us to decide which infections are new.

5.3.3.1 The Method

Let $X \in R^r$ be the vector whose components are the susceptible, recovered, and other classes of noninfected individuals. Let $Y \in R^s$ be the vector of latent and other stages that are infected but do not transmit the disease. Finally, let $Z \in R^n$ be the vector whose components are classes of infected individuals who are infectious and transmit the disease. The epidemic system splits into three groups of equations:

$$\begin{aligned} X' &= f(X,Y,Z), \\ Y' &= g(X,Y,Z), \\ Z' &= h(X,Y,Z). \end{aligned} \tag{5.30}$$

Let $V_0 = (X^*,0,0)$ be the disease-free equilibrium. We have

$$f(X^*,0,0) = g(X^*,0,0) = h(X^*,0,0) = 0.$$

We make the following assumption:

$$h(X,0,0) = 0 \quad \text{for} \quad \text{every} \quad X \geq 0.$$

The method is applied in the following steps:

- Linearize $g(X^*,Y,Z) = 0$ as a function of Y and Z; that is, neglect all higher-order terms in Y and Z. The linearized equation takes the form

$$g_1(X^*,Y,Z) = 0,$$

and it is linear in Y and Z.
- Solve $g_1(X^*,Y,Z) = 0$ in terms of Y:

$$Y = \tilde{g}(X^*,Z).$$

- Substitute Y in the function h to obtain the composite function

$$h(X^*,\tilde{g}(X^*,Z),Z).$$

- Take the Jacobian of this composite function with respect to the Z variables, and evaluate it at $Z = 0$:

$$A = D_Z h(X^*,\tilde{g}(X^*,0),0).$$

- Rewrite the matrix A in the form $A = M - D$, where $M \geq 0$, $D > 0$, and the matrix D is diagonal. If $m(A)$ is the spectral bound of A, and $\rho(A)$ the spectral radius of A, we have the following result:

Theorem 5.3. *The following are equivalent:*

1. $m(A) < 0$ *if and only if* $\rho(MD^{-1}) < 1$.
2. $m(A) > 0$ *if and only if* $\rho(MD^{-1}) > 1$.

This theorem says that $\rho(MD^{-1})$ has the threshold properties of the reproduction number.
- Define the basic reproduction number as the spectral radius of the matrix MD^{-1}:

$$\mathscr{R}_0 = \rho(MD^{-1}).$$

As a consequence of the theorem, we have that if $\mathscr{R}_0 < 1$, the disease-free equilibrium is stable, while if $\mathscr{R}_0 > 1$, it is unstable.

5.3.3.2 An Example: The Treatment and Relapse Model

We consider the treatment and relapse model (5.6). With the van den Driessche and Watmough approach, we obtained two possible reproduction numbers for that system, and we argued that the second one is more sensible. It is interesting to know what reproduction number will result from the Castillo-Chavez, Feng, Huang method. We use model (5.6) to illustrate this latter method.

First, we define the various groups of variables. In particular, we have $X = (S,T)$, $Y = (E)$, and $Z = (I)$, where X, Y, and Z have the same meaning as before. We notice that

$$h(X,0,0) = 0.$$

Furthermore, the disease-free equilibrium is given by $V_0 = (\frac{\Lambda}{\mu},0,0,0)$. As a first step, we consider the equation $g(X^*,Y,Z) = 0$. We obtain

$$\beta_1 S^* I/N - (\mu+\kappa+r_1)E + pr_2 I = 0,$$

where $N = S^* + E + I$. Multiplying both sides of this equation by N, we have

$$\beta_1 S^* I - (\mu+\kappa+r_1)E(S^*+E+I) + pr_2 I(S^*+E+I) = 0.$$

We linearize this equation by neglecting the higher-order terms. The linearized equation becomes

$$\beta_1 S^* I - (\mu+\kappa+r_1)ES^* + pr_2 IS^* = 0.$$

Solving for E gives

$$E = \frac{\beta_1 I + pr_2 I}{\mu+\kappa+r_1} =: \tilde{g}(X^*,Z).$$

The next step is replacing the expression for E in the function $h(X^*,E,I)$. We obtain

$$h(X^*,\tilde{g}(X^*,I),I) = \frac{\kappa\beta_1 I + \kappa pr_2 I}{\mu+\kappa+r_1} - (r_2+\mu)I.$$

Differentiating this function of I with respect to I and setting $I = 0$ to evaluate at the disease-free equilibrium, we have

$$\frac{\partial}{\partial I}h(X^*,\tilde{g}(X^*,0),0) = \frac{\kappa\beta_1 + \kappa pr_2}{\mu+\kappa+r_1} - (r_2+\mu).$$

Here, the matrix A consists of one element only:

$$A = \left(\frac{\kappa\beta_1 + \kappa pr_2}{\mu+\kappa+r_1} - (r_2+\mu)\right).$$

Thus,

$$M = \left(\frac{\kappa\beta_1 + \kappa pr_2}{\mu+\kappa+r_1}\right)$$

and $D = (r_2+\mu)$. Hence,

$$\mathscr{R}_0 = \rho(MD^{-1}) = \frac{\kappa\beta_1 + \kappa pr_2}{(\mu+\kappa+r_1)(r_2+\mu)}.$$

This means that we obtain the first version of the reproduction number, $\mathscr{R}_0^{NG_1}$.

Problems

5.1. Congenital Malaria
Malaria is a vector-borne disease transmitted by mosquitoes. A percentage of the pregnant women who are infected with malaria give birth to malaria-infected newborns. We consider the model of vector-borne disease with temporary immunity and incorporate into it vertical transmission. The model for the vector is as before:

$$\begin{aligned} S_v' &= \mu_v - paS_vI - \mu_v S_v, \\ I_v' &= paS_vI - \mu_v I_v, \end{aligned} \tag{5.31}$$

where the total population size of the vector is $S_v + I_v = 1$. The model for the human population incorporates the vertical transmission:

$$\begin{aligned} S' &= \mu(S + \sigma I + R) - qaSI_v - \mu S + \gamma R, \\ I' &= (1-\sigma)\mu I + qaSI_v - (\mu + \alpha)I, \\ R' &= \alpha I - (\mu + \gamma)R. \end{aligned} \tag{5.32}$$

The model for humans also assumes that the total human population is $S + I + R = 1$. The new parameter σ gives fraction of newborns that are healthy.

(a) Draw a flowchart of the model (5.31)–(5.32).
(b) Use the Jacobian approach to compute the basic reproduction number of the model (5.31)–(5.32). Attempt to interpret $\mathscr{R}_0$ epidemiologically.
(c) Use the next-generation approach to compute the basic reproduction number of the model (5.31)–(5.32). Attempt to interpret $\mathscr{R}_0$ epidemiologically.

5.2. SEIR Model with Asymptomatic Stage
Consider the SEIR model with asymptomatic stage (5.2).

(a) Use van den Driessche and Watmough next-generation approach to compute the basic reproduction number. Interpret the expression for $\mathscr{R}_0$ epidemiologically.
(b) Use the Castillo-Chavez, Feng, and Huang next-generation approach to compute the basic reproduction number. Interpret the expression for $\mathscr{R}_0$ epidemiologically.

5.3. SCIRS Model with Carrier Stage
Consider the SCIRS model given in (5.3).

(a) Use the Jacobian approach to determine the reproduction number and the stability of the disease-free equilibrium. Interpret the expression for $\mathscr{R}_0$ epidemiologically.
(b) Use van den Driessche and Watmough next-generation approach to compute the basic reproduction number. Interpret the expression for $\mathscr{R}_0$ epidemiologically, if different from the above.
(c) Use the Castillo-Chavez, Feng, and Huang next-generation approach to compute the basic reproduction number.

5.4. SIQR Model Isolation
Consider the SIQR model given in (5.5).

(a) Use the Jacobian approach to determine the reproduction number and the stability of the disease-free equilibrium. Interpret the expression for $\mathscr{R}_0$ epidemiologically.
(b) Use van den Driessche and Watmough next-generation approach to compute the basic reproduction number. Interpret the expression for $\mathscr{R}_0$ epidemiologically, if different from the above.
(c) Use the Castillo-Chavez, Feng, and Huang next-generation approach to compute the basic reproduction number.

5.5. SI Model of Avian Influenza
Consider the SI model with wild and domestic birds given in (5.8).

(a) Use van den Driessche and Watmough next-generation approach to compute the basic reproduction number. Can you interpret the expression for $\mathscr{R}_0$ epidemiologically.
(b) Use the Castillo-Chavez, Feng, and Huang next-generation approach to compute the basic reproduction number.
(c) Take $\beta_{11} = 0.0001, \beta_{22} = 0.0005, \mu_w = 0.1, \mu_d = 0.5, \alpha_w = \alpha_d = 365/10, \Lambda_w = 20000, \Lambda_d = 10000$. Plot the reproduction number computed in part (a) as a function of the cross-transmission rate β_{12} for various values of β_{21}. How does cross transmission affect the persistence of the disease?

5.6. Model with Quarantine and Isolation

(a) Based on the (5.5) model considered in this chapter, compose a model with quarantine, isolation, and exposed periods. Explain the meanings of the parameters. Draw a flowchart of the model.
(b) Use van den Driessche and Watmough next-generation approach to compute the basic reproduction number. Can you interpret the expression for $\mathscr{R}_0$ epidemiologically?
(c) Plot the reproduction number as a function of two variables: the quarantine rate and the isolation rate. What epidemiological conclusions can you draw from this plot?

5.7. HIV/AIDS Model with Treatment and Prevention
Consider a model of HIV/AIDS with treatment and prevention:

$$\begin{aligned} S' &= \Lambda - (1-p)\frac{\beta_1 SI + \beta_2 ST + \beta_3 SA}{N} - \mu S, \\ I' &= (1-p)\frac{\beta_1 SI + \beta_2 ST + \beta_3 SA}{N} - (\mu + \rho + \sigma)I, \\ T' &= \rho I - (\mu + \gamma)T, \\ A' &= \sigma I + \gamma T - (\mu + d)A, \end{aligned} \tag{5.33}$$

where S is the number of susceptible individuals, I is the number of HIV-infected but not treated individuals, T is the number of HIV-infected but treated individuals, and A is the number of individuals with AIDS. The meanings of the parameters are as follows: p is the proportion protected by condom use, β_1, β_2, and β_3 are the transmission rates of the infected, treated, and AIDS individuals respectively, ρ is the treatment rate, σ is the progression rate to AIDS without treatment, γ is the progression rate to AIDS with treatment, and d is the disease-induced death rate.

(a) Draw a flowchart of the model.
(b) Use the Jacobian approach to determine the reproduction number and the stability of the disease-free equilibrium. Interpret the expression for $\mathscr{R}_0$ epidemiologically.
(c) Use the van den Driessche and Watmough next-generation approach to compute the basic reproduction number. Interpret the expression for $\mathscr{R}_0$ epidemiologically, if different from the above.
(d) Verify that $\mathscr{R}_0$ decreases with increasing p and ρ. Compute the critical fraction of condom use p_c so that $\mathscr{R}_0(p_c) = 1$. Plot p_c as a function of σ. Interpret what you observe epidemiologically.

Chapter 6
Fitting Models to Data

6.1 Introduction

In the previous chapters, we have learned to develop both simple and complex epidemic models and to perform some partial analysis on them, such as computing the reproduction number. It has become clear that multiple models can be developed to describe a particular epidemic. Which models are good, which are bad, and how can we discriminate among them? Much of the answer to this question is in the realm of statistics, but we will introduce some basic techniques here to address such a question.

After composing a model, perhaps one of the most important steps is to compare the model with data or perform what is often referred to as *validation*. Model validation is the process of determining the degree to which a mathematical model is an accurate representation of the real-world data [133]. An excellent introduction to model validation can be found in [68]. In mathematics, validation is often not used with the majority of models analyzed, which are never connected to data. Linking our models to data is necessary, for it helps us not only to gain more confidence in the model that we have created, but also to obtain realistic estimates of the parameters.

In Chap. 2, we used data on influenza in an English boarding school to estimate the parameters, so the number of cases predicted by an SIR model compared well with the data. This example is a good illustration of how models can be connected to data, but the approach taken relies heavily on the fact that an implicit solution to the SIR model can be obtained. In Chap. 3, we fitted a number of single-equation demographic models to the world population data, but we again used the fact that the solutions to these models could be explicitly obtained. This is not the case with most models that we create or encounter. In this chapter, we approach the problem from a general perspective.

We assume that we have data in the form of a time series for one or more of the classes in the model. Data could be given on the prevalence of the disease, as was the example with influenza in the English boarding school; it may be given on the

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_6

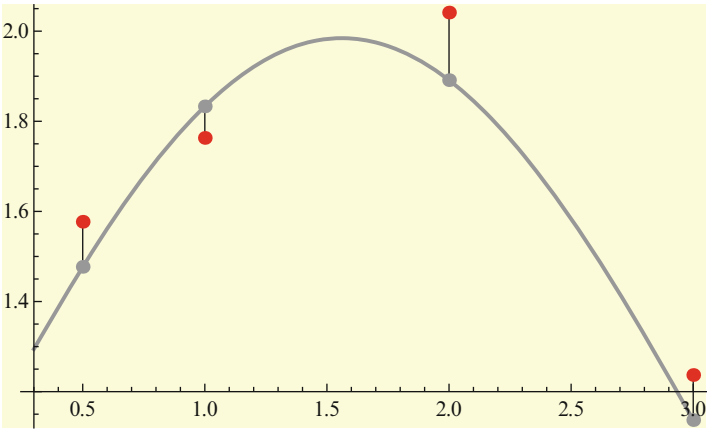

Fig. 6.1 Least-squares residuals. The *red* points are the given data points

incidence; and sometimes, it may be given on the number of recovered individuals. We recall that curve-fitting or calibration is the process of identifying the parameters of the model so that the solution best fits the data. What does it mean for a solution to best fit the data? Clearly, ideally, we would like the solution to pass through all the data points. This type of fit is called *interpolation.* However, interpolation is not always the best approach to fit real data, since the data may contain errors, and capturing every tiny change in them may be impractical. A better way to fit the solution to the data is the *least-squares approach*. In the least-squares approach, we assume that the time coordinates of the data are exact, but their y-coordinates may be noisy or distorted. We fit the solution curve through the data (see Fig. 6.1) so that the sum of the squares of the vertical distances from the data points to the point on the curve is as small as possible. In particular, suppose we are fitting the prevalence $I(t)$, and we are given the data $\{(t_1, Y_1), \ldots, (t_n, Y_n)\}$. Then we consider the *sum-of-squares error*:

$$\text{SSE} = \sum_{j=1}^{n} (Y_j - I(t_j))^2.$$

The sum-of-squares error SSE is a function of the parameters of the model. So the basic problem is to identify the parameters such that the SSE is as small as possible:

$$\text{SSE} \longrightarrow \min.$$

Minimizing the SSE is an optimization problem with its own difficulties. Differential equation epidemic models are typically nonlinear and cannot be solved explicitly. Hence, the resulting minimization problem is also highly nonlinear. As a result, in the general case, this problem is solved numerically with the use of

computer algebra systems such as Mathematica, Matlab, and R. The code requires two basic components: a differential equation solver and a minimization routine. The minimization is typically performed iteratively. The user specifies initial parameter values, and the computer solves the differential equations with those parameter values, evaluates the SSE, and improves the parameter values so that the SSE is reduced. This process is repeated a number of times until the SSE no longer becomes smaller. One important difficulty is that the minimization process is *local*, so depending on the initially specified parameter values, a minimum may occur for different sets of parameter values, and the minimal value of the SSE may be different. In practice, it may be advisable to check several sets of initial parameter values and use the smallest SSE obtained.

6.2 Fitting Epidemic Models to Data: Examples

In this section, we will consider a number of examples of fitting ODE epidemic models to data. Of course, one interesting question is, where do the data come from? There are several ways of acquiring epidemic data. First, a mathematician can work with biologists or epidemiologists who can collect the data. This is typically possible for limited datasets in limited locations. Comprehensive long-term datasets are usually collected by various health organizations such as the World Health Organization (WHO), Centers for Disease Control and Prevention (CDC), and various foundations. Because these datasets are often collected with taxpayer money, they are public and can be obtained by requesting them from the health organization or perhaps obtained online. For instance, if you go to the WHO Data and Statistics website http://www.who.int/research/en/, there are a number of important diseases listed with data and statistics about them. Suppose we are interested in the cholera epidemic that occurred in Haiti after the devastating earthquake of January 2010. The central WHO website gives only the number of yearly cases by country. If we need more resolution, if, for instance, we need monthly or weekly cases, we can google "cholera data monthly." We may find the data on the Pan American Health Organization website http://new.paho.org/hq/images/Atlas_IHR/CholeraHispaniola/atlas.html. The third possible approach is to obtain the data from published articles. Data in articles are often published as plots. Hence, if we want the actual coordinates, we need to extract them from the plots. There are many routines that can be used to extract values for the points in a plot. One is PlotDigitizer http://plotdigitizer.sourceforge.net. Matlab also has capabilities to extract data values from a plot. To use Matlab, download Matlab's `grabit.zip` from the web and unzip it to obtain the Matlab file `grabit.m`. Then follow the instructions at http://extractdata.blogspot.com/ on how to use it. This is the way we obtained the data on influenza in the English boarding school.

6.2.1 Using Matlab to Fit Data for the English Boarding School

As described in Chap. 2, in January–February 1978, an epidemic of influenza occurred in a boarding school in the north of England. The boarding school housed a total of 763 boys, who were at risk during the epidemic. On January 22, three boys were sick. The table below gives the number of boys ill on the nth day after January 22 ($n = 1$).

To fit with Matlab, we do not need to know the final size of the epidemic. Once we have the data (Table 6.1), the first question that we have to answer is, what model we should fit to the data? Since these are outbreak data, we need an epidemic

Table 6.1 Daily number influenza infected boys

Day	No. infected[a]	Day	No. infected
3	25	9	192
4	75	10	126
5	227	11	71
6	296	12	28
7	258	13	11
8	236	14	7

[a] Data taken from "Influenza in a Boarding School," *British Medical Journal*, 4 March 1978

model without demography. As we discussed in Chap. 2, the SIR model without demography is appropriate for this case. We recall the model:

$$\begin{aligned} S'(t) &= -\beta S(t)I(t), \\ I'(t) &= \beta S(t)I(t) - \alpha I(t), \end{aligned} \tag{6.1}$$

where we have omitted the recovered class.

The next question that we need to address is which model parameters we should fit and which we should pre-estimate and fix. Potentially, we can fit α, β, and the initial conditions—four parameters altogether. We can pre-estimate α from the duration of infectiousness, and the two initial conditions from the given data. For instance, we know from the data that $I(3) = 25$, and therefore $S(3) = 738$. The duration of infectiousness is 2–4 days, so we may take $\alpha = 0.3$. Even if we plan to fit all these parameters, pre-estimating what we can is useful with the initial guess of the parameters. In addition, to derive the initial guesses for the remaining parameters before using Matlab to fit, we may use Mathematica's `Manipulate` command. In Mathematica, we can fix $S(3)$, $I(3)$, and α to the above values and "manipulate" β to obtain a good agreement with the data. Say we set the value $\beta = 0.0025$. With these initial values, we may use Matlab to fit. Below is the Matlab code used for the fitting.

```
function BSFluFittingv1
% This function fits the first set of BSfludata to and ...
    SIR model


clear all
close all
clc

load BSfludat.txt  % loading data

format long        % specifying higher precision

tdata = BSfludat(:,1);  % define array with t-coordinates ...
    of the data

qdata = BSfludat(:,2);  % define array with y-coordinates ...
    of the data

tforward = 3:0.01:14;   % t mesh for the solution of the ...
    differential equation

tmeasure = [1:100:1101]'; % selects the points in the ...
    solution
                          % corresponding to the t values ...
                              of tdata


a = 0.3;
b = 0.0025;      % initial values of parameters to be fitted









    function dy = model_1(t,y,k)     % DE


        a = k(1);        % Assignes the parameters in the ...
            DE the current
                         % value of the parameters
        b = k(2);


        dy = zeros(2,1);    % assigns zeros to dy

        dy(1) = - b * y(1) * y(2);            % RHS of ...
            first equation
        dy(2) =  b * y(1) * y(2) - a * y(2);  % RHS of ...
            second equation


    end
```

```
46
47 function error_in_data = moder(k)  % computing the error ...
       in the data
48
49
50
51
52
53
54     [T Y] = ode23s(@(t,y)(model_1(t,y,k)),tforward,[738.0 ...
           25.0]);
55
56                             % solves the DE; output is ...
                                   written in T and Y
57
58
59     q = Y(tmeasure(:),2); % assignts the y-coordinates of ...
           the solution at
60                             % at the t-coordinates of tdata
61
62
63
64     error_in_data = sum((q - qdata).^2) %computes SSE
65
66 end
67
68
69
70  k =  [a b]; % main routine; assigns initial values of ...
        parameters
71
72
73  [T Y] = ode23s(@(t,y)(model_1(t,y,k)),tforward,[738.0  ...
        25.0]);
74
75               % solves the DE with the initial values of ...
                     the parameters
76
77  yint = Y(tmeasure(:),2);
78
79               % assigns the y-coordinates of the solution ...
                     at tdata to yint
80
81  figure(1)
82  subplot(1,2,1);
83  plot(tdata,qdata,'r.');
84  hold on
85  plot(tdata,yint,'b-');       % plotting of solution and ...
        data with initial
86                               % guesses for the parameters
87  xlabel('time in days');
88  ylabel('Number of cases');
89  axis([3 14 0 350]);
90
```

```
91
92
93
94  [k,fval] =  fminsearch(@moder,k); % minimization routine; ...
        assigns the new
95                                    % values of parameters ...
                                          to k and the SSE
96                                    % to fval
97
98
99   disp(k);
100
101  [T Y] = ode23s(@(t,y)(model_1(t,y,k)),tforward,[738.0  ...
        25.0]);
102                          % solving the DE with the final ...
                                 values of the
103                          % parameters
104
105  yint = Y(tmeasure(:),2); % computing the y-coordinates ...
        corresponding to the
106                              % tdata
107
108  subplot(1,2,2)
109  plot(tdata,qdata,'r.');
110  hold on
111  plot(tdata,yint,'b-');
112  xlabel('time in days');          % plotting final fit
113  ylabel('Number of cases');
114  axis([3 14 0 350]);
115
116
117 end
```

We run the Matlab code above. It tells us that the original SSE is equal to $7.2 * 10^4$. After the optimization, the newly computed parameters are $\alpha = 0.465$ and $\beta = 0.00237$. The new SSE is $4 * 10^3$. We can run the code, taking as an initial guess the parameters we computed in Chap. 2, and Matlab will improve on those, too. In general, the use of computer algebra systems such as Mathematica and Matlab is the best approach to obtain a good fit of the model solution to the data. The newly computed value of α gives the duration of the infectious period as $1/\alpha = 2.15$ days. This infectious period is meaningful, since infected students showing symptoms were quarantined. We should always ask ourselves whether the computed parameters have a sensible biological interpretation. If that is not the case, we should refit, using upper and lower bounds for the parameters.

Using Mathematica's `NonlinearModelFit` command, we fitted the model to the data and obtained the same best-fitted parameters as Matlab. One advantage of Mathematica's `NonlinearModelFit` is that it can provide many types of statistics that can help us judge the goodness of the fit. One such statistic is the residuals. The *residuals* are defined as the differences between the y-coordinates of the data points and the corresponding value of the solution. In particular,

$$\text{residuals} = \{Y_j - I(t_j) | j = 1, \ldots, n\}.$$

The best-fitted solution with Mathematica and the data are plotted in Fig. 6.2(left). The residuals are plotted in Fig. 6.2(right).

If the fit is good, the residuals should be randomly distributed. Examining the residuals in Fig. 6.2(right), we can conclude that the fit is reasonably good. Mathematica can also provide 95% confidence intervals. A *95% confidence interval (CI)* is an interval calculated from many observations, in principle different from data set to data set, that 95% of the time will include the parameter of interest if the experiment is repeated. The CI for the above fitting are $[0.4257, 0.5037]$ for α and $[0.0022099, 0.00254]$ for β.

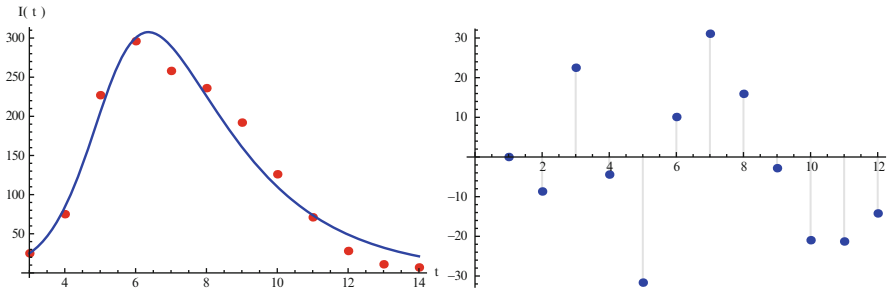

Fig. 6.2 The *left* figure shows the fit of an SIR model with the English boarding school data. The *right* figure shows the distribution of the residuals of the fit in the *left* figure. Residuals are randomly distributed. Mathematica plots residuals in time starting from $t = 1$ rather than starting from $t = 3$

6.2.2 Fitting World HIV/AIDS Prevalence

Human immunodeficiency virus (HIV) infection is a disease of the immune system caused by the HIV virus. HIV is transmitted primarily via unprotected sexual intercourse, contaminated blood transfusions, and from mother to child during pregnancy, delivery, or breastfeeding (vertical transmission). After entering the body, the virus causes acute infection, which often manifests itself with flulike symptoms. The acute infection is followed by a long asymptomatic period. As the illness progresses, it weakens the immune system more and more, making the infected individual much more likely to get other infections, called opportunistic infections, that are atypical for healthy individuals. There is no cure or vaccine against HIV; however, antiretroviral treatment can slow the course of the disease and may lead to a near-normal life expectancy.

Because people with HIV now live longer, even though the incidence of HIV is declining, the number of individuals infected with HIV or having advanced-stage AIDS is still slowly increasing worldwide. The United Nations, in their Millennium Development Goals Report 2010, gives the number of people living with HIV [122]

as well as the incidence and the number of deaths from HIV worldwide. We include the prevalence data in Table 6.2.

Our main objective is to determine a model that can be fitted to the data. The simplest HIV model is the SI epidemic model with disease-induced mortality. However, this model does not fit the data well. The main reason for that, perhaps, is the fact that a simple SI model has an exponentially distributed time spent in the infectious stage, that is, the probability of surviving in the stage declines exponentially. That is not very realistic for HIV, where the infectious stage is long and the duration is subject to significant variation. That requires that the distribution of the waiting time in the infectious class have a nonzero mode. To incorporate this effect, a typical approach is to use Erlang's "method of stages." This approach is primarily applied with stochastic HIV models, but its deterministic variant requires the infectious period to be represented as a series of k stages such that the durations of stay in each stage

Table 6.2 Prevalence (in millions) of HIV worldwide 1990–2011. 1990 gives $t = 0$

Year	Time (in years)	Prevalence	Year	Time (in years)	Prevalence
1990	0	7.3	2001	11	29.0
1991	1	9.2	2002	12	30.0
1992	2	11.3	2003	13	30.8
1993	3	13.5	2004	14	31.4
1994	4	15.9	2005	15	31.9
1995	5	18.3	2006	16	32.4
1996	6	20.6	2007	17	32.8
1997	7	22.7	2008	18	33.4
1998	8	24.6	2009	19	33.3
1999	9	26.3	2010	20	34.0[a]
2000	10	27.8	2011	21	34.0[a]

[a] Other sources

are independent identically distributed exponential variables [79]. To this end, we divide the infectious class $I(t)$ into four subclasses: $I_1(t), I_2(t), I_3(t), I_4(t)$ with an exit rate γ. Individuals in all four stages are infectious and can infect susceptible individuals $S(t)$. Denote by $I(t)$ the sum of all infectious classes:

$$I(t) = I_1(t) + I_2(t) + I_3(t) + I_4(t).$$

We further assume that the force of infection $\lambda(t)$ is nonmonotone and is given by

$$\lambda(t) = \beta e^{-\alpha I(t)/N(t)} I(t)/N(t),$$

where $N(t)$ denotes the total population size:

$$N(t) = S(t) + I_1(t) + I_2(t) + I_3(t) + I_4(t).$$

This force of infection is sensible for HIV, since as the infection spreads, it is likely that the remaining susceptible individuals become more cautious about their contacts and potential exposure to HIV, and the force of infection begins to decline.

The flowchart of the model is given in Fig. 6.3. The model becomes

$$\begin{aligned} S'(t) &= \Lambda - \lambda(t)S(t) - \mu S(t), \\ I_1'(t) &= \lambda(t)S(t) - (\gamma+\mu)I_1(t), \\ I_2'(t) &= \gamma I_1(t) - (\gamma+\mu)I_2(t), \\ I_3'(t) &= \gamma I_2(t) - (\gamma+\mu)I_3(t), \\ I_4'(t) &= \gamma I_3(t) - (\gamma+\mu)I_4(t). \end{aligned} \tag{6.2}$$

The last exit rate γ from the class I_4 is considered to be disease-induced mortality.

To fit the model to the data, we first have to decide in what units to fit. The data are given in millions, and as such, they are neither too large nor too small as numbers. If the numbers we fit are too large or too small, the round-off errors may be large, and the fit may be bad. Therefore, we need to use units that make our numbers reasonable. Furthermore, we will fit in years. After we have decided on the units, we have to decide which parameters to fit, and which to pre-estimate. This decision may have significant impact on the fit. We decide to fix Λ and μ as well as the initial values. The current natural mortality rate of humans can be taken to be $1/70$. Because the current world population size is 7 billion, that is, 7000 million, then if we take that to be the equilibrium population, we have $7000 = \Lambda/\mu$. We estimate that $\Lambda = 100$ million people per year. We further assume that in 1990, all individuals infected with HIV were actually in class I_1. Hence, $S(0) = 6992.7$ and $I_1(0) = 7.3$ million people. We set the remaining initial conditions to zero. In this fitting, we do not fit the initial conditions. We fit α, β, and γ. We fit both with Mathematica and Matlab, but this time, the best-fitted parameters are slightly different. Matlab obtains $\alpha = 260.4972$, $\beta = 0.334547$, and $\gamma = 0.339958755$. The SSE $= 0.47$ with these parameters. Mathematica's best-fitted parameters with their standard errors

Table 6.3 Mathematica's best-fitted parameters with standard errors and 95% CI

Parameter	Estimate	Standard error	95% Confidence interval
α	253.567	5.84757	[241.282,265.853]
β	0.332276	0.00298656	[0.326002,0.338551]
γ	0.349035	0.00700517	[0.334318,0.363753]

and 95% CI are given in Table 6.3. These results from Mathematica are obtained when the initial values of the parameters are the results from Matlab $\alpha = 260$, $\beta = 0.33$, and $\gamma = 0.3399$. The standard errors are small, so the parameters are well identified. The fit obtained by Mathematica is shown in Fig. 6.4 together with the residuals. The residuals do not look random, which suggests that a different model might capture the shape of the data better. Nonetheless, the residuals are small, and the fit is reasonably good. We interpret the best-fitted parameters. The only fitted

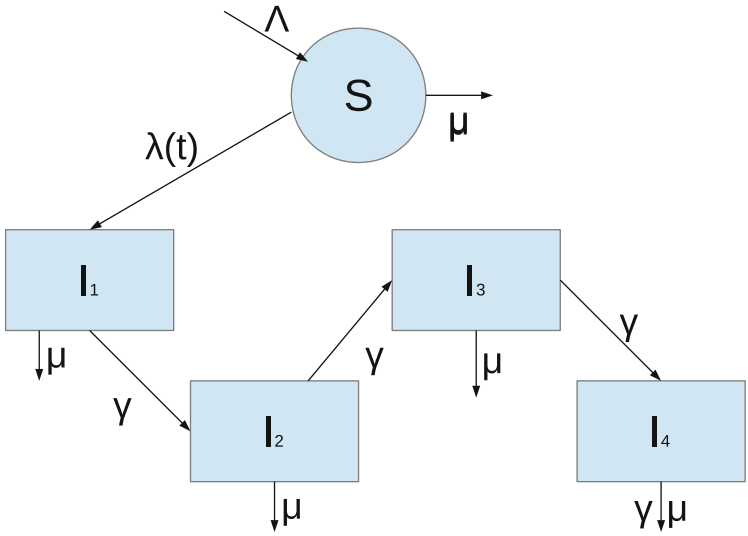

Fig. 6.3 Flowchart of the HIV model

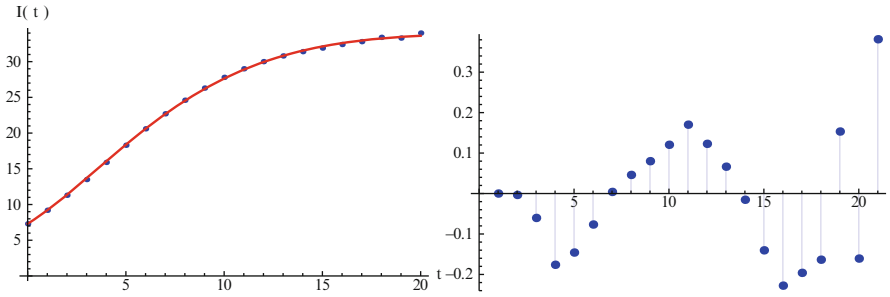

Fig. 6.4 *Left:* the solution with the best-fitted parameters alongside the data for the HIV model. *Right:* the residuals of the fit in the *left* figure

parameter that has biological meaning is γ, where $1/\gamma$ is the time spent in each of the infectious classes. Since $\gamma \approx 0.34$, it follows that $1/\gamma = 2.94$. Hence, the time spent in each class is approximately three years, which is reasonable.

6.3 Summary of Basic Steps

When you prepare to fit a mathematical model to data, think about the following basic steps in the fitting process:

1. Examine your data. Are the values involved too large or too small? If yes, determine units that allow you to work with average-size numbers.

2. Choose your model. Is your model sensible for the disease you are modeling? Should your model include demography? Decide whether your data are epidemic or endemic. What is the time span modeled?
3. Decide which model parameters to fit and which to pre-estimate and fix. Don't forget that the initial conditions for the differential equations are also in the parameter set. Never fit more parameters than the number of data points.
4. Choose initial guesses for the parameters that will be fitted. Use biological sense or prefit using Mathematica's `Manipulate`.
5. Perform the fit. Plot the solution alongside the data and examine the fit. Does the solution agree with the data? Plot the residuals. Are the residuals small and random? If they are not random, you may need a better model.
6. Determine the best-fitted parameters. Interpret them biologically. Do they make sense? If not, refit specifying upper and lower bounds for those parameters.
7. Determine the standard errors and 95% CI. Are they small? If they are not small, that may mean that some of the parameters are unidentifiable. Refit, fixing some more parameters.

There are a number of reference books and manuals that describe guidelines on fitting models to data. Further information on this topic can be found in [119, 68].

6.4 Model Selection

In mathematics, the model is typically postulated. Assuming the model, further analysis and simulations are performed with it. The model is derived from first principles, but how do we know that the model is reasonable? One way to justify our model is to confront it with data. If the model is reasonable and fits the data, then we may accept that is a reasonable model to work with. However, given a biological scenario, multiple models can be created. For instance, in modeling HIV, we can set up a regular SI model, a regular SI model with vertical transmission, an SI model with k stages of infectious individuals, where k can vary, and an SI model with k stages and vertical transmission. We would like to know which model is the best model, so that we may further work with it. If we are given data, we can confront all models with the data and decide which model best describes the data.

Definition 6.1. *Model selection* is the task of selecting a mathematical model from a set of candidate models, given data.

In model selection, we assume the data and look for the model that best describes the data. A set of candidate models has to be determined by the researcher. All candidate models must be reasonable for the epidemiological scenario being modeled. Once the set of candidate models has been selected, the mathematical analysis can be performed to choose the best model. One way that comes first to mind to compare the models is to arrange them by SSE. The model with the least SSE best fits the data. However, it is well known that the more parameters we fit, the better we

can capture the data, but the additional parameters we fit may not represent anything useful. A good selection criterion must balance two points: (1) goodness of fit; (2) simplicity (parsimony) of the model. In other words, a good selection technique must choose the simplest model that best fits the data. There are many statistical criteria that may be used to decide on the best model. Some of these are the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and cross-validation. A good book on model selection is [31]. One of the most frequently used selection criteria is the AIC [4], which we introduce here.

6.4.1 Akaike Information Criterion

The Akaike information criterion (AIC) is a criterion for model selection that compares multiple competing models, taking into account both the SSE and the number of parameters being fitted.

Definition 6.2. The *Akaike information criterion* (AIC) is a measure of the relative goodness of fit of a mathematical model.

The AIC does not tell us, however, whether a models is reasonable or whether it fits the data. Computation of the AIC is not difficult. Mathematica's `NonlinearModelFit` will compute the AIC automatically. It is computed using the number of fitted parameters (including the initial conditions if they are being fitted) and the SSE.

For a given model, the AIC is calculated as

$$\text{AIC} = n\left[\ln\left(\frac{\text{SSE}}{n}\right)\right] + 2k, \tag{6.3}$$

where n is the number of data points in the data set, k is the number of parameters fitted plus one, and SSE is the least-squares error.

Given a set of candidate models, we compute the AIC for each model and each fit. The best model is the one with the smallest AIC. The AIC is smaller if the SSE is smaller, that is, the model fits the data well, and smaller if k is smaller, that is, the number of parameters fitted is smaller. Hence, the AIC penalizes the fitting of too many parameters and discourages *overfitting*.

The AIC has its foundations in information theory. If we assume that the data are generated by some process $\mathscr{P}$ that is unknown and we have J candidate models to represent the process $\mathscr{M}_1 \ldots \mathscr{M}_J$, the AIC is a measure of the loss of information by representing the actual process $\mathscr{P}$ with the model $\mathscr{M}_j$. Hence, the model $\mathscr{M}_j$ that minimizes the information loss, that is, has the smallest AIC, should represent the unknown process. The AIC, however, is only an estimate of the information

loss, and as an estimate, it is valid only asymptotically, that is, only if the number of parameters fitted is much smaller than the number of data points in the data set. As a rule of thumb, the data set is large enough for the AIC to be used if

$$\frac{n}{K} \geq 40, \tag{6.4}$$

where K is the number of parameters in the most complex model among the competing models, and n as before is the number of data points in the data set. If (6.4) does not hold, then a modified version of the AIC, called the corrected AIC, or AICc, is recommended. The AICc was first derived by Hurvich and Tsai [77].

For a given model, the AICc is calculated as

$$\text{AICc} = n\left[\ln\left(\frac{\text{SSE}}{n}\right)\right] + 2k + \frac{2k(k+1)}{n-k-1}, \tag{6.5}$$

where n is the number of data points in the data set, k is the number of parameters fitted plus one, and SSE is the least-squares error.

Mathematica's `NonlinearModelFit` will also compute the AICc automatically. Formula (6.5) suggests that the AICc includes a greater penalty for the number of parameters fitted. Since the AICc converges to the AIC as n gets large, it is recommended that the AICc be used instead of the AIC regardless of the size of the sample [32]. Using the AIC instead of the AICc when the sample size is not much larger than the number of parameters fitted increases the risk of selecting a model with too many parameters, that is, it increases the risk of overfitting. However, the AIC or the AICc must be used consistently for all models in the model set.

The AIC and AICc are of arbitrary scale and are difficult to interpret. Therefore, different measures are considered in the decision process of which models are supported by the data and which are not. Once the model with the minimal AIC (or AICc) is selected, then we can compute the distance of the remaining models to the model with the minimal AIC (or AICc). If $\text{AIC}_{\min}$ is the AIC or AICc of the model with the minimal AIC, then

The distances of the AIC or AICc for any model to the model with minimal AIC or AICc is defined as

$$\Delta_j = \text{AIC}_j - \text{AIC}_{\min}, \tag{6.6}$$

where AIC_j is the AIC or AICc of the jth model $\mathscr{M}_j$, and $\text{AIC}_{\min}$ is the AIC or AICc of the model with minimal AIC.

The values Δ_j are important for interpreting which of the competing models in the model set have relative support in the data. A practical rule of thumb that can be applied is the following [31]:

- Models having $\Delta_j \leq 2$ have substantial support in the data.
- Models having $4 \leq \Delta_j \leq 7$ have considerably less support in the data;
- Models having $\Delta_j > 10$ have essentially no support in the data.

Another useful measure of the support a model has by the data are the *Akaike weights,* w_j. The Akaike weights [31] give the probability of a model, given the data and the models in the model set.

The Akaike weight of the ith model $\mathscr{M}_i$ is defined as follows:

$$w_i = \frac{e^{-\Delta_i/2}}{\sum_{j=1}^{J} e^{-\Delta_j/2}}. \tag{6.7}$$

We compute the Akaike weights for all models in the model set. The Akaike weights must sum to one: $\sum_{i=1}^{J} w_i = 1$. If the best-fitted model has an Akaike weight $w_i > 0.9$, then robust inferences can be made using just the best-fitted model. If none of the models has a weight $w_i > 0.9$, then several models might need to be used for multimodel inferences. See [31] for information on multimodel inference.

We note that both the Δ_j's and the Akaike weights provide support for one model over other models in the model set, but that support depends on the model set and the data at hand.

6.4.2 Example of Model Selection Using AIC

To illustrate model selection, we will use the data for the influenza epidemic in the English boarding school. We postulate six models in our candidate model set for model selection. We want to check whether our preferred fitted SIR model without demography in Sect. 6.2.1 is better or worse then other candidate models. We call the fitted SIR model without demography model number one, $\mathscr{M}_1$. Because influenza has an exposed period of several days, it is sensible to consider an SEIR model. We define the model $\mathscr{M}_2$ as (we omit the recovered class)

$$\mathscr{M}_2 : \begin{cases} S'(t) = -\beta S(t) I(t), \\ E'(t) = \beta S(t) I(t) - \alpha E(t), \\ I'(t) = \alpha E(t) - \gamma I(t). \end{cases} \tag{6.8}$$

We will fit model $\mathscr{M}_2$ two ways. First we will fix all initial conditions and fit β, α, and γ. Then, we increase the number of parameters that we are fitting and we fit not only β, α, and γ, but also the initial number of exposed individuals $E(3)$. Another common class of individuals in the spread of influenza is the class of asymptomatic individuals. Asymptomatic individuals do not get treated but spread the disease at a much lower rate. The coefficient of reduction is denoted by q. The model with asymptomatic individuals, called model $\mathscr{M}_3$, takes the form

$$\mathscr{M}_3 : \begin{cases} S'(t) = -\beta S(t)(I(t) + qA(t)), \\ E'(t) = \beta S(t)(I(t) + qA(t)) - (\alpha + k)E(t), \\ I'(t) = \alpha E(t) - \gamma I(t), \\ A'(t) = kE(t) - \nu A(t), \end{cases} \tag{6.9}$$

where k is the rate at which individuals progress to the asymptomatic stage and ν is the recovery from the asymptomatic class. In the first fitting with this model, we fit α, β, γ, q, k, and ν, that is, six parameters. In the second fitting, we fit α, β, γ, $A(3)$, q, k, and ν. We take $S(3) = 763 - 25.0 - 20.0 - A(3)$, where we have taken $E(3) = 20.0$. As a result of the fitting, we obtain that the estimate of k is very small, so $k \approx 0$. This suggests that in this case of influenza, there were no asymptomatic individuals. In general, these two fits are not good, since the standard errors and the confidence intervals are very large. We may be overfitting by fitting so many parameters (see Table 6.4). The last model that we will consider in the candidate models set includes all stages as before plus an isolated class denoted by $Q(t)$. The model takes the form

$$\mathscr{M}_4 : \begin{cases} S'(t) = -\beta S(t)(I(t) + qA(t)), \\ E'(t) = \beta S(t)(I(t) + qA(t)) - (\alpha + k)E(t), \\ I'(t) = \alpha E(t) - (\gamma + d)I(t), \\ A'(t) = kE(t) - \nu A(t), \\ Q'(t) = dI(t) - \rho Q(t), \end{cases} \tag{6.10}$$

where d is the rate of isolating the infectious individuals and ρ is the recovery rate from the isolated class. Parameters fitted in this model are α, β, γ, d, q, k, ρ and ν, a total of eight parameters.

Table 6.4 Model selection table

Model	Fitted parameters	SSE	AIC	AICc	Δ AICc	w_i
$\mathscr{M}_1$	α,β	4028.26	75.79	78.79	0	0.61
$\mathscr{M}_2$	α,β,γ	2952.62	74.07	79.78	0.99	0.37
$\mathscr{M}_2$	α,β,γ,$E(3)$	3077.81	76.56	85.56	6.77	0.021
$\mathscr{M}_3$	6 param.[a]	4077.19	83.94	111.94	33.15	$3.85*10^{-8}$
$\mathscr{M}_3$	$A(3)$ and 6 param.[a]	3083.26	82.59	130.59	51.8	$3.44*10^{-12}$
$\mathscr{M}_4$	8 param.[b]	3166.42	84.91	174.91	96.12	$8.17*10^{-22}$

[a] α, β, γ, q, k, and ν
[b] α, β, γ, d, q, k, ρ, and ν

We fitted the models using Mathematica. Mathematica computes the SSE. We use formula (6.3) to compute the AIC. Since the ratio of the number of data points to the number of parameters fitted is less than forty, we need to actually use the AICc. We use formula (6.5) to compute the AICc for each model. The results are listed in Table 6.4. Notice that the model with the lowest AIC is $\mathscr{M}_2$, with three parameters fitted, while the model with the lowest AICc is $\mathscr{M}_1$. Further, we compute ΔAICc for all models. Following the rule of thumb, we conclude that models $\mathscr{M}_1$ and $\mathscr{M}_2$ with three parameters fitted have substantial support in the data; model $\mathscr{M}_2$ with four parameters fitted has considerably less support in the data, and models $\mathscr{M}_3$ and $\mathscr{M}_4$ have no support in the data. This does not mean that models $\mathscr{M}_3$ and $\mathscr{M}_4$ are bad models but it simply means that $\mathscr{M}_3$ and $\mathscr{M}_4$ cannot be used to interpret and make inferences from this specific set of data. Looking at the Akaike weights, we see that the probability of model $\mathscr{M}_1$ is 0.61, the probability of model $\mathscr{M}_2$ with three parameters is 0.37, about twice smaller. The sum of the probabilities of these two models is over 0.9, so these are the models that must be used for inference. The remaining models have no support in the data.

6.5 Exploring Sensitivity

Estimation of parameters and initial conditions through fitting is often subject to variation. The pre-fixed parameters are selected from a range, and consequently, the fitted parameters may vary in a range. Varying the parameters varies the output of the model, but which parameters have the most significant impact on that output?

Definition 6.3. The goal of *sensitivity analysis* is to decide qualitatively which parameters are most influential in the model output.

Sensitivity analysis can be performed on a dynamical system or on static quantities such as the reproduction number or equilibrial prevalence.

6.5.1 Sensitivity Analysis of a Dynamical System

To perform sensitivity analysis of a dynamical system, we assume that the differential equations depend on a parameter p:

$$y_i'(t) = f_i(y_1, \ldots, y_n, t, p) \qquad i = 1, \ldots, n.$$

The parameter may be one of the coefficients in the system or one of the initial conditions. The solution of the initial value problem can be thought of as a function of both the time variable t and the parameter p: $y_i(t, p)$, $i = 1, \ldots, n$.

Definition 6.4. A parameter is called *sensitive* if small changes in the value of the parameter produce large changes in the solution of the differential equations.

To consider the change in the solution with respect to the parameter, we look at the derivative of the solution with respect to the parameter. Thus, we introduce the new variables

$$Z_i = \frac{\partial y_i}{\partial p} \qquad i = 1, \ldots, n.$$

The differential equations for the variables Z_i are obtained as follows:

$$Z_i' = \frac{\partial Z_i}{\partial t} = \frac{\partial}{\partial t}\left(\frac{\partial y_i}{\partial p}\right) = \frac{\partial}{\partial p}\left(\frac{\partial y_i}{\partial t}\right).$$

Hence, we have

$$Z_i' = \frac{\partial}{\partial p} f_i(y_1(t,p), \ldots, y_n(t,p), t, p).$$

Applying the chain rule, we have

$$Z_i' = \frac{\partial f_i}{\partial p} + \sum_{j=1}^{n} \frac{\partial f_i}{\partial y_j}\frac{\partial y_j}{\partial p}.$$

Therefore, the differential equations for the Z_i's become

$$Z_i' = \frac{\partial f_i}{\partial p} + \sum_{j=1}^{n} \frac{\partial f_i}{\partial y_j} Z_j. \tag{6.11}$$

To complete the system for the variables Z_i, we have to derive the initial conditions. The quantity $Z_i(0)$ is the following limit:

$$Z_i(0) = \lim_{\Delta p \to 0} \frac{y_i(0, p+\Delta p) - y_i(0,p)}{\Delta p}.$$

There are two distinct cases:

Case 1 The parameter p is not an initial condition. Then $y_i(0, p+\Delta p) - y_i(0,p) = 0$ for all i. Hence, $Z_i(0) = 0$ for $i = 1, \ldots, n$.

Case 2 The parameter p is an initial condition for $y_k(0) = p$. In this case, as in Case 1, we can derive that $Z_i(0) = 0$ for all $i \neq k$. For $i = k$, $y_k(0, p+\Delta p) - y_k(0,p) = \Delta p$. Hence the limit is equal to 1. Therefore, $Z_k(0) = 1$.

To determine the sensitivity of the solution with respect to parameter p, we solve the system

$$\begin{aligned}
&y_i' = f_i(y, t, p), \qquad i = 1, \ldots, n,\\
&y_i(0) = y_0^i,\\
&Z_i' = \frac{\partial f_i}{\partial p} + \sum_{j=1}^{n} \frac{\partial f_i}{\partial y_j} Z_j, \qquad i = 1, \ldots, n,\\
&Z_i(0) = 0,\\
&Z_k(0) = 1, \quad \text{if} \quad y_k(0) = p.
\end{aligned} \tag{6.12}$$

In the approach above, we vary only one parameter. This is the case when we perform local sensitivity analysis.

Definition 6.5. *Local sensitivity analysis* is a sensitivity analysis that examines the change in the output values that result from change in **one** input value (parameter). *Global sensitivity analysis* examines the change in the output values that result from changes in all parameter values over the parameters' ranges.

To perform global sensitivity analysis we have to expand the system above to include all parameter values. This is clearly a very computationally intensive problem, particularly for large models. A good review of local and global sensitivity analysis for differential equations is given in [53].

Instead of solving the differential equation system (6.12), a common approach in practice to local sensitivity is to fix all parameters except p and solve the differential equation model for several values of the parameter p. We took this approach ininvestigating the sensitivity with respect to α and β for the best-fitted model, the

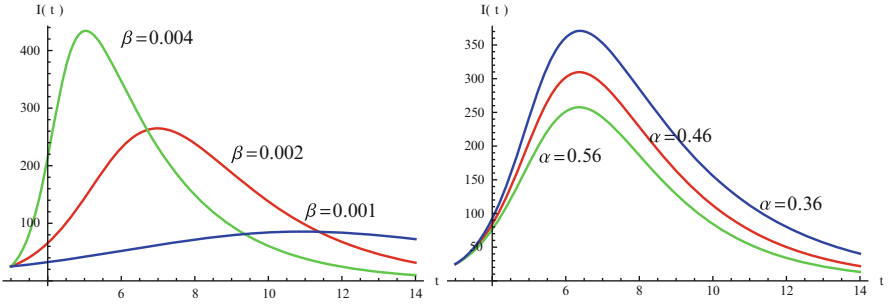

Fig. 6.5 *Left:* sensitivity of the solution of the SIR model with respect to the transmission coefficient β. *Right:* sensitivity of the solutions of the SIR model with respect to the recovery rate α

SIR model, for the case of influenza in the English boarding school. The best-fitted parameters were $\alpha \approx 0.46$ and $\beta \approx 0.00237$. We fix α at 0.46 and we vary β with values $\beta = 0.001, 0.002, 0.004$. The corresponding graphs are plotted in Fig. 6.5.

Figure 6.5 shows that the prevalence is quite sensitive to the variation in the transmission rate β. As β decreases, the peak becomes less pronounced and moves to the right. For β small enough, there will be no peak, and the solution will decrease directly to zero. To obtain the right plot in Fig. 6.5, we fix β at 0.00237 and vary α for values $\alpha = 0.36, 0.46, 0.56$. The figure shows that the prevalence is also sensitive to α. When α increases, the peak becomes less pronounced, and the width of the graph becomes smaller, that is, the epidemic reaches a lower maximum and disappears faster.

6.5.2 Sensitivity and Elasticity of Static Quantities

In epidemiology, important static quantities depend on the parameters of the differential equation model. Such quantities are the reproduction number, the equilibrial prevalence, and the equilibrial incidence. Once we estimate the parameters from the fitting, we can compute the values of these important static quantities. More importantly, we often would like to know how these quantities respond to changes in the parameters. The changes in the output quantity Q with respect to a parameter p is measured by the derivative of this quantity with respect to the parameter.

The *sensitivity* of quantity Q with respect to the parameter p is given by

$$\mathscr{S}_Q^p = \frac{\partial Q}{\partial p}.$$

This definition of sensitivity is *local* because the sensitivity is computed while all parameters, including parameter p, are kept at their estimated values. Despite its simplicity, however, this approach does not fully explore the input space, since it does not take into account the simultaneous variation of input parameters. Another drawback of this definition of sensitivity is that it depends strongly on the magnitude of p and the quantity Q. In this respect, a much more useful concept is *elasticity*, which gives the percentage change in the quantity Q with respect to the percentage change in the parameter p. In particular,

The *elasticity* of quantity Q with respect to the parameter p is given by

$$\varepsilon_Q^p = \frac{\partial Q}{\partial p}\frac{p}{Q} \approx \frac{\%\Delta Q}{\%\Delta p}.$$

The sensitivity or elasticity of Q with respect to p is positive if Q is increasing with respect to p, and negative if Q is decreasing with respect to p.

6.5.2.1 Computing Elasticities of $\mathscr{R}_0$ for the HIV Model

In Sect. 6.2.2, we introduced a model of HIV, and we fitted it to world prevalence data. We will illustrate computing the elasticities by computing the elasticities of $\mathscr{R}_0$ in the model (6.2). We first perform analysis of the model (6.2) to derive $\mathscr{R}_0$. The disease-free equilibrium of this model is $\mathscr{E}_0 = (\frac{\Lambda}{\mu}, 0, 0, 0, 0)$. The Jacobian at the disease-free equilibrium is given by

$$J(\mathscr{E}_0) = \begin{pmatrix} -\mu & -\beta & -\beta & -\beta & -\beta \\ 0 & \beta-(\mu+\gamma) & \beta & \beta & \beta \\ 0 & \gamma & -(\mu+\gamma) & 0 & 0 \\ 0 & 0 & \gamma & -(\mu+\gamma) & 0 \\ 0 & 0 & 0 & \gamma & -(\mu+\gamma) \end{pmatrix}. \tag{6.13}$$

To obtain the characteristic polynomial and the eigenvalues, we subtract λ along the diagonal. We see that one of the eigenvalues is $-\mu$. The other eigenvalues are the solutions of the following characteristic polynomial:

$$(\mu+\gamma+\lambda)^4 = \beta(\mu+\gamma+\lambda)^3 + \beta\gamma(\mu+\gamma+\lambda)^2 + \beta\gamma^2(\mu+\gamma+\lambda) + \beta\gamma^3.$$

Dividing by the left-hand side, we obtain $1 = \mathscr{G}(\lambda)$, where

$$\mathscr{G}(\lambda) = \frac{\beta}{(\mu+\gamma+\lambda)} + \frac{\beta\gamma}{(\mu+\gamma+\lambda)^2} + \frac{\beta\gamma^2}{(\mu+\gamma+\lambda)^3} + \frac{\beta\gamma^3}{(\mu+\gamma+\lambda)^4}.$$

We define the reproduction number $\mathscr{R}_0 = \mathscr{G}(0)$, that is,

$$\mathscr{R}_0 = \frac{\beta}{(\mu+\gamma)} + \frac{\beta\gamma}{(\mu+\gamma)^2} + \frac{\beta\gamma^2}{(\mu+\gamma)^3} + \frac{\beta\gamma^3}{(\mu+\gamma)^4}. \tag{6.14}$$

If $\mathscr{R}_0 > 1$, then the characteristic equation has a real positive root. This is the case because $\mathscr{G}(0) = \mathscr{R}_0 > 1$, and $\mathscr{G}(\lambda)$ is a decreasing function of λ, assumed real, with $\lim_{\lambda\to\infty} G(\lambda) = 0$. If, however, $\mathscr{R}_0 < 1$, then all roots have negative real parts. To see this, assume that there is λ with nonnegative real part: $\Re\lambda \geq 0$. Then

$$\begin{aligned} |\mathscr{G}(\lambda)| &\leq \frac{\beta}{|\mu+\gamma+\lambda|} + \frac{\beta\gamma}{|\mu+\gamma+\lambda|^2} + \frac{\beta\gamma^2}{|\mu+\gamma+\lambda|^3} + \frac{\beta\gamma^3}{|\mu+\gamma+\lambda|^4} \\ &\leq \mathscr{G}(\Re\lambda) \leq \mathscr{G}(0) = \mathscr{R}_0 < 1. \end{aligned} \tag{6.15}$$

Hence, $|\mathscr{G}(\lambda)| < 1$, and such a λ cannot be a solution to the characteristic equation.

Interpreting the reproduction number is simple. The first term gives the number of secondary cases generated by an I_1 infectious individual; the second term has two components: $\gamma/(\mu+\gamma)$ gives the probability of moving to class I_2, and $\beta/(\mu+\gamma)$ gives the number of secondary cases generated by infectious individuals in class I_2. The interpretation of the remaining two terms is similar.

In computing the elasticities, it is not hard to see that

$$\varepsilon^{\beta}_{\mathscr{R}_0} = 1.$$

Computing the derivative of the reproduction number with respect to γ gives us

$$\frac{\partial\mathscr{R}_0}{\partial\gamma} = -\frac{4\beta\gamma^3}{(\mu+\gamma)^5} = -8.9272.$$

Furthermore, the elasticity of $\mathscr{R}_0$ with respect to γ is

$$\varepsilon_{\mathscr{R}_0}^{\gamma} = \frac{-4\gamma^4}{(\mu+\gamma)^4 + \gamma(\mu+\gamma)^3 + \gamma^2(\mu+\gamma)^2 + \gamma^3(\mu+\gamma)} = -0.9.$$

The fact that $\varepsilon_{\mathscr{R}_0}^{\gamma} = -0.9$ means that 1% increase in γ will produce 0.9% decrease in $\mathscr{R}_0$. The elasticities suggest that the magnitudes of the impacts of β and γ on the reproduction number of HIV are approximately the same.

Problems

6.1. Influenza in an English Boarding School
The West Country English Boarding School housed 578 boys. An epidemic of influenza began on January 15, 1978. The epidemiological scenario is described in [139]. The numbers of cases by day are given in Table 6.5, where day 1 is January 16, 1978. Write a computer program to fit an SIR model to the data. Plot the solution with the data and the residuals.

Table 6.5 Influenza in an English boarding school

Time (in days)	Number cases	Time	Number cases	Time	Number cases
1	2	11	53	21	13
2	5	12	55	22	14
3	10	13	58	23	11
4	12	14	–	24	12
5	14	15	52	25	9
6	15	16	42	26	7
7	–	17	40	27	5
8	31	18	30	28	4
9	42	19	23	29	2
10	45	20	19	30	1
–	–	–	–	32	1

6.2. Mumps in Iowa
The United States experienced a multistate mumps outbreak involving predominantly Midwest states in 2006. The outbreak in Iowa began in the week of January 28, 2006, and continued until the week of September 30, 2006. The cases in Iowa are given in Table 6.6*.

Write a computer program to fit an SIR model to the data. Plot the solution with the data and the residuals.

6.3. Influenza in an English Boarding School
The West Country English Boarding School housed 578 boys. An epidemic of

Table 6.6 Mumps in Iowa, 2006

Time (in weeks)	Number cases	Time (in weeks)	Number cases	Time (in weeks)	Number cases
Jan. 28	13	Apr. 22	255	July 15	5
Feb. 5	6	Apr. 29	194	July 22	2
Feb. 11	7	May 6	117	July 29	3
Feb. 18	4	May 13	97	Aug. 5	3
Feb. 25	31	May 20	64	Aug. 12	1
Mar. 4	27	May 27	19	Aug. 19	3
Mar. 11	45	June 3	28	Aug. 26	4
Mar. 18	81	June 10	14	Sep. 2	1
Mar. 25	118	June 17	7	Sep. 9	1
Apr. 1	236	June 24	9	Sep. 16	2
Apr. 8	292	July 1	7	Sep. 23	1
Apr. 15	251	July 8	9	Sep. 30	1

* Data taken from http://www.idph.state.ia.us/adper/common/pdf/mumps (no longer active)

influenza began on January 15 1978. The epidemiological scenario is described in [139]. The numberd of cases by day are given in the Table 6.5, where day 1 is January 16, 1978.

(a) Write a computer program to fit an SEIR model to the data. Plot the solution with the data and the residuals.
(b) Hold all parameters fixed and plot the solutions for three different values of the transmission rate β, recovery rate from the exposed class, and the recovery rate from the infectious class. What conclusions can you draw?

6.4. Fitting the Number of People Living with HIV in the UK
The table below gives the number of people living with HIV in the UK.

Table 6.7 Number of people living with HIV in the UK

Year	Number cases	Year	Number cases	Year	Number cases
1990	21,000	1997	29,000	2004	61,000
1991	22,000	1998	31,000	2005	66,000
1992	23,000	1999	35,000	2006	72,000
1993	23,000	2000	39,000	2007	77,000
1994	24,000	2001	43,000	2008	82,000
1995	25,000	2002	51,000	2009	85,000
1996	26,000	2003	55,000	2010	91,000

(a) Fit model (6.2) to the data above. Recall that Λ and μ have to be pre-fixed in accord with the population of the UK.
(b) Investigate the sensitivity of $\mathscr{R}_0$ with respect to γ. What conclusions can you draw?

(c) Hold β and μ fixed and plot $I_1(t)$, $I_2(t)$, $I_3(t)$, $I_4(t)$, and $I(t)$ for three different values of γ.
(d) Hold γ and μ fixed and plot $I_1(t)$, $I_2(t)$, $I_3(t)$, $I_4(t)$, and $I(t)$ for three different values of β.

6.5. Fitting the Number of People Living with HIV in the UK
Table 6.7 gives the number of people living with HIV in the UK. Consider the following HIV model with treatment [129]:

$$\begin{aligned} S'(t) &= \Lambda - \lambda(t)S(t) - \mu S(t), \\ I'(t) &= \lambda(t)S(t) - (\mu + \alpha + \gamma)I(t), \\ T'(t) &= \alpha I(t) - (\mu + \rho)T(t), \\ A'(t) &= \gamma I(t) + \rho T(t) - (\mu + d)A(t), \end{aligned} \tag{6.16}$$

where

$$\lambda(t) = \frac{\beta_1 I(t) + \beta_2 T(t) + \beta_3 A(t)}{N(t)},$$

and $N(t) = S(t) + I(t) + T(t) + A(t)$ is the total population size, α is the treatment rate, γ is the rate of progression of untreated individuals to AIDS, ρ is the rate of progression of treated individuals to AIDS, d is the disease-induced death rate.

(a) Fit the above model to the data in Table 6.7 by fitting β_i, $i = 1,2,3$, α, γ, ρ, and d. Pre-fix Λ and μ by determining the population of the UK.
(b) Perform sensitivity analysis of $I(t)$ and $T(t)$ with respect to β_i, $i = 1,2,3$, α, γ, and ρ.
(c) Compute $\mathscr{R}_0$. Perform elasticity analysis of $\mathscr{R}_0$ with respect to β_i, $i = 1,2,3$, α, γ, and ρ.

6.6. Fitting the Incidence of TB in the United States
After a mild increase in tuberculosis (TB) cases in the United States in the late 1980s and the beginning of the 1990s, the incidence (number of new cases per year) of TB has been steadily declining. Table 6.8 gives the TB incidence starting in 1990.

Table 6.8 Number of new cases of TB in USA[a]

Year	Number cases	Year	Number cases	Year	Number cases
1990	25,701	1998	18,287	2006	13,732
1991	26,283	1999	17,500	2007	13,286
1992	26,673	2000	16,309	2008	12,905
1993	25,107	2001	15,945	2009	11,537
1994	24,205	2002	15,055	2010	11,182
1995	22,727	2003	14,835	2011	10,528
1996	21,210	2004	14,499	2012	–
1997	19,751	2005	14,068	2013	–

[a] http://www.cdc.gov/tb/statistics/reports/2010/default.htm

(a) Fit the following TB model (6.17) to the data above:

$$\begin{aligned} S'(t) &= \Lambda - \beta_1 SI/N - \mu S, \\ E'(t) &= \beta_1 SI/N + \beta_2 TI/N - (\mu + \kappa + r_1)E + pr_2 I, \\ I'(t) &= \kappa E - (r_2 + \mu)I, \\ T'(t) &= r_1 E + qr_2 I - \beta_2 TI/N - \mu T, \end{aligned} \tag{6.17}$$

where $T(t)$ is the number of treated individuals, $I(t)$ is the number of individuals with active TB, $E(t)$ is the number of exposed, r_1 is the treatment rate of exposed individuals, r_2 is the treatment rate of infectious individuals, and κ is the progression to the infectious state. We assume that $p + q = 1$.

Hint: (1) You should be fitting the incidence $\beta_1 SI/N + \beta_2 TI/N$ to the data. (2) You may fit the initial conditions or take them as follows: $S(0) = 290{,}000{,}000$, $E(0) = 25{,}000$, $I(0) = 25{,}000$, $T(0) = 22{,}000$.

(b) Investigate the sensitivity of $I(t)$ with respect to β_1, β_2, κ, r_1, r_2, p, and q. Interpret your findings epidemiologically.

(c) Compute $\mathscr{R}_0$ (see Chap. 5). Investigate the elasticity of $\mathscr{R}_0$ with respect to the parameters.

6.7. Fitting the Incidence of Malaria in India

The new cases of malaria in India were slowly increasing in the second half of the 1980s. In 1997, India implemented new control strategies, and the number of malaria cases has been decreasing ever since. Table 6.9 gives India's population along with malaria cases for the period 1985–2011.

Table 6.9 Number of new cases of malaria in India[a] (in millions)

Year	Population	Number cases	Year	Population	Number cases
1985	726	1.864	1999	948.66	2.28
1986	737	1.792	2000	970	2.032
1987	753.55	1.66	2001	984.58	2.085
1988	766.92	1.85	2002	1025.56	1.84
1989	769.32	2.05	2003	1027.16	1.87
1990	784.42	2.019	2004	1044.74	1.915
1991	808.1	2.12	2005	1007.2	1.82
1992	824.14	2.13	2006	1064	1.77
1993	833.89	2.21	2007	1089.8	1.51
1994	861.73	2.51	2008	982	1.53
1995	878.96	2.988	2009	943.93	1.56
1996	905.71	3.04	2010	1024.66	1.6
1997	884.72	2.66	2011	1059.8	1.31
1998	907.3	2.223	2012		

[a] http://www.searo.who.int/entity/malaria/data/en/

(a) Fit the logistic equation to India's population size:

$$P'(t) = rP(t)\left(1 - \frac{P(t)}{K}\right).$$

(b) Fit the following malaria model to the data:

$$\begin{aligned} C'(t) &= b(1-\xi H(t-\tau))(P(t)-C(t)-I(t))y(t) - (\nu+\mu)C(t), \\ I'(t) &= \nu C(t) - (\lambda P(t) + \gamma P(t)H(t-\tau) + \mu)I(t), \\ y'(t) &= \rho(1-\xi H(t-\tau))(1-y(t))I(t) - (d+\eta H(t-\tau))y(t), \end{aligned} \tag{6.18}$$

where $C(t)$ is the number of symptomatic cases, $I(t)$ is the number of infectious individuals, and $y(t)$ is the proportion of mosquitoes. The function $H(t-\tau)$ is the Heaviside function. Pre-estimate the value of τ so that τ gives the middle of 1997 (mid-1997). In mid-1997, after the new measures were implemented, b and ρ decreased while d increased.

Hint: You must fit the incidence $b(1-\xi H(t-\tau))(P(t)-C(t)-I(t))y(t)$ to the number of new cases of malaria. It may be easier if you take ξ and η equal to zero, then fit the remaining parameters to the 1985–1996 data. Fix those parameters at their values and fit ξ and η to the 1997–2011 data [106].

6.8. Model Selection with Influenza in Boarding School Data

For the data in Table 6.5, compose SIR, SEIR, SIQR, SEIQR, SEIAR models, where Q is the number of isolated and A is the number of asymptomatic. Fit the five models to the data and perform model selection using AIC and Akaike weights.

6.9. Model Selection with HIV Data

For the data in Table 6.7, compose five models with one through five possible infectious classes (see the model with four infectious classes in (6.2)). Fit the five models to the data and perform model selection using AIC and Akaike weights.

Chapter 7
Analysis of Complex ODE Epidemic Models: Global Stability

7.1 Introduction

In Chap. 2, we introduced the basic SIR and SIS models and developed tools and techniques for their analysis. In Chap. 5, we saw that incorporating more realism into the models leads to more complex epidemic models with multiple compartments. The mathematical tools developed for the global analysis of the SIS and SIR models are unsuitable for models of dimension three or greater. For instance, the Dulac criterion, used to rule out oscillations, and Bendixon's theorem can be applied only to planar systems and are not valid in higher dimensions. In this chapter, we develop new tools that are applicable to any number of dimensions.

Much of the local analysis that we performed for the SIR model can be performed for higher-dimensional systems. That is, we again have to look at equilibria, which are classified into a disease-free equilibrium and endemic equilibria, and then consider the linearization around those equilibria by evaluating the Jacobian J. For higher-dimensional systems, the principle for showing stability, namely that all eigenvalues of the characteristic equation

$$|J - \lambda I| = 0$$

have negative real parts or are negative, still holds. This is guaranteed by the **Hartman–Grobman theorem** [155] and by the fact that zero is an asymptotically stable solution of a linear system with constant coefficients if and only if all eigenvalues have negative real part [28]. These results hold for multidimensional systems of ODEs.

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_7

7.2 Local Analysis of the SEIR Model

We illustrate the analysis, which is somewhat more elaborate, on the SEIR model. The SEIR model is another classical epidemiological model, which incorporates a compartment of exposed individuals, $E(t)$, where the individuals are *infected* but not *infectious*. With $S(t)$ denoting the number of susceptible individuals, $I(t)$ the number of infectious individuals, and $R(t)$ the number of recovered individuals, the model takes the form

$$\begin{aligned} S' &= \Lambda - \beta SI - \mu S, \\ E' &= \beta SI - (\mu + \gamma)E, \\ I' &= \gamma E - (\mu + \alpha)I, \\ R' &= \alpha I - \mu R. \end{aligned} \tag{7.1}$$

To determine the equilibria of this model, we set the derivatives equal to zero and solve the system

$$\begin{aligned} 0 &= \Lambda - \beta SI - \mu S, \\ 0 &= \beta SI - (\mu + \gamma)E, \\ 0 &= \gamma E - (\mu + \alpha)I, \\ 0 &= \alpha I - \mu R. \end{aligned} \tag{7.2}$$

This system has a disease-free equilibrium, which is obtained by setting $I = 0$, $\mathscr{E}_0 = (\frac{\Lambda}{\mu}, 0, 0, 0)$. To find the endemic equilibria, we solve for E in the third equation,

$$E = \frac{\mu + \alpha}{\gamma} I,$$

and we substitute into the second equation to obtain

$$\beta S = \frac{(\mu + \gamma)(\mu + \alpha)}{\gamma}.$$

Hence

$$S = \frac{(\mu + \gamma)(\mu + \alpha)}{\beta\gamma}.$$

From the first equation, we have

$$I = \frac{\Lambda}{\beta S} - \frac{\mu}{\beta} = \frac{\Lambda\gamma}{(\mu + \gamma)(\mu + \alpha)} - \frac{\mu}{\beta} = \frac{\mu}{\beta}(\mathscr{R}_0 - 1).$$

We define the basic reproduction number as

$$\mathscr{R}_0 = \frac{\Lambda\beta\gamma}{(\mu + \gamma)(\mu + \alpha)\mu}. \tag{7.3}$$

The reproduction number is positive; it is zero if there is no transmission, that is, $\beta = 0$, and it can be interpreted as the number of secondary cases. In particular, we can see that Λ/μ is the number of susceptible individuals in a disease-free population, $\beta\Lambda/(\mu(\mu+\alpha))$ is the number of secondary cases produced by one infectious individual during its lifespan as infectious, and $\gamma/(\mu+\gamma)$ is the fraction of new infections that survive the exposed period and actually become infectious.

From the reproduction number, we have the following result:

Proposition 7.1. *The system has a unique disease-free equilibrium $\mathscr{E}_0$. If $\mathscr{R}_0 > 1$, the system also has a unique endemic equilibrium $\mathscr{E}^* = (S^*, E^*, I^*, R^*)$, where*

$$S^* = \frac{(\mu+\gamma)(\mu+\alpha)}{\beta\gamma}, E^* = \frac{\mu+\alpha}{\gamma}\frac{\mu}{\beta}(\mathscr{R}_0 - 1), I^* = \frac{\mu}{\beta}(\mathscr{R}_0 - 1), R^* = \frac{\alpha}{\beta}(\mathscr{R}_0 - 1).$$

To investigate the local stability, we consider the Jacobian of the system (7.1). We have

$$J = \begin{pmatrix} -\beta I - \mu & 0 & -\beta S & 0 \\ \beta I & -(\mu+\gamma) & \beta S & 0 \\ 0 & \gamma & -(\mu+\alpha) & 0 \\ 0 & 0 & \alpha & -\mu \end{pmatrix}. \tag{7.4}$$

To determine the local stability of the disease-free equilibrium, we evaluate the Jacobian at $\mathscr{E}_0$:

$$J(\mathscr{E}_0) = \begin{pmatrix} -\mu & 0 & -\beta S & 0 \\ 0 & -(\mu+\gamma) & \beta S & 0 \\ 0 & \gamma & -(\mu+\alpha) & 0 \\ 0 & 0 & \alpha & -\mu \end{pmatrix}. \tag{7.5}$$

Subtracting λ along the main diagonal, we see that the characteristic equation $|J(\mathscr{E}_0) - \lambda I| = 0$ has two roots equal to $-\mu$. The remaining roots are the solution to the following equation:

$$\begin{vmatrix} -(\mu+\gamma+\lambda) & \beta S \\ \gamma & -(\mu+\alpha+\lambda) \end{vmatrix} = 0. \tag{7.6}$$

This leads to the following quadratic characteristic equation:

$$(\mu+\gamma+\lambda)(\mu+\alpha+\lambda) - \beta\gamma\frac{\Lambda}{\mu} = 0.$$

From here, it is not hard to see that this equation has one positive real root if $\mathscr{R}_0 > 1$, and two negative real roots or two complex conjugate real roots with negative real parts if $\mathscr{R}_0 < 1$. This leads to the following classical result regarding the disease-free equilibrium.

Proposition 7.2. *If $\mathscr{R}_0 < 1$, then the disease-free equilibrium $\mathscr{E}_0$ is locally asymptotically stable. If $\mathscr{R}_0 > 1$, the disease-free equilibrium is unstable.*

To investigate the stability of the endemic equilibrium $\mathscr{E}^*$, we evaluate the Jacobian at the endemic equilibrium. We obtain the following characteristic equation:

$$\begin{vmatrix} -\beta I-\mu-\lambda & 0 & -\beta S & 0 \\ \beta I & -(\mu+\gamma+\lambda) & \beta S & 0 \\ 0 & \gamma & -(\mu+\alpha+\lambda) & 0 \\ 0 & 0 & \alpha & -(\mu+\lambda) \end{vmatrix} = 0. \tag{7.7}$$

Expanding by the last column, we see that the characteristic equation has one root equal to $-\mu$. The remaining roots are solutions to the following equation:

$$\begin{vmatrix} -\beta I-\mu-\lambda & 0 & -\beta S \\ \beta I & -(\mu+\gamma+\lambda) & \beta S \\ 0 & \gamma & -(\mu+\alpha+\lambda) \end{vmatrix} = 0. \tag{7.8}$$

Expanding the determinant, we have, after some simplification, the following polynomial characteristic equation:

$$[\beta I^*+\mu+\lambda][\mu+\gamma+\lambda][\mu+\alpha+\lambda] = \beta S^*\gamma(\mu+\lambda). \tag{7.9}$$

One approach here to showing that all solutions to the above equation have negative real parts is to use the Routh–Hurwitz criterion (Chap. 5). We will apply a somewhat trickier but faster method to establish the same result. The question is whether the above equation has solutions with nonnegative real part. Assume that there is a λ with nonnegative real part, $\Re\lambda \geq 0$. Then, we can divide by $\lambda+\mu$ and take the absolute value of both sides of the equation. We have

$$\frac{|\beta I^*+\mu+\lambda||\mu+\gamma+\lambda||\mu+\alpha+\lambda|}{|\mu+\lambda|} = \beta\gamma S^*. \tag{7.10}$$

Recall the value of S^*. Hence, we have

$$\beta\gamma S^* = (\mu+\gamma)(\mu+\alpha).$$

From the left-hand side in Eq. (7.10), we have that

$$\frac{|\beta I^*+\mu+\lambda|}{|\mu+\lambda|} > 1$$

for all values of λ real or complex, as long as $\Re\lambda \geq 0$. If $\lambda = x+yi$, where i is the imaginary unit, then

$$\begin{aligned} &\frac{|\beta I^*+\mu+\lambda||\mu+\gamma+\lambda||\mu+\alpha+\lambda|}{|\mu+\lambda|} \\ &> |\mu+\gamma+\lambda||\mu+\alpha+\lambda| \geq |\mu+\gamma+x||\mu+\alpha+x| \\ &\geq (\mu+\gamma)(\mu+\alpha) = \beta\gamma S^*. \end{aligned} \tag{7.11}$$

This means that for λ with $\Re\lambda \geq 0$, the left-hand side of Eq. (7.10) is always grater than the right-hand side. Hence, the characteristic equation cannot have such solutions. We have established the following result:

Proposition 7.3. *Assume $\mathscr{R}_0 > 1$. Then the endemic equilibrium $\mathscr{E}^*$ is locally asymptotically stable.*

7.3 Global Stability via Lyapunov Functions

For higher-dimensional systems, there are several techniques that could establish the global stability of an equilibrium. One of the most commonly used is the Lyapunov function. Lyapunov functions are scalar functions that may be used to prove the global stability of an equilibrium. They are named after the Russian mathematician Aleksandr Mikhailovich Lyapunov, who proposed the theory in his doctoral dissertation [98].

7.3.1 Lyapunov–Kasovskii–LaSalle Stability Theorems

Let x^* be an equilibrium of $x' = f(x)$, where $f : \mathbf{R}^n \to \mathbf{R}^n$.

Definition 7.1. A scalar function $V(x)$ such that $V : \mathbf{R}^n \to \mathbf{R}$ is called *radially unbounded* if

$$V(x) \to \infty \qquad \text{if} \qquad ||x|| \to \infty.$$

One significant property of Lyapunov functions is that they are positive definite in the entire space.

Definition 7.2. Let V be a continuous scalar function, that is,

$$V : \mathbf{R}^n \to \mathbf{R}.$$

The function V is called *positive definite* on the entire space if

- $V(x^*) = 0$,
- $V(x) > 0$ for $x \neq x^*$,

where x^* is an equilibrium of the autonomous system $x' = f(x)$. We define the derivative of $V(x)$ along the solutions of the system of differential equations as

$$V'(x) = \frac{d}{dt}V(x(t)) = \frac{\partial V}{\partial x}\frac{dx}{dt}.$$

Now we can state Lyapunov's theorem for global stability of the equilibrium x^*. For a proof of Lyapunov's, theorem see [41].

Theorem 7.1 (Lyapunov's Stability Theorem). *If a function $V(x)$ is globally positively definite and radially unbounded, and its time derivative is globally negative,*

$$V'(x) < 0 \qquad \text{for} \quad \text{all} \quad x \neq x^*,$$

then the equilibrium x^* is globally stable.

Definition 7.3. If a function $V(x)$ exists that satisfies the conditions of Theorem 7.1, then this function is called a *Lyapunov function*.

There are no established rules for finding a Lyapunov function, and often finding a Lyapunov function is tricky and computationally intensive (but see [146]). However, if a Lyapunov function is found, it can establish the global stability of an equilibrium.

Lyapunov's Theorem requires that the derivative of a Lyapunov function with respect to t be strictly negative; however, we can often show only nonpositivity. In this case, an extension of Lyapunov's theorem was given by LaSalle [93] and Krasovskii [91].

Theorem 7.2 (Krasovkii–LaSalle Theorem). *Consider the autonomous system $x' = f(x)$, where x^* is an equilibrium, that is, $f(x^*) = 0$. Suppose there exists a continuously differentiable function $V : \mathbf{R}^n \to \mathbf{R}$ and that this function is positive definite on the entire space and radially unbounded and that it satisfies*

$$V'(x) \leq 0 \qquad \text{for} \quad \text{all} \quad t \quad \text{and} \quad \text{all} \quad x \in \mathbf{R}^n.$$

Define the invariant set

$$\mathscr{S} = \{x \in \mathbf{R}^n | V'(x) = 0\}.$$

If $\mathscr{S}$ contains only the equilibrium x^*, then the equilibrium x^* is globally stable.

7.3.2 Global Stability of Equilibria of the SEIR Model

We illustrate the application of Lyapunov's theorem by showing global stability of the disease-free equilibrium and the endemic equilibrium for the SEIR model in (7.1).

Proposition 7.4. *Assume $\mathscr{R}_0 < 1$. Then the disease-free equilibrium is globally asymptotically stable.*

Proof. We approach the problem by constructing a Lyapunov function. We will consider the SEIR model on the space of the first three variables only (S, E, I). It is clear that if the disease-free equilibrium for the first three equations is globally stable, then $R(t) \to 0$, and the disease-free equilibrium for the full SEIR model is globally stable.

Consider the following candidate for a Lyapunov function on $\mathbf{R}^3_+$. It is sufficient to work with $\mathbf{R}^3_+$, because we are interested in working with only the positive orthant.

$$V = \kappa\left(S - S^* - S^* \ln \frac{S}{S^*}\right) + \frac{1}{\mu+\gamma}E + \frac{1}{\gamma}I, \tag{7.12}$$

where $\kappa > 0$ is to be determined and $S^* = \frac{\Lambda}{\mu}$. First, it is not hard to see that $V = 0$ at the disease-free equilibrium. To establish that $V > 0$ for all $(S,E,I) \neq (\frac{\Lambda}{\mu},0,0)$, it is enough to notice that

$$\kappa S^*\left(\frac{S}{S^*} - 1 - \ln\frac{S}{S^*}\right) > 0,$$

because the function $g(x) = x - 1 - \ln x$ achieves a global minimum at $x = 1$ and $g(1) = 0$. Hence $g(x) > 0$ for all $x > 0$ and $x \neq 1$. So the first term is positive. The remaining two terms are also clearly positive. The coefficients of E and I are chosen in such a way that the negative term in the E' equation cancels with the positive term in the I' equation in (7.1). Furthermore, V is also clearly radially unbounded. We take the derivative of V with respect to t:

$$\begin{aligned}\frac{d}{dt}V &= \kappa\left(1 - \frac{S^*}{S}\right)S' + \frac{1}{\mu+\gamma}E' + \frac{1}{\gamma}I' \\ &= \kappa\left(1 - \frac{S^*}{S}\right)[\Lambda - \beta SI - \mu S] + \frac{1}{\mu+\gamma}(\beta SI - (\mu+\gamma)E) + \frac{1}{\gamma}(\gamma E - (\mu+\alpha)I) \\ &= 2\kappa\Lambda - \beta\kappa SI - \kappa\mu S - \frac{\Lambda^2\kappa}{\mu S} + \frac{\Lambda\beta\kappa}{\mu}I + \frac{\beta}{\mu+\gamma}SI - \frac{\mu+\alpha}{\gamma}I,\end{aligned} \tag{7.13}$$

where after differentiation, we have used equations (7.1) to replace the derivatives with the right-hand sides. The last line is obtained by multiplying out. We have several positive terms for which we need to compensate. If we choose $\kappa = 1/(\mu+\gamma)$, then the SI terms have the same coefficient and opposite signs, so they cancel each other. We combine the two I-terms and the remaining terms as follows:

$$\frac{d}{dt}V = -\kappa\Lambda\left(\frac{\Lambda}{\mu S} + \frac{\mu S}{\Lambda} - 2\right) + \frac{\mu+\alpha}{\gamma}(\mathscr{R}_0 - 1)I. \tag{7.14}$$

Since $\mathscr{R}_0 < 1$, the last term is nonpositive. The first term is more interesting. If we set $a = \Lambda/(\mu S)$, then we have $a + 1/a - 2$. We claim that this expression is positive for every $a > 0$, $a \neq 1$. Indeed,

$$a + \frac{1}{a} - 2 = \frac{a^2 - 2a + 1}{a} = \frac{(a-1)^2}{a} > 0.$$

Hence we have $V' < 0$ for all $(S,E,I) \neq (S^*,0,0)$. Therefore, by Lyapunov's theorem, the disease-free equilibrium is globally asymptotically stable. □

Now we consider the endemic equilibrium. The endemic equilibrium is unique and locally stable whenever it exists. For $\mathscr{R}_0 > 1$, the only other equilibrium is $\mathscr{E}_0$, which is unstable. This suggests that the endemic equilibrium may be globally stable. Indeed, that is the case. The global stability of the endemic equilibrium for the SEIR model was first established by Li and Muldowney [96] using different techniques. The global stability of the endemic equilibrium of the SEIR model via a Lyapunov function was fist established by Korobeinikov and Maini [89]. We will establish this result via a Lyapunov function. Before we state the main result, we include a lemma that is very helpful in establishing the negativity of the time derivative of a candidate Lyapunov function.

Lemma 7.1. *Assume that $x_1,\dots,x_n$ are n positive numbers. Then their arithmetic mean is greater than or equal to their geometric mean. In particular,*

$$\frac{x_1+\cdots+x_n}{n} \geq \sqrt[n]{x_1\dots x_n}.$$

Theorem 7.3. *Assume $\mathscr{R}_0 > 1$. Then the endemic equilibrium is globally asymptotically stable.*

Proof. We consider again only the first three components of system (7.1), (S,E,I). We assume that they belong to the positive orthant $\mathbf{R}^3_+$. We define a Lyapunov function

$$V = \kappa_1\left(S - S^* - S^*\ln\frac{S}{S^*}\right) + \kappa_2\left(E - E^* - E^*\ln\frac{E}{E^*}\right) + \kappa_3\left(I - I^* - I^*\ln\frac{I}{I^*}\right), \tag{7.15}$$

where $\kappa_1 > 0$, $\kappa_2 > 0$, and $\kappa_3 > 0$ will be determined later. Notice that $V = 0$ when $(S,E,I) = (S^*,E^*,I^*)$ and $V > 0$ otherwise; V is also radially unbounded. What remains to be proved is that the derivative of V with respect to t is negative. We differentiate with respect to t and replace S', E', and I' with their equals from (7.1):

$$\begin{aligned}\frac{dV}{dt} &= \kappa_1\left(1-\frac{S^*}{S}\right)S' + \kappa_2\left(1-\frac{E^*}{E}\right)E' + \kappa_3\left(1-\frac{I^*}{I}\right)I' \\ &= \kappa_1\left(1-\frac{S^*}{S}\right)(\Lambda - \beta SI - \mu S) + \kappa_2\left(1-\frac{E^*}{E}\right)(\beta SI - (\mu+\gamma)E) \\ &\quad + \kappa_3\left(1-\frac{I^*}{I}\right)(\gamma E - (\mu+\alpha)I).\end{aligned} \tag{7.16}$$

One of the classical first steps here is to replace Λ with its equal from the equilibrium equations, that is, $\Lambda = \beta S^*I^* + \mu S^*$. Then $\mu S^* - \mu S$ can be combined with the first term in the product to yield a negative term. We multiply out all other products:

$$\begin{aligned}\frac{dV}{dt} = &- \kappa_1 \frac{(S-S^*)^2}{S} + \kappa_1\beta S^* I^* - \kappa_1\beta SI - \kappa_1\beta \frac{S^{*2} I^*}{S} \\ &+ \kappa_1\beta S^* I + \kappa_2\beta SI - \kappa_2(\mu+\gamma)E \\ &- \kappa_2\beta \frac{E^* SI}{E} + \kappa_2(\mu+\gamma)E^* + \kappa_3\gamma E - \kappa_3(\mu+\alpha)I - \kappa_3\gamma\frac{I^* E}{I} + \kappa_3(\mu+\alpha)I^*.\end{aligned} \tag{7.17}$$

Now we can see that if $\kappa_1 = \kappa_2$, then $-\kappa_1\beta SI$ can be canceled with $\kappa_2\beta SI$. Also, where we have fractions, we multiply and divide by the equilibrium value:

$$\begin{aligned}\frac{dV}{dt} = &-\kappa_1 \frac{(S-S^*)^2}{S} + \kappa_1\beta S^* I^* - \kappa_1\beta \frac{S^{*2} I^*}{S} + \kappa_1\beta S^* I - \kappa_2(\mu+\gamma)E \\ &-\kappa_2\beta S^* I^* \frac{E^* SI}{ES^* I^*} + \kappa_2(\mu+\gamma)E^* + \kappa_3\gamma E - \kappa_3(\mu+\alpha)I - \kappa_3\gamma E^* \frac{I^* E}{IE^*} \\ &+\kappa_3(\mu+\alpha)I^*.\end{aligned} \tag{7.18}$$

We want to combine all constant terms with all fractional terms, because all constant terms are positive, and all fractional terms are negative. First, we notice that since $\kappa_1 = \kappa_2$, we have $\beta S^* I^* = (\mu+\gamma)E^*$ from the corresponding equilibrium equation. We need to choose κ_3 such that

$$\kappa_3(\mu+\alpha)I^* = \kappa_2(\mu+\gamma)E^*.$$

Hence, $\kappa_3 = \kappa_2\frac{\mu+\gamma}{\gamma}$. We pull out $\kappa_1\beta S^* I^*$ from all terms. We have

$$\begin{aligned}\frac{dV}{dt} = &- \kappa_1 \frac{(S-S^*)^2}{S} + \kappa_1\beta S^* I^* \left[3 - \frac{S^*}{S} - \frac{E^* SI}{ES^* I^*} - \frac{I^* E}{IE^*}\right] \\ &+ (\kappa_1\beta S^* - \kappa_3(\mu+\alpha))I + (\kappa_3\gamma - \kappa_2(\mu+\gamma))E.\end{aligned} \tag{7.19}$$

Because $\kappa_3 = \kappa_2(\mu+\gamma)/\gamma$, the last two terms in the formula above are zero. We may now choose $\kappa_1 = \kappa_2 = 1$. Then

$$\kappa_3 = \frac{\mu+\gamma}{\gamma}.$$

In this case, the derivative of the Lyapunov function becomes

$$\frac{dV}{dt} = -\frac{(S-S^*)^2}{S} + \beta S^* I^* \left[3 - \frac{S^*}{S} - \frac{E^* SI}{ES^* I^*} - \frac{I^* E}{IE^*}\right]. \tag{7.20}$$

The first term above is clearly negative unless $S = S^*$. We argue that the second term is also negative. Indeed, we will apply Lemma 7.1. Let

$$x_1 = \frac{S^*}{S} \qquad x_2 = \frac{E^*SI}{ES^*I^*}, \qquad x_3 = \frac{I^*E}{IE^*}.$$

Then, notice that $x_1x_2x_3 = 1$. According the the lemma, the arithmetic mean is larger than the geometric mean. Therefore,

$$\frac{S^*}{S} + \frac{E^*SI}{ES^*I^*} + \frac{I^*E}{IE^*} \geq 3.$$

Hence, the second term is nonpositive, and it is zero whenever $(S,E,I) = (S^*,E^*,I^*)$. We have

$$\frac{dV}{dt} \leq 0.$$

We now have to apply the Krasovkii–LaSalle theorem. We consider the set where the Lyapunov function is equal to zero:

$$\mathscr{S} = \{x \in \mathbf{R}^n | V'(x) = 0\}.$$

It is clear that $V' = 0$ if and only if

$$S = S^* \qquad \text{and} \qquad \frac{S^*}{S} + \frac{E^*SI}{ES^*I^*} + \frac{I^*E}{IE^*} = 3.$$

Since $S = S^*$, then $\frac{dS}{dt} = 0$, and from the first equation in system (7.1), we can conclude that $I = I^*$. Finally, from the second equality above, we have

$$\frac{E^*}{E} + \frac{E}{E^*} = 2.$$

It is easy to see that this equality holds if and only if $E = E^*$. Hence, the set $\mathscr{S}$ consists of the singleton (S^*,E^*,I^*). This concludes the proof. □

7.4 Hopf Bifurcation in Higher Dimensions

The Hopf bifurcation theorem that we stated in Chap. 3 is valid in higher dimensions. We will use it here without restating it. The characteristic equation of higher-dimensional ODE systems is a polynomial equation of degree n. The Hopf bifurcation theorem requires that for some value of a parameter, this equation have a purely imaginary root and all other roots have negative real part. The following theorem gives a necessary and sufficient condition for this to happen [57]:

Theorem 7.4. *Let $p(\lambda)$, where $\lambda \in \mathbf{C}$, be a polynomial of degree n with real coefficients:*

$$p(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n,$$

with $a_0 > 0$. Let $\Delta_1, \Delta_2, \ldots, \Delta_n$ be the Hurwitz determinants for the polynomial $p(\lambda)$. Then $p(\lambda)$ has a pair of distinct roots, $-i\omega$ and $i\omega$, on the imaginary axis, and all other roots are in the left half-plane if and only if

$$a_n > 0, \quad \Delta_{n-1} = 0, \quad \Delta_{n-2} > 0, \ldots, \Delta_1 > 0.$$

We recall that the $n \times n$ matrix

$$H = \begin{pmatrix} a_1 & a_3 & a_5 & \ldots & \ldots \\ a_0 & a_2 & a_4 & \ldots & \ldots \\ 0 & a_1 & a_3 & a_5 & \ldots \\ 0 & a_0 & a_2 & a_4 & \ldots \\ & & & & \ddots \end{pmatrix} \tag{7.21}$$

is called a *Hurwitz matrix*. The ith-order principal minor of H is called the *ith Hurwitz determinant* Δ_i.

To illustrate Hopf bifurcation in higher dimensions, we consider a specific example that focuses on recurrent outbreaks in childhood disease. It has been long since observed that the incidence of some childhood diseases like measles and chickenpox experiences oscillations [130]. The question whether these oscillations can be captured by an autonomous ODE model was first investigated by Feng and Thieme [61], who considered an SIQR model to approach the problem. To introduce the model, let $S(t)$, $I(t)$, $Q(t)$, and $R(t)$ be the numbers of susceptible, infected/infectious, isolated, and recovered individuals. We assume that isolated individuals do not mix in the population, and we define the active population $A(t) = S(t) + I(t) + R(t)$. The total population as usual is denoted by $N(t)$. The model is as follows:

$$\begin{aligned} S' &= \Lambda - \frac{\beta SI}{A} - \mu S, \\ I' &= \frac{\beta SI}{A} - (\mu + \gamma) I, \\ Q' &= \gamma I - (\mu + \xi) Q, \\ R' &= \xi Q - \mu R. \end{aligned} \tag{7.22}$$

The model assumes that the isolation is perfect, that is, that everybody who becomes infectious is isolated and that isolated individuals cannot infect. We will show that the endemic equilibrium of this model is not always stable, but it can undergo a Hopf bifurcation that leads to a stable oscillatory solution. That, in particular, means that the endemic equilibrium is not globally stable.

It is not hard to see that the model has two equilibria: a disease-free equilibrium $\mathscr{E}_0 = (\frac{\Lambda}{\mu}, 0, 0, 0)$ and an endemic equilibrium $\mathscr{E}^* = (S^*, I^*, Q^*, R^*)$. To obtain the endemic equilibrium, we introduce the fractions $s = S^*/A^*$, $i = I^*/A^*$, $q = Q^*/A^*$, and $r = R^*/A^*$. Clearly, $s + i + r = 1$ and

$$\frac{N^*}{N^* - Q^*} = \frac{S^*}{A^*} + \frac{I^*}{A^*} + \frac{Q^*}{A^*} + \frac{R^*}{A^*} = 1 + q.$$

The endemic equilibrium is a solution of the following system:

$$\begin{aligned} 0 &= \Lambda - \tfrac{\beta SI}{A} - \mu S, \\ 0 &= \tfrac{\beta SI}{A} - (\mu+\gamma)I, \\ 0 &= \gamma I - (\mu+\xi)Q, \\ 0 &= \xi Q - \mu R. \end{aligned} \tag{7.23}$$

With the new notation, the last three equations from the system become

$$\begin{aligned} 0 &= \beta si - (\mu+\gamma)i, \\ 0 &= \gamma i - (\mu+\xi)q, \\ 0 &= \xi q - \mu r. \end{aligned} \tag{7.24}$$

Hence,

$$s = \frac{\mu+\gamma}{\beta} \qquad q = \frac{\gamma}{\mu+\xi} i. \tag{7.25}$$

From the first equation in (7.23), we have (recall $N^* = \frac{\Lambda}{\mu}$)

$$\frac{\Lambda}{\mu} = \frac{S^*}{\mu}(\mu+\beta i).$$

Subtracting Q^* from both sides, we have

$$\frac{\Lambda}{\mu} - Q^* = \frac{S^*}{\mu}(\mu+\beta i) - Q^*.$$

Dividing by the left-hand side, we have

$$1 = \frac{s}{\mu}(\mu+\beta i) - q.$$

We replace s and q from (7.25) and solve for i to obtain

$$i = \frac{\mu(\mu+\xi)}{\mu^2+\mu\xi+\gamma\xi}\left(1 - \frac{1}{\mathscr{R}_0}\right)$$

and

$$q = \frac{\mu\gamma}{\mu^2+\mu\xi+\gamma\xi}\left(1 - \frac{1}{\mathscr{R}_0}\right),$$

where $\mathscr{R}_0 = \frac{\beta}{\mu+\gamma}$ is the reproduction number of the disease. Define $\kappa = 1+q$. Then from $N^*/(N^*-Q^*) = 1+q = \kappa$, we have

$$Q^* = \frac{\kappa-1}{\kappa} N^*.$$

Hence, $N^* - Q^* = \frac{N^*}{\kappa}$. Therefore,

$$S^* = \frac{sN^*}{\kappa} \qquad I^* = \frac{iN^*}{\kappa}, \qquad R^* = \frac{(1-s-i)N^*}{\kappa}.$$

We summarize the result in the following proposition:

Proposition 7.5. *The system (7.22) has a unique disease-free equilibrium $\mathscr{E}_0$. If $\mathscr{R}_0 > 1$, the system also has a unique endemic equilibrium $\mathscr{E}^*$.*

To determine the stability of equilibria, we consider the Jacobian of the system. System (7.22) is rather complex. We rewrite it in terms of S, I, Q, N:

$$\begin{aligned} S' &= \Lambda - \frac{\beta SI}{N-Q} - \mu S, \\ I' &= \frac{\beta SI}{N-Q} - (\mu+\gamma)I, \\ Q' &= \gamma I - (\mu+\xi)Q, \\ N' &= \Lambda - \mu N. \end{aligned} \tag{7.26}$$

With the notation introduced above, we have

$$J = \begin{pmatrix} -\mu-\beta i & -\beta s & -\beta si & \beta si \\ \beta i & \beta s-(\mu+\gamma) & \beta si & -\beta si \\ 0 & \gamma & -(\mu+\xi) & 0 \\ 0 & 0 & 0 & -\mu \end{pmatrix}. \tag{7.27}$$

The Jacobian computed at the disease-free equilibrium gives

$$J(\mathscr{E}_0) = \begin{pmatrix} -\mu & -\beta s & 0 & 0 \\ 0 & \beta s-(\mu+\gamma) & 0 & 0 \\ 0 & \gamma & -(\mu+\xi) & 0 \\ 0 & 0 & 0 & -\mu \end{pmatrix}. \tag{7.28}$$

The characteristic equation $|J(\mathscr{E}_0) - \lambda I| = 0$ has two eigenvalues equal to $-\mu$ and one eigenvalue equal to $-(\mu+\xi)$. The last eigenvalue is $\beta s - (\mu+\gamma)$. Since $s = 1$ for the disease-free equilibrium, this eigenvalue is negative if and only if $\mathscr{R}_0 < 1$. Therefore, we have the following typical result.

Proposition 7.6. *Assume $\mathscr{R}_0 < 1$. Then the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, the disease-free equilibrium is unstable.*

To determine the stability of the endemic equilibrium $\mathscr{E}^*$, we consider the characteristic equation $|J - \lambda I| = 0$. This characteristic equation has one eigenvalue equal to $-\mu$. The remaining eigenvalues are solutions of the following equation:

$$J = \begin{vmatrix} -\mu-\beta i-\lambda & -\beta s & -\beta si \\ \beta i & -\lambda & \beta si \\ 0 & \gamma & -(\mu+\xi+\lambda) \end{vmatrix} = 0, \tag{7.29}$$

where we have used the fact that $\beta s-(\mu+\gamma)=0$ for the endemic equilibrium. Expanding the determinant, we have

$$\lambda(\mu+\beta i+\lambda)(\mu+\xi+\lambda)+\beta si\gamma\beta i+\beta^2 si(\mu+\xi+\lambda)-\gamma\beta si(\mu+\beta i+\lambda)=0.$$

Expanding along the powers of λ, we have the following cubic polynomial:

$$\lambda^3+a_1\lambda^2+a_2\lambda+a_3=0, \tag{7.30}$$

where

$$\begin{aligned} a_1 &= 2\mu+\beta i+\xi,\\ a_2 &= (\mu+\beta i)(\mu+\xi)+\beta^2 si-\gamma\beta si,\\ a_3 &= \beta^2 si(\mu+\xi)-\gamma\beta si\mu. \end{aligned} \tag{7.31}$$

We show that Hopf bifurcation occurs in the following theorem:

Theorem 7.5. *Assume $\mathscr{R}_0>1$. Then the endemic equilibrium $\mathscr{E}^*$ can become unstable through a Hopf bifurcation leading to an oscillatory solution of system (7.22).*

Proof. First, we use Theorem 7.4 to show the presence of purely imaginary roots. The constant term in a_3 is positive if and only if

$$\beta^2 si(\mu+\xi)>\gamma\beta si\mu.$$

However, since we require $\mathscr{R}_0>1$, this means that $\beta>\mu+\gamma$, and the above condition trivially holds for all parameter values for which $\mathscr{R}_0>1$. Second, $\Delta_1=a_1>0$. In fact, for $\mathscr{R}_0>1$, we have all there coefficients positive:

$$a_1>0,\quad a_2>0,\quad a_3>0.$$

It remains to show that Δ_2 can become zero. We have

$$\Delta_2=\begin{vmatrix} a_1 & a_3\\ 1 & a_2\end{vmatrix}=a_1a_2-a_3. \tag{7.32}$$

Hence,

$$\Delta_2=-\gamma\beta si(\mu+\beta i+\xi)+(\mu+\beta i)(\mu+\xi)(2\mu+\beta i+\xi)+\beta^2 si(\mu+\beta i).$$

To show that for some parameter values, Δ_2 can become zero, we consider Δ_2 as a function of γ. We consider time in days and we set $\mu=0.000039$, $\beta=1.6$, $\xi=1/7$. We then plot Δ_2 as a function of γ in Fig. 7.1. The figure shows that $\Delta_2=0$ for $\gamma_c=0.607825$. We also see that for these parameter values,

$$\frac{\partial\Delta_2}{\partial\gamma}|_{\gamma=\gamma_c}<0.$$

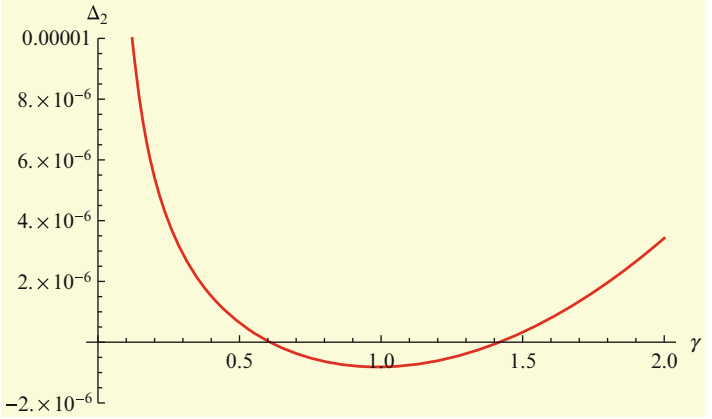

Fig. 7.1 Δ_2 as a function of γ. The figure clearly shows that $\Delta_2(0.607825) = 0$. Thus, the critical value of γ_c is 0.607825. For those parameter values, $\mathscr{R}_0 \approx 2.6$

Next, to satisfy the conditions of the Hopf theorem, we have to show that the real part of the leading root crosses the imaginary axis with nonzero speed. In particular, it remains to show that if $\lambda = \alpha(\gamma) + i\omega(\gamma)$ is a solution to the characteristic equation (7.30), and $\alpha(\gamma_c) = 0$, then

$$\frac{\partial \alpha(\gamma)}{\partial \gamma}|_{\gamma=\gamma_c} \neq 0.$$

To see this, first observe that from the characteristic equation (7.30) for $\gamma = \gamma_c$, we have

$$-i\omega_0^3 - a_1\omega_0^2 + ia_2\omega_0 + a_3 = 0,$$

where $\omega_0 = \omega(\gamma_c)$. Separating the real and the imaginary parts, each of which must be equal to zero, we have

$$\omega_0^2 = a_2.$$

Next, we differentiate the characteristic equation with respect to γ:

$$[3\lambda^2 + 2a_1\lambda + a_2]\frac{d\lambda}{d\gamma} + \frac{\partial a_1}{\partial \gamma}\lambda^2 + \frac{\partial a_2}{\partial \gamma}\lambda + \frac{\partial a_3}{\partial \gamma} = 0.$$

Now we evaluate this equation at $\gamma = \gamma_c$, $\lambda = i\omega_0$:

$$[-3\omega_0^2 + 2a_1 i\omega_0 + a_2]\frac{d\lambda}{d\gamma} - \frac{\partial a_1}{\partial \gamma}\omega_0^2 + \frac{\partial a_2}{\partial \gamma}i\omega_0 + \frac{\partial a_3}{\partial \gamma} = 0.$$

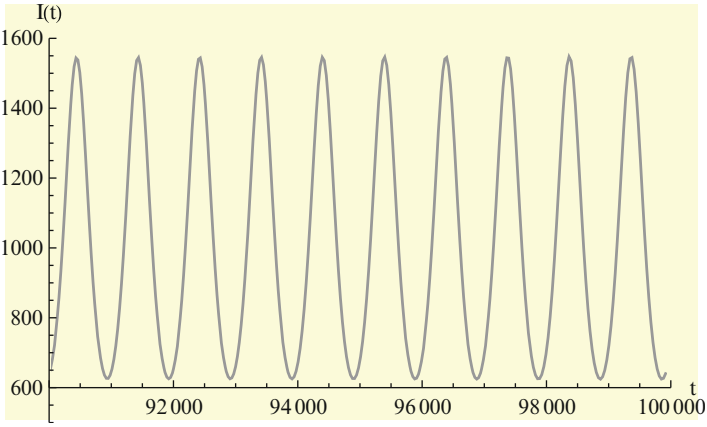

Fig. 7.2 Oscillations of the SIQR model with $\mu = 0.000039$, $\xi = 1/7$, $\Lambda = 1000$, $\beta = 1.6$, and $\gamma = 0.61$

Solving for $d\lambda/d\gamma$ and using $\omega_0^2 = a_2$ to eliminate ω_0, we have

$$\frac{d\lambda}{d\gamma}|_{\gamma=\gamma_c} = \frac{-a_2\frac{\partial a_1}{\partial \gamma} + \frac{\partial a_3}{\partial \gamma} + i\sqrt{a_2}\frac{\partial a_2}{\partial \gamma}}{2(a_2 - ia_1\sqrt{a_2})}. \tag{7.33}$$

Multiplying the denominator by the complex conjugate, we have

$$\frac{d\lambda}{d\gamma}|_{\gamma=\gamma_c} = \frac{[-a_2\frac{\partial a_1}{\partial \gamma} + \frac{\partial a_3}{\partial \gamma} + i\sqrt{a_2}\frac{\partial a_2}{\partial \gamma}][a_2 + ia_1\sqrt{a_2}]}{2(a_2^2 + a_1^2 a_2)}. \tag{7.34}$$

Hence, the real part of the derivative is given by

$$\Re\frac{d\lambda}{d\gamma}|_{\gamma=\gamma_c} = \frac{(-a_2\frac{\partial a_1}{\partial \gamma} + \frac{\partial a_3}{\partial \gamma})a_2 - a_1 a_2\frac{\partial a_2}{\partial \gamma}}{2(a_2^2 + a_1^2 a_2)}.$$

Canceling a_2, we have

$$\Re\frac{d\lambda}{d\gamma}|_{\gamma=\gamma_c} = \frac{-[\frac{\partial a_1 a_2}{\partial \gamma} - \frac{\partial a_3}{\partial \gamma}]}{2(a_2 + a_1^2)} = \frac{-\frac{\partial \Delta_2}{\partial \gamma}}{2(a_2 + a_1^2)} > 0.$$

This completes the proof. □

We illustrate the oscillations in the model in Fig. 7.2.

7.5 Backward Bifurcation

Until now, we have investigated models in which at the critical value of the reproduction number $\mathscr{R}_0 = 1$, an endemic equilibrium bifurcates and exists when $\mathscr{R}_0 > 1$. In this case, if we plot I^* as a function of $\mathscr{R}_0$, the bifurcation is forward. However, there are cases in which the bifurcating endemic equilibrium exists for $\mathscr{R}_0 < 1$. It this case, it is said that *backward bifurcation* occurs. In the case of backward bifurcation, there is a range of the reproduction number

$$\mathscr{R}_0^* < \mathscr{R}_0 < 1,$$

where there are at least two endemic equilibria, typically at least one of which is stable. In this case, the disease-free equilibrium is not globally stable when $\mathscr{R}_0 < 1$, and the infection can persist even if the reproduction number is less than one. The phenomenon of backward bifurcation has serious consequences for disease control.

7.5.1 Example of Backward Bifurcation and Multiple Equilibria

To illustrate the phenomenon, we consider the following SEI model, in which exposed individuals can become reinfected. A version of this model incorporating treatment was proposed in [60] to model tuberculosis (TB), which is caused by the bacterium *mycobacterium tuberculosis* (MTB). The causative agent is spread by aerosol droplets produced by infectious people when they speak, sneeze, or cough. Once the bacterium enters the body, it often forms a granuloma and remains dormant potentially for many years. Infected individuals with granulomas are not infectious, and they form the class of exposed individuals. Exposed individuals can progress to active TB, where they become infectious, either naturally or through a new infection with TB, a process called *reinfection*. The model takes the form

$$\begin{aligned} S' &= \Lambda - \beta \frac{SI}{N} - \mu S, \\ E' &= \beta \frac{SI}{N} - p\beta \frac{EI}{N} - (\mu + \alpha)E, \\ I' &= p\beta \frac{EI}{N} + \alpha E - \mu I, \end{aligned} \tag{7.35}$$

where S, E, and I are susceptible, exposed, and individuals with active TB; α is the rate of progression to active TB; and $p\beta$ is the transmission rate at reinfection. We note that standard incidence is used.

The disease-free equilibrium is obtained from setting $I = 0$. Then, we obtain

$$\mathscr{E}_0 = (\frac{\Lambda}{\mu}, 0, 0).$$

To determine the reproduction number, we look at the stability of the disease-free equilibrium. First, we will derive the generic form of the Jacobian. In the case of standard incidence, one has to keep in mind that N is variable and $N = S + E + I$. Therefore, the Jacobian takes the form

$$J = \begin{pmatrix} -\beta\frac{I}{N} + \beta\frac{SI}{N^2} - \mu & \beta\frac{SI}{N^2} & -\beta\frac{S}{N} + \beta\frac{SI}{N^2} \\ \beta\frac{I}{N} - \beta\frac{SI}{N^2} + p\beta\frac{EI}{N^2} & -\beta\frac{SI}{N^2} - p\beta\frac{I}{N} + p\beta\frac{EI}{N^2} - (\mu+\alpha) & \beta\frac{S}{N} - \beta\frac{SI}{N^2} - p\beta\frac{E}{N} + p\beta\frac{EI}{N^2} \\ -p\beta\frac{EI}{N^2} & p\beta\frac{I}{N} - p\beta\frac{EI}{N^2} + \alpha & p\beta\frac{E}{N} - p\beta\frac{EI}{N^2} - \mu \end{pmatrix}. \tag{7.36}$$

Evaluating the Jacobian at the disease-free equilibrium, we have

$$J(\mathscr{E}_0) = \begin{pmatrix} -\mu & 0 & -\beta \\ 0 & -(\mu+\alpha) & \beta \\ 0 & \alpha & -\mu \end{pmatrix}. \tag{7.37}$$

The Jacobian at the disease-free equilibrium has one eigenvalue equal to $-\mu$. The remaining eigenvalues have negative real part if and only if

$$(\mu+\alpha)\mu - \alpha\beta > 0.$$

This leads to the following reproduction number:

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

In interpreting the reproduction number, we notice that β/μ is the number of secondary infections that one infectious individual will produce in an entirely susceptible population during its lifespan as infectious, and $\alpha/(\mu+\alpha)$ is the probability that a newly infected individual will survive the exposed period and become infectious.

To consider the endemic equilibria, we notice that from the equation for the total population size, we have $\Lambda = \mu N$. The equations for the endemic equilibria are then given by

$$\begin{aligned} 0 &= \mu N - \beta\frac{SI}{N} - \mu S, \\ 0 &= \beta\frac{SI}{N} - p\beta\frac{EI}{N} - (\mu+\alpha)E, \\ 0 &= p\beta\frac{EI}{N} + \alpha E - \mu I. \end{aligned} \tag{7.38}$$

Because we are using standard incidence, the computations can be simplified if one sets

$$s = \frac{S}{N} \qquad e = \frac{E}{N} \qquad i = \frac{I}{N}.$$

With this notation, the system for the equilibria becomes

$$\begin{aligned} 0 &= \mu(1-s) - \beta si, \\ 0 &= \beta si - p\beta ei - (\mu+\alpha)e, \\ 0 &= p\beta ei + \alpha e - \mu i. \end{aligned} \tag{7.39}$$

This system does not necessarily have a unique solution that can be obtained explicitly. The general approach in this case is to express s and e in terms of i. Using the first and the third equations in system (7.39), we have

$$s = \frac{\mu}{\beta i + \mu} \qquad e = \frac{\mu i}{p\beta i + \alpha}. \tag{7.40}$$

Now one needs to replace s and e in equation two in system (7.39); however, equation two can be simplified first if added to equation three:

$$0 = \beta si - \mu e - \mu i.$$

Now we eliminate s and e from the above equation to obtain

$$\frac{\beta\mu i}{\beta i + \mu} - \frac{\mu^2 i}{p\beta i + \alpha} = \mu i.$$

Since we are looking for endemic equilibria, we have $i \neq 0$, and we can divide the above equation by μi:

$$\frac{\beta}{\beta i + \mu} - \frac{\mu}{p\beta i + \mu} = 1. \tag{7.41}$$

Rewriting this equation as a quadratic equation in i, we have

$$p\beta^2 i^2 + (\alpha\beta + \mu p\beta + \mu\beta - p\beta^2)i + \alpha\mu + \mu^2 - \alpha\beta = 0.$$

Recall the reproduction number $\mathscr{R}_0$. Dividing the above equation by β, we can rewrite it in the form

$$p\beta i^2 + (\alpha + \mu + \mu p - p\beta)i + \alpha\left(\frac{1}{\mathscr{R}_0} - 1\right) = 0. \tag{7.42}$$

Since $\mathscr{R}_0 = \alpha\beta/(\mu(\mu+\alpha))$, we want to express the above equation in terms of $\mathscr{R}_0$ and i. To this end, we have to replace one of the parameters that participates in Eq. (7.42) with $\mathscr{R}_0$. This parameter must be a part of $\mathscr{R}_0$, and it is typically taken to be β. Thus, we have

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

Substituting β in Eq. (7.42), we obtain the following equation in $\mathscr{R}_0$ and i:

$$p\frac{\mathscr{R}_0\mu(\mu+\alpha)}{\alpha}i^2+(\alpha+\mu+\mu p-p\frac{\mathscr{R}_0\mu(\mu+\alpha)}{\alpha})i+\alpha\left(\frac{1}{\mathscr{R}_0}-1\right)=0. \quad (7.43)$$

Multiplying by $\alpha/(\mu(\mu+\alpha))$, we obtain the following simplified equation:

$$p\mathscr{R}_0 i^2+\left(\frac{\alpha}{\mu}+\frac{p\alpha}{\mu+\alpha}-p\mathscr{R}_0\right)i+\frac{\alpha^2}{\mu(\mu+\alpha)}\left(\frac{1}{\mathscr{R}_0}-1\right)=0. \quad (7.44)$$

We notice that the above equation has a unique positive solution when $\mathscr{R}_0>1$, since the constant term is negative. We may consider Eq. (7.44) to define a curve in the $(\mathscr{R}_0,i)$ positive quadrant. This curve has one of two distinctive shapes (see Fig. 7.3). In the left figure, the endemic equilibrium curve bifurcates forward, giving a nontrivial endemic equilibrium when $\mathscr{R}_0>1$. This bifurcation is characterized by positive slope, that is, if we consider locally i as a function of $\mathscr{R}_0$, the derivative at the critical value $(\mathscr{R}_0,i)=(1,0)$ is positive:

$$\frac{\partial i}{\partial \mathscr{R}_0}|_{\mathscr{R}_0=1,i=0}>0.$$

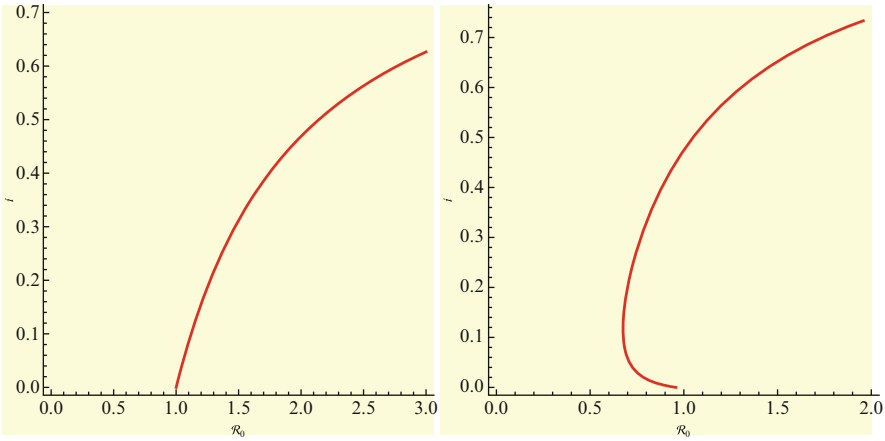

Fig. 7.3 Forward bifurcation and backward bifurcation of endemic equilibria. The *left* figure shows forward bifurcation at the critical value $\mathscr{R}_0=1$, in which case an endemic equilibrium exists if and only if $\mathscr{R}_0>1$. The *right* figure shows backward bifurcation at the critical value $\mathscr{R}_0=1$, in which case there are two endemic equilibria for $\mathscr{R}_0^*<\mathscr{R}_0<1$ and one endemic equilibrium if $\mathscr{R}_0\geq 1$

In the right figure, the endemic equilibrium curve bifurcates backward, giving a nontrivial endemic equilibrium when $\mathscr{R}_0>1$ and two nontrivial equilibria for $\mathscr{R}_0^*<\mathscr{R}_0<1$. This bifurcation is characterized by negative slope, that is, if we

consider locally i as a function of $\mathscr{R}_0$, the derivative at the critical value $(\mathscr{R}_0, i) = (1,0)$ is negative:

$$\frac{\partial i}{\partial \mathscr{R}_0}|_{\mathscr{R}_0=1,i=0} < 0.$$

We can use this inequality to derive a necessary and sufficient condition on the parameters such that backward bifurcation occurs. To obtain this condition, we compute the derivative $\frac{\partial i}{\partial \mathscr{R}_0}|_{\mathscr{R}_0=1,i=0}$ by differentiating Eq. (7.44) implicitly. In particular, differentiating implicitly with respect to $\mathscr{R}_0$, we have

$$pi^2 + p\mathscr{R}_0 2i\frac{\partial i}{\partial \mathscr{R}_0} - pi + \left(\frac{\alpha}{\mu} + \frac{p\alpha}{\mu+\alpha} - p\mathscr{R}_0\right)\frac{\partial i}{\partial \mathscr{R}_0} - \frac{\alpha^2}{\mu(\mu+\alpha)\mathscr{R}_0^2} = 0.$$

Now we evaluate at $\mathscr{R}_0 = 1$ and $i = 0$, and we solve for the derivative to get

$$\frac{\partial i}{\partial \mathscr{R}_0}|_{\mathscr{R}_0=1,i=0} = \frac{\frac{\alpha^2}{\mu(\mu+\alpha)}}{\left(\frac{\alpha}{\mu} + \frac{p\alpha}{\mu+\alpha} - p\right)}.$$

This derivative can be negative if and only if the denominator is negative. Hence, we obtain the following necessary and sufficient condition for backward bifurcation:

$$\frac{\alpha}{\mu} + \frac{p\alpha}{\mu+\alpha} - p < 0,$$

which can be solved for p to obtain

$$p > \frac{\alpha(\mu+\alpha)}{\mu^2}.$$

Recall that $1/\alpha$ is the duration of the exposed period, and $1/\mu$ is the lifespan of humans. Since the duration of the exposed period is smaller than the lifespan of humans, we may expect for realistic parameters to have $\alpha > \mu$. So backward bifurcation occurs only if $p > 1$. In other words, unless the infection at reinfection is more likely than the regular infection, backward bifurcation cannot occur. This is known not to be the case in reality [100]. Thus, backward bifurcation in this model is possible, but not for realistic parameter values.

In the remaining part of this subsection we will establish the stability of equilibria. Stability of equilibria, when multiple equilibria are present, is nontrivial. Typically for $\mathscr{R}_0 > 1$, the unique equilibrium is locally asymptotically stable. However, for $\mathscr{R}_0 < 1$, the two endemic equilibria have different stabilities. Most commonly, the lower one with smaller i is unstable, while the upper one with larger i is locally asymptotically stable. If more than two equilibria are present, typically their stabilities alternate, with the lowest one being unstable when $\mathscr{R}_0 < 1$ and the lowest one being locally stable when $\mathscr{R}_0 > 1$. This requires us to find a property of the equilibria

that separates one of the classes of equilibria from the other. We will return to this question later. Now we consider the characteristic equation of the endemic equilibrium. From (7.36), we have (recall the notation)

$$\begin{vmatrix} -\beta i+\beta si-\mu-\lambda & \beta si & -\beta s+\beta si \\ \beta i-\beta si+p\beta ei & -\beta si-p\beta i+p\beta ei-(\mu+\alpha)-\lambda & \beta s-\beta si-p\beta e+p\beta ei \\ -p\beta ei & p\beta i-p\beta ei+\alpha & p\beta e-p\beta ei-\mu-\lambda \end{vmatrix} = 0. \tag{7.45}$$

Adding the first row to the second row and the third row to the second row, we obtain the following equation:

$$\begin{vmatrix} -\beta i+\beta si-\mu-\lambda & \beta si & -\beta s+\beta si \\ -(\mu+\lambda) & -(\mu+\lambda) & -(\mu+\lambda) \\ -p\beta ei & p\beta i-p\beta ei+\alpha & p\beta e-p\beta ei-\mu-\lambda \end{vmatrix} = 0. \tag{7.46}$$

We can factor $-(\mu+\lambda)$ from the second row, and we see that one of the eigenvalues of the characteristic equation is $-\mu$. This simplification often works with epidemic models that do not involve disease-induced mortality. The remaining eigenvalues are solutions of the equation

$$\begin{vmatrix} -\beta i+\beta si-\mu-\lambda & \beta si & -\beta s+\beta si \\ 1 & 1 & 1 \\ -p\beta ei & p\beta i-p\beta ei+\alpha & p\beta e-p\beta ei-\mu-\lambda \end{vmatrix} = 0. \tag{7.47}$$

Expanding the determinant, we have

$$\begin{aligned} &-(\beta i(1-s)+\mu+\lambda)(p\beta e(1-i)-\mu-\lambda-p\beta i(1-e)-\alpha) \\ &-\beta si(p\beta e(1-i)-\mu-\lambda+p\beta ei) \\ &-\beta s(1-i)(p\beta i(1-e)+\alpha+p\beta ei)=0. \end{aligned} \tag{7.48}$$

Canceling the quadratic terms simplifies the equation to

$$\begin{gathered}(\beta i(1-s)+\mu+\lambda)(p\beta i-p\beta e+\mu+\alpha+\lambda)+\beta si(-p\beta e+\mu+\lambda) \\ -\beta s(1-i)(p\beta i+\alpha)=0.\end{gathered}$$

Further simplifications lead to the following equation:

$$(\beta i-\beta s+\mu+\lambda)(p\beta i+\alpha)+(\beta i+\mu+\lambda)(-p\beta e+\mu+\lambda)=0.$$

Denote by $\mathscr{F}(\lambda)$ the left-hand side of the equation above. Thus, the characteristic equation becomes $\mathscr{F}(\lambda)=0$. Then,

$$\mathscr{F}(0)=(\beta i-\beta s+\mu)(p\beta i+\alpha)+(\beta i+\mu)(-p\beta e+\mu).$$

Now we return to the question of distinguishing between alternative equilibria. Recall that all equilibria are solutions of Eq. (7.41). We define

$$f(i) = \frac{\beta}{\beta i + \mu} - \frac{\mu}{p\beta i + \mu} - 1.$$

The key observation is that alternative equilibria differ by the sign of the slope of $f(i)$ when computed for each specific equilibrium. When $\mathscr{R}_0 > 1$, $f(i)$ is a decreasing function of i with $f(0) > 0$. Thus at the equilibrium value i^* that satisfies $f(i^*) = 0$, we also have $f'(i^*) < 0$. When $\mathscr{R}_0 < 1$, we have $f(0) < 0$. Thus, if two distinct equilibria exist, the function $f(i)$ should increase to cross the x-axis and then decrease to cross it again. This gives two solutions of the equation $f(i) = 0$: i_1^* and i_2^* with $i_1^* < i_2^*$. At the first equilibrium, $f(i)$ is increasing, so we have $f'(i_1^*) > 0$. At the second equilibrium, $f(i)$ is decreasing, so we have $f'(i_2^*) < 0$. The main observation that we have to make is that

$$\mathscr{F}(0) \sim f'(i).$$

Thus, the sign of f' computed at a specific endemic equilibrium determines the sign of $\mathscr{F}(0)$. Deriving the relationship above is the trickiest part of the proof. To simplify matters, we notice that Eq. (7.41) gives

$$\beta(p\beta i + \alpha) - \mu(\beta i + \mu) = (\beta i + \mu)(p\beta i + \alpha).$$

We use this equality to replace the corresponding term in $\mathscr{F}(0)$, and $\mathscr{F}(0)$ becomes

$$\mathscr{F}(0) = (\beta i + \mu)(\mu - p\beta e) + \beta(p\beta i + \alpha) - \mu(\beta i + \mu) - \beta s(p\beta i + \alpha).$$

Then, using the expressions for s and e in (7.40), we have

$$\mathscr{F}(0) = (\beta i + \mu)\frac{\mu\alpha}{p\beta i + \alpha} + \beta(p\beta i + \alpha) - \mu(\beta i + \mu) - \frac{\beta\mu}{\beta i + \mu}(p\beta i + \alpha),$$

where

$$\mu - p\beta e = \mu - \frac{p\beta\mu i}{p\beta i + \alpha} = \frac{\mu\alpha}{p\beta i + \alpha}.$$

Collecting terms in $\mathscr{F}(0)$, we have

$$\begin{aligned}
\mathscr{F}(0) &= \beta(p\beta i + \alpha)\left[1 - \frac{\mu}{\beta i + \mu}\right] - \mu(\beta i + \mu)\left[1 - \frac{\alpha}{p\beta i + \alpha}\right] \\
&= \beta(p\beta i + \alpha)\frac{\beta i}{\beta i + \mu} - \mu(\beta i + \mu)\frac{p\beta i}{p\beta i + \alpha} \\
&= \beta i(p\beta i + \alpha)(\beta i + \mu)\left[\frac{\beta}{(\beta i + \mu)^2} - \frac{p\mu}{(p\beta i + \alpha)^2}\right] \\
&= -i(p\beta i + \alpha)(\beta i + \mu)f'(i). \qquad (7.49)
\end{aligned}$$

In the last equality, we have taken into account that

$$f'(i) = -\beta\left[\frac{\beta}{(\beta i+\mu)^2} - \frac{p\mu}{(p\beta i+\alpha)^2}\right].$$

Now we have that the following statements hold:

- $\mathscr{F}(0) > 0$ iff $f'(i) < 0$.
- $\mathscr{F}(0) = 0$ iff $f'(i) = 0$.
- $\mathscr{F}(0) < 0$ iff $f'(i) > 0$.

We first note that in the case $\mathscr{F}(0) = 0$, the characteristic equation $\mathscr{F}(\lambda) = 0$ has a solution $\lambda = 0$. In this case, the stability of the equilibrium in question cannot be determined. However, the case $\mathscr{F}(0) = 0$ occurs if and only if $f'(i^*) = 0$. Since $f(i^*) = 0$ and $f'(i^*) = 0$, it follows that i^* is a solution of the equation $f(i) = 0$ of higher multiplicity.

Definition 7.4. A root i^* of the equation $f(i) = 0$ is called *simple* if $f'(i^*) \neq 0$. A root i^* of $f(i) = 0$ is called a *root of higher multiplicity* if $f'(i^*) = 0$.

Hence, we conclude that if a root i^* is of higher multiplicity, then we cannot determine its stability. In what follows, we will assume that all roots of the equilibrium equation $f(i) = 0$ are simple roots. In that case, there are two possibilities.

Case 1. Let i^* be an equilibrium for which $f'(i^*) < 0$, that is, $\mathscr{F}(0) > 0$. Then such an equilibrium is locally asymptotically stable. Indeed, rewrite $\mathscr{F}(\lambda)$ in the form

$$\mathscr{F}(\lambda) = (\lambda + \beta i + \mu)(\lambda + \mu - p\beta e + p\beta i + \alpha) - \beta s(p\beta i + \alpha).$$

Suppose $\mathscr{F}(\lambda) = 0$ has a solution λ with $\Re\lambda = x \geq 0$ and $\Im\lambda = y$. Then, noticing that $\mu - p\beta e \geq 0$, we have

$$\begin{aligned} |\mathscr{F}(\lambda)| &\geq |\lambda + \beta i + \mu||\lambda + \mu - p\beta e + p\beta i + \alpha| - \beta s(p\beta i+\alpha) \\ &\geq \sqrt{(x+\beta i+\mu)^2+y^2}\sqrt{(x+\mu - p\beta e + p\beta i+\alpha)^2+y^2} - \beta s(p\beta i + \alpha) \\ &\geq (x+\beta i+\mu)(x+\mu - p\beta e + p\beta i + \alpha) - \beta s(p\beta i + \alpha) \\ &= \mathscr{F}(\Re\lambda) \geq \mathscr{F}(0) > 0. \end{aligned} \tag{7.50}$$

This is a contradiction. Hence, the equation $\mathscr{F}(\lambda) = 0$ cannot have solutions with $\Re\lambda \geq 0$. This implies that all solutions of the equation $\mathscr{F}(\lambda) = 0$ have negative real part. Hence, the equilibrium i^* is locally asymptotically stable.

Case 2. Let i^* be an equilibrium for which $f'(i^*) > 0$, that is, $\mathscr{F}(0) < 0$. Then such an equilibrium is unstable. Indeed, if we think that $\mathscr{F}(\lambda)$ is a function of a real variable λ, then $\mathscr{F}(0) < 0$, and since $\mathscr{F}(\lambda)$ is a quadratic polynomial, we have that

$$\lim_{\lambda\to\infty} \mathscr{F}(\lambda) = \infty.$$

Hence, there must be a real positive λ^* such that $\mathscr{F}(\lambda^*) = 0$. Therefore, the equilibrium i^* is unstable.

We summarize these conclusions in the following theorem.

Theorem 7.6. *If $\mathscr{R}_0 > 1$, then system (7.35) has a unique endemic equilibrium that is locally asymptotically stable. If $\mathscr{R}_0 < 1$ and*

$$p < \frac{\alpha(\mu+\alpha)}{\mu^2},$$

then system (7.35) has no endemic equilibria. If $\mathscr{R}_0^* < \mathscr{R}_0 < 1$ and

$$p > \frac{\alpha(\mu+\alpha)}{\mu^2},$$

then system (7.35) has two endemic equilibria. The one with the smaller number of infecteds, $\mathscr{E}_1$, is unstable, while the other, with a higher number of infecteds, $\mathscr{E}_2$, is locally asymptotically stable.

7.5.2 Castillo-Chavez and Song Bifurcation Theorem

The approach given in the previous subsection to determine the direction of the bifurcation at the critical point $(\mathscr{R}_0, i) = (1,0)$ is applicable only when a unique equation for the number/proportion of infected individuals can be derived, and that equation is reasonably simple to differentiate. As an alternative to this approach, a Theorem by Castillo-Chavez and Song [38] can be used that determines the direction of the bifurcation at the critical value of a parameter.

Theorem 7.7 (Castillo-Chavez and Song). *Consider the following general system of ODEs with a parameter ϕ:*

$$\frac{dx}{dt} = f(x,\phi), \qquad f : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}^n, \qquad f \in \mathbf{C}^2(\mathbf{R}^n \times \mathbf{R}), \tag{7.51}$$

where 0 is an equilibrium point of the system, that is, $f(0,\phi) \equiv 0$ for all ϕ. Assume the following:

A1. $\mathscr{A} = D_x f(0,0) = \left(\frac{\partial f_i}{\partial x_j}(0,0)\right)$ is the linearization matrix of system (7.51) around the equilibrium 0 with ϕ evaluated at 0. Zero is a simple eigenvalue of $\mathscr{A}$, and other eigenvalues have negative real parts.

A2. The matrix $\mathscr{A}$ has a nonnegative right eigenvector w and a left eigenvector v each corresponding to the zero eigenvalue.

Let f_k be the kth component of f and

$$a = \sum_{k,i,j=1}^{n} v_k w_i w_j \frac{\partial^2 f_k}{\partial x_i \partial x_j}(0,0)$$
$$b = \sum_{k,i=1}^{n} v_k w_i \frac{\partial^2 f_k}{\partial x_i \partial \phi}(0,0). \tag{7.52}$$

The local dynamics of the system around 0 *are completely determined by the signs of a and b:*

i. $a > 0, b > 0$. *When* $\phi < 0$ *with* $|\phi| \ll 1$, 0 *is locally asymptotically stable, and there exists a positive unstable equilibrium; when* $0 < \phi \ll 1$, 0 *is unstable, and there exists a negative and locally asymptotically stable equilibrium.*

ii. $a < 0, b < 0$. *When* $\phi < 0$ *with* $|\phi| \ll 1$, 0 *is unstable; when* $0 < \phi \ll 1$, 0 *is locally asymptotically stable, and there exists a positive unstable equilibrium;*

iii. $a > 0, b < 0$. *When* $\phi < 0$ *with* $|\phi| \ll 1$, 0 *is unstable, and there exists a locally asymptotically stable negative equilibrium; when* $0 < \phi \ll 1$, 0 *is stable, and a positive unstable equilibrium appears;*

iv. $a < 0, b > 0$. *When* ϕ *changes from negative to positive,* 0 *changes its stability from stable to unstable. Correspondingly, a negative unstable equilibrium becomes positive and locally asymptotically stable.*

Particularly, if $a > 0$ *and* $b > 0$, *then a backward bifurcation occurs at* $\phi = 0$.

Remark 7.1. In practice, the following two observations are important.

1. The equilibrium 0 is actually the disease-free equilibrium, ϕ is one of the parameters in the reproduction number, and the critical value of ϕ is the value of the parameter that makes the reproduction number equal to one.
2. Since the disease-free equilibrium has positive entries, the right eigenvector w need not be nonnegative. Components of the right eigenvector w that correspond to positive entries in the disease-free equilibrium could be negative. However, components that correspond to zero entries in the disease-free equilibrium should be nonnegative.

To illustrate this theorem, we apply it to determine the condition for backward bifurcation in model (7.35). We set $x_1 = S$, $x_2 = E$, and $x_3 = I$. System (7.35) becomes

$$\begin{aligned} x_1' &= \Lambda - \beta \frac{x_1 x_3}{x_1 + x_2 + x_3} - \mu x_1, \\ x_2' &= \beta \frac{x_1 x_3}{x_1 + x_2 + x_3} - p\beta \frac{x_2 x_3}{x_1 + x_2 + x_3} - (\mu + \alpha) x_2, \\ I' &= p\beta \frac{x_2 x_3}{x_1 + x_2 + x_3} + \alpha x_2 - \mu x_3. \end{aligned} \tag{7.53}$$

The role of the parameter ϕ is played by β with a critical value obtained from $\mathscr{R}_0 = 1$, which is given by

$$\beta^* = \frac{\mu(\mu+\alpha)}{\alpha}.$$

The disease-free equilibrium is given by $[x_1^* = \frac{\Lambda}{\mu}, x_2^* = 0, x_3^* = 0]$. The linearization around the disease-free equilibrium evaluated at β^* is given by (see (7.37))

$$\mathscr{A} = \begin{pmatrix} -\mu & 0 & -\beta^* \\ 0 & -(\mu+\alpha) & \beta^* \\ 0 & \alpha & -\mu \end{pmatrix}. \tag{7.54}$$

The characteristic equation is $|\mathscr{A} - \lambda I| = 0$, which expanded gives $-(\mu+\lambda)[(\mu+\alpha+\lambda)(\mu+\lambda) - \beta^*\alpha = 0$. The solutions of this equation, taking into account that $\beta^*\alpha = \mu(\mu+\alpha)$, are $\lambda_1 = -\mu$, $\lambda_2 = -(2\mu+\alpha)$, and $\lambda_3 = 0$. Hence 0 is a simple eigenvalue of $D_x f$. To compute a right eigenvector w, we consider the system $\mathscr{A}w = 0$. Assume $w = (x, y, z)$. The system becomes

$$\begin{aligned} -\mu x - \beta^* z &= 0, \\ -(\mu+\alpha)y + \beta^* z &= 0, \\ \alpha y - \mu z &= 0. \end{aligned} \tag{7.55}$$

We see that from the last two equations, we have

$$y = \frac{\mu}{\alpha} z.$$

The last two equations hold for every value of z and for y computed in terms of z from above. From the first equation, we have

$$x = -\frac{\mu+\alpha}{\alpha} z,$$

where we have replaced β^* with its equal. Hence the right eigenvector is $w = [-(\mu+\alpha), \mu, \alpha]$, where we have taken $z = \alpha$ to simplify the expressions for the components. We notice that the first component of w is negative but that is acceptable, since it corresponds to the first entry of the disease-free equilibrium, which is strictly positive. Next, we compute the left eigenvector. We have to solve $v\mathscr{A} = 0$. We set $v = (x, y, z)$, and the system becomes

$$\begin{aligned} -\mu x &= 0, \\ -(\mu+\alpha)y + \alpha z &= 0, \\ -\beta^* x + \beta^* y - \mu z &= 0. \end{aligned} \tag{7.56}$$

From the first equation, we have that $x = 0$. One solution of the remaining system is $v = [0, \alpha, (\mu+\alpha)]$, where we have taken into account the expression for β^*.

Since the first component of v is zero, we don't need the derivatives of f_1. From the derivatives of f_2 and f_3, the only ones that are nonzero are the following:

$$\frac{\partial^2 f_2}{\partial x_2 \partial x_3} = -(1+p)\frac{\beta^*}{x_1^*}, \qquad \frac{\partial^2 f_2}{\partial x_3^2} = -2\frac{\beta^*}{x_1^*},$$
$$\frac{\partial^2 f_3}{\partial x_2 \partial x_3} = p\frac{\beta^*}{x_1^*}, \qquad \frac{\partial^2 f_2}{\partial x_3 \partial \beta^*} = 1. \tag{7.57}$$

Hence, for backward bifurcation we need the following conditions to hold:

$$a = w_3\left[2v_2w_2\frac{\partial^2 f_2}{\partial x_2 \partial x_3} + v_2w_3\frac{\partial^2 f_2}{\partial x_3^2} + 2v_3w_2\frac{\partial^2 f_3}{\partial x_2 \partial x_3}\right] > 0,$$
$$b = v_2w_3 = \alpha^2 > 0. \tag{7.58}$$

The second condition always holds. From $a > 0$, we have

$$2v_2w_2\frac{\partial^2 f_2}{\partial x_2 \partial x_3} + v_2w_3\frac{\partial^2 f_2}{\partial x_3^2} + 2v_3w_2\frac{\partial^2 f_3}{\partial x_2 \partial x_3} > 0,$$

which gives

$$2\frac{\beta^*}{x_1^*}\left[-\alpha\mu(1+p) - \alpha^2 + \mu(\mu+\alpha)p\right] > 0.$$

This last inequality holds if and only if

$$p > \frac{\alpha(\mu+\alpha)}{\mu^2},$$

which is exactly the same condition for backward bifurcation that we previously obtained.

Problems

7.1. Lyapunov Function for an SIS Model
The following SIS model with disease-induced mortality has been studied in [50]:

$$\begin{aligned} S' &= \Lambda - \beta SI - \mu S + \phi I, \\ I' &= \beta SI - (\phi + \mu + \alpha)I, \end{aligned} \tag{7.59}$$

where ϕ is the recovery rate and α is the disease-induced mortality.

(a) The disease-free equilibrium is $(S^0, 0)$, where $S^0 = \Lambda/\mu$. The reproduction number of the model is

$$\mathscr{R}_0 = \frac{\Lambda\beta}{\mu(\phi+\mu+\alpha)}.$$

Use the following Lyapunov function to show that the disease-free equilibrium is globally stable whenever $\mathscr{R}_0 \leq 1$:

$$V(S,I) = \frac{1}{2}[(S-S^0)+I]^2 + \frac{(\alpha+2\mu)}{\beta} I.$$

(b) Show that the model has a unique endemic equilibrium $\mathscr{E}^* = (S^*, I^*)$, given by

$$S^* = \frac{\Lambda}{\mu \mathscr{R}_0} \qquad I^* = \frac{\mu(\phi+\alpha+\mu)}{\beta(\alpha+\mu)}(\mathscr{R}_0 - 1).$$

Show that the endemic equilibrium is locally asymptotically stable.

(c) Use the following Lyapunov function to show that the endemic equilibrium is globally stable:

$$V(S,I) = \frac{1}{2}[(S-S^*)+(I-I^*)]^2 + \frac{\alpha+2\mu}{\beta}\left(I - I^* - I^* \ln \frac{I}{I^*}\right).$$

Hint: You have to use the equations for the equilibrium.

7.2. Lyapunov Function for an SIRS Model

The following SIRS model with disease-induced mortality has been studied in [50]:

$$\begin{aligned} S' &= \Lambda - \beta SI - \mu S + \gamma I, \\ I' &= \beta SI - (\kappa + \mu + \alpha) I, \\ R' &= \kappa I - (\mu + \gamma) R, \end{aligned} \tag{7.60}$$

where κ is the recovery rate, α is the disease-induced mortality, and γ is the loss of immunity rate.

(a) The disease-free equilibrium is $(S^0, 0, 0)$, where $S^0 = \Lambda/\mu$. The reproduction number of the model is

$$\mathscr{R}_0 = \frac{\Lambda \beta}{\mu(\kappa + \mu + \alpha)}.$$

Use the following Lyapunov function to show that the disease-free equilibrium is globally stable whenever $\mathscr{R}_0 \leq 1$:

$$V(S,I,R) = \frac{1}{2}[(S-S^0)+I+R]^2 + \frac{\alpha+2\mu}{\beta} I + \frac{\alpha+2\mu}{2\kappa} R^2.$$

(b) Show that the model has a unique endemic equilibrium $\mathscr{E}^* = (S^*, I^*, R^*)$, given by

$$\begin{aligned} S^* &= \frac{\Lambda}{\mu \mathscr{R}_0}, \\ I^* &= \frac{\mu(\kappa+\alpha+\mu)(\mu+\gamma)}{\beta[\kappa\mu + (\alpha+\mu)(\mu+\gamma)]}(\mathscr{R}_0 - 1), \\ R^* &= \frac{\kappa\mu(\kappa+\alpha+\mu)}{\beta[\kappa\mu + (\alpha+\mu)(\mu+\gamma)]}(\mathscr{R}_0 - 1). \end{aligned} \tag{7.61}$$

Show that the endemic equilibrium is locally asymptotically stable.

(c) Use the following Lyapunov function to show that the endemic equilibrium is globally stable:

$$V(S,I,R) = \frac{1}{2}[(S-S^*)+(I-I^*)+(R-R^*)]^2 + \frac{\alpha+2\mu}{\beta}\left(I-I^*-I^*\ln\frac{I}{I^*}\right) + \frac{\alpha+2\mu}{2\kappa}(R-R^*)^2. \tag{7.62}$$

Hint: You have to use the equations for the equilibrium.

7.3. Lyapunov Function for an SEIS Model

The following SEIS model has been studied in [88]:

$$\begin{aligned} S' &= \mu-\beta SI-\mu S+\gamma I,\\ E' &= \beta SI-(\mu+\alpha)E,\\ I' &= \alpha I-(\mu+\gamma)I, \end{aligned} \tag{7.63}$$

where γ is the recovery rate, and α is the rate of progression to infectiousness.

(a) The disease-free equilibrium is $(1,0,0)$. The reproduction number of the model is

$$\mathscr{R}_0 = \frac{\alpha\beta}{(\gamma+\mu)(\mu+\alpha)}.$$

Use the following Lyapunov function to show that the disease-free equilibrium is globally stable whenever $\mathscr{R}_0 \le 1$:

$$V(S,I,R) = (S-1-\ln S)+E+\frac{\alpha+\mu}{\alpha}I.$$

(b) Show that the model has a unique endemic equilibrium $\mathscr{E}^* = (S^*,E^*,I^*)$, given by

$$\begin{aligned} S^* &= \frac{1}{\mathscr{R}_0},\\ E^* &= \frac{(\mu+\gamma)}{(\mu+\gamma+\alpha)}\left(1-\frac{1}{\mathscr{R}_0}\right),\\ I^* &= \frac{\alpha}{(\mu+\gamma+\alpha)}\left(1-\frac{1}{\mathscr{R}_0}\right). \end{aligned} \tag{7.64}$$

Show that the endemic equilibrium is locally asymptotically stable.

(c) Use the following Lyapunov function to show that the endemic equilibrium is globally stable:

$$V(S,E,I) = \left(S-S^*-S^*\ln\tfrac{S}{S^*}\right)+\left(E-E^*-E^*\ln\tfrac{E}{E^*}\right) \tfrac{\alpha+\mu}{\alpha}\left(I-I^*-I^*\ln\tfrac{I}{I^*}\right). \tag{7.65}$$

Hint: You have to use the equations for the equilibrium.

7.4. Lyapunov Function for a Vector–Host Model

The following vector–host model with disease-induced mortality has been studied in [51]:

$$\begin{aligned} S_h' &= \mu N_h - \kappa S_h I_v - \mu S_h + \gamma I_h, \\ I_h' &= \kappa S_h I_v - (\mu+\gamma) I_h, \\ S_v' &= \eta N_v - \beta S_v I_h - \eta S_v, \\ I_v' &= \beta S_v I_h - \eta I_v, \end{aligned} \tag{7.66}$$

where γ is the recovery rate, μ is the birth/death rate of humans, and η is the birth/death rate for vectors.

(a) The disease-free equilibrium is $(N_h, 0, N_v, 0)$. The reproduction number of the model is

$$\mathscr{R}_0 = \frac{\kappa\beta N_h N_v}{\eta(\mu+\gamma)}.$$

Use the following Lyapunov function to show that the disease-free equilibrium is globally stable whenever $\mathscr{R}_0 \le 1$.

$$\begin{aligned} V(S_h, I_h, S_v, I_v) &= \left(S_h - N_h - N_h \ln \frac{S_h}{N_h}\right) + I_h \\ &\quad + \frac{(\gamma+\mu)}{\beta N_v}\left(S_v - N_v - N_v \ln \frac{S_v}{N_v}\right) + I_v. \end{aligned} \tag{7.67}$$

(b) Show that the model has a unique endemic equilibrium $\mathscr{E}^* = (S^*, I^*, R^*)$, given by

$$\begin{aligned} S_h^* &= \frac{N_h((\mu+\gamma)\mathscr{R}_0 + \kappa N_v)}{(\kappa N_v + \mu + \gamma)\mathscr{R}_0}, \\ I_h^* &= \frac{\eta(\mu+\gamma)}{\beta[\kappa N_v + \mu + \gamma]}(\mathscr{R}_0 - 1), \\ S_v^* &= \frac{N_v(\kappa N_v + \mu + \gamma)}{(\kappa N_v + (\mu+\gamma)\mathscr{R}_0}, \\ I_v^* &= \frac{N_v(\mu+\gamma)}{\beta[\kappa N_v + (\mu+\gamma)\mathscr{R}_0]}(\mathscr{R}_0 - 1). \end{aligned} \tag{7.68}$$

Show that the endemic equilibrium is locally asymptotically stable.

(c) Use the following Lyapunov function to show that the endemic equilibrium is globally stable:

$$\begin{aligned} V(S_h, I_h, S_v, I_v) &= c_1\left(S_h - S_h^* - S_h^* \ln \frac{S_h}{S_h^*}\right) + c_2\left(I_h - I_h^* - I_h^* \ln \frac{I_h}{I_h^*}\right) \\ &\quad + c_3\left(S_v - S_v^* - S_v^* \ln \frac{S_v}{S_v^*}\right) + c_4\left(I_v - I_v^* - I_v^* \ln \frac{I_v}{I_v^*}\right), \end{aligned} \tag{7.69}$$

where $c_1 = c_2 = \beta S_v^* I_h^*$ and $c_3 = c_4 = \kappa S_h^* I_v^*$.

Hint: You have to use the equations for the equilibrium.

7.5. Oscillations in a Malaria Vector Model
Let $U(t)$ be the class of fertilized nourished reproducing female vectors, $W(t)$ the class of fertilized but nonreproducing vectors in quest of a blood meal, and let $V(t)$ represent the vectors that have just laid eggs and are resting [125]. The total vector population is given by $N_v = V + U + W$. Let H be the population of humans at the breeding site, assumed constant. Consider the following model:

$$\begin{aligned} U' &= p\tau HW - (a+\mu)U, \\ V' &= \frac{a\rho U}{Q+U} + aU - \left(\mu + \frac{bH}{H+K}\right)V, \\ W' &= \left(\frac{bH}{H+K}\right)V - (\mu + \tau H)W, \end{aligned} \tag{7.70}$$

where $0 \leq p \leq 1$ is the probability that a mosquito will succeed in taking a blood meal and τ is the contact rate; a, b, K, Q, and, ρ are positive constants.

(a) Determine the equilibria of the system.
(b) Show that the unique nontrivial equilibrium can become destabilized, and oscillations are possible.

7.6. Backward Bifurcation in an SIS Model with Education
The following SIS model with education was proposed in [67].

$$\begin{aligned} S_1' &= \mu(1-\kappa) - \psi S_1 + \delta S_2 - \beta_1 S_1 I + \alpha(1-\sigma)I - \mu S_1, \\ S_2' &= \mu\kappa + \psi S_1 - \delta S_2 - \beta_2 S_2 I + \alpha\sigma I - \mu S_2, \\ I' &= (\beta_1 S_1 + \beta_2 S_2)I - (\alpha+\mu)I, \end{aligned} \tag{7.71}$$

where S_1 are the usual susceptibles, while S_2 are the educated susceptibles, I are the infectious individuals, μ is the fertility–mortality parameter with fraction κ going into the educated class, ψ is the transmission rate from the normal to the educated state, δ is the transition rate from the educated to normal state, β_1 and β_2 are the transmission rates to normal and educated susceptibles.

(a) Compute the disease-free equilibrium and the reproduction number of the model. Show that the disease-free equilibrium is locally stable if $\mathscr{R}_0 < 1$ and unstable if $\mathscr{R}_0 > 1$.
(b) Reduce the equations for the endemic equilibria to a single equation in I^*. Show that backward bifurcation can occur. Derive a necessary and sufficient condition for this to happen.
(c) Use Theorem 7.7 to show that backward bifurcation occurs.

7.7. Backward Bifurcation in an SIR Model with Saturating Treatment
The following SIR model with education was proposed in [170].

$$
\begin{aligned}
S' &= \Lambda - \frac{\beta SI}{1+\kappa I} - \mu S,\\
I' &= \frac{\beta SI}{1+\kappa I} - (\mu+\gamma+\varepsilon)I - \frac{\alpha I}{\omega+I},\\
R' &= \gamma I + \frac{\alpha I}{\omega+I} - \mu R,
\end{aligned}
\tag{7.72}
$$

where S are the susceptibles, I are the infectious individuals, and R are the recovered; μ is the mortality parameter, γ is the recovery rate, ε is the disease-induced death rate, α, ω, and κ are positive constants.

(a) Compute the disease-free equilibrium and the reproduction number of the model. Show that the disease-free equilibrium is locally stable if $\mathscr{R}_0 < 1$ and unstable if $\mathscr{R}_0 > 1$.
(b) Reduce the equations for the endemic equilibria to a single equation in I^*. Show that backward bifurcation can occur. Derive a necessary and sufficient condition for this to happen.
(c) Use Theorem 7.7 to show that backward bifurcation occurs.

7.8. Backward Bifurcation in an SEI Risk-Structured Model
The following SEI model with education was proposed in [66].

$$
\begin{aligned}
S_l' &= (1-p)\Lambda + \psi_h S_h - \frac{\beta S_l I}{N} - (\mu+\psi_l)S_l,\\
S_h' &= p\Lambda + \psi_l S_l - \frac{\kappa\beta S_h I}{N} - (\mu+\psi_h)S_h,\\
E' &= \frac{\beta(S_l+\kappa S_h)I}{N} - (\mu+\sigma)E,\\
I' &= \sigma E - (\mu+\gamma+\delta)I,
\end{aligned}
\tag{7.73}
$$

where S_l are the low-risk susceptibles, S_h are the high-risk susceptibles, E are the exposed individuals, and I are the infectious individuals; μ is the mortality parameter, γ is the recovery rate, ψ_l and ψ_h are transition rates between low- and high-risk susceptible classes, δ is the disease-induced death rate, σ is the transition rate to infectiousness, and κ is a positive constant.

(a) Compute the disease-free equilibrium and the reproduction number of the model. Show that the disease-free equilibrium is locally stable if $\mathscr{R}_0 < 1$ and unstable if $\mathscr{R}_0 > 1$.
(b) Use Theorem 7.7 to show that backward bifurcation occurs. Derive a condition for backward bifurcation.

Chapter 8
Multistrain Disease Dynamics

8.1 Competitive Exclusion Principle

The causative agents of diseases, such as viruses and bacteria, may be represented by multiple variants, called by the general name *strains* (or subtypes). The presence of multiple strains of a pathogen complicates our ability to combat the disease. For instance, in influenza, it is believed that each strain imparts permanent immunity, but drift evolution creates new strains, and in a new flu season, we can contract the disease again. There are other diseases whose causative agents are represented by multistrain pathogens. For example, Haemophilus influenzae, which is responsible for a range of infections, is represented by six serotypes: a, b, c, d, e, and f, as well as some variants that are not typeable. *Streptococcus pneumonae* is represented by 90 serotypes. Dengue virus has four serotypes. In this chapter, we study models with multiple strains.

In ecology, the competitive exclusion principle is one of the main principles that govern species competition. The principle was first formulated as a law in the 1930s by the Russian ecologist Georgy Gause [64], who discovered it based on laboratory experiments. In short, the principle can be stated as *complete competitors cannot coexist*. In this section, we introduce through models the epidemiological context of this principle.

8.1.1 A Two-Strain Epidemic SIR Model

To account for the genetic variability of the pathogen, we investigate an SIR epidemic model with multiple strains. The the two-strain model has the following compartments: $S(t)$ is the number of susceptible individuals at time t, $I_1(t)$ is the number of individuals infected by strain one, $I_2(t)$ is the number of individuals infected by strain two, and $R(t)$ is the number of recovered individuals. We assume

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_8

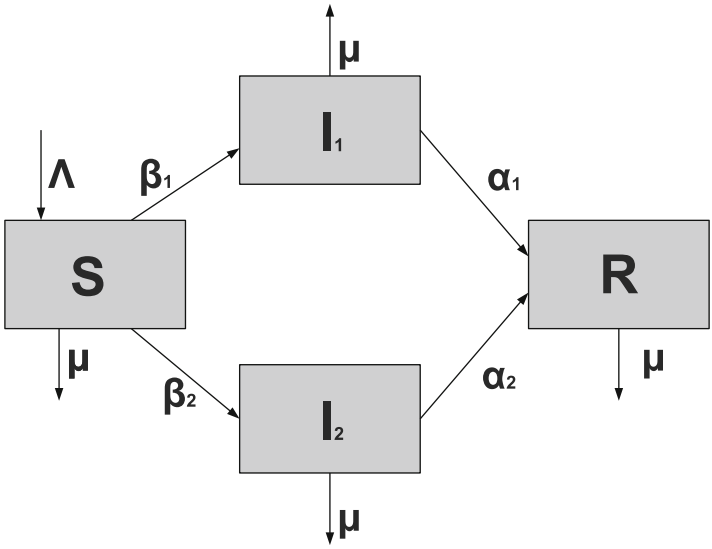

Fig. 8.1 A flowchart of an SIR model with two strains and perfect immunity

that each strain, once contracted, imparts permanent immunity to itself and to the other strain.

Susceptible individuals can become infected either by strain one at a transmission rate β_1 or by strain two at a transmission rate β_2. Those infected with strain one recover at a rate α_1, and those infected with strain two recover at a rate α_2. Susceptible individuals are recruited at a rate Λ. Individuals in all classes die at a natural death rate μ. The model is given below. The flowchart of the model is given in Fig. 8.1:

$$\begin{aligned} S' &= \Lambda - \beta_1 \frac{SI_1}{N} - \beta_2 \frac{SI_2}{N} - \mu S, \\ I_1' &= \beta_1 \frac{SI_1}{N} - (\mu + \alpha_1) I_1 \\ I_2' &= \beta_2 \frac{SI_2}{N} - (\mu + \alpha_2) I_2 \\ R' &= \alpha_1 I_1 + \alpha_2 I_2 - \mu R \end{aligned} \tag{8.1}$$

The total population size N is given by $N(t) = S(t) + I_1(t) + I_2(t) + R(t)$. Adding the equations above, we see that the equation of the total population size is the simplified logistic $N'(t) = \Lambda - \mu N$.

The next step will be to compute the equilibria. Working with a standard incidence allows us to consider the equilibrium equations of the proportions. For the equilibrial values of the variables S, I_1, I_2, and R, we set $s = S/N$, $i_1 = I_1/N$, $i_2 = I_2/N$, $r = R/N$. Since N satisfies the equation $0 = \Lambda - \mu N$, and hence $\Lambda = \mu N$, the equations for the equilibrium proportions are given by

$$\begin{aligned} 0 &= \mu - \beta_1 s i_1 - \beta_2 s i_2 - \mu s, \\ 0 &= \beta_1 s i_1 - (\mu + \alpha_1) i_1, \\ 0 &= \beta_2 s i_2 - (\mu + \alpha_2) i_2, \\ 0 &= \alpha_1 i_1 + \alpha_2 i_2 - \mu r. \end{aligned} \tag{8.2}$$

Model (8.1) has three equilibria. The model has a disease-free equilibrium in which neither strain one nor strain two is present. At the disease-free equilibrium, we have $s = 1, i_1 = 0, i_2 = 0, r = 0$. Hence, the disease-free equilibrium in the original variables is given by $\mathscr{E}_0 = \left(\frac{\Lambda}{\mu}, 0, 0, 0\right)$. In contrast to single-strain models, multi-strain models have multiple reproduction numbers, one for each strain. To define the reproduction numbers associated with strain one and strain two, we have to look at the local stability of the disease-free equilibrium. Computing the Jacobian at the disease-free equilibrium yields

$$J = \begin{pmatrix} -\mu & -\beta_1 & -\beta_2 & 0 \\ 0 & \beta_1 - (\mu + \alpha_1) & 0 & 0 \\ 0 & 0 & \beta_2 - (\mu + \alpha_2) & 0 \\ 0 & \alpha_1 & \alpha_2 & -\mu \end{pmatrix}. \tag{8.3}$$

The characteristic equation $|J - \lambda I| = 0$ has one double eigenvalue $\lambda_1 = \lambda_2 = -\mu$, and the following two eigenvalues:

$$\begin{aligned} \lambda_3 &= \beta_1 - (\mu + \alpha_1), \\ \lambda_4 &= \beta_2 - (\mu + \alpha_2). \end{aligned} \tag{8.4}$$

The disease-free equilibrium is locally asymptotically stable if $\lambda_3 < 0$ and $\lambda_4 < 0$. We notice that the eigenvalue λ_3 is associated with strain one and gives rise to the reproduction number of strain one, $\mathscr{R}_1$. The eigenvalue λ_4 is associated with strain two and gives rise to the reproduction number of strain two, $\mathscr{R}_2$. Thus, we define

$$\mathscr{R}_1 = \frac{\beta_1}{\mu + \alpha_1}, \qquad \mathscr{R}_2 = \frac{\beta_2}{\mu + \alpha_2}. \tag{8.5}$$

We have the following result:

Proposition 8.1. *The disease-free equilibrium of system (8.1) is locally asymptotically stable if both reproduction numbers are less than* 1*, that is if*

$$\mathscr{R}_1 < 1 \qquad \mathscr{R}_2 < 1.$$

The disease-free equilibrium is unstable if at least one of the above inequalities is reversed.

8.1.2 The Strain-One- and Strain-Two-Dominance Equilibria and Their Stability

A strain-one-dominance equilibrium is a boundary equilibrium in which strain one is present $i_1 \neq 0$, while strain two is not present $i_2 = 0$. From the second equation in (8.2), we have

$$s = \frac{\mu + \alpha_1}{\beta_1} = \frac{1}{\mathscr{R}_1}.$$

We need $s < 1$ for it to be a proper fraction. Thus, for a strain-one-dominance equilibrium to be meaningful, we need $\mathscr{R}_1 > 1$. To compute the value of the infected individuals with strain one, we start from the first equation:

$$\frac{\mu}{s} = \beta_1 i_1 + \mu.$$

Replacing s with $1/\mathscr{R}_1$ and solving for i_1, we have

$$\begin{aligned} i_1 &= \frac{\mu}{(\mu + \alpha_1)\mathscr{R}_1} = \frac{\mu}{\mu + \alpha_1}\left(1 - \frac{1}{\mathscr{R}_1}\right), \\ r &= \frac{\alpha_1}{\mu} i_1 = \frac{\alpha_1}{\mu + \alpha_1}\left(1 - \frac{1}{\mathscr{R}_1}\right). \end{aligned} \tag{8.6}$$

The strain-one-dominance equilibrium is given by

$$\mathscr{E}_1 = \left(\frac{1}{\mathscr{R}_1}\frac{\Lambda}{\mu}, \frac{\mu}{\mu + \alpha_1}\left(1 - \frac{1}{\mathscr{R}_1}\right)\frac{\Lambda}{\mu}, 0, \frac{\alpha_1}{\mu + \alpha_1}\left(1 - \frac{1}{\mathscr{R}_1}\right)\frac{\Lambda}{\mu}\right).$$

A strain-two-dominance equilibrium is a boundary equilibrium in which strain two is present $i_2 \neq 0$, while strain one is not present $i_1 = 0$. A strain-two-dominance equilibrium exists if and only if $\mathscr{R}_2 > 1$. The strain-two-dominance equilibrium is given by

$$\mathscr{E}_2 = \left(\frac{1}{\mathscr{R}_2}\frac{\Lambda}{\mu}, 0, \frac{\mu}{\mu + \alpha_2}\left(1 - \frac{1}{\mathscr{R}_2}\right)\frac{\Lambda}{\mu}, \frac{\alpha_2}{\mu + \alpha_2}\left(1 - \frac{1}{\mathscr{R}_2}\right)\frac{\Lambda}{\mu}\right).$$

An equilibrium for which both strain one and strain two are present, that is, $i_1 \neq 0$ and $i_2 \neq 0$, is called a *coexistence equilibrium*. The equilibrium value of s in a coexistence equilibrium satisfies the following two equations, which are obtained from (8.2):

$$\begin{aligned} 0 &= \beta_1 s - (\mu + \alpha_1), \\ 0 &= \beta_2 s - (\mu + \alpha_2). \end{aligned} \tag{8.7}$$

The first equation requires

$$s = \frac{\mu + \alpha_1}{\beta_1} = \frac{1}{\mathscr{R}_1}.$$

The second equation requires

$$s = \frac{\mu + \alpha_2}{\beta_2} = \frac{1}{\mathscr{R}_2}.$$

These two expressions for s can be consistent if and only if $\mathscr{R}_1 = \mathscr{R}_2$. Hence, in the generic case in which the two strains have different reproduction numbers, a coexistence equilibrium does not exist.

To investigate the stability of the dominance equilibria, we consider the Jacobian at an equilibrium $\mathscr{E} = (s, i_1, i_2, r)$. Because $N(t)$ is asymptotically constant, we treat it as constant. It can be checked that this does not change the results.

$$J = \begin{pmatrix} -\beta_1 i_1 - \beta_2 i_2 - \mu & -\beta_1 s & -\beta_2 s & 0 \\ \beta_1 i_1 & \beta_1 s - (\mu + \alpha_1) & 0 & 0 \\ \beta_2 i_2 & 0 & \beta_2 s - (\mu + \alpha_2) & 0 \\ 0 & \alpha_1 & \alpha_2 & -\mu \end{pmatrix}. \tag{8.8}$$

To determine the local stability of a strain-one-dominance equilibrium, we consider the Jacobian at that equilibrium:

$$J(s, i_1, 0, r) = \begin{pmatrix} -\beta_1 i_1 - \mu & -\beta_1 s & -\beta_2 s & 0 \\ \beta_1 i_1 & \beta_1 s - (\mu + \alpha_1) & 0 & 0 \\ 0 & 0 & \beta_2 s - (\mu + \alpha_2) & 0 \\ 0 & \alpha_1 & \alpha_2 & -\mu \end{pmatrix}. \tag{8.9}$$

The Jacobian has one eigenvalue $\lambda_1 = -\mu$, and another

$$\lambda_2 = \beta_2 s - (\mu + \alpha_2) = (\mu + \alpha_2)\left(\frac{\mathscr{R}_2}{\mathscr{R}_1} - 1\right).$$

The eigenvalue λ_2 is called the *growth rate of strain two when strain one is at equilibrium*. The remaining eigenvalues of the strain-one-dominance equilibrium are the eigenvalues of the 2×2 matrix

$$\begin{pmatrix} -\beta_1 i_1 - \mu & -\beta_1 s \\ \beta_1 i_1 & \beta_1 s - (\mu + \alpha_1) \end{pmatrix}. \tag{8.10}$$

We notice that the entry in the second row, second column in the matrix above is $\beta_1 s - (\mu + \alpha_1) = 0$, since $s = 1/\mathscr{R}_1$. Hence, the above matrix has Tr $= -\beta_1 i_1 - \mu < 0$ and Det $= \beta_1 s \beta_1 i_1 > 0$. Thus, the eigenvalues of the above matrix are negative or have negative real part. We conclude that the local stability of the dominance equilibrium $\mathscr{E}_1$ depends on the sign of the eigenvalue λ_2. A dominance equilibrium $\mathscr{E}_1$ is locally asymptotically stable if and only if $\lambda_2 < 0$. Consequently, the dominance equilibrium $\mathscr{E}_1$ is locally asymptotically stable if and only if $\mathscr{R}_2 < \mathscr{R}_1$, that is, when strain one has a larger reproduction number than strain two. By symmetry,

a dominance equilibrium $\mathscr{E}_2$ is locally asymptotically stable if and only if $\mathscr{R}_2 > \mathscr{R}_1$, that is, when strain two has a larger reproduction number than strain one. We summarize these results in the following theorem:

Theorem 8.1. *A strain-j-dominance equilibrium exists if and only if $\mathscr{R}_j > 1$. If $\mathscr{R}_1 > \mathscr{R}_2$, then the strain-one-dominance equilibrium is locally asymptotically stable. If $\mathscr{R}_1 < \mathscr{R}_2$, it is unstable. If $\mathscr{R}_1 < \mathscr{R}_2$, then the strain-two-dominance equilibrium is locally asymptotically stable. If $\mathscr{R}_1 > \mathscr{R}_2$, it is unstable. Coexistence is not possible outside of the degenerate case $\mathscr{R}_1 = \mathscr{R}_2$.*

Figure 8.2 gives the competitive outcomes for the two strains. The competitive outcomes are also listed in Table 8.1.

Table 8.1 Competitive outcomes for the two-strain model (8.1)

Region	Long-term behavior	Competitive outcome
$\mathscr{R}_1 < 1, \mathscr{R}_2 < 1$	$I_1(t) \to 0, I_2(t) \to 0$	Both strains die out
$\mathscr{R}_1 > 1, \mathscr{R}_2 < 1$	$I_1(t)$ persists, $I_2(t) \to 0$	Strain 1 dominates
$\mathscr{R}_1 < 1, \mathscr{R}_2 > 1$	$I_1(t) \to 0, I_2(t)$ persists	Strain 2 dominates
$\mathscr{R}_1 > 1, \mathscr{R}_2 > 1, \mathscr{R}_1 > \mathscr{R}_2$	$I_1(t)$ persists, $I_2(t) \to 0$	Strain 1 dominates
$\mathscr{R}_1 > 1, \mathscr{R}_2 > 1, \mathscr{R}_1 < \mathscr{R}_2$	$I_1(t) \to 0, I_2(t)$ persists	Strain 2 dominates

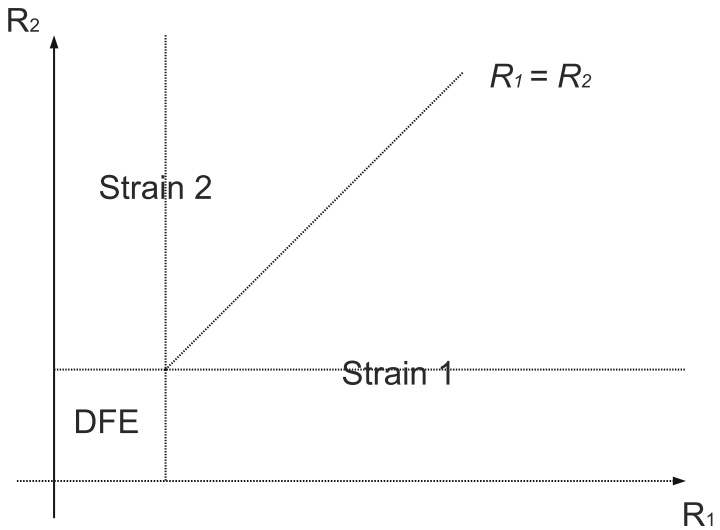

Fig. 8.2 Strain-one and strain-two dominance regions. The 45° line is the line $\mathscr{R}_1 = \mathscr{R}_2$. Above that line is the region $\mathscr{R}_2 > \mathscr{R}_1$, where strain two outcompetes strain one. Below this line is the region $\mathscr{R}_2 < \mathscr{R}_1$, where strain one outcompetes strain two

8.1.3 The Competitive Exclusion Principle

Theorem 8.1 and Table 8.1 state that based on *local results*, we can conclude that when two strains in the population compete, the strain with the larger reproduction number outcompetes the other strain and drives it to extinction. This local result is the foundation of the **competitive exclusion principle**. However, for the competitive exclusion principle to hold, this result has to be *global*, that is, it has to hold for all values of the initial conditions. In this subsection, we will establish the global validity of the competitive outcomes in Table 8.1. First, we formulate the competitive exclusion principle for n strains.

Competitive Exclusion Principle:
When n strains compete in a population, the strain with the largest reproduction number outcompetes the other strains and drives them to extinction.

The global results that support the competitive exclusion principle are formulated and established in the following theorem [29]:

Theorem 8.2. *If $\mathscr{R}_1 < 1$ and $\mathscr{R}_2 < 1$, then the DFEis globally asymptotically stable. If $\mathscr{R}_1 > 1$ and/or $\mathscr{R}_2 > 1$, then the strain with the largest reproduction number persists, and the other one dies out. Coexistence is not possible outside of the degenerate case $\mathscr{R}_1 = \mathscr{R}_2$.*

Proof. First, assume that $\mathscr{R}_1 < 1$. Then from Eq. (8.1), we have

$$I'(t) \leq \beta_1 I_1 - (\mu + \alpha_1) I_1 = (\mu + \alpha_1)(\mathscr{R}_1 - 1) I_1,$$

where we have used the fact that $S/N \leq 1$. It is clear from the above inequality that if $\mathscr{R}_1 < 1$, then $I_1(t) \to 0$ as $t \to \infty$. A similar result holds if $\mathscr{R}_2 < 1$. This establishes the global stability of the disease-free equilibrium.

Now if at least one of the reproduction numbers is larger than one, we proceed as follows. Assume $\mathscr{R}_1 > \mathscr{R}_2$. Since the system is symmetric with respect to strain one and strain two, we can derive results similar to those in the case $\mathscr{R}_1 < \mathscr{R}_2$. To see this, set

$$\xi(t) = \frac{I_1^{\beta_2}}{I_2^{\beta_1}}.$$

Differentiating ξ with respect to time yields

$$\xi' = \frac{\beta_2 I_1^{\beta_2 - 1} I_1' I_2^{\beta_1} - \beta_1 I_1^{\beta_2} I_2' I_2^{\beta_1 - 1}}{[I_2^{\beta_1}]^2}.$$

Substituting I_1' and I_2' from the model equation (8.1), the numerator of ξ' becomes

$$\begin{aligned}
&\beta_2 I_1^{\beta_2-1} I_2^{\beta_1}[\beta_1 I_1 s - (\mu+\alpha_1)I_1] - \beta_1 I_1^{\beta_2} I_2^{\beta_1-1}[\beta_2 I_2 s - (\mu+\alpha_2)I_2] \\
&= \beta_2 I_1^{\beta_2} I_2^{\beta_1}[\beta_1 s - (\mu+\alpha_1)] - \beta_1 I_1^{\beta_2} I_2^{\beta_1}[\beta_2 s - (\mu+\alpha_2)] \\
&= I_1^{\beta_2} I_2^{\beta_1}[\beta_1\beta_2 s - \beta_2(\mu+\alpha_1) - \beta_1\beta_2 s + \beta_1(\mu+\alpha_2)] \\
&= I_1^{\beta_2} I_2^{\beta_1}(\mu+\alpha_1)(\mu+\alpha_2)[\mathscr{R}_1 - \mathscr{R}_2]. \qquad (8.11)
\end{aligned}$$

Thus, the differential equation for ξ becomes

$$\xi'(t) = \frac{I_1^{\beta_2} I_2^{\beta_1}(\mu+\alpha_1)(\mu+\alpha_2)[\mathscr{R}_1 - \mathscr{R}_2]}{[I_2^{\beta_1}]^2}.$$

Hence, ξ satisfies the differential equation $\xi' = \nu\xi$, where $\nu = (\mu+\alpha_1)(\mu+\alpha_2)[\mathscr{R}_1 - \mathscr{R}_2]$. The solution to this equation is given by $\xi(t) = \xi(0)e^{\nu t}$. Therefore, $\xi(t) \to \infty$ as $\nu > 0$. Since I_1 is bounded, the only way that could happen is if $I_2(t) \to 0$ as $t \to \infty$.

□

Question: What do strains compete for? Answer: The strains compete for susceptible individuals.

Ecological Interpretation of the Competitive Exclusion Principle: From an ecological perspective, the two strains can be viewed as two consumers competing for a common "resource"—the susceptible individuals. The competitive exclusion principle in this case states that only the consumer that can persist on the lower value of the resource persists; the other one is excluded. The resource that each strain needs to persist is $s = S/N$. For strain one, the value of the resource needed for persistence is $s = 1/\mathscr{R}_1$. For strain two, the value of the resource needed for persistence is $s = 1/\mathscr{R}_2$. Thus, the strain with the larger reproduction number can persist on a lower value of the resource.

8.2 Multistrain Diseases: Mechanisms for Coexistence

In various natural environments, many species of microorganisms stably coexist for long periods of time by interacting with each other. For example, in tuberculosis, drug-sensitive and drug-resistant variants of the causative agent have been around for a while. Dengue's four serotypes also coexist in nature. So, if the outcome of the simplest multistrain model (8.1) is competitive exclusion, what causes the long-term coexistence of pathogen variants? Several mechanisms have been identified as

causing stable coexistence of pathogens in epidemic models. Such mechanisms are called *trade-off mechanisms*.

8.2.1 Mutation

Mutations are changes in the DNA or RNA sequence of a microorganism. Mutations are caused by errors that occur during DNA or RNA replication. Microorganisms (such as viruses) that use RNA as their genetic material have rapid mutation rates, which can be an advantage, since those pathogens evolve constantly and rapidly, developing different antigenic characteristics and thus evading the defensive responses of the human immune system.

Mutation is accounted for in epidemic models through a term that transfers individuals infected with one of the strains into individuals infected with the other. To illustrate how mutation is treated in epidemic models, we introduce a two-strain SIR epidemic model with mutation. The model is very similar to the competitive

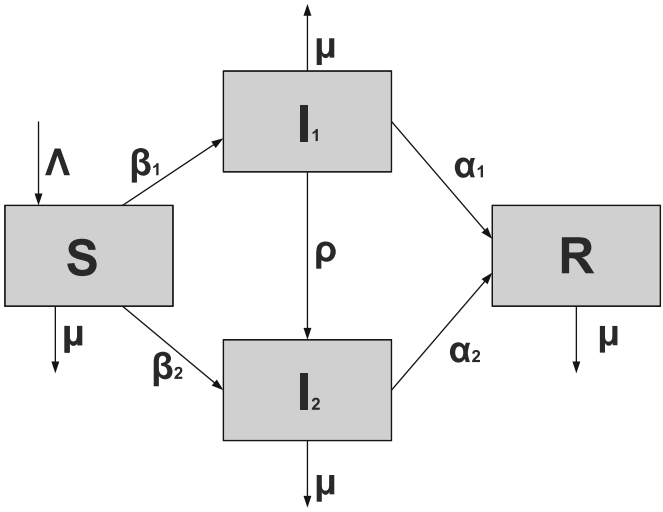

Fig. 8.3 Flowchart of a two-strain SIR epidemic model with mutation

exclusion model (8.1) but includes the mutation of strain-one-infected individuals into strain-two-infected individuals at a mutation rate ρ. The flowchart of the model is given in Fig. 8.3.

The model, first introduced in [25], is given below. The notation is the same as in model (8.1):

$$
\begin{aligned}
S' &= \Lambda - \beta_1 \frac{SI_1}{N} - \beta_2 \frac{SI_2}{N} - \mu S, \\
I_1' &= \beta_1 \frac{SI_1}{N} - (\mu + \alpha_1 + \rho) I_1, \\
I_2' &= \beta_2 \frac{SI_2}{N} - (\mu + \alpha_2) I_2 + \rho I_1, \\
R' &= \alpha_1 I_1 + \alpha_2 I_2 - \mu R.
\end{aligned} \tag{8.12}
$$

We mention that mutation incorporated in this way is modeled as a continuous event.

8.2.2 Superinfection

Superinfection is the process by which an individual that has previously been infected by one pathogen variant becomes infected with a different strain of the pathogen, or another pathogen at a later point in time. The second strain is assumed to "take over" the infected individual immediately. Thus this individual becomes infected with the second strain of the pathogen.

Superinfection is accounted for in epidemic models through a term that transfers individuals infected with one of the strains into individuals infected with the other.

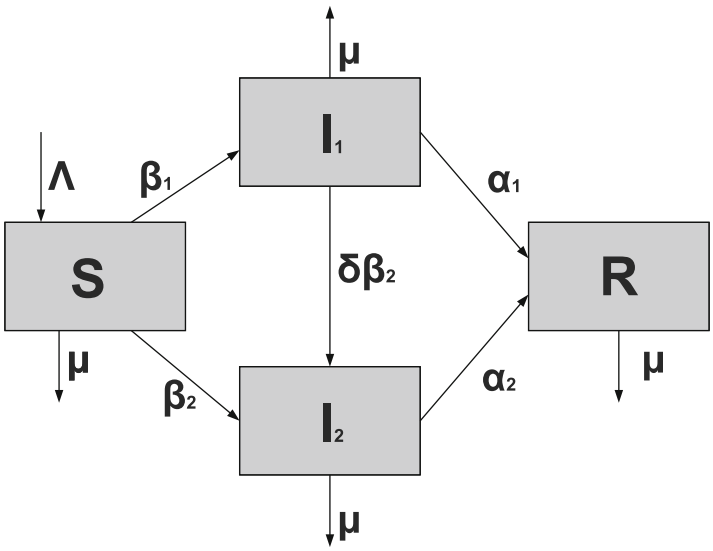

Fig. 8.4 Flowchart of a two-strain SIR epidemic model with superinfection

To illustrate how superinfection is treated in epidemic models, we introduce a two-strain SIR epidemic model with superinfection. The model is very similar to the competitive exclusion model (8.1) but includes the superinfection of individuals

infected with strain one by individuals infected with strain two. The transmission rate β_2 at superinfection is reduced or enhanced by δ. If $\delta < 1$, then the transmission rate β_2 is reduced; if $\delta > 1$, then the transmission rate is enhanced. The flowchart of the model is given in Fig. 8.4.

The model, first introduced in [126], is given below. Notation is the same as in model (8.1):

$$\begin{aligned} S' &= \Lambda - \beta_1 \frac{SI_1}{N} - \beta_2 \frac{SI_2}{N} - \mu S, \\ I_1' &= \beta_1 \frac{SI_1}{N} - \delta\beta_2 \frac{I_1 I_2}{N} - (\mu + \alpha_1) I_1, \\ I_2' &= \beta_2 \frac{SI_2}{N} + \delta\beta_2 \frac{I_1 I_2}{N} - (\mu + \alpha_2) I_2, \\ R' &= \alpha_1 I_1 + \alpha_2 I_2 - \mu R. \end{aligned} \tag{8.13}$$

In the above model, individuals infected with strain two can superinfect individuals infected with strain one. That is, individuals infected with strain one who come into contact with individuals infected with strain two can become immediately infected with strain two. One open question with the superinfection model (8.13) is what happens if the superinfection goes in both directions. To investigate this option, suppose in addition to strain two superinfecting strain one in the model above, we have that strain one also superinfects strain two. Hence model (8.13) takes the form

$$\begin{aligned} S' &= \Lambda - \beta_1 \frac{SI_1}{N} - \beta_2 \frac{SI_2}{N} - \mu S, \\ I_1' &= \beta_1 \frac{SI_1}{N} - \delta\beta_2 \frac{I_1 I_2}{N} + \delta_1\beta_1 \frac{I_1 I_2}{N} - (\mu + \alpha_1) I_1, \\ I_2' &= \beta_2 \frac{SI_2}{N} + \delta\beta_2 \frac{I_1 I_2}{N} - \delta_1\beta_1 \frac{I_1 I_2}{N} - (\mu + \alpha_2) I_2, \\ R' &= \alpha_1 I_1 + \alpha_2 I_2 - \mu R. \end{aligned} \tag{8.14}$$

It can be seen that the two superinfection terms in each of the equations for I_1' and I_2' are the same except for their coefficients. This means that they can be combined. For instance,

$$-\delta\beta_2 \frac{I_1 I_2}{N} + \delta_1\beta_1 \frac{I_1 I_2}{N} = \left(-\delta\beta_2 + \delta_1\beta_1\right) \frac{I_1 I_2}{N}.$$

The constant coefficient is either positive or negative. If it is negative, it can be written as

$$-\delta\beta_2 + \delta_1\beta_1 = -\hat{\delta}\beta_2.$$

Hence, the equation for I_2' takes the same form as in system (8.13). The expression in the equation for I_2' in (8.14) is the same but with the opposite sign. We conclude that the symmetric system (8.14) is mathematically equivalent to the asymmetric system (8.13). For that reason, typically only the asymmetric system (8.13) is investigated.

8.2.3 Coinfection

Coinfection is the process of infection of a single host with two or more pathogen variants (strains) or with two or more distinct pathogen species. Coinfection with multiple pathogen strains is particularly common in HIV, but it occurs in many other diseases. Coinfection with multiple pathogen species is also thought to be a very common occurrence. Particularly widely distributed combinations are HIV and tuberculosis, HIV and hepatitis, HIV and malaria, and others. Coinfection is of significant importance because it may have negative effect both on the health of the coinfected individuals as well as on the public health in general. For instance, a coinfection of a human or a pig with human influenza strain and H5N1 strain may result in a pandemic strain, causing a widespread deadly pandemic.

To model coinfection, we need to introduce a new dependent variable, namely $J(t)$, the number of coinfected individuals in the population. The model again is built on the basis of the competitive exclusion model (8.1), but with a coinfected class J. Individuals infected with strain one can become coinfected with strain two and move to the coinfected class, and similarly for individuals originally infected with strain two. A typical assumption is that the probability of a susceptible individual getting infected with both strains simultaneously is too small and can be neglected. Both infected individuals with strain one and coinfected individuals can infect with strain one, and similarly with strain two. Recovery in a coinfection model is complex. Jointly infected individuals can recover from strain one, thereby moving to the class I_2, or they can recover from strain two and move to the class I_1. The possibility that jointly infected individuals recover from both classes also exists. In that case, they move to the recovered class R.

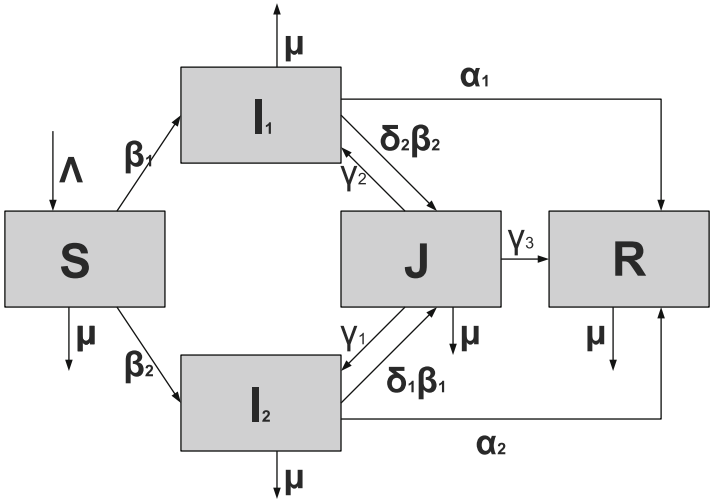

Fig. 8.5 Flowchart of a two-strain SIR coinfection model

The flowchart of the model is given in Fig. 8.5. The model, previously introduced in [113], takes the form

$$\begin{aligned}
S' &= \Lambda - \beta_1 \frac{S(I_1+J)}{N} - \beta_2 \frac{S(I_2+J)}{N} - \mu S,\\
I_1' &= \beta_1 \frac{S(I_1+J)}{N} - \delta_2\beta_2 \frac{I_1(I_2+J)}{N} - (\mu+\alpha_1)I_1 + \gamma_2 J,\\
I_2' &= \beta_2 \frac{S(I_2+J)}{N} - \delta_1\beta_1 \frac{(I_1+J)I_2}{N} - (\mu+\alpha_2)I_2 + \gamma_1 J,\\
J' &= \delta_1\beta_1 \frac{(I_1+J)I_2}{N} + \delta_2\beta_2 \frac{I_1(I_2+J)}{N} - (\mu+\gamma_1+\gamma_2+\gamma_3)J,\\
R' &= \alpha_1 I_1 + \alpha_2 I_2 + \gamma_3 J - \mu R.
\end{aligned} \tag{8.15}$$

Here some of the new parameters have the following meanings: δ_i is a coefficient of reduction/enhancement of infection during coinfection, α_i is the recovery rate of strain i during a single-strain infection, γ_i is the recovery rate of strain i during coinfection, and γ_3 is the recovery of jointly infected individuals from both strains simultaneously.

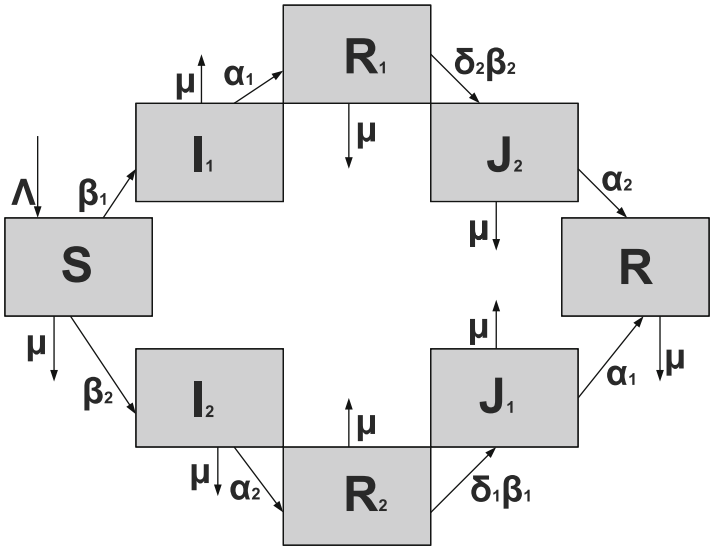

Fig. 8.6 A flowchart of a two-strain model with cross-immunity

8.2.4 Cross-Immunity

Cross-immunity is a form of immunity in which prior infection with one variant of the pathogen renders partial protection against another variant of the same pathogen or a different pathogen. Cross immunity, just like coinfection, can apply to strains of the same microorganism or can apply to different pathogen species. In the first

case, the most notable example is influenza, where infection with one strain often provides some sort of immunity to other influenza strains. Mathematical models also have suggested that short-lived cross-immunity may exist among the four dengue serotypes [2]. In terms of cross-protective immunity between two distinct pathogens, mathematical models have suggested that such may exist between leprosy and tuberculosis [99]. Another example of cross-reactivity that has been confirmed in humans is one that involves the influenza virus and the hepatitis C virus [165].

To introduce the model, besides the traditional classes S, I_1, and I_2, there are also the classes of the recovered individuals from strain i, denoted by R_i, the class J_2 of individuals recovered from strain one and now infected with strain two, and symmetrically, the class J_1 recovered from strain two and now infected with strain one, and finally, the class R of individuals recovered from both strains. The cross-immunity that strain i provides to strain j is incorporated in σ_j.

The flowchart of the model is given in Fig. 8.6. The model, previously introduced in [37], takes the form

$$\begin{aligned}
S' &= \Lambda - \beta_1 \frac{S(I_1+J_1)}{N} - \beta_2 \frac{S(I_2+J_2)}{N} - \mu S,\\
I_1' &= \beta_1 \frac{S(I_1+J_1)}{N} - (\mu+\alpha_1) I_1,\\
R_1' &= \alpha_1 I_1 - \sigma_2 \beta_2 \frac{R_1(I_2+J_2)}{N} - \mu R_1,\\
J_1' &= \sigma_1 \beta_1 \frac{R_2(I_1+J_1)}{N} - (\mu+\alpha_1) J_1,\\
I_2' &= \beta_2 \frac{S(I_2+J_2)}{N} - (\mu+\alpha_2) I_2,\\
R_2' &= \alpha_2 I_2 - \delta_1 \beta_1 \frac{(I_1+J) I_2}{N} - \mu R_2,\\
J_2' &= \sigma_2 \beta_2 \frac{R_1(I_2+J_2)}{N} - (\mu+\alpha_2) J_2,\\
R' &= \alpha_1 J_1 + \alpha_2 J_2 - \mu R.
\end{aligned} \tag{8.16}$$

There are other trade-off mechanisms that are known to induce coexistence of multiple pathogens. What causes coexistence of pathogens is still an open question of particular interest. Several articles review the literature devoted to multistrain interaction [109, 152]. *Multistrain interactions play a particularly important role, since most diseases that are still a public health concern today are caused by very mutable pathogens.*

8.3 Analyzing Two-Strain Models with Coexistence: The Case of Superinfection

In this section, we consider a simple model of a disease without recovery where the strains interact through superinfection. An example of a disease that can be described by the model is HIV. In modeling HIV, we need to include disease-induced mortality and standard incidence to account for the fact that the contacts do not grow linearly with the population size. The model with superinfection takes the form

$$\begin{aligned} S' &= \Lambda - \beta_1 \frac{SI_1}{N} - \beta_2 \frac{SI_2}{N} - \mu S, \\ I_1' &= \beta_1 \frac{SI_1}{N} - \delta\beta_2 \frac{I_1 I_2}{N} - (\mu + \alpha_1) I_1, \\ I_2' &= \beta_2 \frac{SI_2}{N} + \delta\beta_2 \frac{I_1 I_2}{N} - (\mu + \alpha_2) I_2. \end{aligned} \tag{8.17}$$

8.3.1 Existence and Stability of the Disease-Free and Two Dominance Equilibria

Most of the analysis of multistrain models follows the analysis of single-strain models. In particular, multistrain models also have a disease-free equilibrium, that is, an equilibrium in which none of the strains is present. If we set the derivatives to zero, the system for the equilibria satisfies

$$\begin{aligned} \Lambda - \beta_1 \frac{SI_1}{N} - \beta_2 \frac{SI_2}{N} - \mu S &= 0, \\ \beta_1 \frac{SI_1}{N} - \delta\beta_2 \frac{I_1 I_2}{N} - (\mu + \alpha_1) I_1 &= 0, \\ \beta_2 \frac{SI_2}{N} + \delta\beta_2 \frac{I_1 I_2}{N} - (\mu + \alpha_2) I_2 &= 0. \end{aligned} \tag{8.18}$$

For the disease-free equilibrium, we have $I_1 = 0$ and $I_2 = 0$. This gives $S = \frac{\Lambda}{\mu}$. Hence, the disease-free equilibrium is $\mathscr{E}_0 = \left(\frac{\Lambda}{\mu}, 0, 0\right)$. To compute the reproduction numbers of each strain, we look at the stability of the disease-free equilibrium. The Jacobian takes the form

$$J = \begin{pmatrix} -\mu & -\beta_1 & -\beta_2 \\ 0 & \beta_1 - (\mu + \alpha_1) & 0 \\ 0 & 0 & \beta_2 - (\mu + \alpha_2) \end{pmatrix}. \tag{8.19}$$

This gives the following two reproduction numbers of strain one and strain two:

$$\mathscr{R}_1 = \frac{\beta_1}{\mu+\alpha_1}, \qquad \mathscr{R}_2 = \frac{\beta_2}{\mu+\alpha_2}. \tag{8.20}$$

Setting $I_j = 0$, we can compute the dominance equilibrium of strain i for $i \neq j = 1,2$. We obtain the following two dominance equilibria that correspond to each strain:

$$\mathscr{E}_1 = \left(\frac{\Lambda}{\mu}\frac{1}{\mathscr{R}_1}, \frac{\Lambda}{\mu}\left(1-\frac{1}{\mathscr{R}_1}\right), 0\right), \qquad \mathscr{E}_2 = \left(\frac{\Lambda}{\mu}\frac{1}{\mathscr{R}_2}, 0, \frac{\Lambda}{\mu}\left(1-\frac{1}{\mathscr{R}_2}\right)\right). \tag{8.21}$$

It is clear from the expressions for the equilibria that the dominance equilibrium for strain i exists and is positive if and only if $\mathscr{R}_i > 1$.

We summarize our observations thus far in the following two theorems:

Theorem 8.3. *The disease-free equilibrium $\mathscr{E}_0$ always exists. In addition, there is a dominance equilibrium corresponding to strain one, $\mathscr{E}_1$, if and only if $\mathscr{R}_1 > 1$. Analogously, there is a dominance equilibrium corresponding to strain two, $\mathscr{E}_2$, if and only if $\mathscr{R}_2 > 1$.*

Regarding the stability of the disease-free equilibrium, we have the following result:

Proposition 8.2. *If $\mathscr{R}_1 < 1$ and $\mathscr{R}_2 < 1$, then the disease-free equilibrium is locally stable. If $\mathscr{R}_1 > 1$ and/or $\mathscr{R}_2 > 1$, then the disease-free equilibrium is unstable.*

Because this result always holds, it is somewhat customary to define an overall reproduction number. In particular, $\mathscr{R}_0$ is defined as follows:

$$\mathscr{R}_0 = \max\{\mathscr{R}_1, \mathscr{R}_2\}. \tag{8.22}$$

This definition is consistent with the next-generation approach for defining the reproduction number as the maximum element in the spectrum of the next-generation matrix.

To investigate the local stability of the dominance equilibria, we consider the Jacobian of system (8.17). Since this model uses standard incidence and the denominator is strictly nonconstant, we need to be careful in computing the Jacobian. In particular, we need to keep in mind that $N = S + I_1 + I_2$ and differentiate the denominator too. The generic form of the Jacobian is given by

$$J = \begin{pmatrix} j_{11} & -\beta_1 s + \beta_1 s i_1 + \beta_2 s i_2 & \beta_1 s i_1 - \beta_2 s + \beta_2 s i_2 \\ \beta_1 i_1 - \beta_1 s i_1 + \delta\beta_2 i_1 i_2 & j_{22} & -\beta_1 s i_1 - \delta\beta_2 i_1 + \delta\beta_2 i_1 i_2 \\ \beta_2 i_2 - \beta_2 s i_2 - \delta\beta_2 i_1 i_2 & -\beta_2 s i_2 + \delta\beta_2 i_2 - \delta\beta_2 i_1 i_2 & j_{33} \end{pmatrix}, \tag{8.23}$$

where

$$\begin{aligned} j_{11} &= -\beta_1 i_1 + \beta_1 s i_1 - \beta_2 i_2 + \beta_2 s i_2 - \mu, \\ j_{22} &= \beta_1 s - \beta_1 s i_1 - \delta\beta_2 i_2 + \delta\beta_2 i_1 i_2 - (\mu+\alpha_1), \\ j_{33} &= \beta_2 s - \beta_2 s i_2 + \delta\beta_2 i_1 - \delta\beta_2 i_1 i_2 - (\mu+\alpha_2), \end{aligned} \tag{8.24}$$

and $s = S/N$, $i_1 = I_1/N$, and $i_2 = I_2/N$ evaluated at the respective equilibria.

The local stability of the equilibrium $\mathscr{E}_1$ is given by the Jacobian evaluated at $\mathscr{E}_1$:

$$J_1 = \begin{pmatrix} -\beta_1 i_1 + \beta_1 s i_1 - \mu & -\beta_1 s + \beta_1 s i_1 & \beta_1 s i_1 - \beta_2 s \\ \beta_1 i_1 - \beta_1 s i_1 & \beta_1 s - \beta_1 s i_1 - (\mu + \alpha_1) & -\beta_1 s i_1 - \delta\beta_2 i_1 \\ 0 & 0 & \beta_2 s + \delta\beta_2 i_1 - (\mu + \alpha_2) \end{pmatrix}. \tag{8.25}$$

Two entries in the last row are zeros, while the third one gives an eigenvalue of the 3×3 matrix J_1:

$$\lambda_1 = \beta_2 s + \delta\beta_2 i_1 - (\mu + \alpha_2).$$

The other two eigenvalues are the eigenvalues of the submatrix

$$J_{11} = \begin{pmatrix} -\beta_1 i_1 + \beta_1 s i_1 - \mu & -\beta_1 s + \beta_1 s i_1 \\ \beta_1 i_1 - \beta_1 s i_1 & \beta_1 s - \beta_1 s i_1 - (\mu + \alpha_1) \end{pmatrix}. \tag{8.26}$$

We notice that since $s = 1/\mathscr{R}_1$, we have $\beta_1 s - (\mu + \alpha_1) = 0$. This kind of simplification is typical for epidemic models, and one should always remember to simplify the Jacobian. The matrix J_{11} takes the form

$$J_{11} = \begin{pmatrix} -\beta_1 i_1 + \beta_1 s i_1 - \mu & -\beta_1 s + \beta_1 s i_1 \\ \beta_1 i_1 - \beta_1 s i_1 & -\beta_1 s i_1 \end{pmatrix}. \tag{8.27}$$

To determine the stability of the matrix J_{11}, we apply Theorem 3.2 from Chap. 3. Since this is a 2×2 matrix, its eigenvalues will have negative real parts if $\mathrm{Tr}J_{11} < 0$ and $\mathrm{Det}J_{11} > 0$. We can easily see that $\mathrm{Tr}J_{11} = -\beta_1 i_1 + \beta_1 s i_1 - \mu - \beta_1 s i_1 = -\beta_1 i_1 - \mu < 0$. Furthermore, we note that $-\beta_1 i_1 + \beta_1 s i_1 = -\beta_1 i_1 (1 - s) < 0$. Hence, the j_{11} entry in J_{11} is negative. Similarly, the j_{12} entry is negative, the j_{21} entry is positive, and $j_{22} < 0$. This implies that $\mathrm{Det}J_{11} = j_{11} j_{22} - j_{12} j_{21} > 0$. Thus the eigenvalues of submatrix J_{11} have negative real parts. Therefore, stability of $\mathscr{E}_1$ is determined by the sign of the eigenvalue λ_1. The equilibrium $\mathscr{E}_1$ will be locally asymptotically stable if and only if $\lambda_1 < 0$, that is, if and only if

$$\beta_2 s + \delta\beta_2 i_1 - (\mu + \alpha_2) < 0. \tag{8.28}$$

Replacing s and i_1 with their respective values, we have

$$\beta_2 \frac{1}{\mathscr{R}_1} + \delta\beta_2 \left(1 - \frac{1}{\mathscr{R}_1}\right) < \mu + \alpha_2.$$

As with the reproduction number, we rewrite this inequality as a quantity smaller than one:

$$\frac{\mathscr{R}_2}{\mathscr{R}_1} + \delta\mathscr{R}_2 \left(1 - \frac{1}{\mathscr{R}_1}\right) < 1.$$

The quantity on the left-hand side of the above inequality is called the **invasion reproduction number** or **invasion number** of strain two at the equilibrium of strain one, and it is denoted by $\hat{\mathscr{R}}_2^1$. Hence, we define the invasion reproduction number as

$$\hat{\mathscr{R}}_2^1 = \frac{\mathscr{R}_2}{\mathscr{R}_1} + \delta \mathscr{R}_2 \left(1 - \frac{1}{\mathscr{R}_1}\right). \tag{8.29}$$

Mathematically, the invasion number gives a threshold for the stability of a dominance equilibrium. Since the eigenvalue λ_1 gives the growth rate of strain two when strain one is at equilibrium, and we have

$$\lambda_1 < 0 \qquad \Longleftrightarrow \qquad \hat{\mathscr{R}}_2^1 < 1,$$

it follows that equilibrium $\mathscr{E}_1$ is locally asymptotically stable if and only if $\hat{\mathscr{R}}_2^1 < 1$. Thus, strain two cannot grow when strain one is at equilibrium if and only if $\hat{\mathscr{R}}_2^1 < 1$. In this case, we say that strain two cannot invade the equilibrium of strain one.

Epidemiologically, the invasion number of strain two at the equilibrium of strain one gives the number of secondary infections one individual infected with strain two will produce in a population in which strain one is at equilibrium during its lifetime as infectious.

Just like the reproduction number, the invasion reproduction number must be positive and interpretable. Sometimes, meeting these two conditions simultaneously may be a challenge. However, $\hat{\mathscr{R}}_2^1$ is clearly positive. For the interpretation, we consider the invasion number in the form (see (8.28))

$$\hat{\mathscr{R}}_2^1 = \frac{\beta_2 s}{\mu + \alpha_2} + \frac{\delta \beta_2 i_1}{\mu + \alpha_2}.$$

$\beta_2 \dfrac{SI_2}{N} + \delta \beta_2 \dfrac{I_1 I_2}{N}$ Gives the number of secondary infections that I_2 individuals can produce per unit of time.

$\beta_2 s + \delta \beta_2 i_1$ Gives the number of secondary infections that *one* strain-two-infected individual can produce per unit of time when the proportion of susceptibles is s and the proportion of those infected with strain one is i_1 at equilibrium.

$\dfrac{\beta_2 s + \delta \beta_2 i_1}{\mu + \alpha_2}$ Gives the number of secondary infections that one strain-two-infected individual can produce during its lifetime as infectious in a population where strain one is at equilibrium.

We note that $\dfrac{\beta_2 s}{\mu + \alpha_2}$ gives the number of secondary infections of susceptible individuals, while $\dfrac{\delta \beta_2 i_1}{\mu + \alpha_2}$ gives the number of secondary infections of strain-one-infected individuals. Therefore, the invasion number counts also the infections of already infected individuals.

Next we consider the stability of the dominance equilibrium of strain two, $\mathscr{E}_2$. This stability is governed by the **invasion number** of strain one at the equilibrium of strain two. Using (8.23) we derive the Jacobian at $\mathscr{E}_2$:

$$J_2 = \begin{pmatrix} -\beta_2 i_2 + \beta_2 s i_2 - \mu & -\beta_1 s + \beta_2 s i_2 & -\beta_2 s + \beta_2 s i_2 \\ 0 & \beta_1 s - \delta\beta_2 i_2 - (\mu + \alpha_1) & 0 \\ \beta_2 i_2 - \beta_2 s i_2 & -\beta_2 s i_2 + \delta\beta_2 i_2 & \beta_2 s - \beta_2 s i_2 - (\mu + \alpha_2) \end{pmatrix}. \tag{8.30}$$

Looking at the second row of J_2, we see that one of the eigenvalues is

$$\lambda_2 = \beta_1 s - \delta\beta_2 i_2 - (\mu + \alpha_1).$$

The other two eigenvalues are the eigenvalues of the matrix obtained by deleting the second row and second column of J_2:

$$J_{22} = \begin{pmatrix} -\beta_2 i_2 + \beta_2 s i_2 - \mu & -\beta_2 s + \beta_2 s i_2 \\ \beta_2 i_2 - \beta_2 s i_2 & \beta_2 s - \beta_2 s i_2 - (\mu + \alpha_2) \end{pmatrix}. \tag{8.31}$$

Since $s = 1/\mathscr{R}_2$, as before we have $\beta_2 s - (\mu + \alpha_2) = 0$. This simplifies the matrix to the form

$$J_{22} = \begin{pmatrix} -\beta_2 i_2 + \beta_2 s i_2 - \mu & -\beta_2 s + \beta_2 s i_2 \\ \beta_2 i_2 - \beta_2 s i_2 & -\beta_2 s i_2 \end{pmatrix}. \tag{8.32}$$

Similar reasoning as before shows that the entries of that matrix have the following signs: $j_{11} < 0$, $j_{12} < 0$, $j_{21} > 0$ and $j_{22} < 0$. From here, it is easy to see that $\mathrm{Tr}J < 0$ and $\mathrm{Det}J > 0$. Hence, Theorem 3.2 in Chap. 3 implies that the eigenvalues of the matrix J_{22} have negative real parts. Therefore, the stability of $\mathscr{E}_2$ is determined by the sign of the eigenvalue λ_2. The equilibrium $\mathscr{E}_2$ is locally asymptotically stable if and only if $\lambda_2 < 0$, that is, if and only if

$$\beta_1 s - \delta\beta_2 i_2 - (\mu + \alpha_1) < 0. \tag{8.33}$$

Solving the inequality so that we have 1 on the RHS, we obtain

$$\frac{\beta_1 s}{\delta\beta_2 i_2 + \mu + \alpha_1} < 1. \tag{8.34}$$

Replacing the values of s and i_2 from $\mathscr{E}_2$, we obtain the following invasion number:

$$\hat{\mathscr{R}}_1^2 = \frac{\beta_1 \frac{1}{\mathscr{R}_2}}{\delta\beta_2 \left(1 - \frac{1}{\mathscr{R}_2}\right) + \mu + \alpha_1}. \tag{8.35}$$

As before, the epidemiological interpretation of $\hat{\mathscr{R}}_1^2$ is that it gives the number of secondary infections that one strain-one-infected individual can produce in a population where strain two is at equilibrium. Rewriting inequality (8.33) in the form

(8.34) means that we treat the rate of infection of the unique strain-one infectious individual by individuals infected by strain two as a decrease of the "lifespan" of that individual infected with strain one. Equilibrium $\mathscr{E}_2$ is locally asymptotically stable if and only if

$$\lambda_2 < 0 \quad \Longleftrightarrow \quad \hat{\mathscr{R}}_1^2 < 1.$$

That is, equilibrium $\mathscr{E}_2$ is stable when the growth rate λ_2 of strain one when strain two is at equilibrium is negative, that is $\lambda_2 < 0$, or equivalently, when the invasion number of strain one, $\hat{\mathscr{R}}_1^2$, is less than 1, that is, when strain one cannot invade the equilibrium of strain two. These results can be summarized in the following theorem:

Theorem 8.4. *Let $\mathscr{R}_1 > 1$ and $\mathscr{R}_2 > 1$.*

1. *The dominance equilibrium of strain one $\mathscr{E}_1$ is locally asymptotically stable if and only if $\hat{\mathscr{R}}_2^1 < 1$, that is, if and only if strain two cannot invade the equilibrium of strain one.*
2. *The dominance equilibrium of strain two $\mathscr{E}_2$ is locally asymptotically stable if and only if $\hat{\mathscr{R}}_1^2 < 1$, that is, if and only if strain one cannot invade the equilibrium of strain two.*

In the competitive exclusion case considered in Sect. 8.1, if $\mathscr{R}_1 > 1$ and $\mathscr{R}_2 > 1$, then exactly one of the dominance equilibria was stable. Here, if $\hat{\mathscr{R}}_1^2 > 1$, $\hat{\mathscr{R}}_2^1 > 1$, then they are both unstable. In this case, we will see that there may also be an equilibrium in which both strains are present. Such an equilibrium is called an interior equilibrium or a coexistence equilibrium.

8.3.2 Existence of the Coexistence Equilibrium

Under some conditions, system (8.17) has equilibria in which both strains are present. If $\mathscr{E}^* = (S^*, I_1^*, I_2^*)$ is an equilibrium, then if $I_1^* \neq 0$ *and* $I_2^* \neq 0$, this equilibrium is called a **coexistence equilibrium**. Coexistence equilibria are nontrivial solutions of the system (8.18). To solve this system, we first rewrite it in terms of the proportions. Let $s = S/N$, $i_1 = I_1/N$, and $i_2 = I_2/N$. From system (8.18), we can compute the equation of the total population size:

$$\Lambda - \mu N - \alpha_1 I_1 - \alpha_2 I_2 = 0.$$

Thus,

$$\frac{\Lambda}{N} = \mu + \alpha_1 i_1 + \alpha_2 i_2.$$

Dividing each equation in (8.18) by N and using the above expression, we arrive at the following system for the proportions:

$$\begin{aligned} \mu + \alpha_1 i_1 + \alpha_2 i_2 - \beta_1 s i_1 - \beta_2 s i_2 - \mu s &= 0, \\ \beta_1 s i_1 - \delta\beta_2 i_1 i_2 - (\mu + \alpha_1) i_1 &= 0, \\ \beta_2 s i_2 + \delta\beta_2 i_1 i_2 - (\mu + \alpha_2) i_2 &= 0. \end{aligned} \tag{8.36}$$

Since $i_1 \neq 0$ and $i_2 \neq 0$, we can cancel them in the second and third equations. Then we use the second and third equations to express i_1 and i_2 in terms of s:

$$i_1 = \frac{(\mu + \alpha_2) - \beta_2 s}{\delta\beta_2} \qquad i_2 = \frac{\beta_1 s - (\mu + \alpha_1)}{\delta\beta_2}. \tag{8.37}$$

From the equation for the total population size $N = S + I_1 + I_2$, we have a corresponding equation for the proportions $s + i_1 + i_2 = 1$. We substitute expressions (8.37) in $s + i_1 + i_2 = 1$, take a common denominator, and obtain the following equation for s:

$$\delta\beta_2 s + (\mu + \alpha_2) - \beta_2 s + \beta_1 s - (\mu + \alpha_1) = \delta\beta_2.$$

This equation can be solved for s, and it gives the following value:

$$s = \frac{\delta\beta_2 + (\mu + \alpha_1) - (\mu + \alpha_2)}{\delta\beta_2 - \beta_2 + \beta_1}.$$

At this point, we do not know whether the expression for s is positive or less than 1. We also want $i_1 > 0$ and $i_2 > 0$. To see this, for i_1, we have

$$i_1 = \frac{\mu + \alpha_2}{\delta\beta_2} - \frac{1}{\delta}\frac{\delta\beta_2 + (\mu + \alpha_1) - (\mu + \alpha_2)}{\delta\beta_2 - \beta_2 + \beta_1}.$$

We want $i_1 > 0$. Consequently, canceling δ, which is positive, we want

$$\frac{1}{\mathscr{R}_2} - \frac{\delta\beta_2 + (\mu + \alpha_1) - (\mu + \alpha_2)}{\delta\beta_2 - \beta_2 + \beta_1} > 0.$$

Now we take a common denominator:

$$\frac{\frac{1}{\mathscr{R}_2}(\delta\beta_2 - \beta_2 + \beta_1) - (\delta\beta_2 + \mu + \alpha_1) + (\mu + \alpha_2)}{\delta\beta_2 - \beta_2 + \beta_1} > 0.$$

Combining terms in the numerator, we have

$$\frac{-\delta\beta_2(1 - \frac{1}{\mathscr{R}_2}) - (\mu + \alpha_1) - \frac{1}{\mathscr{R}_2}\beta_2 + \frac{1}{\mathscr{R}_2}\beta_1 + (\mu + \alpha_2)}{\delta\beta_2 - \beta_2 + \beta_1} > 0.$$

We notice that $-\frac{1}{\mathscr{R}_2}\beta_2 + (\mu + \alpha_2) = 0$, so the above expression simplifies to

$$i_1 = \frac{(\delta\beta_2(1 - \frac{1}{\mathscr{R}_2}) + (\mu + \alpha_1))[\hat{\mathscr{R}}_1^2 - 1]}{\delta\beta_2 - \beta_2 + \beta_1}\frac{1}{\delta}.$$

We conclude that $i_1 > 0$ if and only if one of the following holds:

Case 1. $\hat{\mathscr{R}}_1^2 > 1$ and $\delta\beta_2 - \beta_2 + \beta_1 > 0$.
Case 2. $\hat{\mathscr{R}}_1^2 < 1$ and $\delta\beta_2 - \beta_2 + \beta_1 < 0$.

Next, we want $i_2 > 0$, that is, we want

$$\frac{\beta_1}{\delta\beta_2}\frac{\delta\beta_2 + (\mu+\alpha_1) - (\mu+\alpha_2)}{\delta\beta_2 - \beta_2 + \beta_1} - \frac{\mu+\alpha_1}{\delta\beta_2} > 0.$$

Factoring out $(\mu+\alpha_1)/\delta\beta_2$, we have

$$\frac{\mu+\alpha_1}{\delta\beta_2}\left(\mathscr{R}_1\frac{\delta\beta_2 + (\mu+\alpha_1) - (\mu+\alpha_2)}{\delta\beta_2 - \beta_2 + \beta_1} - 1\right) > 0.$$

Here we note that if we could show that $i_2 > 0$, that would imply that $s > 1/\mathscr{R}_1$, and consequently s would also be positive. In addition, if $s > 0$, $i_1 > 0$, and $i_2 > 0$, then they are all necessarily less than 1, since $s + i_1 + i_2 = 1$. Taking again a common denominator and noticing that $\mathscr{R}_1(\mu+\alpha_1) - \beta_1 = 0$, we see that the expression in the parentheses above becomes

$$\frac{\delta\beta_2(\mathscr{R}_1 - 1) + \beta_2 - \mathscr{R}_1(\mu+\alpha_2)}{\delta\beta_2 - \beta_2 + \beta_1} > 0.$$

Hence,

$$\frac{(\mu+\alpha_2)\mathscr{R}_1[\hat{\mathscr{R}}_2^1 - 1]}{\delta\beta_2 - \beta_2 + \beta_1} > 0.$$

We conclude that $i_2 > 0$ if and only if one of the following conditions holds:

Case 1. $\hat{\mathscr{R}}_2^1 > 1$ and $\delta\beta_2 - \beta_2 + \beta_1 > 0$.
Case 2. $\hat{\mathscr{R}}_2^1 < 1$ and $\delta\beta_2 - \beta_2 + \beta_1 < 0$.

We may summarize these results in the following theorem:

Theorem 8.5.

1. *Let $\delta\beta_2 - \beta_2 + \beta_1 > 0$. Then a unique coexistence equilibrium exists if and only if $\hat{\mathscr{R}}_2^1 > 1$ and $\hat{\mathscr{R}}_1^2 > 1$, that is, if each strain can invade the equilibrium of the other.*
2. *Let $\delta\beta_2 - \beta_2 + \beta_1 < 0$. Then a unique coexistence equilibrium exists if and only if $\hat{\mathscr{R}}_2^1 < 1$ and $\hat{\mathscr{R}}_1^2 < 1$, that is, if neither strain can invade the equilibrium of the other.*

Simulations suggest that the coexistence equilibrium in the case $\hat{\mathscr{R}}_2^1 > 1$ and $\hat{\mathscr{R}}_1^2 > 1$ is locally asymptotically stable, while the one in the case $\hat{\mathscr{R}}_2^1 < 1$ and $\hat{\mathscr{R}}_1^2 < 1$ is unstable.

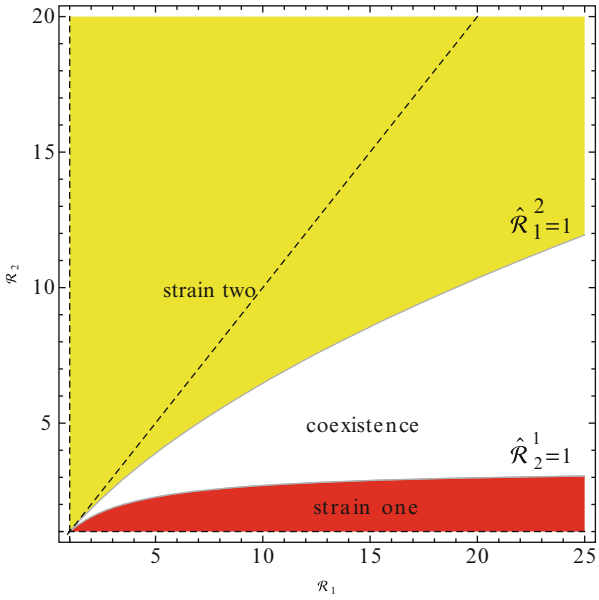

Fig. 8.7 Plot of the invasion regions in the $(\mathscr{R}_1, \mathscr{R}_2)$-plane. Parameters are taken as $\mu = 0.1$, $\alpha_1 = 0.5$, $\alpha_2 = 0.1$, $\delta = 0.3$

8.3.3 Competitive Outcomes, Graphical Representation, and Simulations

The invasion reproduction numbers play a critical role in determining the competitive outcomes between the two strains. The potential competitive outcomes have been summarized in Table 8.2.

Table 8.2 Competitive outcomes when coexistence is possible

Region	Long-term behavior	Competitive outcome
$\hat{\mathscr{R}}_1^2 > 1, \hat{\mathscr{R}}_2^1 < 1$	$I_1(t)$ persists, $I_2(t) \to 0$	Strain 1 dominates
$\hat{\mathscr{R}}_1^2 < 1, \hat{\mathscr{R}}_2^1 > 1$	$I_1(t) \to 0$, $I_2(t)$ persists	Strain 2 dominates
$\hat{\mathscr{R}}_1^2 > 1, \hat{\mathscr{R}}_2^1 > 1$	$I_1(t)$ persists, $I_2(t)$ persists	Strains coexist
$\hat{\mathscr{R}}_1^2 < 1, \hat{\mathscr{R}}_2^1 < 1$	$I_1(t)$ persists **or** $I_2(t)$ persists	Depends on initial conditions

It can be seen that $\hat{\mathscr{R}}_2^1$ and $\hat{\mathscr{R}}_1^2$ are functions of $\mathscr{R}_1$ and $\mathscr{R}_2$. Since the curves $\hat{\mathscr{R}}_2^1 = 1$ and $\hat{\mathscr{R}}_1^2 = 1$ in the $(\mathscr{R}_1, \mathscr{R}_2)$-plane separate the different regions of competitive outcomes, it is customary to graph them to illustrate the zones of coexistence and competitive exclusion. Figure 8.7 illustrates the case with a coexistence region.

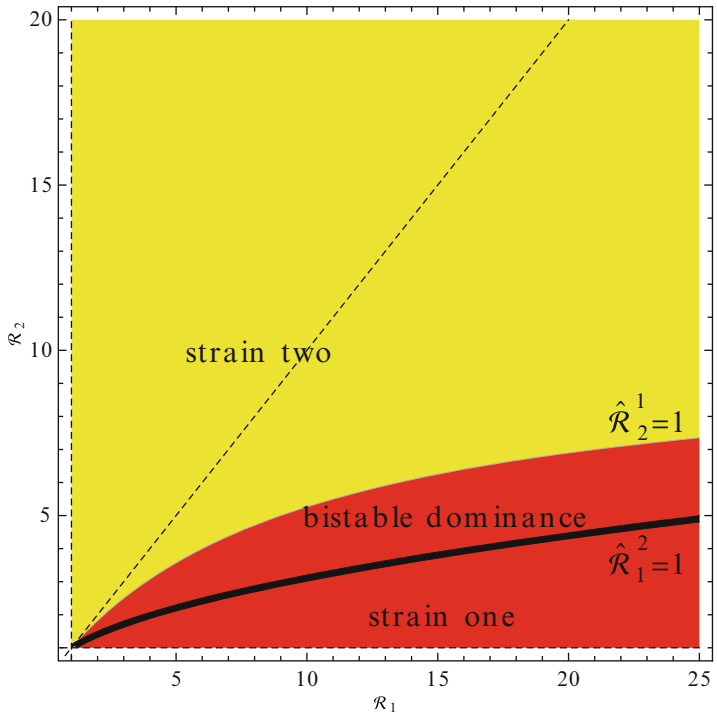

Fig. 8.8 Plot of the invasion regions in the $(\mathscr{R}_1, \mathscr{R}_2)$-plane. Parameters are taken as $\mu = 0.1$, $\alpha_1 = 0.5$, $\alpha_2 = 0.1$, $\delta = 0.3$

Remarks:

1. Notice that in Fig. 8.7, the area of dominance of strain two is much larger than that of strain one. This is to be expected, since strain two can infect even individuals infected with strain one.
2. Furthermore, notice that strain one eliminates strain two only if $\mathscr{R}_1 > \mathscr{R}_2$. In contrast, strain two can eliminate strain one if $\mathscr{R}_2 > \mathscr{R}_1$ but also if $\mathscr{R}_2 < \mathscr{R}_1$. Thus, in the case of superinfection, the principle that the strain with the maximal reproduction number excludes the one with the lower reproduction number is no longer valid. Thus a strain (strain two) with suboptimal reproduction number can dominate in the population. This occurs in the case with superinfection, because the two reproduction numbers do not take into account the cases of strain-two infection that are produced through superinfection. Thus, in the absence of superinfection, strain one will eliminate strain two, because it has a larger reproduction number. In the presence of superinfection, however, strain two superinfects individuals infected with strain one, and this gives the competitive advantage to strain two even if its reproduction number is lower.
3. When $\delta\beta_2 - \beta_2 + \beta_1 < 0$, the plot of the curves has a similar appearance (see Fig. 8.8). The difference between Figs. 8.7 and 8.8 is that the higher curve is $\hat{\mathscr{R}}_2^1 = 1$, and the lower curve is $\hat{\mathscr{R}}_1^2 = 1$. Thus, in the region between the curves,

we have

$$\hat{\mathscr{R}}_2^1 < 1 \qquad \text{and} \qquad \hat{\mathscr{R}}_1^2 < 1.$$

In that region, there is a coexistence equilibrium, but it is unstable. However, both of the boundary equilibria $\mathscr{E}_1$ and $\mathscr{E}_2$ are locally asymptotically stable. What is observed in simulations is that the solution converges either to $\mathscr{E}_1$ or to $\mathscr{E}_2$, depending on the initial conditions. Thus, depending on the initial conditions, either strain one persists or strain two persists. This effect is sometimes called *bistable dominance*. In ecology, the same effect is called *priority effects* or *founder control*, since the species that is in the better position originally dominates.

8.4 Computing the Invasion Numbers Using the Next-Generation Approach

In the previous section, we used the Jacobian approach to compute the invasion numbers. As with the reproduction number, the Jacobian approach can be difficult or impossible to use. The next-generation approach, originally developed for the computation of the reproduction number, can be adapted to facilitate the computation of the invasion numbers.

8.4.1 General Description of the Method

The method provides a technique for the derivation of the next-generation matrix of strain i from ordinary differential equation compartmental models for disease transmission. The system is divided into compartments with two broad categories: infected with strain-i compartments and noninfected/infected with strain-j compartments. A compartment is called an **infected with strain i compartment** if the individuals in that compartment are infected with strain i. Notice that we do not require these individuals to be infectious. The remaining compartments in which the individuals are not infected or are infected with strain j are the noninfected/infected with strain j compartments. Assume that there are n infected with strain i compartments and m noninfected/infected with train j compartments, so the entire ordinary differential equation model has $m+n$ dependent variables. Let x be the vector of dependent variables in the infected with strain i compartments, and let y be the vector of variables in the noninfected/infected with strain j compartments. We have $x \in R^n$ and $y \in R^m$. We then proceed as follows:

1. First, we arrange the equations so that the first n components of the ODE system correspond to the infected with strain i compartments. Thus, we write the original ODE system as

$$\begin{aligned} x_k' &= f_k(x,y), \qquad k = 1,\dots n, \\ y_j' &= g_j(x,y), \qquad j = 1,\dots,m. \end{aligned} \tag{8.38}$$

2. Second, we split the right-hand side in the infected with strain i compartments in the following way:

$$\begin{aligned} x_k' &= \mathscr{F}_k(x,y) - \mathscr{V}_k(x,y), \qquad k = 1,\dots,n, \\ y_j' &= g_j(x,y), \qquad j = 1,\dots,m, \end{aligned} \tag{8.39}$$

   where

   - $\mathscr{F}_k(x,y)$ is the rate of appearance of **new** infections with strain i in compartment k;
   - $\mathscr{V}_k(x,y)$ incorporates the remaining transitional terms, namely, births, deaths, disease progression, and recovery.

   We note that just as in the case with the reproduction number, this decomposition into infected with strain i and noninfected/infected with strain j compartments as well as the decomposition into $\mathscr{F}$ and $\mathscr{V}$ may not be unique. Different decompositions may correspond to different interpretations of the disease process and may lead to somewhat different expressions for the invasion number of strain i. The decomposition should satisfy the following *properties*:

   - $\mathscr{F}_k(0,y) = 0$ and $\mathscr{V}_k(0,y) = 0$ for $y \geq 0$ and $k = 1,\dots,n$. The first condition says that all new infections with strain i are secondary infections arising from infected with strain i hosts. The second conditions says that there is no immigration of susceptible individuals/individuals infected with strain j into the infected with strain i compartments.
   - $\mathscr{F}_k(x,y) \geq 0$ for all $x,y \geq 0$.
   - $\mathscr{V}_k(x,y) \leq 0$ whenever $x_k = 0$, for $k = 1,\dots,n$. Each component $\mathscr{V}_k$ represents the net outflow of an infected with strain i compartment and must give inflow only (that is, be negative) if the compartment is empty.
   - $\sum_{k=1}^{n} \mathscr{V}_k(x,y) \geq 0$ for all $x,y \geq 0$. The total outflow of all infected with strain i compartments is positive.

3. Assume that the system in the absence of strain i,

$$y' = g(0,y),$$

   has a unique nontrivial strain-j-dominance equilibrium $\mathscr{E}_{0j} = (0,y^0)$ such that all solutions with initial conditions of the form $(0,y)$ approach $(0,y^0)$ as $t \to \infty$. Determine the strain-j-dominance equilibrium $\mathscr{E}_{0j}$.
4. Determine the matrices F and V with components

$$F = \left[\frac{\partial \mathscr{F}_k(0,y^0)}{\partial x_j}\right] \qquad \text{and} \qquad V = \left[\frac{\partial \mathscr{V}_k(0,y^0)}{\partial x_j}\right].$$

These matrices appear from the linearization of the system (8.39) around the equilibrium of strain j. It can be shown that

$$\frac{\partial \mathscr{F}_k(0,y^0)}{\partial y_j} = \frac{\partial \mathscr{V}_k(0,y^0)}{\partial y_j} = 0$$

for every pair (k,j). This implies that the linearized equations for the infected with strain i compartments x computed at the equilibrium of strain j are decoupled from the remaining equations. The linearized system for the infected with strain i compartments can be written as

$$x' = (F-V)x,$$

where the matrices F and V are defined above.

5. The next-generation matrix of strain i is defined as

$$K = FV^{-1}$$

and

$$\hat{\mathscr{R}}_i^j = \rho(FV^{-1}),$$

where $\rho(A)$ denotes the spectral radius of A.

As before, it can be shown that V is a nonsingular M-matrix. Since V is an M-matrix, $V^{-1} \geq 0$, that is, V^{-1} has only nonnegative entries. Since F also has only nonnegative entries, the next-generation matrix of strain i, $K = FV^{-1}$, is also nonnegative. This implies that the next-generation matrix of strain i has its spectral radius given by $\hat{\mathscr{R}}_i^j$, $\hat{\mathscr{R}}_i^j$ is an eigenvalue of K, and there is no other eigenvalue with larger modulus. Hence,

> The invasion number $\hat{\mathscr{R}}_i^j$ is computed as the largest positive eigenvalue of the next-generation matrix of strain i.

Remark: Condition 3 above is needed to guarantee that if $\hat{\mathscr{R}}_i^j < 1$, the strain-$j$-dominance equilibrium is locally asymptotically stable. However, on many occasions that is not true. Even if $\hat{\mathscr{R}}_i^j < 1$, the strain-$j$ equilibrium may become unstable, say through Hopf bifurcation. For that reason, in defining the invasion numbers, condition 3 is relaxed. If condition 3 is not valid, then the only inference that can be made from the invasion number is that if $\hat{\mathscr{R}}_i^j > 1$, then the equilibrium of strain j is unstable.

8.4.2 Example

As an example we consider a two-strain model of influenza with isolation [128]. Influenza strains interact through cross-immunity. To introduce the model, let S denote susceptible individuals, I_j individuals infected with strain j, Q_j individuals infected with strain j and quarantined, R_j individuals recovered from strain j, J_k individuals infected with strain k after recovery from strain j, and W individuals recovered from both strains. The model is a modification of model (8.16). It takes the form

$$\begin{aligned}
S' &= \Lambda - \beta_1 \frac{S(I_1+J_1)}{A} - \beta_2 \frac{S(I_2+J_2)}{A} - \mu S,\\
I_1' &= \beta_1 \frac{S(I_1+J_1)}{A} - (\mu+\alpha_1+\delta_1)I_1,\\
Q_1' &= \delta_1 I_1 - (\mu+\gamma_1)Q_1,\\
R_1' &= \alpha_1 I_1 + \gamma_1 Q_1 - \sigma_2\beta_2 \frac{R_1(I_2+J_2)}{A} - \mu R_1,\\
J_1' &= \sigma_1\beta_1 \frac{R_2(I_1+J_1)}{A} - (\mu+\alpha_1)J_1,\\
I_2' &= \beta_2 \frac{S(I_2+J_2)}{A} - (\mu+\alpha_2+\delta_2)I_2,\\
Q_2' &= \delta_2 I_2 - (\mu+\gamma_2)Q_2,\\
R_2' &= \alpha_2 I_2 + \gamma_2 Q_2 - \delta_1\beta_1 \frac{(I_1+J)I_2}{A} - \mu R_2,\\
J_2' &= \sigma_2\beta_2 \frac{R_1(I_2+J_2)}{A} - (\mu+\alpha_2)J_2,\\
W' &= \alpha_1 J_1 + \alpha_2 J_2 - \mu W,
\end{aligned} \tag{8.40}$$

where A denotes the active population $A = N - Q_1 - Q_2$. We define the reproduction numbers as

$$\mathscr{R}_1 = \frac{\beta_1}{\mu+\delta_1+\alpha_1} \qquad \text{and} \qquad \mathscr{R}_2 = \frac{\beta_2}{\mu+\delta_2+\alpha_2}.$$

The model is symmetric with respect to strains one and two, so the two invasion numbers will also be symmetric. First, we need to compute the strain-one-dominance equilibrium. All infected classes associated with strain two will be set to zero. That is, $I_2 = Q_2 = R_2 = J_2 = W = 0$. From the last equation, we deduce that $J_1 = 0$. The last six equations in system (8.40) are trivially satisfied. The strain-one-dominance equilibrium satisfies the system

$$\begin{aligned}
\Lambda - \beta_1 \frac{SI_1}{A} - \mu S &= 0,\\
\beta_1 \frac{SI_1}{A} - (\mu+\alpha_1+\delta_1)I_1 &= 0,\\
\delta_1 I_1 - (\mu+\gamma_1)Q_1 &= 0,\\
\alpha_1 I_1 - \mu R_1 &= 0.
\end{aligned} \tag{8.41}$$

From the last two equations above, we can express Q_1 and R_1 in terms of I_1:

$$Q_1 = \frac{\delta_1 I_1}{\mu + \gamma_1} = \kappa_1 I_1 \qquad \text{and} \qquad \mathrm{R}_1 = \frac{\alpha_1 \mathrm{I}_1}{\mu}. \tag{8.42}$$

From the first two equations, we obtain the following system:

$$\begin{aligned} \Lambda - \mu S - (\mu + \alpha_1 + \delta_1) I_1 &= 0, \\ \beta_1 \frac{S}{A} - (\mu + \alpha_1 + \delta_1) &= 0. \end{aligned} \tag{8.43}$$

From these two equations, and $\frac{S}{A} + \frac{I_1}{A} + \frac{R_1}{A} = 1$ we obtain

$$\begin{aligned} \frac{S}{A} &= \frac{1}{\mathscr{R}_1}, \\ \frac{I_1}{A} &= \frac{\mu}{\mu + \alpha_1} \left(1 - \frac{1}{\mathscr{R}_1} \right). \end{aligned} \tag{8.44}$$

Clearly, the strain-one-dominance equilibrium exists if and only if $\mathscr{R}_1 > 1$. To apply the next-generation approach and compute the invasion number of strain two, we notice that the variables associated with strain-two infection are variables I_2, Q_2, and J_2. Thus, we have $x = (I_2, Q_2, J_2)$; y denotes the remaining variables. We rewrite the equations for these three variables as a difference of a new-infections term and an outflow term: $x' = \mathscr{F}(x,y) - \mathscr{V}(x,y)$, where

$$\mathscr{F} = \begin{pmatrix} \beta_2 \dfrac{S(I_2 + J_2)}{A} \\ 0 \\ \sigma_2 \beta_2 \dfrac{R_1 (I_2 + J_2)}{A} \end{pmatrix} \qquad \text{and} \qquad \mathscr{V} = \begin{pmatrix} (\mu + \alpha_2 + \delta_2) I_2 \\ -\delta_2 I_2 + (\mu + \gamma_2) Q_2 \\ (\mu + \alpha_2) J_2 \end{pmatrix}. \tag{8.45}$$

We define the matrices F and V:

$$\begin{aligned} F &= \begin{pmatrix} \beta_2 \dfrac{S^*}{A^*} & 0 & \beta_2 \dfrac{S^*}{A^*} \\ 0 & 0 & 0 \\ \sigma_2 \beta_2 \dfrac{R_1^*}{A^*} & 0 & \sigma_2 \beta_2 \dfrac{R_1^*}{A^*} \end{pmatrix} \qquad \text{and} \\ V &= \begin{pmatrix} \mu + \alpha_2 + \delta_2 & 0 & 0 \\ -\delta_2 & \mu + \gamma_2 & 0 \\ 0 & 0 & \mu + \alpha_2 \end{pmatrix}. \end{aligned} \tag{8.46}$$

We first notice that V is a M-matrix. It has the Z sign pattern. Also, the transpose of V, V^T, satisfies $V^T u > 0$, where $u = (1,1,1)^T$. Hence, $V^{-1} > 0$. The matrix V is not diagonal, and it is a 3×3 matrix, so it is most easily inverted with the help of a computer algebra system. We obtain

$$V^{-1} = \begin{pmatrix} \dfrac{1}{\mu+\alpha_2+\delta_2} & 0 & 0 \\ v_{21} & \dfrac{1}{\mu+\gamma_2} & 0 \\ 0 & 0 & \dfrac{1}{\mu+\alpha_2} \end{pmatrix}, \tag{8.47}$$

where

$$v_{21} = \frac{\delta_2(\mu+\alpha_2)}{(\mu+\alpha_2+\delta_2)(\mu+\alpha_2)(\mu+\gamma_2)}.$$

Hence,

$$FV^{-1} = \begin{pmatrix} \dfrac{\beta_2 S^*}{(\mu+\alpha_2+\delta_2)A^*} & 0 & \dfrac{\beta_2 S^*}{(\mu+\alpha_2)A^*} \\ 0 & 0 & 0 \\ \dfrac{\sigma_2\beta_2 R_1^*}{(\mu+\alpha_2+\delta_2)A^*} & 0 & \dfrac{\sigma_2\beta_2 R_1^*}{(\mu+\alpha_2)A^*} \end{pmatrix}. \tag{8.48}$$

We know that $\hat{\mathscr{R}}_2^1 = \rho(FV^{-1})$. It is not hard to compute that the invasion number is given by

$$\hat{\mathscr{R}}_2^1 = \frac{\beta_2 S^*}{(\mu+\alpha_2+\delta_2)A^*} + \frac{\sigma_2\beta_2 R_1^*}{(\mu+\alpha_2)A^*},$$

where S^*/A^* and R_1^*/A^* are the values of the strain-one equilibrium.

Problems

8.1. Show that the definition of the reproduction number $\mathscr{R}_0$ in the two-strain case given by (8.22) follows from the next-generation approach (see Chap. 5).

8.2. Write a code to simulate model (8.17).

(a) Using Fig. 8.7 and the parameters given in it, choose values of $\mathscr{R}_1$ and $\mathscr{R}_2$ such that strain two has a lower reproduction number but outcompetes strain one.
(b) Using Fig. 8.8 and the parameters given in it, choose values of $\mathscr{R}_1$ and $\mathscr{R}_2$ such that strain one outcompetes strain two or vice versa depending on the initial conditions.

8.3. Saturating Incidence
Consider a model of two strains with saturated incidence [97]:

$$\begin{aligned} S' &= \Lambda - \frac{\beta_1 SI}{1+a_1 N} - \frac{\beta_2 SJ}{1+a_2 N} - \mu S, \\ I' &= \frac{\beta_1 SI}{1+a_1 N} - (\mu+\alpha_1)I, \\ J' &= \frac{\beta_2 SJ}{1+a_2 N} - (\mu+\alpha_2)J. \end{aligned} \tag{8.49}$$

(a) Compute the disease-free equilibrium and the reproduction numbers of the two strains.
(b) Compute the two dominance equilibria. Investigate their stability and define the two invasion numbers.
(c) Show that if both invasion numbers are greater than one, there is a coexistence equilibrium.

8.4. Saturating Per Capita Treatment Rates
Consider a model of two strains with saturated treatment rate:

$$\begin{aligned} S' &= \Lambda - \frac{\beta_1 SI}{N} - \beta_2 \frac{SJ}{N} - \mu S, \\ I' &= \frac{\beta_1 SI}{N} - \mu I - \frac{\alpha_1 I^2}{A+I+J}, \\ J' &= \frac{\beta_2 SJ}{N} - \mu J - \frac{\alpha_2 J^2}{B+I+J}. \end{aligned} \tag{8.50}$$

(a) Compute the disease-free equilibrium and the reproduction numbers of the two strains.
(b) Compute the two dominance equilibria. Investigate their stability, and define the two invasion numbers.
(c) Show that if both invasion numbers are greater than one, there is a coexistence equilibrium.

8.5. Two Strains with Mutation
Consider a model of two strains with mutation:

$$\begin{aligned} S' &= \Lambda - \frac{\beta_1 SI}{N} - \frac{\beta_2 SJ}{N} - \mu S, \\ I' &= \frac{\beta_1 SI}{N} - (\mu + \alpha_1 + m)I, \\ J' &= \frac{\beta_2 SJ}{N} - (\mu + \alpha_2)J + mI. \end{aligned} \tag{8.51}$$

(a) Compute the disease-free equilibrium and the reproduction numbers of the two strains.
(b) Compute the dominance equilibrium of strain J. Show that there is no dominance equilibrium of strain I. Investigate local stability of the dominance equilibrium. Define the appropriate invasion number.
(c) Show that there is a coexistence equilibrium.

8.6. Consider the model (8.15).

(a) Compute the reproduction numbers of the two strains.
(b) Compute the two dominance equilibria.
(c) Use the next-generation approach to compute the invasion reproduction numbers.
(d) Draw the regions of dominance and coexistence in the $(\mathscr{R}_1, \mathscr{R}_2)$-plane.

8.7. Consider the model (8.16).

(a) Compute the reproduction numbers of the two strains.
(b) Compute the two dominance equilibria.
(c) Use the next-generation approach to compute the invasion reproduction numbers.
(d) Draw the regions of dominance and coexistence in the $(\mathscr{R}_1, \mathscr{R}_2)$-plane.

8.8. Two-Strain SIS Model with Delays
Consider an SIS model with two strains. Assume that each strain has its own incubation period. This gives the following two-strain epidemic model with two delays:

$$\begin{aligned} S' &= \Lambda - \beta_1 SI - \beta_2 SJ - \mu S + \alpha_1 I + \alpha_2 J, \\ I' &= e^{-\mu\tau_1}\beta_1 S(t-\tau_1)I(t-\tau_1) - (\mu+\alpha_1)I, \\ J' &= e^{-\mu\tau_2}\beta_2 S(t-\tau_2)J(t-\tau_2) - (\mu+\alpha_2)J. \end{aligned} \tag{8.52}$$

(a) Compute the disease-free equilibrium and the reproduction numbers of the two strains. Determine the stability of the disease-free equilibrium.
(b) Compute the dominance equilibria of strain I and strain J. Investigate local stability of the dominance equilibria. Define the appropriate invasion numbers.
(c) Show that there is no coexistence equilibrium.

8.9. Two-Species Model of Malaria
Malaria is caused by four species of the genus *Plasmodium*. Consider the following model, which includes two of the species:

$$\begin{aligned} I_1' &= \frac{\alpha_1\beta_1 N_v I_1}{\beta_1 I_1 + \beta_2 I_2 + \mu_V}(N_H - I_1 - I_2) - (\delta + r_1)I_1, \\ I_2' &= \frac{\alpha_2\beta_2 N_v I_2}{\beta_1 I_1 + \beta_2 I_2 + \mu_V}(N_H - I_1 - I_2) - (\delta + r_2)I_2, \end{aligned} \tag{8.53}$$

where N_v is the total vector population and N_H is the total human population (assume constant), μ_v is the vector natural mortality, δ is the human natural mortality, α_i for $i = 1,2$ are the transmission rates from vector to human, and β_i for $i = 1,2$ are the transmission rates from human to vector. The r_i are the appropriate recovery rates.

(a) Compute the disease-free equilibrium and the reproduction numbers of the two malaria species.
(b) Compute the dominance equilibria of species I_1 and species I_2. Investigate local stability of the dominance equilibria. Define the appropriate invasion numbers.
(c) Determine whether there exists a coexistence equilibrium.

Chapter 9
Control Strategies

9.1 Introduction

Measures for prevention and control of infectious diseases include vaccination, treatment, quarantine, isolation, and prophylaxis.

Prophylaxis is the series of measures taken to prevent a specific infectious disease. These measures can be as simple as hand-washing with soap and water, or wearing protective gear, or taking a medication to prevent a disease. **Treatment** is the use of an agent, procedure, or regimen, such as a drug, or bed rest in an attempt to cure or mitigate a disease. Nowadays, for most infectious diseases, medications exist that can cure or lessen the impact of the diseases, while improving the life of the patients. Diseases for which medications can offer a cure include malaria and tuberculosis. Diseases for which medications offer relief but not a cure include HIV and genital herpes.

Vaccination is the process through which killed (inactivated) or weakened microorganisms are placed into the body. Our immune system recognizes vaccine agents as foreign. That triggers an immune response, and antibodies against them are developed. As a result, if the same types of microorganisms enter the body again, they will be destroyed much faster by the antibodies. Thus, an individual that is immunized is protected against the disease. If a large majority of people are vaccinated, it is much more difficult for an outbreak of disease to occur, let alone spread. This effect is called **herd immunity**.

Vaccination is one of the greatest achievements of public health. Vaccination has led to the complete eradication of smallpox worldwide, and a near eradication of polio. Table 9.1 gives the reduction of disease load in the United States as a result of widespread vaccination campaigns.

Vaccines do not guarantee complete protection from a disease. There remains the possibility that a vaccinated person may get the disease. Even if the host develops antibodies, some pathogens can mutate (the common cold and influenza viruses are highly efficient at this), and in any case, the immune system might still not be able

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_9

Table 9.1 Achievements of vaccination in the United States[a]

Disease	Baseline years	Cases/year	Cases in 1998	% Decrease
Smallpox	1900–1904	48,164	0	100
Diphtheria	1920–1922	175,885	1	100
Pertussis	1922–1925	147,271	6,279	95.7
Tetanus	1922–1926	1,314	34	97.4
Poliomyelitis	1951–1954	16,316	0	100
Measles	1958–1962	503,282	89	100
Mumps	1968	152,209	606	99.6
Rubella	1966–1968	47,745	345	99.3
Hib	1985	20,000	54+71	99.7

[a]**Source:** CDC, Morbidity and Mortality Weekly Report (MMWR) 48(12), 1999. *Achievements of Public Health, 1900–1999: Impact of Vaccines Universally Recommended for Children—US, 1990–1998*

to defeat the infection. The degree to which vaccinated individuals are protected against the disease is called *efficacy of the vaccine*.

Quarantine and **isolation** are two measures by which exposed or infectious individuals are removed from the population to prevent further spread of the infection. Quarantine is applied to seemingly healthy but potentially infected individuals, while isolation is applied to already infectious individuals. Isolation has been used and is being used to control many dangerous diseases. Quarantine is applied less often. It is one of the first response methods that can be used in an extreme emergency. Quarantine was implemented during the SARS epidemic of 2002–2003.

The reproduction number, computed for mathematical models involving control strategies, depends on the control strategies, and it is often called a *controlled reproduction number*.

9.2 Modeling Vaccination: Single-Strain Diseases

There are two points in which vaccination models can differ from one another. The first is that some models assume that vaccination is equivalent to going through the disease and treats vaccinated individuals as recovered individuals. Thus an SIR model can include vaccinated individuals without an additional class. Other models assume that vaccinated individuals have to be separated into a vaccinated class V. The second point of distinction is that some classes of models assume that individuals enter the system at a point of their life when they either get vaccinated or skip vaccination and enter the system as susceptibles. This is more or less accurate for school children. Other models allow for continuous vaccination of individuals while in the system.

9.2.1 A Model with Vaccination at Recruitment

Assume that we have a perfect vaccine, whereby everybody who is vaccinated is completely protected. Suppose we vaccinate at recruitment into the system a fraction p of individuals. So if μN is the recruitment term, a fraction $p\mu N$ goes directly into the recovered class, and a proportion $q\mu N$, where $q = 1 - p$, enters the susceptible class. Thus the SIR model with vaccination becomes

$$\begin{aligned} \frac{dS}{dt} &= q\mu N - \beta SI - \mu S, \\ \frac{dI}{dt} &= \beta SI - (\mu + \alpha)I, \\ \frac{dR}{dt} &= p\mu N + \alpha I - \mu R. \end{aligned} \tag{9.1}$$

The equation of the total population size here is $N'(t) = 0$, and the total population size is constant, $N = S_0 + I_0 + R_0$. The disease-free equilibrium, obtained from setting the derivatives equal to zero and $I = 0$, is given by $\mathscr{E}_0 = (qN, 0, pN)$. Thus, if

$$\mathscr{R}_0 = \frac{\beta N}{\alpha + \mu}$$

is the reproduction number in the absence of vaccination ($p = 0$), then $q\mathscr{R}_0$ is the reproduction number of the disease in the presence of vaccination. Consequently, vaccination has reduced the original reproduction number by the fraction q.

Question: What fraction, p, of the population must be vaccinated so that the reproduction number of the disease is reduced below 1?

To answer this question, we need $q\mathscr{R}_0 < 1$. Replacing q with $1 - p$ and solving the inequality for p, we obtain that $p > \hat{p}$, where

$$\hat{p} = 1 - \frac{1}{\mathscr{R}_0}.$$

Consequently, if a fraction $\hat{p}$ of the population is successfully vaccinated, then the disease will not spread in the population. In effect, the whole population will be protected. This is a manifestation of the herd immunity.

9.2.2 A Model with Continuous Vaccination

Most diseases for which vaccination is successful have a recovered (immune) stage. After all, vaccination works with the immune system more or less as the disease does, so if the disease does not provide immunity, how could vaccination? However, there are diseases for which it is justified to consider vaccination in addition to an SIS model, that is, a model where recovery brings the individual back to

the susceptible class. One such disease is tuberculosis, which imparts very short-lived immunity. Another situation occurs with bacterial infections with *Neisseria meningitidis* and *Streptococcus pneumoniae*. Both these bacteria can exist in the host without causing disease, a scenario, called *carriage*. Both carriers and infected (sick) people can transmit the microorganism, so from the point of view of disease transmission, they can be considered indistinguishable and modeled with one class. Carriage and disease impart immunity against the disease but probably not so much against carriage. Thus individuals who become completely pathogen-free can be counted as susceptible (at least for carriage). In both cases, there are vaccines, at least against some variants of the microorganisms, and an SIS model with vaccination may be appropriate.

9.2.2.1 An SIS Model with Vaccination

Let $V(t)$ denote the number of vaccinated individuals, and ψ the per capita vaccination rate. Vaccination is applied only to healthy individuals, so only susceptible individuals get vaccinated. In this model, we also take into account the fact that vaccines are rarely perfect, and some of the vaccinated individuals can become infected and infectious even though they have been vaccinated. That happens at a reduced transmission rate $\beta\delta$, where $0 \leq \delta \leq 1$ is the reduction coefficient. If $\delta = 0$, then vaccinated individuals cannot get infected, and the vaccine is perfect. This implies that the vaccine efficacy is $\varepsilon = 1$. If $\delta = 1$, then vaccinated individuals get infected just like susceptible individuals, and the vaccine plays no protective role. In that case, the vaccine efficacy is $\varepsilon = 0$.

We list the parameters and the variables in the Table 9.2.

Table 9.2 List of parameters, variables, and their meanings

Notation	Meaning
Λ	Birth/recruitment rate into the population
μ	Per capita natural death rate
β	Per capita transmission rate
γ	Per capita recovery rate
χ	Proportion of individuals who recover to the vaccinated class
$1-\chi$	Proportion of individuals who recover to the susceptible class
ψ	Per capita vaccination rate
$\varepsilon = 1-\delta$	Vaccine efficacy
$S(t)$	Number of susceptible individuals
$I(t)$	Number of infected individuals
$V(t)$	Number of vaccinated individuals

The model takes the form

$$\begin{aligned}\frac{dS}{dt} &= \Lambda - \frac{\beta SI}{N} - (\mu+\psi)S + \chi\gamma I,\\ \frac{dI}{dt} &= \frac{\beta SI}{N} + \frac{\beta\delta VI}{N} - (\mu+\gamma)I,\\ \frac{dV}{dt} &= \psi S - \frac{\beta\delta VI}{N} + (1-\chi)\gamma I - \mu V.\end{aligned} \tag{9.2}$$

The flowchat of the model is given in Fig. 9.1.

The disease-free equilibrium is given by

$$\mathscr{E}_0 = \left(\frac{\Lambda}{\mu+\psi}, 0, \frac{\Lambda\psi}{\mu(\mu+\psi)}\right).$$

Since the equation of the total population is $N'(t) = \Lambda - \mu N$, the equilibrium total population size is $N = \frac{\Lambda}{\mu}$. Thus the proportions of susceptible and vaccinated in the disease-free population are given by

$$s^0 = \frac{\mu}{\mu+\psi}, \qquad v^0 = \frac{\psi}{\mu+\psi}.$$

9.2.2.2 The Reproduction Number and the Critical Vaccination Proportion

To compute the reproduction number, we compute the Jacobian at the disease-free equilibrium:

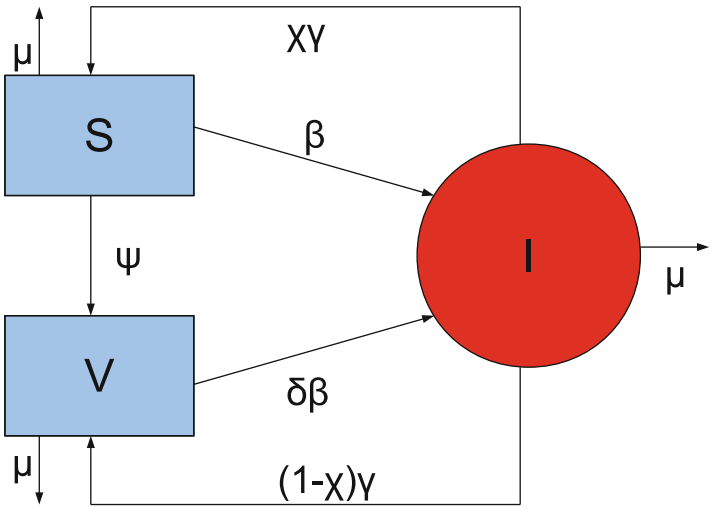

Fig. 9.1 Flowchart of the model with continuous vaccination with imperfect vaccine

$$\mathscr{J}(\mathscr{E}_0) = \begin{pmatrix} -(\mu+\psi) & -\beta s^0 + \chi\gamma & 0 \\ 0 & \beta s^0 + \beta\delta v^0 - (\mu+\gamma) & 0 \\ \psi & -\beta\delta v^0 + (1-\chi)\gamma & -\mu \end{pmatrix}.$$

The Jacobian has two negative eigenvalues, $-\mu$ and $-(\mu+\psi)$. The third eigenvalue is given by $\beta s^0 + \beta\delta v^0 - (\mu+\gamma)$. Thus we define the reproduction number in the presence of vaccination as

$$\mathscr{R}(\psi) = \frac{\beta(\mu+\delta\psi)}{(\mu+\gamma)(\mu+\psi)}.$$

The reproduction number of the disease in the absence of vaccination is obtained by letting $\psi = 0$, and is given by

$$\mathscr{R}_0 = \frac{\beta}{\mu+\gamma}.$$

In interpreting the reproduction number, we notice that $\frac{\beta SI}{N}$ gives the number of secondary infections of susceptible individuals per unit of time. The number of secondary infections of susceptible individuals per unit of time for one infectious individual will be $\frac{\beta S}{N}$. The proportion of susceptibles in a disease-free population is $\frac{S}{N} = s^0 = \frac{\mu}{\mu+\psi}$. Since $\frac{1}{\mu+\gamma}$ is the time spent as an infectious individual, the first term in $\mathscr{R}(\psi)$, given by $\frac{\beta\mu}{(\mu+\gamma)(\mu+\psi)}$, gives the number of secondary infections of susceptible individuals that one infected individual can produce in a disease-free population. Similarly, $\frac{\beta\delta VI}{N}$ gives the number of secondary infections of vaccinated individuals per unit of time. The number of secondary infections of vaccinated individuals per unit of time for one infectious individual will be $\frac{\beta\delta V}{N}$. The proportion of vaccinated individuals in a disease-free population is $\frac{V}{N} = v^0 = \frac{\psi}{\mu+\psi}$. Since $\frac{1}{\mu+\gamma}$ is the time spent as an infectious individual, the second term in $\mathscr{R}(\psi)$, given by $\frac{\beta\delta\psi}{(\mu+\gamma)(\mu+\psi)}$, gives the number of secondary infections of vaccinated individuals that one infected individual can produce in a disease-free population.

One can see that the reproduction number in the presence of vaccination is a decreasing function of the vaccination rate ψ. Thus, the higher the vaccination rate, the smaller the reproduction number. Furthermore,

$$\lim_{\psi\to\infty} \mathscr{R}(\psi) = \delta\mathscr{R}_0.$$

Thus, if the vaccine efficacy ε is not high enough (that is, δ is not small enough), then even if we vaccinate everybody, we may not be able to eradicate the disease. In other words, we cannot bring $\mathscr{R}(\psi)$ below 1, since the vaccinated individuals can become infected.

Question: What is the critical proportion of individuals that should be vaccinated if the vaccine is continuously applied and imperfect?

A critical vaccination proportion $\hat{p}_\varepsilon$ for the eradication of a disease with imperfect vaccination exists only if $\delta\mathscr{R}_0 < 1$, that is, if the vaccine efficacy satisfies

$$\varepsilon > \left(1 - \frac{1}{\mathscr{R}_0}\right). \tag{9.3}$$

If $\delta\mathscr{R}_0 < 1$, then there exists a critical vaccination level ψ^* such that $\mathscr{R}(\psi^*) = 1$. This critical vaccination level for eradication of the disease is given by

$$\psi^* = \frac{(\mathscr{R}_0 - 1)\mu}{1 - \delta\mathscr{R}_0}.$$

The proportion vaccinated in the population is given by $\psi/(\mu + \psi)$. We conclude that

The critical proportion of the population that needs to be vaccinated with vaccine with efficacy ε is given by

$$\hat{p}_\varepsilon = \frac{1}{\varepsilon}\left(1 - \frac{1}{\mathscr{R}_0}\right).$$

In words, the critical proportion of the population that needs to be vaccinated with imperfect vaccine is the critical population that needs to be vaccinated with perfect vaccine divided by the vaccine efficacy.

We note that the formula above is an extension of the critical vaccination proportion to imperfect vaccines. If the vaccine is perfect, that is, if $\varepsilon = 1$, then we obtain the customary formula for the critical vaccination proportion for perfect vaccines.

Table 9.3 gives the estimates of $\mathscr{R}_0$ before the introduction of vaccination. Most data on the reproduction number before vaccination are from England, Wales, and the USA [10]. The table gives the critical vaccination fraction with perfect vaccines, vaccine efficacies of the most common vaccines used in the USA, and the critical vaccination fractions with imperfect vaccines. It can be seen from the table that the current vaccines are incapable of eliminating pertussis, and may be useful in eliminating polio and diphtheria if a sufficient proportion of the population is vaccinated. In fact, polio has been eliminated in the developed countries for which the reproduction number before vaccination and vaccine efficacies are most accurate.

9.2.2.3 Backward Bifurcation in the Imperfect Vaccination Model

The critical threshold above gives only the proportion that has to be vaccinated so that the reproduction number in the presence of vaccination is below one. However, imperfect vaccines have the disadvantage that they lead to backward bifurcation, and endemic equilibria exist and are stable even when the reproduction number in the presence of vaccination is below one. The main reason for the backward bifurcation is the fact that imperfect vaccination creates two classes of susceptible individuals with different susceptibilities—the naive susceptible and the vaccinated susceptible.

Table 9.3 Diseases and their eradication vaccination levels

Disease	$\mathscr{R}_0$	$\hat{p}$,%	Vaccine efficacy[a]	$\hat{p}_\varepsilon$,%
Smallpox	3–5	67–80	0.75[b]	89–100
Measles	12–13	92	0.75–0.95	97–100
Mumps	4–7	75–86	0.75–0.95	79–100
Rubella	6–7	83–86	0.75–0.95	87–100
Chickenpox	9–10	89–90	0.8–0.95	94–100
Pertussis	13–17	92–94	0.8–0.9	–
Poliomyelitis	6	83	0.9–0.99	84–92
Diphtheria	4–6	75–83	0.87–0.96	78–95

[a] http://www.whale.to/vaccines/efficacy.html
[b] Vaccine efficacy never measured in clinical trials

To obtain a necessary and sufficient condition for backward bifurcation, we compute the endemic equilibria. First, we consider the equations for the proportions ($s = \frac{S}{N}$, $i = \frac{I}{N}$, $v = \frac{V}{N}$):

$$\begin{aligned} 0 &= \mu - \beta si - (\mu + \psi)s + \chi\gamma i, \\ 0 &= \beta si + \beta\delta vi - (\mu + \gamma)i, \\ 0 &= \psi s - \beta\delta vi + (1-\chi)\gamma i - \mu v. \end{aligned} \tag{9.4}$$

Expressing s from the first equation and v from the third equation yields

$$s = \frac{\mu + \chi\gamma i}{\beta i + \mu + \psi}, \qquad v = \frac{\psi s + (1-\chi)\gamma i}{\beta\delta i + \mu},$$

and substituting them in the second equation, we obtain a quadratic equation in i:

$$\begin{aligned} &\beta(\mu + \chi\gamma i)(\beta\delta i + \mu + \delta\psi) + \beta\delta(1-\chi)\gamma i(\beta i + \mu + \psi) \\ &\quad = (\mu + \gamma)(\beta\delta i + \mu)(\beta i + \mu + \psi). \end{aligned} \tag{9.5}$$

If we think of β as a function of i, that is, $\beta(i)$, and we differentiate implicitly the above equation, we obtain for β' at the critical value $i = 0$ the following expression:

$$\begin{aligned} &\beta'(0) \\ &= \frac{\beta\{\delta(\mu+\gamma)(\mu+\psi) + \mu(\mu+\gamma) - \chi\gamma(\mu+\delta\psi) - \beta\delta\mu - \delta(1-\chi)\gamma(\mu+\psi)\}}{\mu(\mu+\delta\psi)}. \end{aligned}$$

The bifurcation at the critical value $i = 0$ ($\mathscr{R}(\psi) = 1$) is backward if and only if $\beta'(0) < 0$, that is, if and only if the parameters satisfy the following condition:

$$\begin{aligned} \delta(\mu+\gamma)(\mu+\psi) + \mu(\mu+\gamma) < \chi\gamma(\mu+\delta\psi) + \frac{(\mu+\gamma)(\mu+\psi)\delta\mu}{\mu+\delta\psi} \\ +\delta(1-\chi)\gamma(\mu+\psi). \end{aligned}$$

To plot the dependence of i on $\mathscr{R}(\psi)$, we rewrite the equation for i as a quadratic equation in i, $Ai^2 + Bi + C = 0$, where after dividing by β in (9.5), the coefficients are

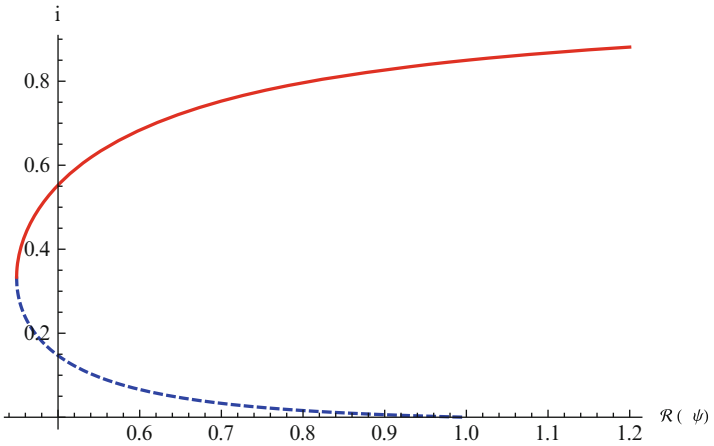

Fig. 9.2 The graph shows that the equilibrium value of i exhibits backward bifurcation as a function of the vaccine-dependent reproduction number. The parameters are taken as follows: $\mu = 0.01$, $\gamma = 3$, $\chi = 1$, $\delta = 0.1$, $\psi = 1$

$$\begin{aligned} A &= \beta\delta\mu, \\ B &= \mu(\mu+\gamma) + \delta(\mu+\gamma)(\mu+\psi) - \beta\mu\delta - (\mu+\delta\psi)\chi\gamma - \delta(1-\chi)\gamma(\mu+\psi), \\ C &= \mu(\mu+\psi)(1-\mathscr{R}(\psi)). \end{aligned} \tag{9.6}$$

We express these coefficients as functions of $\mathscr{R}(\psi)$ and eliminate β:

$$\begin{aligned} A &= \mathscr{R}(\psi)\eta\delta\mu, \\ B &= \mu(\mu+\gamma) + \delta(\mu+\gamma)(\mu+\psi) - \mathscr{R}(\psi)\eta\mu\delta - (\mu+\delta\psi)\chi\gamma - \delta(1-\chi)\gamma(\mu+\psi), \\ C &= \mu(\mu+\psi)(1-\mathscr{R}(\psi)), \end{aligned} \tag{9.7}$$

where $\eta = \frac{(\mu+\gamma)(\mu+\psi)}{(\mu+\delta\psi)}$. We illustrate the backward bifurcation in Fig. 9.2.

Imperfect vaccines lead to backward bifurcation. It is not hard to see that in the model above, backward bifurcation does not occur if the vaccine is perfect, $\delta = 0$. Also, if there is no vaccination $\psi = 0$, then backward bifurcation does not occur. In this case, it can be seen that if $\mathscr{R}_0 < 1$, the disease-free equilibrium is globally stable.

The presence of backward bifurcation means that in practice, if we vaccinate with imperfect vaccine, we may need to reduce the vaccine reproduction number not below one but below a much smaller value under which there are no endemic equilibria. Thus, it may appear that vaccinating with imperfect vaccine makes the task of controlling the disease harder rather than easier. However, it must be noted that at the same time, vaccination increases the parameter space of the remaining parameters where the vaccine-dependent reproduction number is below one, and

the disease-free equilibrium is locally stable. To illustrate this idea, assume that μ, ψ, and δ are given and fixed. Then in the absence of vaccination, the region in the (γ,β)-plane where the disease-free equilibrium is stable is given by $\mu+\gamma>\beta$, since there, $\mathscr{R}_0<1$. In the presence of vaccination, the region of local stability of the disease-free equilibrium is given by

$$\frac{\mu+\psi}{\mu+\delta\psi}(\mu+\gamma)>\beta,$$

which is a larger region, since the fraction $(\mu+\psi)/(\mu+\delta\psi)$ is greater than one.

9.3 Vaccination and Genetic Diversity of Microorganisms

When a pathogen is represented by several variants, they may not all be included in the vaccine. The strains that are included in the vaccine are called *vaccine strains*. The number of strains included in the vaccine is called vaccine *valency*. For instance, the flu vaccine is trivalent, that is, it contains three strains.

The immunity that a vaccine creates is specific to those strains that are included in the vaccine. The vaccine may provide partial immunity, or no immunity at all, to strains that are not included in the vaccine. That makes impossible the eradication of diseases whose causative agents mutate and that are represented by multiple variants.

Biologists report an increase of genetic diversity after the introduction of vaccination [142]. In terms of modeling, this says that vaccination should cause coexistence of pathogen variants, in other words, vaccination is a coexistence mechanism. To see this, we consider the model above with two strains. We assume that one of the strains is a vaccine strain with respect to which the vaccine is perfect. With respect to the other strain, the vaccine offers only partial protection. The model with two strains and vaccination becomes

$$\begin{aligned}
\frac{dS}{dt} &= \Lambda-\frac{\beta_1 SI}{N}-\frac{\beta_2 SJ}{N}-(\mu+\psi)S+\chi\gamma I+\alpha J,\\
\frac{dI}{dt} &= \frac{\beta_1 SI}{N}+\frac{\beta_1\delta VI}{N}-(\mu+\gamma)I,\\
\frac{dJ}{dt} &= \frac{\beta_2 SJ}{N}-(\mu+\alpha)J,\\
\frac{dV}{dt} &= \psi S-\frac{\beta_1\delta VI}{N}+(1-\chi)\gamma I-\mu V,
\end{aligned} \tag{9.8}$$

where $I(t)$ is the number infected with the first strain, and $J(t)$ is the number infected with the second strain. The parameter α is the per capita recovery rate from the second strain. Recovered individuals from the second strain go to the susceptible class, because only susceptible individuals can become infected with the second strain. The second strain is assumed to be the vaccine strain. The reproduction number of

the first strain is as before:

$$\mathscr{R}_1(\psi) = \frac{\beta_1(\mu + \delta\psi)}{(\mu+\gamma)(\mu+\psi)}.$$

The reproduction number of the second strain is

$$\mathscr{R}_2(\psi) = \frac{\beta_2\mu}{(\mu+\alpha)(\mu+\psi)}.$$

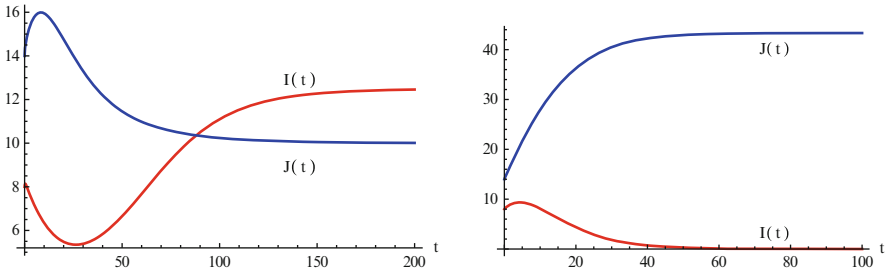

Fig. 9.3 The *left figure* illustrates that the number infected with strain one, $I(t)$, and the number infected with strain two, $J(t)$, may tend toward a coexistence equilibrium when $\psi = 0.5$. The *right figure* illustrates that if $\psi = 0$, strain two eliminates strain one. The remaining parameters used for these figures are $\beta_1 = 6$, $\beta_2 = 4.5$, $\gamma = 0.8$, $\alpha = 0.5$, $\mu = 0.1$, $\chi = 1.0$, $\delta = 0.04$, $\Lambda = 5$. The corresponding reproduction numbers are given by $\mathscr{R}_1(\psi) = 1.333$ and $\mathscr{R}_2(\psi) = 1.25$. The reproduction numbers in the absence of vaccination are $\mathscr{R}_1 = 6.66667$ and $\mathscr{R}_2 = 7.5$

Proving the existence of a unique coexistence equilibrium is possible but not trivial. So to see the coexistence, we do a simulation. Figure 9.3 illustrates the coexistence.

Question: What causes the coexistence? We can answer this question by examining the parts for the model that cause the coexistence. In particular, we examine the equations for the coexistence equilibrium:

$$\begin{aligned} 0 &= \mu - \beta_1 si - \beta_2 sj - (\mu+\psi)s + \chi\gamma i + \alpha j, \\ 0 &= \beta_1 si + \beta_1\delta vi - (\mu+\gamma)i, \\ 0 &= \beta_2 sj - (\mu+\alpha)j, \\ 0 &= \psi s - \beta_1\delta vi + (1-\chi)\gamma i - \mu v, \end{aligned} \tag{9.9}$$

where as before, s, i, j, v denote the proportions. If $\delta = 0$, then from the second and third equations, we have

$$s = \frac{\mu + \gamma}{\beta_1}, \qquad s = \frac{\mu + \alpha}{\beta_2}.$$

Clearly these two expressions for s are equal in very special cases, but not in general. So coexistence does not occur. Thus a necessary condition for coexistence is the imperfection of the vaccine. If there is no vaccination, that is, $\psi = 0$ and $\chi = 1$ (no recovery to the vaccinated class), then $v = 0$, and s must satisfy the same two expressions. So coexistence does not occur. Thus vaccination, and particularly vaccine imperfections, are the cause of coexistence.

When a disease is caused by a pathogen of multiple variants, not all of them are included in a vaccine (for various reasons). Vaccination is carried out under several scenarios:

1. Vaccination is carried against the dominant subtype. For instance, *Haemophilus influenzae* is represented by six serotypes: a, b, c, d, e, f, but before vaccination was instituted, serotype b caused most disease. Vaccination is now carried out against serotype b.
2. Vaccination is carried out against several strains that account for most cases. For instance, *Streptococcus pneumoniae* is represented by more than 90 serotypes, but only 23 of the most common ones are included in the polysaccharide vaccine.
3. When possible, vaccination is carried out against all subtypes (possibly one by one). For instance, poliomyelitis (caused by poliovirus, PV) is represented by three serotypes. Vaccination against each one is necessary, but polio has been nearly eradicated.

When vaccination is carried out against only one or more but not all of the pathogen variants, what is observed is decline in the number of disease cases caused by those variants included in the vaccine. At the same time, disease cases caused by other pathogen variants not included in the vaccine rise. This phenomenon is called *strain (serotype) replacement* (Table 9.4). The main mechanism by which serotype replacement occurs is that the vaccine has *differential effectiveness*: it is very effective with respect to some strains, and very little effective, or not effective at all, with respect to other strains. Thus vaccinated individuals are removed from the susceptible pool of the vaccine strains but effectively added to the susceptible pool of the nonvaccine strains, since the vaccine strains can no longer infect them.

That differential effectiveness of the vaccine leads to strain replacement can be seen from model (9.8). We illustrate this in Fig. 9.4. We note that the overall prevalence before vaccination is greater than the prevalence after vaccination. Thus replacement cannot completely "erase" what is being gained from vaccination. However, strain replacement is undesirable, because it still takes from what could have been gained.

Since differential effectiveness of the vaccine leads to replacement, vaccine developers have tried to make vaccines less differentially effective. One way to

Table 9.4 Reported increases in nonvaccine strains after vaccination [109]

Disease	Vaccine	Increase in	Region
H. influenzae	Hib	Nontype b	Alaska
	Hib	Type f	m. states, US
	conj. Hib	Type a	Brazil
	conj. Hib	Noncapsulated	UK
S. pneumoniae	PCV-7	NVT	Finland
	PCV-7	NVT (carriage)	US
	PCV-7	Serogroups 15 and 33	US PMPSG, US
	PCV-7	NVT (AOM)	Pittsburgh
	PPV-23	12F*, 7F, 22F, 7C	Alaska
N. meningitidis	A-C vaccine	Serogroup B	Austria
	A-C vaccine	Serogroup B	Europe
	A-C vaccine	Serogroup B	Cuba

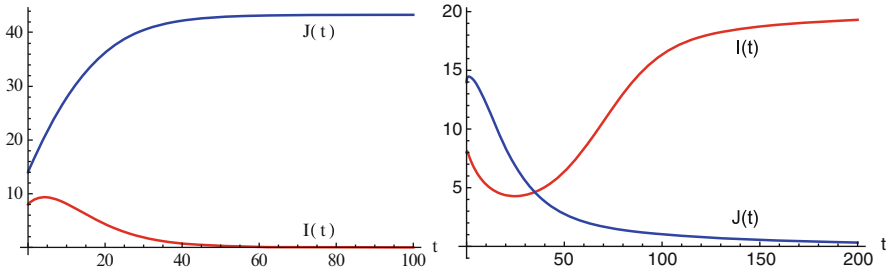

Fig. 9.4 The *left figure* illustrates that the number infected with strain one $I(t)$ tends to zero and the number infected with strain two $J(t)$ tends toward an endemic equilibrium when $\psi = 0$. The *right figure* illustrates that if $\psi = 0.7$, then strain one eliminates strain two. The remaining parameters used for these figures are $\beta_1 = 6$, $\beta_2 = 4.5$, $\gamma = 0.8$, $\alpha = 0.5$, $\mu = 0.1$, $\chi = 1.0$, $\delta = 0.04$, $\Lambda = 5$. The corresponding reproduction numbers are given by $\mathscr{R}_1(\psi) = 1.06667$ and $\mathscr{R}_2(\psi) = 0.9375$. The reproduction numbers in the absence of vaccination are $\mathscr{R}_1 = 6.66667$ and $\mathscr{R}_2 = 7.5$

do that is to include (if possible) more strains in the vaccine. That has been the case with pneumococcal polysaccharide vaccine, which originally contained very few serotypes of *Streptococcus pneumoniae* but now contains 23. That is still many fewer than the 90 serotypes that exist. A new approach is to target surface proteins that are common in all 90 serotypes.

Question: Suppose we can produce a vaccine that is perfect with respect to all strains. Will we eliminate strain replacement?

The answer is expected to be affirmative if differential effectiveness is the mechanism behind strain replacement. Although such perfect vaccines do not yet exist, we can address this question with mathematical models. Consider the model of superinfection. We add vaccination with a perfect vaccine to this model. Thus, the model becomes

$$\begin{aligned}
\frac{dS}{dt} &= \Lambda - \beta_1 \frac{SI}{N} - \beta_2 \frac{SJ}{N} - (\mu + \psi)S + \gamma_1 I + \gamma_2 J, \\
\frac{dI}{dt} &= \beta_1 \frac{SI}{N} + \beta_1 \delta \frac{IJ}{N} - (\mu + \gamma_1)I, \\
\frac{dJ}{dt} &= \beta_2 \frac{SJ}{N} - \beta_1 \delta \frac{IJ}{N} - (\mu + \gamma_2)J, \\
\frac{dV}{dt} &= \psi S - \mu V,
\end{aligned} \tag{9.10}$$

where $N = S + I + J + V$ is the total population. Notice that vaccinated individuals cannot become infected with any of the strains. It turns out, however, that strain replacement still occurs. We illustrate this in Fig. 9.5. What is causing it? If the vaccine is not "differentially effective," how does it differentiate between the strains? In what follows, we address these questions.

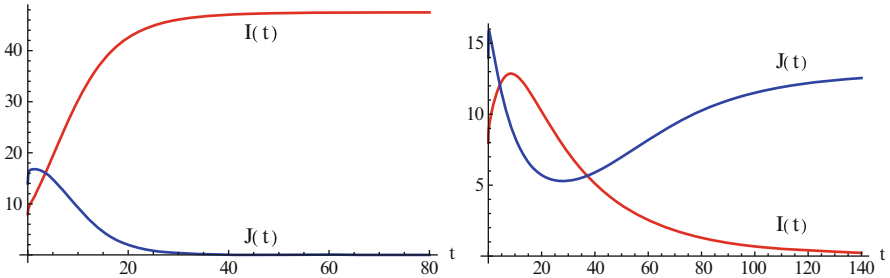

Fig. 9.5 The *left figure* illustrates that the number infected with strain one, $I(t)$, tends to a nonzero equilibrium, and the number infected with strain two, $J(t)$, tends to zero when $\psi = 0$. The *right figure* illustrates that if $\psi = 1.75$, then strain two eliminates strain one. The remaining parameters used for these figures are $\beta_1 = 12$, $\beta_2 = 15$, $\gamma_1 = 0.5$, $\gamma_2 = 0.5$, $\mu = 0.1$, $\delta = 0.03$, $\Lambda = 5$

The reproduction numbers of the two strains are given by

$$\mathscr{R}_i(\psi) = \frac{\beta_i \mu}{(\mu + \gamma_i)(\mu + \psi)}, \qquad i = 1, 2.$$

Note that they are both decreasing functions of the vaccination rate ψ. In addition, they do not depend on superinfection, and particularly on the coefficient of reduction or enhancement δ, since superinfection does not lead to infection of susceptible individuals.

The corresponding invasion reproduction numbers, however, are not independent of the superinfection process, since they measure the number of secondary infections one strain-i infected individual will produce in a population in which strain j is at equilibrium. To compute the invasion numbers, we first compute the two dominance equilibria. The dominance equilibrium of strain one is given by $\mathscr{E}_1 = (s, i, 0, v)$ with

$$\mathscr{E}_1 = \left(\frac{1}{\mathscr{R}_1}, 1 - \frac{1}{\mathscr{R}_1(\psi)}, 0, \frac{\psi}{\mu} \frac{1}{\mathscr{R}_1} \right),$$

where $\mathscr{R}_i = \mathscr{R}_i(0)$. The dominance equilibrium of strain two is given by $\mathscr{E}_2 = (s,0,j,v)$ with

$$\mathscr{E}_2 = \left(\frac{1}{\mathscr{R}_2}, 0, 1 - \frac{1}{\mathscr{R}_2(\psi)}, \frac{\psi}{\mu}\frac{1}{\mathscr{R}_2}\right).$$

The invasion reproduction number of strain one is obtained from differentiating the right-hand side of the equation for I with respect to I (to get the respective diagonal entry in the Jacobian). We get $\beta_1 s + \beta_1 \delta j - (\mu + \gamma_1)$. We substitute s and j from $\mathscr{E}_2$. Therefore, the invasion reproduction number of the first strain is given by

$$\hat{\mathscr{R}}_1 = \frac{\mathscr{R}_1}{\mathscr{R}_2} + \delta \mathscr{R}_1 \left(1 - \frac{1}{\mathscr{R}_2(\psi)}\right).$$

An important observation here is that as the vaccination rate ψ increases, $\hat{\mathscr{R}}_1$ decreases. Thus, vaccination decreases the invasion capabilities of the first strain.

To obtain the invasion reproduction number of strain two, we differentiate the right-hand side of the equation for J with respect to J (to get the respective diagonal

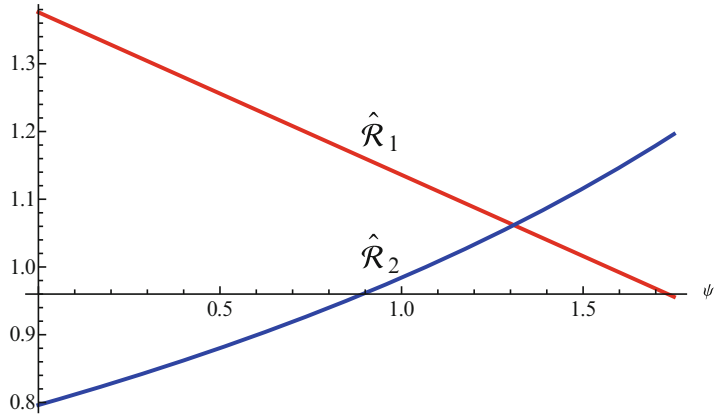

Fig. 9.6 The graph shows the invasion reproduction numbers $\hat{\mathscr{R}}_1$ and $\hat{\mathscr{R}}_2$ as functions of ψ. Clearly, $\hat{\mathscr{R}}_1$ is a decreasing function ofψ. In contrast, $\hat{\mathscr{R}}_2$ is an increasing function ofψ. The parameters are taken as follows: $\beta_1 = 12$, $\beta_2 = 15$, $\mu = 0.1$, $\gamma_1 = 0.5$, $\gamma_2 = 0.5$, $\delta = 0.03$

entry in the Jacobian). We get $\beta_2 s - \beta_1 \delta i - (\mu + \gamma_2)$. We substitute s and i from $\mathscr{E}_1$. Therefore, the invasion reproduction number of the second strain is given by

$$\hat{\mathscr{R}}_2 = \frac{(\mu + \gamma_2)\frac{\mathscr{R}_2}{\mathscr{R}_1}}{(\mu + \gamma_2) + (\mu + \gamma_1)\mathscr{R}_1 \delta \left(1 - \frac{1}{\mathscr{R}_1(\psi)}\right)}.$$

In contrast, the invasion reproduction number of the second strain is an increasing function of the vaccination rate. Hence, vaccination increases the invasion capabilities of the second strain. The reason for this effect is that when the two strains coexist, increasing the vaccination rate decreases the number of those infected with strain 1. That, in turn, reduces the superinfections, which take away from

the infections with the second strain. This produces an overall effect of increase in infections with the second strain.

We illustrate the trend with increasing ψ in the two invasion reproduction numbers in Fig. 9.6. Figure 9.6 also shows that there is a vaccination level ψ_1^* such that for $\psi < \psi_1^*$, the following conditions are satisfied: $\hat{\mathscr{R}}_2 < 1$ and $\hat{\mathscr{R}}_1 > 1$ (while $\mathscr{R}_1(\psi) > 1$ and $\mathscr{R}_2(\psi) > 1$). In this case, strain one, which can invade the equilibrium of strain two, dominates, since strain two cannot invade the equilibrium of strain one. Then there is a vaccination level ψ_2^* such that for $\psi_1^* < \psi < \psi_2^*$, the following conditions are satisfied: $\hat{\mathscr{R}}_2 > 1$ and $\hat{\mathscr{R}}_1 > 1$. In this case, both strains can invade each other's equilibrium, and therefore, they coexist. For vaccination levels $\psi > \psi_2^*$, the following conditions are satisfied: $\hat{\mathscr{R}}_2 > 1$ and $\hat{\mathscr{R}}_1 < 1$. In this case, strain two, which can invade the equilibrium of strain one, dominates, since strain one cannot invade the equilibrium of strain two. Thus, replacement of strain one, which dominated without vaccination, has occurred. The replacing strain is strain two.

9.4 Modeling Quarantine and Isolation

Quarantine and isolation are typically modeled by introducing separate classes into the model. Isolation is more often employed as a control strategy in epidemic models than quarantine. Isolated infected individuals move to a separate class Q. A simple extension of the SIR model with isolation will take the form

$$\begin{aligned}
\frac{dS}{dt} &= \Lambda - \beta \frac{SI}{N-Q} - \mu S, \\
\frac{dI}{dt} &= \beta \frac{SI}{N-Q} - (\mu + \sigma + r_2) I, \\
\frac{dQ}{dt} &= \sigma I - (\mu + r_1) Q, \\
\frac{dR}{dt} &= r_2 I + r_1 Q - \mu R.
\end{aligned} \tag{9.11}$$

We note here that in standard incidence, the total active population is $N - Q$.

Epidemic models with isolation have been considered with respect to different diseases. Isolation has been found to destabilize the dynamics and lead to oscillations [61, 73] (see Chap. 7). As a result, isolation has been suggested as a potential intrinsic mechanism responsible for the recurrent outbreaks of childhood diseases [61].

The controlled reproduction number of model (9.11) is given by

$$\mathscr{R}_c = \frac{\beta}{\mu + \sigma + r_2}.$$

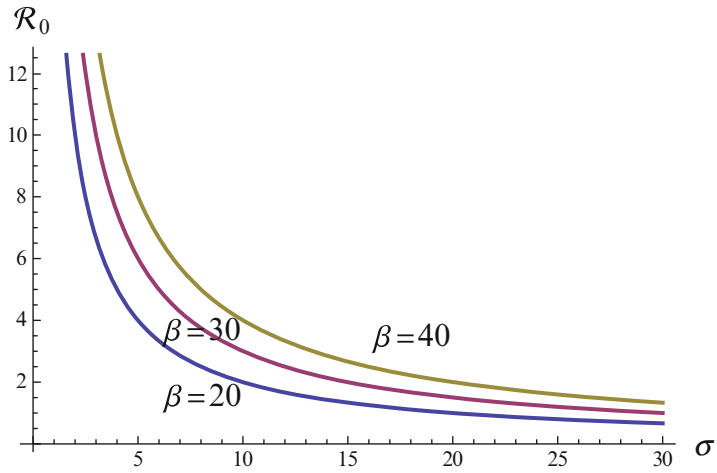

Fig. 9.7 The figure shows the controlled reproduction number as a decreasing, concave up, function of σ for several values of β. The smaller the β, the steeper the decline in $\mathcal{R}_c$

The reproduction number is a decreasing function of the isolation rate σ. The critical isolation rate that gives $\mathcal{R}_0 = 1$ is given by $\sigma^* = \beta - \mu - r_2$. The reproduction number as a function of σ is plotted in Fig. 9.7.

The disease prevalence at equilibrium is given by

$$I^* = \frac{\Lambda \mathcal{R}_c(\mathcal{R}_c - 1)}{\beta(\mathcal{R}_c - 1) + \mu(1+p)\mathcal{R}_c},$$

where

$$p = \left(\frac{r_2}{\mu} + \frac{r_1\sigma}{\mu(\mu + r_1)}\right).$$

The prevalence is also a decreasing function of σ, at least when $\mathcal{R}_c > 1$. However, Fig. 9.8 suggests that the nonzero endemic equilibrium exists even if $\mathcal{R}_c < 1$. It can be shown that this equilibrium is unstable and that for $\mathcal{R}_c < 1$ the disease-free equilibrium is globally stable.

To design a model with quarantine and isolation, we need to express in terms of equations the events that happen in reality. Susceptible individuals S come into a contact with infectious I and exposed E individuals and move to the exposed class E. At the same time, the contacts of infectious individuals are traced. Some of the traced individuals happen to be susceptible, and others happen to be exposed. Traced susceptible individuals move to the quarantine class Q_1 at a rate ρ. Traced exposed individuals move to the quarantine class at a rate ρ. Quarantined individuals either show no symptoms and after the end of the quarantine return to the susceptible class, or they become sick and move to the isolated class Q_2:

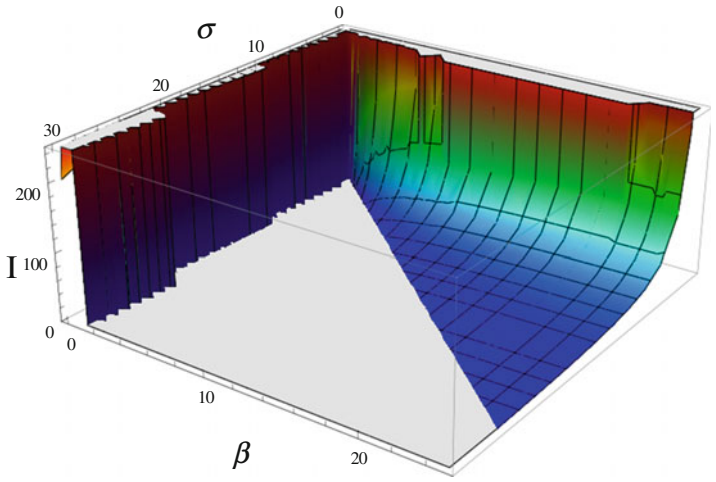

Fig. 9.8 The figure shows the prevalence as a decreasing, concave up function of σ and an increasing function of β. The smaller the β, the steeper the decline in prevalence

$$
\begin{aligned}
\frac{dS}{dt} &= \Lambda - \beta \frac{S(I+qE)}{N-Q} - \rho S - \mu S + \eta_1 Q_1, \\
\frac{dE}{dt} &= \beta \frac{S(I+qE)}{N-Q} - \rho E - (\mu+\gamma)E, \\
\frac{dQ_1}{dt} &= \rho S + \rho E - (\mu+\eta_1+\eta_2)Q_1, \\
\frac{dI}{dt} &= \gamma E - (\mu+\sigma+r_2)I, \\
\frac{dQ_2}{dt} &= \sigma I + \eta_2 Q_1 - (\mu+r_1)Q_2, \\
\frac{dR}{dt} &= r_2 I + r_1 Q_2 - \mu R.
\end{aligned}
\tag{9.12}
$$

Quarantined and isolated individuals do not participate in the total active population, so the total active population in the denominator of the standard incidence is given by $N - Q_1 - Q_2 = N - Q$, where $Q = Q_1 + Q_2$. Infectious and isolated individuals recover and move to the recovered class R. The meaning and the values of the parameters are given in Table 9.5.

The controlled reproduction number is given by

$$\mathscr{R}_c = \frac{\beta\gamma}{(\gamma+\rho+\mu)(r_2+\sigma+\mu)} + \frac{q\beta}{\gamma+\rho+\mu}.$$

In interpreting the controlled reproduction number, we notice that the first term is the number of secondary infections generated by the infectious individuals, and the

Table 9.5 Parameter meanings and parameter values [120]

Parameter	Parameter meaning	Value
Λ	Recruitment rate	240 people/day
β	Transmission rate	0.25 per day
ρ	Quarantine rate	1/10 per day
μ	Natural death rate	1/(70*365) per day
γ	Rate of developing symptoms	1/6 per day
σ	Isolation rate	1/5 per day
η_1	Rate of return to susceptible class	1/10 per day
η_2	Rate of progression to infectiousness	1/6.5 per day
r_1	Recovery rate for isolated individuals	1/20 per day
r_2	Recovery rate for infectious individuals	1/25 per day
q	Reduction of infectivity of exposed individuals	0.8 (variable)

second term is the number of secondary infections generated by exposed individuals; $\gamma/(\gamma+\rho+\mu)$ is the proportion of exposed individuals who move to the infectious class.

We plot the region $\mathscr{R}_c > 1$ for two values of $q = 0.5$ and $q = 0.8$ in Fig. 9.9. The region for $q = 0.8$ is larger and asymmetric. We plot the point with coordinates given in Table 9.5 in red. That point belongs to the region $\mathscr{R}_c > 1$; hence for the quarantine and isolation rates in Table 9.5, the disease will not be eradicated. We would like to compute the values of quarantine and isolation rates that will represent the smallest change from the values in Table 9.5 but will lead to eradication of the disease. For that reason, we compute the point on the curve $\mathscr{R}_c = 1$ that is closest to the red point. To do that, let the black point have coordinates (x,y). The square of the distance between the two points is given by

$$(x-0.1)^2+(y-0.2)^2,$$

where $(0.1,0.2)$ are the coordinates of the red point. Furthermore, we replace ρ with x and σ with y in $\mathscr{R}_c$. From the equation $\mathscr{R}_c = 1$, we express σ (or y) as a function of ρ (or x): $y = f(x)$. Substituting y in the distance formula, we obtain the square of the distance as a function of x:

$$(x-0.1)^2+(f(x)-0.2)^2.$$

To minimize that function, we differentiate with respect to x and set the derivative to zero. This leads to the equation

$$(x-0.1)+(f(x)-0.2)f'(x)=0.$$

In the case $q = 0.5$, the black point has coordinates $(0.12214,0.214265)$; in the case of $q = 0.8$, the black point has coordinates $(0.180881,0.242281)$. From these coordinates, we need to see how we need to change the quarantine and isolation

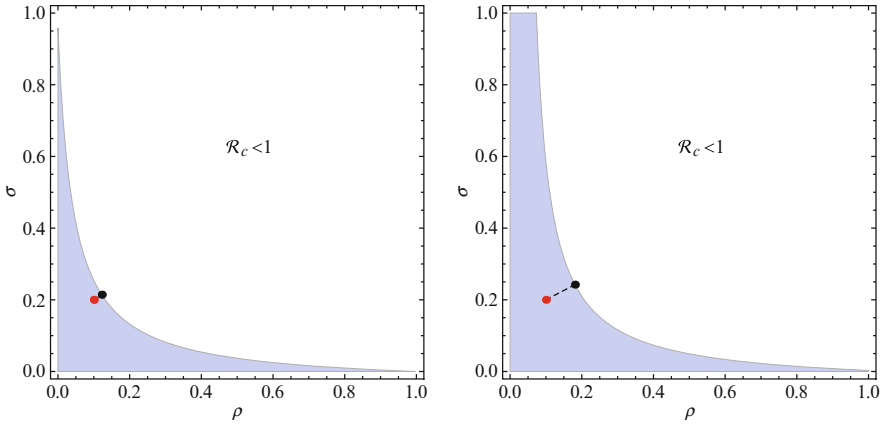

Fig. 9.9 Both figures illustrate the region $\mathcal{R}_c > 1$, the epidemic situation with a *red point*, and the closest point on the curve $\mathcal{R}_c = 1$ in *black*. The *left figure* does so for $q = 0.5$, while the *right figure* gives the same scenario for $q = 0.8$

rates to achieve elimination. The optimal new periods for quarantine and isolation are given by $1/c$, where c is a coordinate of a black point. These optimal periods that will lead to elimination are listed in Table 9.6.

Table 9.6 Optimal periods for quarantine and isolation

Strategy	$q = 0.5$	$q = 0.8$
$1/\rho$	8.19 days	5.53 days
$1/\sigma$	4.67 days	4.13 days

This table suggests that in the case $q = 0.8$, the contact tracing and quarantining should improve dramatically from 10 days to 5.5 days, while isolation should improve from 5 days to 4 days, in order for the disease to be eliminated.

9.5 Optimal Control Strategies

In previous sections, we considered control strategies to be constant in time, but in reality, control strategies are variable in time. The mathematical theory used to derive optimal control strategies that vary in time is called *optimal control theory*. In this section, we introduce the basic principles and illustrate them with examples.

9.5.1 Basic Theory of Optimal Control

Optimal control is applied to differential equation models in normal form. Here we will be concerned with ordinary differential equation models. We consider a system of ODEs

$$\begin{aligned} \mathbf{x}'(t) &= \mathbf{f}(\mathbf{x}(t)), \\ \mathbf{x}(0) &= \mathbf{x}_0, \end{aligned} \tag{9.13}$$

where the given initial condition is $\mathbf{x}_0 \in \mathbb{R}^n$, and $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$. The unknown vector is $\mathbf{x} : [0,\infty) \to \mathbb{R}^n$.

Now we generalize the setup and suppose that the right-hand side depends on a parameter $\mathbf{u} : [0,\infty) \to A$, where $A \subset \mathbb{R}^m$, that is allowed to depend on time $u(t)$. Thus the system above becomes

$$\begin{aligned} \mathbf{x}'(t) &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \\ \mathbf{x}(0) &= \mathbf{x}_0, \\ \mathbf{x}(T) &\quad \text{free}. \end{aligned} \tag{9.14}$$

The variable $\mathbf{u}(t)$ is called *control*, and in the presence of the control, the solution $\mathbf{x}(t)$ depends on the control. The trajectory that corresponds to the control $\mathbf{u}(t)$ is called a corresponding *response* of the system.

To make this presentation more specific, we recast some of the models for vaccination and isolation from this chapter in the framework of control. For instance, in model (9.2), the "control" is the vaccination, given by the vaccination rate ψ. Hence, the right-hand side of (9.2) depends on the dependent variables and the control parameter ψ. Now we let ψ vary with time, and we replace it with $u(t)$. We obtain the following problem with control:

$$\begin{aligned} \frac{dS}{dt} &= \Lambda - \frac{\beta SI}{N} - (\mu + u(t))S + \chi\gamma I, \\ \frac{dI}{dt} &= \frac{\beta SI}{N} + \frac{\beta\delta VI}{N} - (\mu+\gamma)I, \\ \frac{dV}{dt} &= u(t)S - \frac{\beta\delta VI}{N} + (1-\chi)\gamma I - \mu V. \end{aligned} \tag{9.15}$$

Here the control is given by $u : [0,\infty) \to \mathbb{R}_+$.

We introduce the set of *admissible controls*

$$\mathscr{A} = \{\mathbf{u}(t) \in L^1(0,T) | \mathbf{u}(t) \in A\}.$$

As posed, problem (9.14) does not have a solution, since the control may be arbitrary. We need to find the best control in some sense. For disease-control models, we need to find the control in such a way that we minimize the prevalence and/or minimize the cost of controlling the disease. To make this more specific, we define a *payoff functional*:

$$\mathscr{C}[\mathbf{u}] := \int_0^T g(\mathbf{x}(t), \mathbf{u}(t))dt, \tag{9.16}$$

where $\mathbf{x}(t)$ solves (9.14) for the specified control $\mathbf{u}(t)$. The function $g : \mathbb{R}^n \times A \to \mathbb{R}$ is given. The terminal time T is given as well. The function g is called the *running payoff*. We need to solve the following optimal control problem: find a control $\mathbf{u}^*(t)$ that minimizes the payoff functional, that is,

$$\mathscr{C}[\mathbf{u}^*] = \min_{u \in \mathscr{A}} \mathscr{C}[u].$$

If such a control $\mathbf{u}^*(t)$ exists, it is called an *optimal control*. The optimal control together with the corresponding solution gives the optimal control pair $(\mathbf{x}^*, \mathbf{u}^*)$.

The first question that must be addressed is whether an optimal control pair $(\mathbf{x}^*(t), \mathbf{u}^*(t))$ exists. The question of existence is settled by the following theorem [143]:

Theorem 9.1 (Filippov–Cesari Existence Theorem). *For all $(t, \mathbf{x}) \in \mathbb{R}^{n+1}$, define the set*

$$N(t, \mathbf{x}) = \{(g(\mathbf{x}, \mathbf{u}) + \xi, \mathbf{f}(\mathbf{x}, \mathbf{u})) : \xi \leq 0, \mathbf{u} \in A\}.$$

Suppose that

1. *$N(t, \mathbf{x})$ is convex for every $(t, \mathbf{x})$.*
2. *A is compact.*
3. *There exists a constant $K > 0$ such that $||\mathbf{x}(t)|| \leq K$ for all $t \in (0, T)$ and all admissible pairs $(\mathbf{x}, \mathbf{u})$.*

Then there exists an optimal pair $(\mathbf{x}^*(t), \mathbf{u}^*(t))$, where $\mathbf{u}^*(t) \in \mathscr{A}$.

If a solution exists, it can be found with Pontryagin's minimum principle [143]. First, one introduces a time-varying Lagrange multiplier vector $\lambda(t)$, whose elements are called the *adjoint variables* of the system. Next, the Hamiltonian H is defined for all $t \in [0, T]$ by

$$H(\mathbf{x}(t), \mathbf{u}(t), \lambda(t)) = g(\mathbf{x}(t), \mathbf{u}(t)) + \sum_{i=1}^{n} \lambda_i(t) f_i(\mathbf{x}(t), \mathbf{u}(t)). \tag{9.17}$$

The Pontryagin minimum principle is as follows.

Theorem 9.2 (Pontryagin's Minimum Principle). *For the optimality of control $\mathbf{u}^*(t)$ and corresponding trajectory $\mathbf{x}^*(t)$ with $t \in [0, T]$, it is necessary that there exist a nonzero adjoint vector function $\lambda^*(t)$ that is a solution to the adjoint system*

$$\begin{aligned} \lambda'(t) &= -\frac{\partial H(\mathbf{x}(t), \mathbf{u}(t), \lambda(t))}{\partial x} \\ \lambda(T) &= 0, \end{aligned} \tag{9.18}$$

so that

$$H(\mathbf{x}^*(t), \mathbf{u}^*(t), \lambda^*(t)) = \min_{\mathbf{u} \in \mathscr{A}} H(\mathbf{x}^*(t), \mathbf{u}(t), \lambda^*(t)).$$

Thus, the necessary conditions for optimizing the Hamiltonian are [117]:

$$\begin{aligned}
&\frac{\partial H}{\partial u} = 0 \Longrightarrow g_u + \sum_{i=1}^{n} \lambda_i(t)(f_i)_u = 0, \quad \text{optimality equation,} \\
&\lambda_i'(t) = -\frac{\partial H(\mathbf{x}(t), \mathbf{u}(t), \lambda(t))}{\partial x_i} \Longrightarrow \lambda_i'(t) = -g_{x_i} - \sum_{i=1}^{n} \lambda_i(t)(f_i)_{x_i}, \text{ adjoint equation,} \\
&\lambda(T) = 0, \quad \text{transversality condition.}
\end{aligned} \tag{9.19}$$

We note that for minimization, we must also have

$$\frac{\partial^2 H}{\partial u^2} \geq 0 \quad \text{at} \quad \mathbf{u}^*.$$

The following theorem gives sufficient conditions for the existence and uniqueness of the optimal pair [143]:

Theorem 9.3 (Mangasarian Theorem). *Suppose*

1. *A is convex.*
2. *The partial derivative $\partial g/\partial u_j$ and $\partial f_i/\partial u_j$ all exist and are continuous.*
3. *The pair $(\mathbf{x}^*(t), \mathbf{u}^*(t))$ satisfies all conditions of the Pontryagin minimum principle.*
4. *$H(t, \mathbf{x}, \mathbf{u})$ is concave down in $(\mathbf{x}, \mathbf{u})$ for all $t \in [0, T]$.*

Then the pair $(\mathbf{x}^*(t), \mathbf{u}^*(t))$ solves the problem. If $H(t, \mathbf{x}, \mathbf{u})$ is strictly concave down in $(\mathbf{x}, \mathbf{u})$, then the solution is unique.

There are several excellent books that introduce optimal control theory applied to biological systems [13, 94]. We illustrate the application of the existence theorem and Pontryagin's minimum principle to finding the optimal control in the next subsection.

9.5.2 Examples

In this subsection we consider two examples of application of optimal control to epidemic models. The first example is an SIS model with treatment.

9.5.2.1 SIS Model with Treatment

The model assumes constant total population size N. In this case, the susceptible individuals can be represented as $S = N - I$, and the 2×2 system can be reduced to a single equation:

$$\begin{aligned}
&I'(t) = \beta(N - I)I - (\mu + \gamma)I - u(t)I, \\
&I(0) = I_0, \\
&I(T) \quad \text{free,}
\end{aligned} \tag{9.20}$$

where β is the transmission rate, μ is the natural death rate, and γ is the natural recovery rate without treatment. The term $u(t)I$ models the additional recovery rate due to treatment. The set of admissible controls is

$$\mathscr{A} = \{u(t) \in L^1(0,T) | 0 \le u(t) \le U_{\max}\},$$

where $U_{\max} < \infty$ is a positive constant. We are applying optimal control theory to determine the "best" treatment regime that will minimize the prevalence and the cost of applying the treatment. In particular, we seek a control u^* that minimizes the payoff functional

$$\mathscr{C}[u^*] = \min_{u \in \mathscr{A}} \int_0^T (w_1 I(t) + u^2(t))dt, \tag{9.21}$$

where w_1 is a constant cost of minimizing prevalence, and u^2 requires us to minimize the treatment, and also the cost of applying it. We assume that the cost of treatment is nonlinear and takes a quadratic form.

We first prove the existence of an optimal control pair. We use the Filippov–Cesari theorem.

Proposition 9.1. *The optimal control problem (9.20)–(9.21) has a solution.*

Proof. Let $N(t,\mathbf{x})$ be defined as in Theorem 9.1. Let $y_1, y_2 \in N(t,\mathbf{x})$. To show that $N(t,\mathbf{x})$ is convex for each $(t,\mathbf{x})$, we will show that the line connecting y_1 and y_2 lies entirely in $N(t,\mathbf{x})$. Hence, we have to show that

$$\alpha y_1 + (1-\alpha)y_2 \in N(t,\mathbf{x}) \quad \text{for every } \alpha \in [0,1].$$

The fact that $y_i \in N(t,\mathbf{x})$ implies that there exist $\xi_1, \xi_2 \le 0$ and control vectors $\mathbf{u}_1(t), \mathbf{u}_2(t) \in A$ such that

$$y_i = \{g(\mathbf{x},\mathbf{u}_i) + \xi_i, \mathbf{f}(\mathbf{x},\mathbf{u}_i)\} \quad \text{for} \quad i = 1,2.$$

Then, we have

$$\begin{aligned}
&\alpha(g(\mathbf{x},\mathbf{u}_1) + \xi_1) + (1-\alpha)(g(\mathbf{x},\mathbf{u}_2) + \xi_2) \\
&\quad = \alpha(w_1 I(t) + u_1^2(t)) + (1-\alpha)(w_1 I(t) + u_2^2(t)) + \alpha\xi_1 + (1-\alpha)\xi_2 \\
&\quad = w_1 I(t) + \alpha u_1^2 + (1-\alpha)u_2^2 + \alpha\xi_1 + (1-\alpha)\xi_2.
\end{aligned} \tag{9.22}$$

Letting $u_3 = \sqrt{\alpha u_1^2 + (1-\alpha)u_2^2}$, we notice that $u_3 \in A$. Furthermore, letting $\xi_3 = \alpha\xi_1 + (1-\alpha)\xi_2$, we notice that $\xi_3 \le 0$. Thus, the first component of the convex combination belongs to $N(t,\mathbf{x})$. Next, we check the second component:

$$\begin{aligned}
&\alpha(f(\mathbf{x},\mathbf{u}_1) + (1-\alpha)f(\mathbf{x},\mathbf{u}_2) \\
&\quad = \alpha(\beta(N-I)I - (\mu+\gamma)I - u_1(t)I) + (1-\alpha)(\beta(N-I)I - (\mu+\gamma)I - u_2(t)I) \\
&\quad = \beta(N-I)I - (\mu+\gamma)I - (\alpha u_1(t) + (1-\alpha)u_2)I.
\end{aligned} \tag{9.23}$$

Letting $u_4 = \alpha u_1(t) + (1-\alpha)u_2$, we notice that $u_4 \in A$. We conclude that the convex combination $\alpha y_1 + (1-\alpha)y_2$ is in $N(t,\mathbf{x})$. Clearly, A is compact. Next, we show that the solution of (9.20) is bounded. Indeed,

$$I'(t) \leq \beta(N-I)I.$$

We have that $I(t) \leq \sup_t \hat{I}$, where $\hat{I}$ is the solution of the equation $\hat{I}'(t) = \beta(N-\hat{I})\hat{I}$. Thus, $\sup_t I(t) \leq \max\{I_0, N\}$. If $I_0 \leq N$, then $\max_t\{I(t)\} \leq N$. This concludes the proof. □

To apply Pontryagin's minimum principle, we define the Hamiltonian:

$$H(I(t),u(t),\lambda(t)) = w_1 I(t) + u^2(t) + \lambda(t)(\beta(N-I(t))I(t) - (\mu+\gamma)I(t) - u(t)I(t)).$$

Posing the necessary conditions from Pontryagin's principle, we have first that u^* must be a critical point of the Hamiltonian, that is, we must have $\partial H/\partial u = 0$. This leads to the following condition on the optimal control: $2u - \lambda(t)I(t) = 0$. Hence, we have

$$u^*(t) = \frac{\lambda(t)I(t)}{2}.$$

Next, we check that the critical point is indeed a minimum: $\partial^2 H/\partial u^2 = 2 > 0$. The adjoint system is given by

$$\begin{aligned} &\lambda'(t) = -w_1 - \lambda(t)(\beta(N-I(t)) - \beta I(t) - (\mu+\gamma) - u(t)), \\ &\lambda(T) = 0. \end{aligned} \tag{9.24}$$

Since u^* must belong to $\mathscr{A}$, we must have

$$u^*(t) = \min\left\{U_{\max}, \max\left\{0, \frac{\lambda(t)I(t)}{2}\right\}\right\}. \tag{9.25}$$

To find the optimal control and the prevalence that corresponds to it, we must solve the system

$$\begin{aligned} &I'(t) = \beta(N-I)I - (\mu+\gamma)I - u^*(t)I, \\ &I(0) = I_0, \\ &\lambda'(t) = -w_1 - \lambda(t)(\beta(N-I(t)) - \beta I(t) - (\mu+\gamma) - u^*(t)), \\ &\lambda(T) = 0, \end{aligned} \tag{9.26}$$

where u^* is given by (9.25). System (9.26) cannot be solved by hand, and numerical methods must be used. Both Mathematica and Matlab can be used. Mathematica's `NDSolve` can take in boundary conditions, and system (9.26) can be directly input into it. The optimal control and the respective solution are plotted in Fig. 9.10.

Matlab requires use of numerical methods to solve the system of differential equations. The forward–backward sweep method [94] is often employed in this case. It combines the forward application of a fourth-order Runge–Kutta method

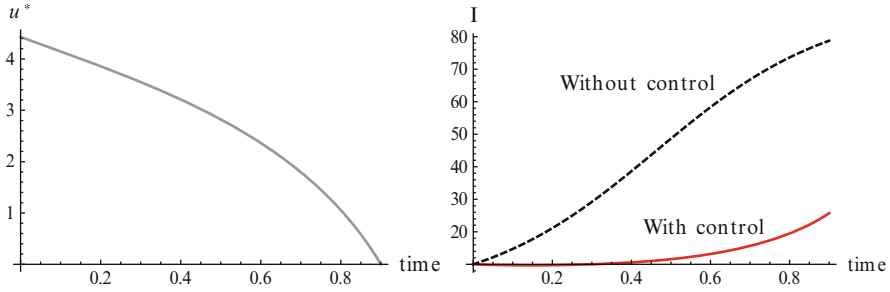

Fig. 9.10 The *left figure* shows the optimal control $u^*(t)$. The *right figure* shows the controlled prevalence $I^*(t)$ and the original prevalence $I(t)$

for the original system with the backward application of a fourth-order Runge–Kutta method for the adjoint system. The Matlab code for system (9.26) is included in the appendix.

9.5.2.2 Two-Strain Model with Vaccination

The second model is the model with two strains and vaccination given in Eq. (9.10). The control $u(t)$ replaces the vaccination rate ψ. The model with control becomes

$$\begin{aligned}
\frac{dS}{dt} &= \Lambda - \beta_1 \frac{SI}{N} - \beta_2 \frac{SJ}{N} - (\mu + u(t))S + \gamma_1 I + \gamma_2 J, \\
\frac{dI}{dt} &= \beta_1 \frac{SI}{N} + \beta_1 \delta \frac{IJ}{N} - (\mu + \gamma_1)I, \\
\frac{dJ}{dt} &= \beta_2 \frac{SJ}{N} - \beta_1 \delta \frac{IJ}{N} - (\mu + \gamma_2)J, \\
\frac{dV}{dt} &= u(t)S - \mu V.
\end{aligned} \tag{9.27}$$

The set of admissible controls is

$$\mathscr{A} = \{u(t) \in L^1(0,T) | 0 \leq u(t) \leq U_{\max}\}.$$

We are applying optimal control theory to determine the "best" vaccination regime that will minimize the the prevalence and the cost of applying the vaccination. In particular, we seek a control u^* that minimizes the payoff functional

$$\mathscr{C}[u^*] = \min_{u \in \mathscr{A}} \int_0^T (w_1 I(t) + w_2 J(t) + w_3 uS(t) + u^2(t))dt,$$

where w_1, w_2 are constant costs of minimizing prevalence, and the term uS intends to minimize the number of vaccines used with constant weight w_3. Finally, u^2 requires us to minimize the vaccination rate, and also the cost of vaccination. We assume that the cost of vaccination is nonlinear and takes a quadratic form.

To apply Pontryagin's minimum principle, we define the Hamiltonian:

$$\begin{aligned} H = {} & w_1 I + w_2 J + w_3 uS + u^2 + \lambda_S \left(\Lambda - \beta_1 \frac{SI}{N} - \beta_2 \frac{SJ}{N} - (\mu + u(t))S + \gamma_1 I + \gamma_2 J \right) \\ & + \lambda_I \left(\beta_1 \frac{SI}{N} + \beta_1 \delta \frac{IJ}{N} - (\mu + \gamma_1) I \right) \\ & + \lambda_J \left(\beta_2 \frac{SJ}{N} - \beta_1 \delta \frac{IJ}{N} - (\mu + \gamma_2) J \right) + \lambda_V (uS - \mu V). \end{aligned} \tag{9.28}$$

Again, applying the necessary conditions from Pontryagin's principle, we have first that u^* must be a critical point of the Hamiltonian, that is, we must have $\partial H / \partial u = 0$. This leads to the following condition on the optimal control: $2u - \lambda_S(t)S(t) + \lambda_V(t)S(t) + w_3 S(t) = 0$. This leads to the following expression for the control:

$$u^*(t) = \frac{(\lambda_S(t) - \lambda_V(t) - w_3)S(t)}{2}.$$

Next, we check that critical point is indeed a minimum: $\partial^2 H / \partial u^2 = 2 > 0$. The adjoint system is given by

$$\begin{aligned} \lambda_S'(t) = {} & -w_3 u - \lambda_S \left(\beta_1 \frac{I}{N} + \beta_1 \frac{SI}{N^2} - \beta_2 \frac{J}{N} + \beta_2 \frac{SJ}{N^2} - (\mu + u) \right) \\ & - \lambda_I \left(\beta_1 \frac{I}{N} - \beta_1 \frac{SI}{N^2} - \beta_1 \delta \frac{IJ}{N^2} \right) \\ & - \lambda_J \left(\beta_2 \frac{J}{N} - \beta_2 \frac{SJ}{N^2} + \beta_1 \delta \frac{IJ}{N^2} \right) - \lambda_V u \\ \lambda_I'(t) = {} & -w_1 - \lambda_S \left(\beta_1 \frac{S}{N} + \beta_1 \frac{SI}{N^2} + \beta_2 \frac{SJ}{N^2} + \gamma_1 \right) \\ & \lambda_I \left(\beta_1 \frac{S}{N} - \beta_1 \frac{SI}{N^2} + \beta_1 \delta \frac{J}{N} - \beta_1 \delta \frac{IJ}{N^2} - (\mu + \gamma_1) \right) \\ & \lambda_J \left(-\beta_2 \frac{IJ}{N^2} - \beta_1 \delta \frac{J}{N} + \beta_1 \delta \frac{IJ}{N^2} \right) \\ \lambda_J'(t) = {} & -w_2 - \lambda_S \left(\beta_1 \frac{SI}{N^2} - \beta_2 \frac{S}{N} + \beta_2 \frac{SJ}{N^2} + \gamma_2 \right) \\ & - \lambda_I \left(-\beta_1 \frac{SI}{N^2} + \beta_1 \delta \frac{I}{N} - \beta_1 \delta \frac{IJ}{N^2} \right) \\ & - \lambda_J \left(\beta_2 \frac{S}{N} - \beta_2 \frac{SJ}{N^2} - \beta_1 \delta \frac{I}{N} + \beta_1 \delta \frac{IJ}{N^2} - (\mu + \gamma_2) \right) \\ \lambda_V'(t) = {} & \mu \lambda_V \\ \lambda(T) = {} & 0; \quad \lambda_I(T) = 0; \quad \lambda_J(T) = 0; \quad \lambda_V(T) = 0. \end{aligned} \tag{9.29}$$

From the equation for λ_V and its boundary condition, we see that $\lambda_V = 0$. Hence, the optimal control is characterized by the following formula:

$$u^*(t) = \min\left\{U_{\max}, \max\left\{0, \frac{(\lambda_S(t) - w_3)S(t)}{2}\right\}\right\}.$$

The optimal control and the solution with and without control are plotted in Fig. 9.11. We note that in the case $w_3 \neq 0$, the control is zero for some of the control interval.

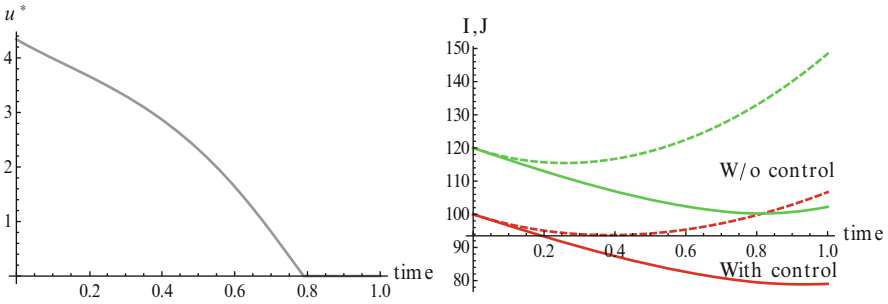

Fig. 9.11 The *left figure* shows the optimal control $u^*(t)$. The *right figure* shows the controlled prevalence $I^*(t)$ and $J^*(t)$ and the original prevalences $I(t)$ and $J(t)$ in *dashed*. Parameters are $\beta_1 = 12; \beta 2 = 15; \gamma 1 = 0.5; \gamma 2 = 0.5; \mu = 0.1; \delta = 0.03; \Lambda = 500; w1 = 1; w2 = 1; w3 = 0.01$

Appendix

In this appendix we include the Matlab code that executes the forward–backward sweep for system (9.26) [94].

```
1
2  function ocmodel1
3  % This function computes the optimal control
4  %  and the corresponding solution using forward-backward ...
       sweep
5  clc;
6  clear all;
7
8
9  test = -1;
10
11 ∆ = 0.001;    %set tolerance
12 N = 100;          %number of subdivisions
13 h = 1/N;          %step
14 t = 0:h:1;        % t-variable mesh
```

```
15
16 u = zeros(1,length(t));      %initialization
17 x = zeros(1,length(t));
18 lam = zeros(1,length(t));
19
20 x(1) = 10;        %initial value assigned to x(0)
21
22 beta = 0.05;      %parameters
23 mu = 0.01;
24 gamma = 0.5;
25 P = 100;
26 w1 = 1;
27
28
29 while (test<0)    % while the tolerance is reached, repeat
30     oldu = u;
31     oldx = x;
32     oldlam = lam;
33
34     for i=1:N     %loop that solve the forward ...
           differential equation
35         k1 = beta*(P-x(i))*x(i) -(mu + gamma)*x(i) - ...
               u(i)*x(i);
36         k2 = beta*(P-x(i)-0.5*k1*h)*(x(i)+0.5*k1*h) - ...
               (mu+gamma)*(x(i)+0.5*k1*h)...
37                         -0.5*(u(i)+u(i+1))*(x(i)+0.5*k1*h);
38         k3 = beta*(P-x(i)-0.5*k2*h)*(x(i)+0.5*k2*h) - ...
               (mu+gamma)*(x(i)+0.5*k2*h)...
39                         -0.5*(u(i)+u(i+1))*(x(i)+0.5*k2*h);
40         k4 = beta*(P-x(i)-k3*h)*(x(i)+k3*h) - ...
               (mu+gamma)*(x(i)+k3*h)...
41                         -u(i+1)*(x(i)+k3*h);
42
43         x(i+1) = x(i) + (h/6)*(k1+2*k2+2*k3+k4);
44
45     end
46
47     for i=1:N     %loop that solves the backward ...
           differential equation of the adjoint system
48         j = N + 2 -i;
49         k1 = ...
               -w1-lam(j)*(beta*(P-x(j))-beta*x(j)-(mu+gamma) ...
               - u(j));
50         k2 = ...
               -w1-(lam(j)-0.5*k1*h)*(beta*(P-x(j)+0.5*k1*h) ...
               -(mu+gamma) -0.5*(u(j)+u(j-1)));
51         k3 = ...
               -w1-(lam(j)-0.5*k2*h)*(beta*(P-x(j)+0.5*k2*h) ...
               -(mu+gamma) -0.5*(u(j)+u(j-1)));
52         k4 = -w1 -(lam(j)-k3*h)*(beta*(P-x(j)+k3*h) ...
               -(mu+gamma) - u(j-1));
53
54         lam(j-1) = lam(j) - (h/6)*(k1+2*k2+2*k3+k4);
55
```

```
    end

    u1 = min(100,max(0,lam.*x/2));
    u = 0.5*(u1 + oldu);

    temp1 = Δ*sum(abs(u)) - sum(abs(oldu - u));
    temp2 = Δ*sum(abs(x)) - sum(abs(oldx - x));
    temp3 = Δ*sum(abs(lam)) - sum(abs(oldlam -lam));

    test = min(temp1,min(temp2,temp3));

end

figure(1)           %plotting
plot(t,u)


figure(2)
plot(t,x)

end
```

Problems

9.1. Consider the model with perfect vaccination

$$\begin{aligned} \frac{dS}{dt} &= \Lambda - \beta SI - (\mu + \psi)S, \\ \frac{dI}{dt} &= \beta SI - (\mu + \gamma)I, \\ \frac{dV}{dt} &= \psi S - \mu V + \gamma I. \end{aligned} \tag{9.30}$$

(a) Compute the reproduction number and investigate the stability of the disease-free equilibrium.
(b) Compute the endemic equilibrium. Does backward bifurcation occur?
(c) Determine the stability of the endemic equilibrium.
(d) Compute the fraction of the population p_c that needs to be vaccinated to eradicate the disease.

9.2. Consider the model with perfect vaccination

$$\begin{aligned} \frac{dS}{dt} &= (1-p)\pi - \beta SI - \mu S, \\ \frac{dI}{dt} &= \beta SI - (\mu + \gamma)I, \\ \frac{dV}{dt} &= p\pi - \mu V + \gamma I, \end{aligned} \tag{9.31}$$

where π is the recruitment rate.

(a) Compute the reproduction number and investigate the stability of the disease-free equilibrium.
(b) Compute the endemic equilibrium.
(c) Determine the stability of the endemic equilibrium.
(d) Compute the fraction of the population p_c that needs to be vaccinated to eradicate the disease.

9.3. Consider the model with perfect vaccination

$$\begin{aligned}\frac{dS}{dt} &= (1-p)\pi - \beta SI - \mu S,\\ \frac{dI}{dt} &= \beta SI + \delta\beta VI - (\mu+\gamma)I,\\ \frac{dV}{dt} &= p\pi - \delta\beta VI - \mu V + \gamma I,\end{aligned} \tag{9.32}$$

where π is the recruitment rate.

(a) Compute the reproduction number and investigate the stability of the disease-free equilibrium.
(b) Compute the endemic equilibrium.
(c) Compute the fraction of the population p_c that needs to be vaccinated to eradicate the disease.

9.4. Consider the model with imperfect vaccination

$$\begin{aligned}\frac{dS}{dt} &= \Lambda - \beta SI - (\mu+\psi)S,\\ \frac{dI}{dt} &= \beta SI + \sigma\beta VI - (\mu+\gamma)I,\\ \frac{dV}{dt} &= \psi S - \sigma\beta VI - \mu V + \gamma I.\end{aligned} \tag{9.33}$$

(a) Compute the reproduction number and investigate the stability of the disease-free equilibrium.
(b) Compute the equation for the endemic equilibria. Derive the condition for backward bifurcation to occur.
(c) Simulate the model and show that even if $\mathscr{R}_0(\psi) < 1$, the solution may converge to an endemic equilibrium.
(d) Consider the model

$$\begin{aligned}\frac{dS}{dt} &= \Lambda - \beta(1-\eta H(t-\tau_1)+\eta H(t-\tau_2))SI - (\mu+\psi)S,\\ \frac{dI}{dt} &= \beta(1-\eta H(t-\tau_1)+\eta H(t-\tau_2))SI\\ &\qquad +\sigma\beta(1-\eta H(t-\tau_1)+\eta H(t-\tau_2))VI - (\mu+\gamma)I,\\ \frac{dV}{dt} &= \psi S - \sigma\beta(1-\eta H(t-\tau_1)+\eta H(t-\tau_2))VI - \mu V + \gamma I,\end{aligned} \tag{9.34}$$

where $H(t-\tau)$ is the Heaviside function. The added term $(1-\eta H(t-\tau_1)+\eta H(t-\tau_2))$ models temporary control measures such as movement restriction, which are adopted at time τ_1 and lifted at time τ_2. Show that with the parameters from part (c), after the lifting of the control measures, the solution may converge to the disease-free equilibrium.

9.5. Vaccine Strain in the Case of Mutation
Consider the following model with mutation:

$$\begin{aligned} S' &= \Lambda - \frac{\beta_1 SI}{N} - \frac{\beta_2 SJ}{N} - \mu S, \\ I' &= \frac{\beta_1 SI}{N} - (\mu+\alpha_1+m)I, \\ J' &= \frac{\beta_2 SJ}{N} - (\mu+\alpha_2)J + mI. \end{aligned} \tag{9.35}$$

Assume $\mathscr{R}_2 < 1$. A vaccine is being designed, but it may include only one of the strains. In that case, the vaccine will be perfect with respect to the vaccine strain and not effective at all with respect to the other. Which of the strains should be the vaccine strain so that the vaccine eliminates both strains?

9.6. Asymptomatic Spread of Avian Influenza
Consider the following model of avian influenza with vaccination and asymptomatic stage:

$$\begin{aligned} \frac{dS}{dt} &= \Lambda - \beta S(I+qA) - (\mu+\psi)S, \\ \frac{dI}{dt} &= \beta S(I+qA) - (\mu+\nu)I, \\ \frac{dV}{dt} &= \psi S - \eta V(I+qA) - \mu V + \gamma A, \\ \frac{dA}{dt} &= \eta V(I+qA) - (\mu+\gamma)A, \end{aligned} \tag{9.36}$$

where A are the asymptomatic individuals infected with avian influenza after imperfect vaccination, and V are the vaccinated individuals.

(a) Compute the disease-free equilibrium and the reproduction number $\mathscr{R}_0(\psi)$. Determine the stability of the disease-free equilibrium based on the reproduction number.
(b) Is the reproduction number an increasing, decreasing, or nonmonotone function of ψ. What is the epidemiological significance of your observation?

9.7. Backward Bifurcation with Perfect Vaccination
Consider the following model of vaccination in a disease with vertical transmission:

$$\begin{aligned}
\frac{dS}{dt} &= (1-p)\pi + (r_1 S + r_2 \eta I)\left(1 - \frac{S+I}{K}\right) - \beta SI - \mu S,\\
\frac{dI}{dt} &= r_2(1-\eta)I\left(1 - \frac{S+I}{K}\right) + \beta SI - (\mu+\alpha)I,\\
\frac{dV}{dt} &= p\pi - \mu V,
\end{aligned} \tag{9.37}$$

where the vaccine is applied at the entry point to the population, and a fraction p is being vaccinated; r_1 and r_2 are the reproduction rates of susceptible and infected individuals respectively, η is the fraction of the progeny of infected individuals that are susceptible.

(a) Compute the disease-free equilibrium and the reproduction number $\mathscr{R}_0(p)$. Determine the stability of the disease-free equilibrium based on the reproduction number.
(b) Derive an equation for the endemic equilibrium. Show that backward bifurcation may occur, even though the vaccine is perfect.

9.8. Saturating Treatment Rates and Vaccination
Consider a model of two strains with saturated per capita treatment rate:

$$\begin{aligned}
S' &= \Lambda - \frac{\beta_1 SI}{N} - \frac{\beta_2 SJ}{N} - (\mu+\psi)S,\\
I' &= \frac{\beta_1 SI}{N} - \mu I - \frac{\alpha_1 I^2}{A+I+J},\\
J' &= \frac{\beta_2 SJ}{N} - \mu J - \frac{\alpha_2 J^2}{B+I+J},\\
V' &= \psi S - \mu V.
\end{aligned} \tag{9.38}$$

(a) Compute the disease-free equilibrium and the reproduction numbers of the two strains.
(b) Show that there is a unique dominance equilibrium corresponding to each strain. Investigate the stability of the dominance equilibria and define the two invasion numbers.
(c) How does the vaccination rate ψ affect the invasion numbers?

9.9. Saturating Incidence
Consider a model of two strains with saturated incidence and perfect vaccination:

$$\begin{aligned}
S' &= \Lambda - \frac{\beta_1 SI}{1+a_1 N} - \frac{\beta_2 SJ}{1+a_2 N} - (\mu+\psi)S,\\
I' &= \frac{\beta_1 SI}{1+a_1 N} - (\mu+\alpha_1)I,\\
J' &= \frac{\beta_2 SJ}{1+a_2 N} - (\mu+\alpha_2)J,\\
V' &= \psi S - \mu V.
\end{aligned} \tag{9.39}$$

(a) Compute the disease-free equilibrium and the reproduction numbers of the two strains.

(b) Compute the two dominance equilibria. Investigate their stability and define the two invasion numbers.
(c) How does vaccination rate ψ affect the invasion numbers?

9.10. Cross-Immunity
Consider a model of two strains with cross-immunity and vaccination:

$$\begin{aligned}
S' &= \Lambda - \beta_1 \frac{S(I_1+J_1)}{A} - \beta_2 \frac{S(I_2+J_2)}{A} - (\mu+\psi)S,\\
I_1' &= \beta_1 \frac{S(I_1+J_1)}{A} - (\mu+\alpha_1+\delta_1)I_1,\\
Q_1' &= \delta_1 I_1 - (\mu+\gamma_1)Q_1,\\
R_1' &= \alpha_1 I_1 + \gamma_1 Q_1 - \sigma_2\beta_2 \frac{R_1(I_2+J_2)}{A} - \mu R_1,\\
J_1' &= \sigma_1\beta_1 \frac{R_2(I_1+J_1)}{A} - (\mu+\alpha_1)J_1,\\
I_2' &= \beta_2 \frac{S(I_2+J_2)}{A} - (\mu+\alpha_2+\delta_2)I_2,\\
Q_2' &= \delta_2 I_2 - (\mu+\gamma_2)Q_2,\\
R_2' &= \alpha_2 I_2 + \gamma_2 Q_2 - \delta_1\beta_1 \frac{(I_1+J)I_2}{A} - \mu R_2,\\
J_2' &= \sigma_2\beta_2 \frac{R_1(I_2+J_2)}{A} - (\mu+\alpha_2)J_2,\\
W' &= \alpha_1 J_1 + \alpha_2 J_2 - \mu W,\\
V' &= \psi S - \mu S.
\end{aligned} \tag{9.40}$$

(a) Compute the disease-free equilibrium and the reproduction numbers of the two strains.
(b) Compute the two dominance equilibria.
(c) Use the next-generation approach to compute the two invasion numbers.
(d) Are the invasion numbers increasing, decreasing, or nonmonotone functions of ψ? What are the epidemiological consequences of this observation?

9.11. Optimal Control
Create an optimal control analogue of model (9.8).

(a) Prove that the optimal control problem has a solution.
(b) Derive the equations for application of Pontryagin's minimum principle.
(c) Write a Matlab code to find the optimal control solution and the optimal control.

9.12. Optimal Control
Create an optimal control analogue of model (9.11).

(a) Prove that the optimal control problem has a solution.
(b) Derive the equations for application of Pontryagin's minimum principle.
(c) Write a Matlab code to find the optimal control solution and the optimal control.

Chapter 10
Ecological Context of Epidemiology

10.1 Infectious Diseases in Animal Populations

Infectious disease pathogens affect numerous animal populations. An animal **population**, by definition, is a collection of individuals of the same species occupying the same habitat. A *species* is a group of individuals who generally breed among themselves and do not naturally interbreed with members of other groups. Fortunately, familiar baseline models of infectious diseases in humans such as the SI, SIS, SIR, SIRS models also can be used to model diseases in animal populations. There are some very important distinctions, however, that we will discuss below.

Why is it important to study diseases in animal populations?

- The simplest and most obvious reason is that a nonhuman species, such as species of fish, birds, and other mammals, represents a much simpler biological system for scientific study than the human species. These species can often be manipulated for better understanding of its properties and dynamics. One such example, in which population dynamics principles were validated through experiments with a beetle population is given in [47].
- There is also a practical reason to study animal diseases. Historically, human diseases have been inextricably linked to epidemics in animal populations. Rapid expansion of civilization in the last few millennia has increased the contact at the human–animal interface, through urban and agricultural expansion that encroaches on wildlife habitats, domestication of cattle and other livestock, or simply by keeping pets. Because of such close intimacy between humans and nonhuman animals, viruses and bacteria that cause various animal diseases continuously "jump across the species barrier" and infect humans. Recent examples include HIV, whict came from monkeys; SARS, which came from bats; and avian flu H5N1, which was first found in wild birds. In fact, about 60% of all human infectious diseases have their origin in animal species, and about three-quarters

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_10

of all *emerging* infectious diseases of humans—those that are occurring for the first time in humans—have been traced back to nonhuman species. Thus, understanding disease dynamics and ways to control disease in animal populations has tremendous human health implications.
- Another practical reason to study disease in natural populations is that the disease can be a regulator of a population, leading to control of dangerous pests [11, 114]. Furthermore, parasites can have substantial effect on community composition [76].

Epidemic models of human populations assume that the population being considered is a closed population, completely separate from the rest of the world. This is rarely true for animal populations. Each animal species is a part of an intricate web of ecological interactions. Interacting populations that share the same habitat form a *community*. Two fundamental community interactions are *competition* for resources and *predation*. Such interspecies interactions impart complex feedback on the dynamics of the diseases in the population that is studied. This necessitates the integration of the principles of infectious disease epidemiology and community ecology.

There are several ways in which the disease dynamics within a population are influenced by the interaction with other species in the community. Here are some of the major ways in which the community interactions influence the host–pathogen dynamics:

- Often, a single pathogen simultaneously infects many different species in the community. Even if the pathogen is completely removed from a given host, it may survive in the other hosts and can reinfect the original host. This is of particular concern if one of the species being infected is humans, while another species is an animal species.
- Intense competition of the focal host species with the other species in the community can bring down the host population size to threateningly small levels, and a pathogen attack can eventually drive it to extinction.
- Predation of the host by other species in the community can regulate pathogen outbreaks in the host species.

Conversely, host–pathogen interactions can feed back and impact the community by weakening the competitive abilities of the affected species.

In this chapter, we will consider the interrelation between infectious diseases and the two principal community interactions: predation and competition. *Predation*, or *predator–prey interactions*, refers to the feeding on the individuals of a *prey* species by the individuals of a *predator* species for the survival and growth of the predator. Some common examples of predator–prey interactions are big fish eating small fish, birds feeding on insects, and cats eating rats and mice. The predator itself can be of two types—a *specialist* predator, whose entire feeding choice is restricted to a single prey species, and a *generalist* predator that feeds on many different prey species. The dynamics of a specialist are strongly coupled to those of the prey, so that any

rise or fall of the population size of the prey triggers a consequent rise or fall in the population size of its predator. The interaction between a specialist predator and its prey is modeled by the familiar Lotka–Volterra predator–prey model. By contrast, the dynamics of a generalist predator are only weakly coupled to any particular focal prey species and may not be influenced by the dynamics of the prey. The predation of a focal species by a generalist predator can be modeled by a predator-added mortality on the prey.

10.2 Generalist Predator and SI-Type Disease in Prey

We can begin studying the effect of predation on disease growth in a host population by considering a host to the disease that is also a prey to a generalist predator. We will model the disease dynamics in the prey by the familiar SI epidemic model that we have used for humans. We will also assume that the predator is a "complete generalist," so that any change of the predator's population size P is caused by other factors and is independent of the dynamics of the focal prey that we are studying. In mathematical terms, we do not need another equation for the rate of change of P, and P is not a dynamic variable but enters the SI model via the death rate of the prey population as a free parameter. Then we have the following system modeling the dynamics of a host–pathogen interaction with predation:

$$\begin{aligned} S'(t) &= \Lambda - \beta SI - \mu(P)S, \\ I'(t) &= \beta SI - \mu(P)I, \end{aligned} \tag{10.1}$$

where Λ is the prey birth rate, β is the transmission parameter, and $\mu(P)$ is the prey death rate, which is split into two components:

$$\mu(P) = \mu_0 + \mu_P(P).$$

The constant μ_0 is a parameter that lumps death rate from the disease, other natural causes, accidental factors, and so on. The function $\mu_P(P)$ is the predation-added mortality of the prey that depends on the predator abundance P, taken as a parameter. For simplicity, we assume that μ_P increases linearly with P, that is, $\mu_P = aP$, where the parameter a denotes the rate at which the predator captures its prey. In general, there should be two different capture rates: a_S and a_I, for susceptible and infected individuals. For instance, the predator may be able to capture physically weak infected prey more easily than susceptible prey, that is, $a_I > a_S$. Here we shall consider two extreme kinds of predation: *selective predation*, in which the predator attacks only one prey type, leaving the other one alone, and *indiscriminate predation*, in which the predator attacks all prey types nonpreferentially. For instance, in selective predation of infected prey alone, we have $a_I > 0$, $a_s = 0$, and in indiscriminate predation, we have $a_S = a_I = a$.

10.2.1 Indiscriminate Predation

In this subsection, we assume that the predator preys indiscriminately on all prey types. Hence,

$$\mu(P) = \mu_0 + aP.$$

We obtain the familiar equation $N'(t) = \Lambda - \mu(P)N$ for the total population size of the prey. From this equation, the equilibrium total population size of the prey becomes

$$N^* = \frac{\Lambda}{\mu} = \frac{\Lambda}{\mu_0 + aP}.$$

So N^* decreases with increasing predator numbers P, as expected. Furthermore, we can solve for the equilibrium number of infected individuals:

$$I^* = \frac{\beta\Lambda - (\mu_0 + aP)^2}{\beta(\mu_0 + aP)}. \tag{10.2}$$

Throughout this chapter, we will call I^* the *disease load*, and the fraction of the infected individuals in the total prey population size will be called the *prevalence*:

$$p = \frac{I^*}{N^*}.$$

From the expressions above for the disease load and the total population size, we have that the prevalence is given by

$$p^* = \frac{\beta\Lambda - (\mu_0 + aP)^2}{\beta\Lambda}. \tag{10.3}$$

From (10.2) and (10.3), we see that both the disease load and the prevalence are decreasing functions of the predator numbers P. The condition for pathogen establishment requires that the numerators of (10.2) and (10.3) be positive, which gives the familiar threshold condition

$$\mathscr{R}_0 = \frac{\beta\Lambda}{(\mu_0 + aP)^2} > 1. \tag{10.4}$$

We see again that the reproduction number of the disease is a decreasing function of the predation level P. We can derive a minimum equilibrium host population size $N^*_{\min}$ that is needed to support the pathogen in the host. To do this, we replace the term $\mu_0 + aP$ in the expression for $\mathscr{R}_0$ above by Λ/N^* and obtain the inequality

$$N^* > N^*_{\min} = \sqrt{\frac{\Lambda}{\beta}}.$$

Thus, $I^* = 0$ and $p^* = 0$ when $N^* < N^*_{\min}$. We have the following interesting result:

While the direct effect of the predator is to harm the prey by increasing its mortality rate, predation can indirectly benefit the prey by keeping its population size low and thereby ruling out epidemic outbreaks.

10.2.2 Selective Predation

In the previous subsection, we assumed indiscriminate predation, whereby the predator's capture rate a is the same for both susceptible and infected prey. In this subsection, we want to see what happens under selective predation, that is, when the predator attacks either the susceptible prey alone $a_S > 0, a_I = 0$, or the infected prey alone, $a_S = 0, a_I > 0$. Even though the first scenario is less likely, it can occur in certain situations such as the infected prey hiding in burrows to avoid predation, or the predator may deliberately avoid infected prey to prevent infecting itself. In the case of selective predation of susceptible prey alone, the model (10.1) becomes

$$\begin{aligned} S'(t) &= \Lambda - \beta SI - (\mu_0 + a_S P)S, \\ I'(t) &= \beta SI - \mu_0 I. \end{aligned} \tag{10.5}$$

The equilibrium disease load is given by

$$I^* = \frac{\beta\Lambda - \mu_0(\mu_0 + a_S P)}{\beta\mu_0}.$$

We see that the equilibrium disease load decreases linearly with increasing predator numbers P. To obtain the equilibrium prey population size, we solve for S^* from the second equation above to obtain

$$S^* = \frac{\mu_0}{\beta}$$

and then use the fact that $N^* = S^* + I^*$. We obtain

$$N^* = \frac{\beta\Lambda - \mu_0 a_S P}{\beta\mu_0}.$$

Hence, the equilibrium prevalence is given by

$$p^* = \frac{I^*}{N^*} = \frac{\beta\Lambda - \mu_0(\mu_0 + a_S P)}{\beta\Lambda - \mu_0 a_S P} = 1 - \frac{\mu_0^2}{\beta\Lambda - \mu_0 a_S P}.$$

From the last expression, we see that p^* is also decreasing with increasing P. The threshold predation level for which $p^* = 0$ is

$$P_{\text{crit}} = \frac{\beta\Lambda - \mu_0^2}{\mu_0 a_S}.$$

We note that the denominator of p^* is positive for all $P < P_{\text{crit}}$.

For the case of selective predation on infected prey alone, the system becomes

$$\begin{aligned} S'(t) &= \Lambda - \beta SI - \mu_0 S, \\ I'(t) &= \beta SI - (\mu_0 + a_I P)I. \end{aligned} \tag{10.6}$$

Following steps similar to those have taken previously, we get an expression for the prevalence p^* as follows:

$$p^* = \frac{I^*}{N^*} = \frac{\beta\Lambda - \mu_0(\mu_0 + a_I P)}{\beta\Lambda + a_I P(\mu_0 + a_I P)}.$$

We see again that the prevalence is decreasing with increasing P. So the qualitative nature of the result—predation lowers epidemic outbreaks in the prey—is similar to the case of indiscriminate predation discussed in the previous subsection.

In summary, with an SI model, we see that predation reduces disease load and prevalence in the prey population, irrespective of whether the predator selectively attacks either susceptible or infected prey, or indiscriminately preys on both prey types. The main reason is that the incidence rate βSI, which gives the rate at which new infections appear in the host, depends bilinearly on both the susceptible and infected prey. Therefore, whether the predator eats selectively or indiscriminately, it always reduces the incidence βSI, and hence the disease level in the prey population. In the next section, we discuss disease in prey with permanent recovery, and investigate the impact of the predator on the disease load and prevalence of the disease in the prey population.

10.3 Generalist Predator and SIR-Type Disease in Prey

In this section, we consider an SIR model with predation by a generalist predator. The presence of a recovered class of individuals makes a difference. Since the number of recovered individuals R does not directly contribute to the disease transmission, the predation of the recovered prey can have only an indirect effect on disease growth. The addition of a recovered class will give very different conclusions based on the different dietary choices of the predator. We consider the following SIR model with predation of a generalist predator:

$$\begin{aligned} S'(t) &= \Lambda - \beta SI - \mu(P)S, \\ I'(t) &= \beta SI - (\alpha + \mu(P))I, \\ R'(t) &= \alpha I - \mu(P)R. \end{aligned} \tag{10.7}$$

The prey death rate $\mu(P)$ is defined as before, $\mu(P) = \mu_0 + aP$. This is the death rate for indiscriminate predation. We will consider two cases of selective predation: predation on infected prey alone, $a_I > 0, a_S = a_R = 0$, and predation on recovered prey alone, $a_R > 0, a_S = a_I = 0$.

10.3.1 Selective Predation

We begin with selective predation on infected prey:

$$\begin{aligned} S'(t) &= \Lambda - \beta SI - \mu_0 S, \\ I'(t) &= \beta SI - (\alpha + \mu_0 + a_I P)I, \\ R'(t) &= \alpha I - \mu_0 R. \end{aligned} \tag{10.8}$$

Solving for the disease load in the endemic equilibrium, we have

$$I^* = \frac{\beta\Lambda - \mu_0(\alpha + \mu_0 + a_I P)}{\beta(\alpha + \mu_0 + a_I P)}.$$

We compute the basic reproduction number from the requirement that $I^* > 0$,

$$\mathscr{R}_0 = \frac{\beta\Lambda}{\mu_0(\alpha + \mu_0 + a_I P)}, \tag{10.9}$$

and the condition $\mathscr{R}_0 > 1$. It is clear that both the disease load and the basic reproduction number decrease with predation level P. We need to know the total prey population size at equilibrium N^* to find out the prevalence $p^* = I^*/N^*$. We solve for S^* and R^* and compute $N^* = S^* + I^* + R^*$:

$$N^* = \frac{\beta\Lambda(\alpha + \mu_0) + \mu_0 a_I P(\alpha + \mu_0 + a_I P)}{\beta\mu_0(\alpha + \mu_0 + a_I P)}.$$

The equilibrium prevalence is then given by

$$p^* = \frac{I^*}{N^*} = \frac{\mu_0[\beta\Lambda - \mu_0(\alpha + \mu_0 + a_I P)]}{\beta\Lambda(\alpha + \mu_0) + \mu_0 a_I P(\alpha + \mu_0 + a_I P)}.$$

It can be seen that p^* decreases with increasing predator numbers P at a rate faster than linear. Similar results can be obtained in the case of selective predation on susceptible individuals only, that is, $a_S > 0, a_I = a_R = 0$. Selective predation on susceptible prey decreases the incidence of the disease βSI and therefore also decreases the disease load and the prevalence.

Now we turn to selective predation of the predator on the recovered individuals, that is, $a_R > 0, a_S = a_I = 0$. The model becomes

$$\begin{aligned} S'(t) &= \Lambda - \beta SI - \mu_0 S, \\ I'(t) &= \beta SI - (\alpha + \mu_0) I, \\ R'(t) &= \alpha I - (\mu_0 + a_R P) R. \end{aligned} \tag{10.10}$$

The equilibrium disease load is

$$I^* = \frac{\beta\Lambda - \mu_0(\alpha + \mu_0)}{\beta(\alpha + \mu_0)}.$$

This is an interesting result—the disease load is independent of the predation level P and remains constant with change of P. This happens because both susceptible and infected prey are ignored by the predator, and therefore the incidence rate βSI is unaffected by predation, and so is the disease load I^*. Furthermore, the condition $I^* > 0$ gives the basic reproduction number $\mathscr{R}_0$:

$$\mathscr{R}_0 = \frac{\beta\Lambda}{\mu_0(\alpha + \mu_0)}, \tag{10.11}$$

which is also independent of the predation level P. An implication of this result is that if the inequality $\mathscr{R}_0 > 1$ holds in the absence of predation $P = 0$, it does not change thereafter with increasing P, and thus the pathogen never becomes extinct as a result of predation. The equilibrium prey population size N^* is obtained as the sum $N^* = S^* + I^* + R^*$,

$$N^* = \frac{\beta\Lambda(\alpha + \mu_0 + a_R P) + \alpha a_R(\alpha + \mu_0)}{\beta(\alpha + \mu_0)(\mu_0 + a_R P)},$$

which then gives equilibrium prevalence:

$$p^* = \frac{I^*}{N^*} = \frac{(\mu_0 + a_R P)[\beta\Lambda - \mu_0(\alpha + \mu_0)]}{\beta\Lambda(\alpha + \mu_0 + a_R P) + \alpha a_R P(\alpha + \mu_0)}.$$

Unlike I^* and $\mathscr{R}_0$, the prevalence p^* depends on P, because the equilibrium total prey population size N^* depends on P. The behavior of p^* with increasing P^* is not so obvious, so we consider the derivative of p^* with respect to P:

$$\frac{dp^*}{dP} = \frac{\alpha a_R[\beta\Lambda - \mu_0(\alpha + \mu_0)]^2}{[\beta\Lambda(\alpha + \mu_0 + a_R P) + \alpha a_R P(\alpha + \mu_0)]^2}.$$

This shows that $dp^*/dP > 0$, and therefore, if the prevalence is positive in the absence of predation, i.e., $\mathscr{R}_0 > 1$, it increases continuously with P. This result is not entirely surprising in light of the fact that the disease load I^* remains constant with increasing prevalence. Since we can expect that the total population size N^* decreases with increasing P, we should expect that the ratio, the prevalence p^*, is increasing with P.

Therefore, once the pathogen is established in the prey, $\mathscr{R}_0 > 1$, the prevalence decreases with increasing predation level P when the predator selectively attacks

susceptible or infected prey only. In contrast, prevalence *increases* with increasing predation level P when the predator attacks preferentially the recovered prey. This last result, although counterintuitive, can be explained by the fact that by attacking recovered prey alone, the predator decreases total population size without impacting the incidence of the disease βSI.

10.3.2 Indiscriminate Predation

As a last case, we consider indiscriminate predation, that is, $a_S = a_I = a_R = a$. The system in this case becomes

$$\begin{aligned} S'(t) &= \Lambda - \beta SI - (\mu_0 + aP)S, \\ I'(t) &= \beta SI - (\alpha + \mu_0 + aP)I, \\ R'(t) &= \alpha I - (\mu_0 + aP)R. \end{aligned} \tag{10.12}$$

Solving for the equilibrium level of the disease load, we obtain

$$I^* = \frac{\beta\Lambda - (\mu_0 + aP)(\alpha + \mu_0 + aP)}{\beta(\alpha + \mu_0 + aP)}.$$

The condition that the disease load has to be positive, $I^* > 0$, gives the following basic reproduction number of the disease:

$$\mathscr{R}_0 = \frac{\beta\Lambda}{(\mu_0 + aP)(\alpha + \mu_0 + aP)}.$$

We see that the reproduction number is decreasing with increasing predation level P. In this case, we can obtain an equation for the total population size. Adding all equations in (10.12), we obtain

$$N'(t) = \Lambda - (\mu_0 + aP)N.$$

From the above equation, we obtain the equilibrium total prey population size

$$N^* = \frac{\Lambda}{\mu_0 + aP}. \tag{10.13}$$

We can derive the minimum prey population size N^*_{min} needed for persistence of the pathogen by plugging (10.13) into the expression for $\mathscr{R}_0$ and using the condition $\mathscr{R}_0 > 1$:

$$N^* > N^*_{\text{min}} = \frac{1}{\beta}(\alpha + \mu_0 + aP).$$

Thus, the pathogen is extinct, $I^* = 0$, when $\mathscr{R}_0 \leq 1$ and equivalently $N^* \leq N^*_{\min}$. We can obtain the equilibrium prevalence:

$$p^* = \frac{I^*}{N^*} = \frac{(\mu_0 + aP)[\beta\Lambda - (\mu_0 + aP)(\alpha + \mu_0 + aP)]}{\beta\Lambda(\alpha + \mu_0 + aP)}. \tag{10.14}$$

It is clear from this expression that the prevalence p^* is not monotone with respect to the predation level P. Looking at the derivative dp^*/dP,

$$\frac{dp^*}{dP} = \frac{\beta\Lambda\alpha - 2(\mu_0 + aP)(\alpha + \mu_0 + aP)^2}{\beta\Lambda(\alpha + \mu_0 + aP)^2},$$

we see that no clear conclusion can be drawn about the sign, and it will depend on the value of the predator population size P. Since for P large enough, the numerator and the whole derivative dp^*/dP are negative, if for $P = 0$ we have

$$\beta\Lambda\alpha > 2\mu_0(\alpha + \mu_0)^2,$$

then $dp^*/dP > 0$ for small P and $dp^*/dP < 0$ for large P. Hence, we can expect a humped-shaped plot of the prevalence p^*, increasing for small P and decreasing for large P.

Thus, we are faced with an intriguing situation. The basic reproduction number $\mathscr{R}_0$ and the disease load I^* decrease with increasing predation pressure P, which tends to give the impression that the epidemic in the prey is weakened in the presence of predation. However, prevalence p^*, which gives another measure of the disease in the population, increases, at least for low predation level P. The reason is that even though both I^* and N^* both decrease with increasing P, N^* falls faster than I^* for low values of P, and therefore the ratio I^*/N^*, which gives the prevalence, increases for those values of P. In other words, while both the population size of the prey and the number of infective hosts are depressed under indiscriminate predation, the proportion of infective individuals in the population increases for low levels of P. The epidemiological significance of this result is that the overall disease burden in the population is reduced, but the risk that a randomly chosen individual is infected can increase with predation, at least at small predation levels.

10.4 Specialist Predator and SI Disease in Prey

Specialist predators prey on a given species, which is our focal species. The dynamics of a specialist predator are closely linked to those of the prey. These dynamics are those described by the Lotka–Volterra predator–prey model. In the next subsection, we discuss various types of predator–prey models.

10.4.1 Lotka–Volterra Predator–Prey Models

The Lotka–Volterra predator–prey model was initially proposed by Alfred J. Lotka in the theory of autocatalytic chemical reactions in 1910. In 1925, Lotka used the equations to analyze predator–prey interactions in his book *Elements of Physical Biology*, deriving the equations that we know today. Vito Volterra, who was interested in statistical analysis of fish catches in the Adriatic, independently investigated the equations in 1926. The equations are based on the observation that the predator–prey dynamics are often oscillatory. The Lotka–Volterra model makes a number of assumptions about the environment and evolution of the predator and prey populations:

- The prey population finds ample food at all times.
- The food supply of the predator population depends entirely on the prey population.
- The rate of change of population is proportional to its size.
- During the process, the environment does not change in favor of one species, and genetic adaptation is sufficiently slow.

The prey population size $N(t)$ grows exponentially in the absence of the predator and dies only as a result of predation. Thus the prey equation takes the form

$$N'(t) = rN - \gamma NP,$$

where $P(t)$ is the predator population size, and r is the growth rate of the prey. The rate of predation on the prey is assumed to be proportional to the rate at which the predators and the prey meet, and it is described by a mass action term γNP. The predator depends entirely for food on the supply of prey and will die exponentially if no prey is available at a natural death rate d. The equation for the predator takes the form

$$P'(t) = \varepsilon\gamma NP - dP,$$

where ε is the conversion efficiency of the predator, that is, ε is a measure of the predator metabolic efficiency by which the biomass of the prey eaten is converted into biomass of the predator. The Lotka–Volterra predator–prey model takes the form

$$\begin{aligned} N'(t) &= rN - \gamma NP, \\ P'(t) &= \varepsilon\gamma NP - dP. \end{aligned} \tag{10.15}$$

The term γNP is called the *predation term*. The per capita rate at which the predator consumes the prey, γN in this model, is called the *predator's functional response*. The predator's functional response term in this model is linear in the prey population size N. This model also assumes that there is no interference among predators in finding prey. In other words, encounters of predators do not reduce the efficiency of search for prey. Mathematically, this is expressed in the fact that the predation term is linear in P. The combination of the assumption of linear functional response and

that of no interference between predators leads to a term proportional to the product NP, which is the mass action predation term. The predator–prey model (10.15) has an extinction equilibrium $\mathscr{E}_0 = (0,0)$. The model also has one coexistence equilibrium

$$\mathscr{E} = \left(\frac{d}{\varepsilon\gamma}, \frac{r}{\gamma}\right).$$

It can be shown that the orbits of the predator–prey system (10.15) are closed curves around the coexistence equilibrium (see Fig. 10.1). Since the solutions are closed orbits, they are periodic. From the direction of the vector field, it can be seen that the solution curves in the (N,P)-plane run counterclockwise. Thus, the maximum prey population comes about one-quarter of a cycle before the maximum predator population. The predator population's fluctuations follow those of the prey population through time. That is, the prey population begins to increase while the predator population is still decreasing, and the prey population decreases while the predator population is still increasing. The classic (and simplest) explanation of these cycles is that the predator drives the changes in the prey population by catching and killing its members, and the prey as the predator's sole food supply drives the predator's population changes, but a lag between the population responses

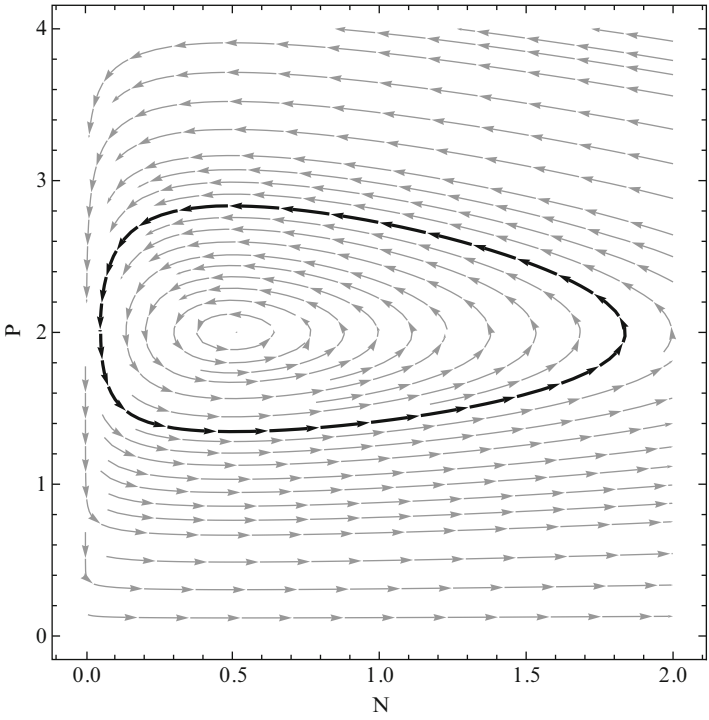

Fig. 10.1 Predator–prey cycles in the (N,P)-plane. The vector field shows that the orbits are traversed counterclockwise

of predator and prey cause the two cycles to be out of phase. However, this explanation has been challenged, and it may not be the only viable explanation for the pattern.

The Lotka–Volterra model represents one of the early triumphs of mathematical modeling, because it captures the oscillatory behavior observed in natural predator–prey systems with a specialist predator. Unfortunately, the model cannot explain these oscillations, because the oscillations in the model are **structurally unstable**, that is, small changes to the model can significantly change the qualitative behavior of the model, e.g., it can stabilize the oscillations. Ideally, we would like the oscillations in the model to be structurally stable, that is, if we make small changes to the model to better reflect reality, the qualitative predictions of the model remain the same, and in particular, the model continues to exhibit oscillations.

The first modification in the Lotka–Volterra model that is natural to be considered is the possibility that the prey population is self-limiting, that is, the prey in the absence of the predator grows logistically. The model becomes

$$\begin{aligned} N'(t) &= rN\left(1-\frac{N}{K}\right)-\gamma NP, \\ P'(t) &= \varepsilon\gamma NP - dP, \end{aligned} \tag{10.16}$$

where K is the carrying capacity of the prey in the absence of the predator. The dynamical behavior of model (10.16) is very different from that of model (10.15).

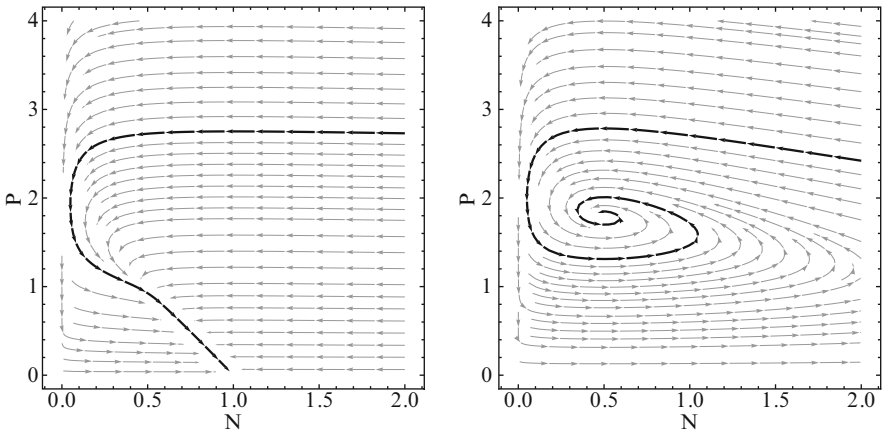

Fig. 10.2 Predator–prey dynamics in the (N,P)-plane. The vector field shows convergence toward the prey-only equilibrium (*left*) or convergence to the predator–prey coexistence equilibrium (*right*)

Model (10.16) has three equilibria. The first equilibrium corresponds to the extinction of both predator and prey, and is called the *extinction equilibrium*, given by $\mathscr{E}_0 = (0,0)$. The Jacobian of the extinction equilibrium, also called the *community matrix*, has one positive and one negative eigenvalue, signifying that the extinction

equilibrium is always a saddle and therefore always unstable. The second equilibrium corresponds to the extinction of the predator only and persistence of the prey population alone. This equilibrium is called the *predator-extinction equilibrium* or semitrivial (boundary) equilibrium, and is given by $\mathscr{E}_1 = (K, 0)$. The community matrix of the prey-only equilibrium is upper triangular and has $\lambda_1 = -r$ and $\lambda_2 = \varepsilon\gamma K - d$. Hence, the prey-only equilibrium is locally asymptotically stable if and only if

$$\frac{\varepsilon\gamma K}{d} < 1. \tag{10.17}$$

The third equilibrium corresponds to a predator–prey coexistence equilibrium. The equilibrium is given by

$$\mathscr{E}^* = \left(\frac{d}{\varepsilon\gamma}, \frac{r}{\gamma}\left(1 - \frac{d}{\varepsilon\gamma K}\right)\right),$$

which exists if and only if $\frac{\varepsilon\gamma K}{d} > 1$. The Jacobian around the coexistence equilibrium has negative trace and positive determinant. Hence, the coexistence equilibrium is locally asymptotically stable. Dulac's criterion can be used to rule out oscillations for this model. It can be shown that if condition (10.17) holds, then the prey-only equilibrium globally stable. If condition (10.17) does not hold, then the predator–prey coexistence equilibrium is globally stable (see Fig. 10.2). Lotka–Volterra predator–prey models have been discussed in multiple texts [90, 27].

10.4.2 Lotka–Volterra Model with SI Disease in Prey

The predator–prey dynamics described by the above models can be impacted by the presence of a disease. The disease may affect the prey, it may affect the predator, or it may affect both the predator and the prey if it is caused by a pathogen that can jump the species barrier. We will consider here a prey population that is subject to predation and impacted by a disease. As a baseline model of the predator–prey dynamics we use the Lotka–Volterra model with self-limiting prey population size and linear functional response. This model exhibits simple dynamics—global convergence to a prey-only equilibrium or to a predator–prey coexistence equilibrium. We will see that the introduction of disease in the prey can lead to much more complex dynamics, namely oscillation and even chaos.

In addition to the assumptions for the predator–prey Lotka–Volterra model with self-limiting prey population size, we also assume the following:

- The disease is transmitted only in the prey and does not affect the predator.
- Infected prey do not recover from the disease—the disease is of SI type for the prey.
- Attack rates of the predator for healthy and infected prey may be different.

- Infected prey does not reproduce but participates in the competition for resources, so it participates in self-limitation.

Assuming that the prey population size N is divided into susceptible S and infective prey I, the Lotka–Volterra model (10.16) with disease in the prey becomes

$$\begin{aligned} S'(t) &= rN\left(1-\frac{N}{K}\right)-\gamma_S SP-\beta SI,\\ I'(t) &= \beta SI-\gamma_I IP-\mu_0 I,\\ P'(t) &= \varepsilon(\gamma_S S+\gamma_I I)P-dP, \end{aligned} \tag{10.18}$$

where γ_S and γ_I are the predation rates of susceptible and infected prey, β is the transmission rate of the disease in the prey, and μ_0 is the natural or disease-induced death rate of infected prey. The natural death rate of susceptible prey is incorporated into the self-limiting logistic term.

Model (10.18) is difficult to analyze, so for analysis, we introduce a slightly simplified model:

$$\begin{aligned} S'(t) &= rS\left(1-\frac{N}{K}\right)-\gamma_S SP-\beta SI,\\ I'(t) &= \beta SI-\gamma_I IP-\mu_0 I,\\ P'(t) &= \varepsilon(\gamma_S S+\gamma_I I)P-dP. \end{aligned} \tag{10.19}$$

Below, we list the equilibria of the system (10.19). The stability of these equilibria depends on the Jacobian (community matrix) evaluated at the corresponding equilibrium. The community matrix at a generic equilibrium is given by

$$J=\begin{pmatrix} r\left(1-\frac{N}{K}\right)-r\frac{S}{K}-\gamma_S P-\beta I & -r\frac{S}{K}-\beta S & -\gamma_S S\\ \beta I & \beta S-\gamma_I P-\mu_0 & -\gamma_I I\\ \varepsilon\gamma_S P & \varepsilon\gamma_I P & \varepsilon(\gamma_S S+\gamma_I I)-d \end{pmatrix}. \tag{10.20}$$

The system (10.19) has four equilibria, three boundary and one interior equilibrium:

1. Extinction or trivial equilibrium: $\mathscr{E}_0=(0,0,0)$. The trivial equilibrium always exists, but it is always unstable, since the community matrix has an eigenvalue $r>0$.
2. Disease-free and predator-free equilibrium: $\mathscr{E}_1=(K,0,0)$. This equilibrium also always exists. If we define a disease reproduction number in the absence of predator

$$\mathscr{R}_0=\frac{\beta K}{\mu_0}$$

and predator invasion number in the absence of disease

$$\mathscr{R}_P^0=\frac{\varepsilon\gamma_S K}{d},$$

then the equilibrium $\mathscr{E}_1$ is locally asymptotically stable if

$$\mathscr{R}_0 < 1 \qquad \text{and} \qquad \mathscr{R}_P^0 < 1$$

and unstable if either inequality is reversed.

3. Predator-free endemic equilibrium in which the disease persists in the prey but the predator dies out, $\mathscr{E}_2 = (S_2, I_2, 0)$, where

$$S_2 = \frac{\mu_0}{\beta} \qquad I_2 = \frac{r\left(1 - \frac{1}{\mathscr{R}_0}\right)}{\frac{r}{K} + \beta}.$$

The equilibrium $\mathscr{E}_2$ exists if and only if $\mathscr{R}_0 > 1$. The equilibrium $\mathscr{E}_2$ is locally asymptotically stable if and only if the invasion number of the predator in the presence of disease is less than one, $\mathscr{R}_p < 1$, where the invasion number of the predator in the presence of the disease is

$$\mathscr{R}_P = \frac{\varepsilon}{d}\left(\gamma_S \frac{\mu_0}{\beta} + \gamma_I \frac{r\left(1 - \frac{1}{\mathscr{R}_0}\right)}{\frac{r}{K} + \beta}\right).$$

4. Predator–prey disease-free equilibrium, where the disease dies out and the predator and the prey coexist disease-free: $\mathscr{E}_3 = (S_3, 0, P_3)$, where

$$S_3 = \frac{d}{\varepsilon\gamma_S} \qquad P_3 = \frac{r}{\gamma_S}\left(1 - \frac{1}{\mathscr{R}_P^0}\right).$$

The disease-free predator–prey coexistence equilibrium $\mathscr{E}_3$ exists if and only if the predator invasion number in the absence of disease satisfies $\mathscr{R}_P^0 > 1$. This equilibrium is locally asymptotically stable if and only if the disease reproduction number in the presence of the predator is less than one: $\mathscr{R}_1 < 1$, where

$$\mathscr{R}_1 = \frac{\beta d}{\varepsilon\gamma_S(\gamma_I P_3 + \mu_0)}.$$

5. Predator–prey–disease coexistence equilibrium: $\mathscr{E}^* = (S^*, I^*, P^*)$. The system for the coexistence equilibrium is given by

$$\begin{aligned} r\left(1 - \frac{N}{K}\right) - \gamma_S P - \beta I &= 0, \\ \beta S - \gamma_I P - \mu_0 &= 0, \\ \varepsilon(\gamma_S S + \gamma_I I) - d &= 0. \end{aligned} \tag{10.21}$$

The system for the interior equilibrium is a linear system, and if it has a nonnegative solution, that solution is unique under the assumption that the determinant

is not zero. The system can be solved, but the expressions obtained do not offer much insight. So we take a different approach. The conditions for existence of a positive equilibrium are stated in the following theorem:

Theorem 10.1. *Assume* $\mathscr{R}_0 > 1$ *and* $\mathscr{R}_P^0 > 1$. *Assume also that*

$$\mathscr{R}_p > 1, \qquad \mathscr{R}_1 > 1, \qquad \text{and} \qquad \mathscr{R}_P^0 < \mathscr{R}_0.$$

Then there exists a unique positive interior equilibrium $\mathscr{E}^* = (S^*, I^*, P^*)$.

Proof. To see the claim, notice that from the second and third equations, we can express I and P as functions of S:

$$P(S) = \frac{\beta S - \mu_0}{\gamma_I}, \qquad I(S) = \frac{d - \varepsilon\gamma_S S}{\varepsilon\gamma_I}. \tag{10.22}$$

We use the first equation in system (10.21) to define the following function:

$$f(S) = r\left(1 - \frac{S + I(S)}{K}\right) - \gamma_S P(S) - \beta I(S).$$

Since $\mathscr{R}_P > 1$, we have $I(S_2) \le I_2$ and $P(S_2) = 0$. Hence,

$$f(S_2) = r\left(1 - \frac{S_2 + I(S_2)}{K}\right) - \gamma_S P(S_2) - \beta I(S_2) \ge r\left(1 - \frac{S_2 + I_2}{K}\right) - \beta I_2 = 0. \tag{10.23}$$

Similarly, since $\mathscr{R}_1 > 1$, we have $I(S_3) = 0$, $P(S_3) = \frac{\beta S_3 - \mu_0}{\gamma_I} \ge P_3$. Hence,

$$f(S_3) = r\left(1 - \frac{S_3}{K}\right) - \gamma_S P(S_3) \le r\left(1 - \frac{S_2}{K}\right) - \gamma_S P_3 = 0. \tag{10.24}$$

Therefore, $f(S_2) \ge 0$, $f(S_3) \le 0$, and there must be a unique solution S^* in the interval (S_2, S_3). The condition $\mathscr{R}_P^0 < \mathscr{R}_0$ implies that $S_2 < S_3$ and $S_2 < S^* < S_3$. Since $I(S)$ is a decreasing function of S, we have $I(S^*) > I(S_3) = 0$. At the same time, $P(S)$ is an increasing function of S, so we have $P(S^*) > P(S_2) = 0$. Hence, $I^* = I(S^*) > 0$ and $P^* = P(S^*) > 0$. This completes the proof. □

We note that an interior equilibrium may exist if

$$\mathscr{R}_p < 1, \qquad \mathscr{R}_1 < 1, \qquad \text{and} \qquad \mathscr{R}_P^0 > \mathscr{R}_0.$$

Concerning the stability of the coexistence equilibrium, we consider that the characteristic equation of the Jacobian J is given by $|J - \lambda I| = 0$. We obtain the following cubic polynomial:

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

where

$$\begin{aligned} a_1 &= rS^*/K, \\ a_2 &= \gamma_S S^* \varepsilon \gamma_S P^* + \varepsilon \gamma_I P^* \gamma_I I^* + \beta I^* (rS^*/K + \beta S^*), \\ a_3 &= \varepsilon P^* \gamma_I I^* r S^*/K(\gamma_I - \gamma_S). \end{aligned} \tag{10.25}$$

The following result is immediate:

Theorem 10.2. *The interior equilibrium is locally stable if $\gamma_I > \gamma_S$. If $\gamma_I < \gamma_S$, the interior equilibrium is unstable.*

We highlight the main conclusion:

> The presence of a disease in the prey and a preferential predation of susceptible individuals can destabilize otherwise stable predator–prey dynamics.

In the case $\gamma_S > \gamma_I$, model (10.18) can exhibit very complex behavior. Since our premise is that the disease destabilizes the otherwise stable predator–prey dynamics, we investigate how the dynamics of the predator–prey-disease system change when the transmission rate is varied while all other parameters are kept fixed. The fixed parameters are listed in Table 10.1.

Table 10.1 Fixed parameter values used in simulations

Parameter	Interpretation	Value
r	Prey growth rate	2
K	Prey carrying capacity	1,000
μ_0	Disease-induced death rate	0.001
ε	Predator conversion efficiency	0.2
γ_I	Attack rate on infectious prey	1.0
γ_S	Attack rate on susceptible prey	9.1
d	Predator death rate	0.2

For $\beta = 6.2$, we observe a closed periodic orbit of period one, which signifies the fact that the periodic orbit makes one loop around the central point (equilibrium) before starting to repeat itself. We illustrate this situation in Fig. 10.3.

As the transmission rate of the disease increases, the dynamics of the system become more and more complex. The system undergoes a process called period-doubling. A period-one periodic solution bifurcates into a period-two periodic solution, that is, a solution that loops twice around the central point. A period-two solution bifurcates into a period-four solution. Such a solution is exhibited in Fig. 10.4. Period-doubling is a common route to chaotic behavior. The period-doubling bifurcations are usually depicted by a bifurcation diagram in which all

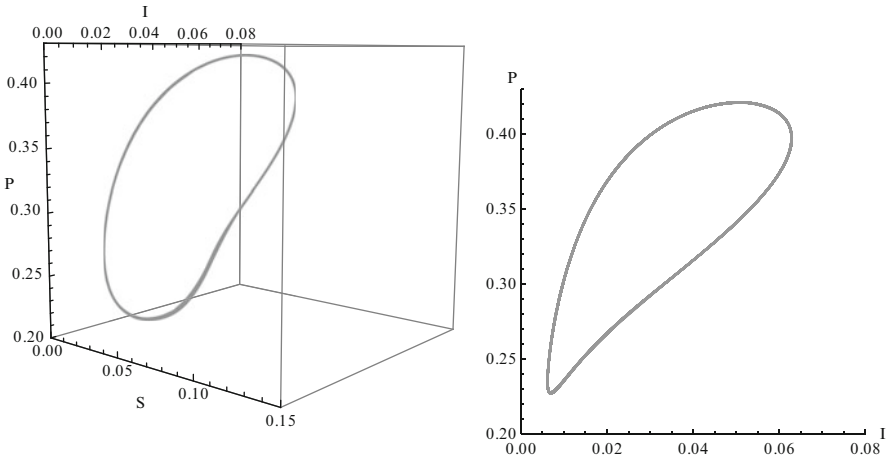

Fig. 10.3 Period-one cycle. The *left figure* shows (S,I,P)-space. The *right figure* shows the (I,P)-plane. Parameters taken from Table 10.1; $\beta = 6.2$

parameters are held fixed except one. Bifurcation diagrams are a common tool for analyzing the behavior of dynamical systems. They are created by running the equations of the system, holding all but one of the variables constant and varying the last one. Then a graph is plotted of the points that a particular value for the changed variable visits after transient factors have been neutralized. Chaotic regions are indicated by filled-in regions of the plot. A bifurcation diagram for the predator–prey–disease dynamical system is shown in Fig. 10.5. For β large enough, the system

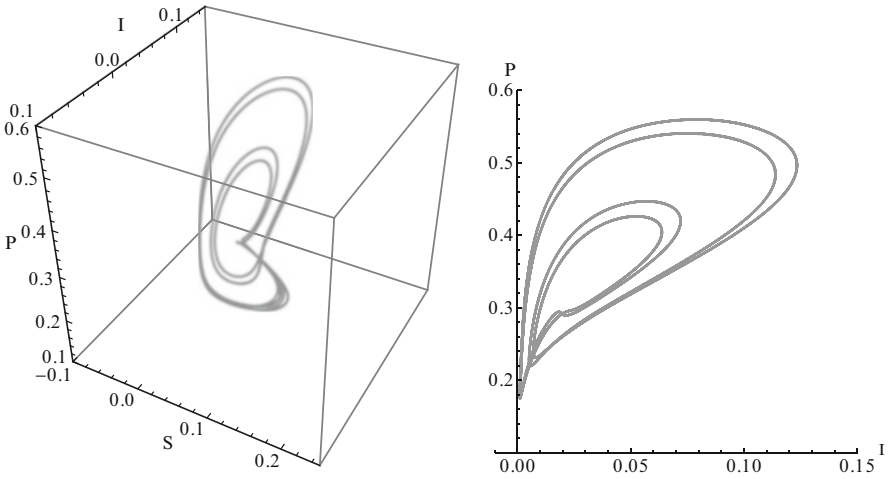

Fig. 10.4 Period-four cycle. The *left figure* shows the (S,I,P)-space. The *right figure* shows the (I,P)-plane. Parameters taken from Table 10.1; $\beta = 6.54$

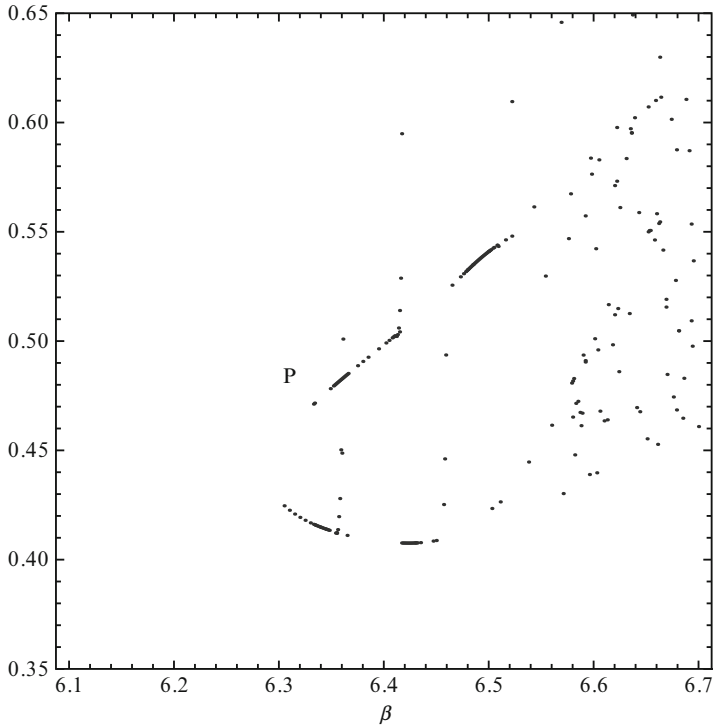

Fig. 10.5 Bifurcation diagram for the system (10.18). Parameter β is plotted on the horizontal axis; P is plotted on the vertical axis. Parameters taken from Table 10.1

exhibits chaos, which is characterized by aperiodic behavior and sensitive dependence on the initial data (Chap. 4). Orbits converge to a chaotic attractor, which is plotted in Fig. 10.6. The attractor has two wings, one of which largely resides in the (S,P)-plane and the other in the (I,P)-plane. Projection of the attractor on the (S,I)-plane is minimal. The dynamics of infected individuals exhibit random spikes modeling outbreak disease, rather than endemicity. The predator persists but also exhibits spikes that coincide with the spikes of the infected prey.

10.5 Competition of Species and Disease

Predation, which we considered in the previous sections, is one of the interactions in the ecological community, and it is certainly the most dynamic interaction. For many years, however, *competition* has been thought to be the main mode of interaction. There is no question that competition is a very important community interaction.

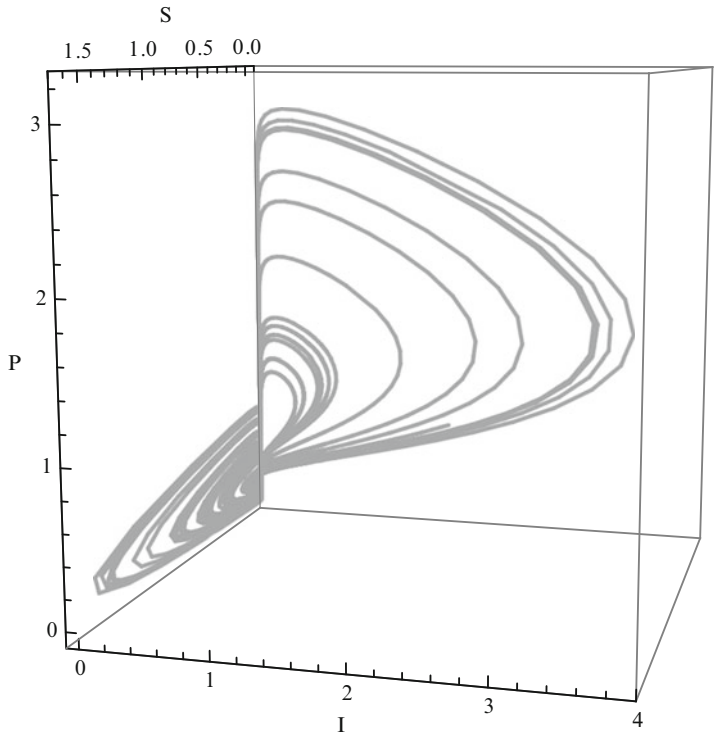

Fig. 10.6 Strange attractor for the system (10.18). Parameter $\beta = 9.5$. Other parameters taken from Table 10.1

10.5.1 Lotka–Volterra Interspecific Competition Models

Lotka and Volterra also developed competition models. Lotka–Volterra competition models describe the competition between two or more species for limited resources. Such competition is called *interspecific competition*, which is contrasted with *intraspecific competition*, which is competition among individuals of one species for limited resources. Lotka–Volterra models are representatives of the **interference competition models**, whereby the increase in the size of one species is assumed to decrease the other species per capita growth rate [90].

We review here the classical Lotka–Volterra competition model considered in [90]. Readers are directed to that source for more detailed discussion of mathematical models of ecology. To introduce the classical Lotka–Volterra model, consider two species with population size $N_1(t)$ and $N_2(t)$ respectively. Each species is assumed to grow logistically in the absence of the other with a growth rate r_i and carrying capacity K_i, $i = 1, 2$:

$$\begin{aligned} N_1'(t) &= r_1 N_1 \left(1 - \frac{N_1 + \alpha_{12} N_2}{K_1}\right), \\ N_2'(t) &= r_2 N_2 \left(1 - \frac{N_2 + \alpha_{21} N_1}{K_2}\right). \end{aligned} \tag{10.26}$$

In these equations, individuals of the second species decrease the per capita growth rate of the individuals of the first species and vice versa. Because the two species are different, the effect of the second species on the first may be stronger or weaker than the effect of the first species on the second. To account for this effect, a pair of competition coefficients α_{12} and α_{21} that describe the strength of the effect of species two on species one and vice versa are introduced.

The system has at most four equilibria. Clearly, it has the extinction equilibrium $\mathscr{E}_0 = (0,0)$ and two semitrivial equilibria corresponding to the dominance of each species: $\mathscr{E}_1 = (K_1,0)$ and $\mathscr{E}_2 = (0,K_2)$. Finally, under appropriate conditions, there is a unique coexistence equilibrium satisfying the system

$$\begin{aligned} N_1 + \alpha_{12} N_2 &= K_1, \\ \alpha_{21} N_1 + N_2 &= K_2. \end{aligned} \tag{10.27}$$

This is a linear system in the unknowns $N_1 \neq 0$ and $N_2 \neq 0$. The solution is given by

$$N_1^* = \frac{\alpha_{12}K_2 - K_1}{\alpha_{12}\alpha_{21} - 1}, \qquad N_2^* = \frac{\alpha_{21}K_1 - K_2}{\alpha_{12}\alpha_{21} - 1}.$$

This solution exists and is positive under appropriate conditions, which we will discuss later. Local stability of equilibria is determined from the community matrix, which in the general case at arbitrary equilibrium is given by the matrix

$$J = \begin{pmatrix} \frac{r_1}{K_1}(K_1 - N_1 - \alpha_{12}N_2) - \frac{r_1}{K_1}N_1 & -\alpha_{12}\frac{r_1}{K_1}N_1 \\ -\alpha_{21}\frac{r_2}{K_2}N_2 & \frac{r_2}{K_2}(K_2 - N_2 - \alpha_{21}N_1) - \frac{r_2}{K_2}N_2 \end{pmatrix}. \tag{10.28}$$

The community matrix for the extinction equilibrium has the eigenvalues $\lambda_1 = r_1$ and $\lambda_2 = r_2$, which are both real and positive. Thus the extinction equilibrium is always an unstable node. The community matrix at the dominance equilibrium of species one is

$$J(K_1,0) = \begin{pmatrix} -r_1 & -\alpha_{12}r_1 \\ 0 & \frac{r_2}{K_2}(K_2 - \alpha_{21}K_1) \end{pmatrix}. \tag{10.29}$$

Thus, one of the eigenvalues is $\lambda_1 = -r_1$, and it is negative. The other eigenvalue is $\lambda_2 = \frac{r_2}{K_2}(K_2 - \alpha_{21}K_1)$. The sign of this eigenvalue depends on the sign of $K_2 - \alpha_{21}K_1$. Thus, the equilibrium $\mathscr{E}_1$ is a stable node or a saddle. Symmetrically, the eigenvalues of the community matrix of the equilibrium $\mathscr{E}_2$ are $\lambda_1 = \frac{r_1}{K_1}(K_1 - \alpha_{12}K_2)$ and $\lambda_2 = -r_2$. The sign of the first eigenvalue depends on the sign of $K_1 - \alpha_{12}K_2$. The equilibrium $\mathscr{E}_2$ is also either a stable node or a saddle. Finally, to simplify the community matrix of the coexistence equilibrium, we notice that at the coexistence equilibrium, we have

$$K_1 - N_1^* - \alpha_{12}N_2^* = 0, \qquad K_2 - N_2^* - \alpha_{21}N_1^* = 0.$$

Thus, the community matrix becomes

$$J(N_1^*, N_2^*) = \begin{pmatrix} -\dfrac{r_1}{K_1}N_1 & -\alpha_{12}\dfrac{r_1}{K_1}N_1 \\ -\alpha_{21}\dfrac{r_2}{K_2}N_2 & -\dfrac{r_2}{K_2}N_2 \end{pmatrix}. \tag{10.30}$$

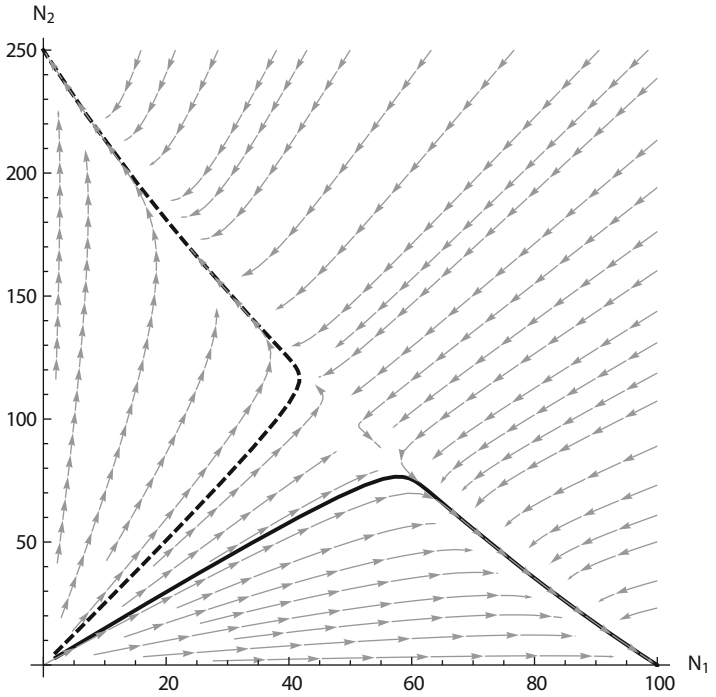

Fig. 10.7 Phase portrait of two competing species. Species one has $K_1 = 100$. Species two has $K_2 = 250$. The *black continuous* trajectory converges to equilibrium $\mathscr{E}_1 = (100, 0)$. The *black dashed* trajectory starts from a different initial condition and converges to equilibrium $\mathscr{E}_2 = (0, 250)$

The trace of the above matrix is clearly negative. The determinant is given as follows:

$$\det J = \frac{r_1}{K_1}\frac{r_2}{K_2}N_1^*N_2^*(1 - \alpha_{12}\alpha_{21}).$$

Thus, the sign of the determinant depends on the sign of the expression $1 - \alpha_{12}\alpha_{21}$. If $1 - \alpha_{12}\alpha_{21} > 0$, then the coexistence equilibrium is a stable node or a stable focus. It can be checked that the discriminant of the characteristic polynomial $\lambda^2 + p\lambda + q = 0$, is given by

$$\Delta = p^2 - 4q = \left(\frac{r_1}{K_1}N_1^* + \frac{r_2}{K_2}N_2^*\right)^2 - 4\frac{r_1}{K_1}\frac{r_2}{K_2}N_1^*N_2^*(1 - \alpha_{12}\alpha_{21}) > 0.$$

Hence, the coexistence equilibrium in this case is a stable node. If $1-\alpha_{12}\alpha_{21}<0$, then the coexistence equilibrium is a saddle. There are four distinct cases that encompass all possibilities. They are centered on the position of the nullclines. Species-one nullclines (x-nullclines) are $N_1=0$ and $N_1+\alpha_{12}N_2=K_1$. Symmetrically, species-two nullclines (y-nullclines) are $N_2=0$ and $\alpha_{21}N_1+N_2=K_2$. There are four cases.

Case 1. $K_1>\alpha_{12}K_2$ and $K_2<\alpha_{21}K_1$. In this case, there is no interior equilibrium, since the numerators of N_1^* and N_2^* have opposite signs. The boundary equilibrium $\mathscr{E}_1$ is a stable node, while the boundary equilibrium $\mathscr{E}_2$ is a saddle (unstable). All orbits tend to $(K_1,0)$ as $t\to\infty$. Thus, species one persists at carrying capacity, while species two becomes extinct.

Case 2. $K_1<\alpha_{12}K_2$ and $K_2>\alpha_{21}K_1$. This is a symmetric case to Case 1. In this case, again there is no interior equilibrium. The boundary equilibrium $\mathscr{E}_1$ is a saddle (unstable), while the boundary equilibrium $\mathscr{E}_2$ is a stable node. All orbits tend to $(0,K_2)$ as $t\to\infty$. Thus, species two persists at carrying capacity, while species one becomes extinct.

Case 3. $K_1<\alpha_{12}K_2$ and $K_2<\alpha_{21}K_1$. These two inequalities imply that $1<\alpha_{12}\alpha_{21}$. Thus the interior equilibrium exists. However, since the determinant of the community matrix evaluated at the interior equilibrium is negative,

$$\mathrm{Det}J(N_1^*,N_2^*)<0,$$

the community matrix has two real eigenvalues of opposite sign ($q<0$). Therefore, the coexistence equilibrium is a saddle. At the same time, both semitrivial equilibria $\mathscr{E}_1$ and $\mathscr{E}_2$ are stable nodes. In this case, the coexistence of the two species is again impossible. One of the species always outcompetes and eliminates the other. However, the winner of the competition is determined by the initial conditions. We recall that this dependence on the initial conditions is called the *founder effect*. This is another example of bistability. Solution orbits that start from the upper part of the plane converge to the equilibrium $\mathscr{E}_2$, while those that start from the lower part converge to the equilibrium $\mathscr{E}_2$ (see Fig. 10.7).

Case 4. $K_1>\alpha_{12}K_2$ and $K_2>\alpha_{21}K_1$. In this case, $1>\alpha_{12}\alpha_{21}$, and the interior equilibrium also exists. The community matrix at the interior equilibrium has negative trace and

$$\mathrm{Det}J(N_1^*,N_2^*)>0.$$

Therefore, the coexistence equilibrium is locally asymptotically stable. It can be further shown that it is a stable node. The community matrices of the two semitrivial equilibria have one positive eigenvalue and one negative eigenvalue. Hence, the two semitrivial equilibria are saddle points. In this case, every orbit that starts from the interior tends to the coexistence equilibrium as $t\to\infty$.

In Cases 1 and 2, we say that *competitive exclusion* occurs. That means that one of the species excludes the other and dominates by itself. We recall that the principle of competitive exclusion was first formulated by Gause in 1934 [64] on the basis of experimental evidence.

10.5.2 Disease in One of the Competing Species

How will the outcome of the competition between two species be influenced if one of the species is plagued by a disease? Models of two competing species with disease have been considered before and they have shown that the presence of the disease tends to destabilize the dynamics of the interaction of the species [160]. We consider model (10.26), and we assume that a disease is spreading among species one. Then the population of species one $N_1(t)$ is split into the number of susceptible individuals $S(t)$ and the number of infected individuals $I(t)$. We have $N_1(t) = S(t) + I(t)$. We assume that the growth rate of species one, r_1, is the growth rate of the susceptible individuals but that infected individuals have a different, possibly lower, growth rate r_I ($r_I < r_1$). The transmission of the disease happens at a rate βSI. Model (10.26) can be modified as follows:

$$\begin{aligned} S'(t) &= S\left(r_1 - \frac{r_1 N_1 + r_1 \alpha_{12} N_2}{K_1}\right) - \beta SI, \\ I'(t) &= I\left(r_I - \frac{r_1 N_1 + r_1 \alpha_{12} N_2}{K_1}\right) + \beta SI, \\ N_2'(t) &= r_2 N_2 \left(1 - \frac{N_2 + \alpha_{21} N_1}{K_2}\right). \end{aligned} \tag{10.31}$$

We notice that in the equation for S, we have multiplied through by r_1. In the equation for I, we first multiply by r_1 and then replace the first occurrence by the reproduction rate r_I. However, the intraspecies interference term remains the same. To simplify the appearance of system (10.31), we rewrite it in the form

$$\begin{aligned} S'(t) &= S\left(r_1 - a_{11} N_1 - a_{12} N_2\right) - \beta SI, \\ I'(t) &= I\left(r_I - a_{11} N_1 - a_{12} N_2\right) + \beta SI, \\ N_2'(t) &= N_2\left(r_2 - a_{22} N_2 - a_{21} N_1\right), \end{aligned} \tag{10.32}$$

where we have set $a_{11} = r_1/K_1$, $a_{12} = r_1\alpha_{12}/K_1$, $a_{21} = r_2\alpha_{21}/K_2$, and $a_{22} = r_2/K_2$.

System (10.32) is a three-dimensional competitive Lotka–Volterra system. Notice that if we set $I = 0$, the second equation is trivially satisfied, and system (10.31) reduces to system (10.26). We consider the equilibria of the system (10.32). First, there is an extinction equilibrium of the system $\mathscr{E}_0 = (0,0,0)$. Next, there are three vertex equilibria: $\mathscr{E}_1 = (K_1,0,0)$, $\mathscr{E}_2 = (0, r_I K_1/r_1, 0)$, and $\mathscr{E}_3 = (0,0,K_2)$. Next we investigate equilibria that have one component equal to zero:

- Equilibrium $\mathscr{E}_{12}$ is an equilibrium in which the disease is present in species one but species two is absent:

$$\mathscr{E}_{12} = \frac{1}{\beta^2}((r_1 - r_I)a_{11} - r_I\beta, (r_I - r_1)a_{11} + r_1\beta, 0).$$

  This equilibrium exists when the following inequalities are satisfied:

$$r_I\beta < (r_1 - r_I)a_{11} < r_1\beta.$$

  See Problem 10.4 for further details.
- Equilibrium $\mathscr{E}_{13}$ is an equilibrium in which species one is present with susceptible individuals only and species two is also present:

$$\mathscr{E}_{13} = \frac{1}{\Delta}(r_1a_{22} - r_2a_{12}, 0, r_2a_{11} - r_1a_{21}) = \left(\frac{K_1 - \alpha_{12}K_2}{1 - \alpha_{12}\alpha_{21}}, 0, \frac{K_2 - \alpha_{21}K_1}{1 - \alpha_{12}\alpha_{21}}\right),$$

  where $\Delta = a_{11}a_{22} - a_{12}a_{21}$ is the determinant of the matrix of coefficients. Hence, the existence and stability of equilibrium $\mathscr{E}_{13}$ is exactly the same as the existence and stability of the coexistence equilibrium of the two species.
- Equilibrium $\mathscr{E}_{23}$ is an equilibrium in which species one is present with infected individuals only and species two is also present:

$$\mathscr{E}_{23} = \frac{1}{\Delta}(0, r_Ia_{22} - r_2a_{12}, 0, r_2a_{11} - r_Ia_{21}) = \left(0, \frac{\frac{r_I}{r_1}K_1 - \alpha_{12}K_2}{1 - \alpha_{12}\alpha_{21}}, \frac{K_2 - \frac{r_I}{r_1}\alpha_{21}K_1}{1 - \alpha_{12}\alpha_{21}}\right),$$

  where $\Delta = a_{11}a_{22} - a_{12}a_{21}$ has the same meaning as above. Hence, the existence and stability of equilibrium $\mathscr{E}_{23}$ can be derived in a similar way to that of the existence and stability of the coexistence equilibrium of the two species. See Problem 10.5.

The system has a unique interior equilibrium $\mathscr{E}^* = (S^*, I^*, N_2^*)$, where (see [160])

$$\begin{aligned} S^* &= \frac{1}{a_{22}\beta^2}(\Delta(r_1 - r_I) - \beta(a_{22}r_I - a_{12}r_2)), \\ I^* &= \frac{1}{a_{22}\beta^2}(\beta(a_{22}r_1 - a_{12}r_2) - \Delta(r_1 - r_I)), \\ N_2^* &= \frac{1}{a_{22}\beta}(\beta r_2 - a_{21}(r_1 - r_I)). \end{aligned} \tag{10.33}$$

We notice that $S^* + I^* = (r_1 - r_I)/\beta$. The interior equilibrium is feasible if the reproduction number of the second species satisfies

$$\mathscr{R}_2 = \frac{\beta r_2}{a_{21}(r_1 - r_I)} > 1.$$

In addition, we need the following double inequality to hold:

$$a_{22}r_1 > \Delta\frac{r_1 - r_I}{\beta} + a_{12}r_2 > a_{22}r_I.$$

To investigate the stability of the interior equilibrium, we consider the community matrix at the interior equilibrium:

$$J = \begin{pmatrix} -a_{11}S^* & -a_{11}S^* - \beta S^* & -a_{12}S^* \\ -a_{11}I^* + \beta I^* & -a_{11}I^* & -a_{12}I^* \\ -a_{21}N_2^* & -a_{21}N_2^* & -a_{22}N_2^* \end{pmatrix}. \tag{10.34}$$

Here we have used the equations of the equilibria to simplify the Jacobian. Namely, we have used

$$\begin{aligned} (r_1 - a_{11}N_1 - a_{12}N_2) - \beta I &= 0, \\ (r_I - a_{11}N_1 - a_{12}N_2) + \beta S &= 0, \\ r_2 - a_{22}N_2 - a_{21}N_1 &= 0. \end{aligned} \tag{10.35}$$

Next, we consider the characteristic equation $|J - \lambda I| = 0$. Expanding the determinant, we obtain the following cubic characteristic polynomial:

$$\lambda^3 + c_1\lambda^2 + c_2\lambda + c_3 = 0,$$

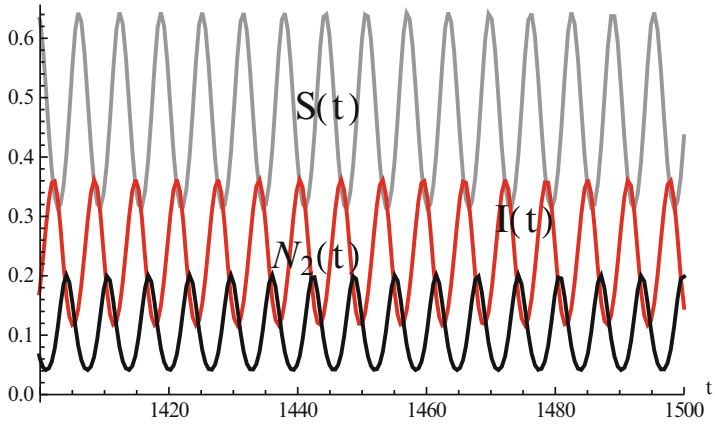

Fig. 10.8 Oscillatory simulation of susceptible species one, $S(t)$, infected species one, $I(t)$, and species two, $N_2(t)$. The susceptible species one have been shifted in the plot with 1.5 units down, that is, what is plotted is $S(t) - 1.5$. Parameters are $r_1 = 25$, $r_I = 14$, $r_2 = 36$, $a_{11} = 10$, $a_{12} = 17$, $a_{21} = 15.3$, $a_{22} = 21.6$, $\beta = 5$ [160]. The oscillatory solution is stable for some initial conditions. Those used to produce the figure are $S(0) = 2$, $I(0) = 0.231667$, $N_2(0) = 0.108333$

where

$$c_1 = a_{22}N_2^* + a_{11}\frac{r_1 - r_I}{\beta},$$
$$c_2 = \Delta N_2^* \frac{r_1 - r_I}{\beta} + \beta^2 S^* I^*,$$
$$c_3 = \beta^2 S^* I^* a_{22} N_2^*. \tag{10.36}$$

Clearly, if $\Delta > 0$, then $c_1 > 0$, $c_2 > 0$, and $c_3 > 0$. Furthermore, it is not hard to see that $c_1 c_2 > c_3$. The Routh–Hurwitz criterion then implies that all roots of the characteristic polynomial have negative real part. Hence the interior equilibrium is locally asymptotically stable. However, if $\Delta < 0$, there is a possibility of Hopf bifurcation and the emergence of a periodic solution. The locally stable periodic solution is illustrated in Fig. 10.8.

Animal populations are subject to a number of ecological interactions. Introducing disease in one or more of the interacting populations is an interesting area of exploration often referred to as **ecoepidemiology**.

Acknowledgements The first part of this chapter is based on lecture notes that Manojit Roy developed and delivered in the Biomathematics Seminar.

Problems

10.1. Competition of Strains under Predation
Consider that a generalist predator is feeding on a prey infected by a pathogen represented by two strains. The model takes the form

$$S'(t) = \Lambda - \beta_1 S I_1 - \beta_2 S I_2 - (\mu_0 + a_S P)S,$$
$$I_1'(t) = \beta_1 S I_1 - (\mu_1 + a_1 P) I_1,$$
$$I_2'(t) = \beta_2 S I_2 - (\mu_2 + a_2 P) I_2. \tag{10.37}$$

(a) Is coexistence of the strains possible in this model? Show competitive exclusion.
(b) Determine which strain dominates depending on the predation level P.

10.2. Competition of Strains under Predation
Consider that a specialist predator is feeding on a prey infected by a pathogen represented by two strains. The model takes the form

$$\begin{aligned} S'(t) &= \Lambda - \beta_1 S I_1 - \beta_2 S I_2 - (\mu_0 + \gamma_S P)S, \\ I_1'(t) &= \beta_1 S I_1 - (\mu_1 + \gamma_1 P) I_1, \\ I_2'(t) &= \beta_2 S I_2 - (\mu_2 + \gamma_2 P) I_2, \\ P'(t) &= \varepsilon(\gamma_S P S + \gamma_1 P I_1 + \gamma_2 P I_2) - dP. \end{aligned} \tag{10.38}$$

(a) Is coexistence of the strains possible in this model? Determine the coexistence equilibrium and the conditions under which it exists.
(b) Use a computer algebra system to simulate the coexistence of the strains. How does changing the predator's predation rates γ_1 and γ_2 affect the competition of the strains?

10.3. Specialist Predator with Disease in Predator
Consider the following model of a specialist predator with disease in the predator. The number of prey is given by $N(t)$. The susceptible predators are given by $S(t)$, and the infected predators by $I(t)$:

$$\begin{aligned} N'(t) &= rN\left(1-\frac{N}{K}\right)-\gamma_S NS-\gamma_I NI,\\ S'(t) &= \varepsilon_S\gamma_S NS-\beta SI-d_S S,\\ I'(t) &= \varepsilon_I\gamma_I NI+\beta SI-d_I I. \end{aligned} \tag{10.39}$$

(a) Find the equilibria of the system.
(b) Compute the reproduction number of the disease in the predator. Determine the stability of the semitrivial equilibria.
(c) Compute the interior equilibrium. When is the interior equilibrium locally asymptotically stable? Does Hopf bifurcation occur?

10.4. Epidemic Model with Vertical Transmission
Consider model (10.32) with species two absent. Assume $r_1 > r_I$:

$$\begin{aligned} S'(t) &= S(r_1-a_{11}N_1)-\beta SI,\\ I'(t) &= I(r_I-a_{11}N_1)+\beta SI. \end{aligned} \tag{10.40}$$

(a) Find the equilibria of model (10.40). Under what conditions does each equilibrium exist?
(b) Determine the local stabilities of each equilibrium.
(c) Use a computer algebra system to draw the phase portrait in each of the cases above.
(d) Explain how the vertical transmission is incorporated in the model.

10.5. Equilibria of Lotka–Volterra Competition Model with Disease
Consider the model (10.32). Assume $r_1 > r_I$.

(a) Find the trivial and semitrivial equilibria of model (10.32). Under what conditions does each equilibrium exist?
(b) Determine the local stabilities of each equilibrium. Determine the corresponding conditions for stability/instability.
(c) Use a computer algebra system to simulate model (10.32). Set $\beta = 0$. Determine parameter values such that $\mathscr{E}_{13}$ is locally stable. Start increasing β. How does the increase in the prevalence of the disease affect the competitive ability of species one?

10.6. Mutualism

Mutualism is an interaction between species that is mutually beneficial for the species involved [90]. Consider the following Lotka–Volterra mutualism model:

$$\begin{aligned} N_1'(t) &= r_1 N_1 \left(1 - \frac{N_1 - \alpha_{12} N_2}{K_1}\right), \\ N_2'(t) &= r_2 N_2 \left(1 - \frac{N_2 - \alpha_{21} N_1}{K_1}\right). \end{aligned} \tag{10.41}$$

Assume $r_1 > 0, r_2 > 0, K_1 > 0, K_2 > 0$. In this case, it is said that the mutualism is facultative.

(a) Find the equilibria of model (10.41). Under what conditions does each an equilibrium exist?
(b) Show that in the coexistence equilibrium, species persist at densities larger than their respective carrying capacities. What does that mean biologically?
(c) Determine the local stabilities of each equilibrium. Determine the corresponding conditions for stability/instability.
(d) Draw a phase portrait in each of the cases $\alpha_{12}\alpha_{21} < 1$ and $\alpha_{12}\alpha_{21} > 1$. Show that in the case $\alpha_{12}\alpha_{21} > 1$, orbits may become unbounded.

10.7. Mutualism with Disease

Consider the mutualism model with disease in one of the species [161]. Let $N_1(t)$ be the density of species one, $S(t)$ the number of susceptible individuals, and $I(t)$ the number of infected individuals of species two:

$$\begin{aligned} N_1'(t) &= N_1(r_1 - a_{11}N_1 + a_{12}(S+I)), \\ S'(t) &= S(r_S + a_{21}N_1 - a_{22}(S+I)) - \beta SI, \\ I'(t) &= I(r_I + a_{21}N_1 - a_{22}(S+I)) + \beta SI, \end{aligned} \tag{10.42}$$

where $r_S > r_I$, and all parameters are positive.

(a) Find the trivial and semitrivial equilibria of model (10.42). Under what conditions does each an equilibrium exist?
(b) Determine the local stabilities of each equilibrium. Determine the corresponding conditions for stability/instability.
(c) Determine the interior equilibrium of the system.
(d) Determine the stability of the interior equilibrium.

10.8. Specialist Predator with a Two-Strain Disease in Predator

Consider the following model of a specialist predator with disease in predator represented by two strains. The number of prey is given by $N(t)$. The susceptible predators are given by $S(t)$, the predators infected by strain one are denoted by $I_1(t)$, and the predators infected by strain two are denoted by $I_2(t)$:

$$
\begin{aligned}
N'(t) &= rN\left(1-\frac{N}{K}\right) - \gamma_S NS - \gamma_{I1} NI_1 - \gamma_{I2} NI_2,\\
S'(t) &= \varepsilon_S \gamma_S NS - \beta_1 SI_1 - \beta_2 SI_2 - d_S S,\\
I_1'(t) &= \varepsilon_{I1} \gamma_{I1} NI_1 + \beta_1 SI_1 - d_{I1} I_1,\\
I_2'(t) &= \varepsilon_{I2} \gamma_{I2} NI_2 + \beta_2 SI_2 - d_{I2} I_2.
\end{aligned} \tag{10.43}
$$

(a) Find the semitrivial equilibria of the system above.
(b) Compute the reproduction numbers of the two strains in the predator. Determine the stability of the semitrivial equilibria.
(c) Is there a coexistence equilibrium for that system?

Chapter 11
Zoonotic Disease, Avian Influenza, and Nonautonomous Models

11.1 Introduction

Zoonotic diseases are contagious diseases that are transmitted between animals and humans. These diseases are caused by bacteria, viruses, parasites, fungi, and prions that are carried by animals and insects.

Zoonotic diseases include vector-borne diseases but also diseases transmitted by vertebrate animals. For some zoonotic diseases, humans are a host that can pass on the pathogen to the animals or to the environment, while for others, the humans are a dead-end host. Examples of the first type are cholera, ebola, and malaria. Examples of the second kind are Rift Valley fever, hantavirus, West Nile virus, and avian influenza.

Zoonotic diseases play a very important role among human communicable diseases. In a review of more than 1400 pathogens known to infect humans, it was found that more than 61% were zoonotic [149]. Zoonotic diseases often serve as a starting point of many pathogens that jump the species barrier and become effectively human-to-human transmissible. Such diseases are called *emergent diseases*.

One of the most dangerous zoonotic pathogens is avian influenza H5N1. Avian influenza is transmitted from birds to humans. As of May 2014, H5N1 has infected more of 600 humans, 60% of whom have died. Besides the high mortality, what makes H5N1 dangerous is the possibility for the pathogen to mutate or reassort into a highly human-to-human transmissible flu pathogen with high mortality. In such a case, a world pandemic would occur, but since humans have no prior exposure to the H5 subtype of influenza A, mortality may be higher than it was in the 2009 H1N1 pandemic, caused by the "swine flu."

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_11

11.2 Modeling Avian Influenza

Modeling zoonotic diseases in general, and avian influenza in particular, involves modeling transmission in several species. Many of the zoonotic diseases involve multiple species. For instance, the avian influenza pathogen infects wild birds, domestic birds, humans, and many animal species such as pigs and cats. One has to decide which species are important to the transmission.

In the case of avian influenza, we typically want to address questions related to human health. Humans are a dead-end host for H5N1, but H5N1 transmits to humans from domestic birds. It transmits effectively in birds and is endemic in the poultry populations in some countries.

11.2.1 Simple Bird–Human Avian Influenza Model

One of the simplest models of avian influenza (AI) was proposed by Iwami [80], following whom, we model transmission in poultry and the spillover to humans. H5N1 is very deadly for chickens. Mortality reaches 90–100% typically within 48 h [40]. This suggests that a simple SI model with disease-induced mortality is a good tool to model the transmission within poultry.

$$\textbf{Domestic Birds:} \begin{cases} S_d'(t) = \Lambda_d - \beta_d S_d I_d - \mu_d S_d, \\ I_d'(t) = \beta_d S_d I_d - (\mu_d + \nu_d) I_d, \end{cases} \tag{11.1}$$

where S_d is the number of susceptible birds, I_d is the number of infected birds, Λ_d is the recruitment rate, μ_d is the poultry natural death rate, ν_d is the disease-induced death rate, and β_d is the transmission coefficient among poultry. Humans become infected from touching infected uncooked poultry products. The spillover model for humans takes the form

$$\textbf{Humans:} \begin{cases} S'(t) = \Lambda - \beta S I_d - \mu S, \\ I'(t) = \beta S I_d - (\mu + \nu) I, \end{cases} \tag{11.2}$$

where S is the number of susceptible humans, I is the number of infected humans, Λ is the recruitment rate for humans, μ is the human natural death rate, ν is the disease-induced death rate for humans, and β is the transmission coefficient from infected poultry to humans.

The dynamics of the solutions to model (11.1)–(11.2) are not very different from those of the SI poultry model. The model has a reproduction number

$$\mathscr{R}_d = \frac{\Lambda_d \beta_d}{\mu_d(\mu_d + \nu_d)}.$$

If $\mathscr{R}_d < 1$, then all solutions approach the disease-free equilibrium: $\mathscr{E}_0 = (\frac{\Lambda_d}{\mu_d}, 0, \frac{\Lambda}{\mu}, 0)$. If $\mathscr{R}_d > 1$, then all solutions approach the endemic equilibrium $\mathscr{E}^* = (S_d^*, I_d^*, S^*, I^*)$,

where

$$S_d^* = \frac{\mu_d + \nu_d}{\beta_d}, \qquad I_d^* = \frac{\mu_d}{\beta_d}(\mathscr{R}_d - 1), \qquad S^* = \frac{\Lambda}{\beta I_d^* + \mu} \qquad I^* = \frac{\beta S^* I_d^*}{\mu + \nu}.$$

We see that if $\beta \neq 0$, then the outcome of the disease in humans is a direct consequence of the outcome of the disease in poultry.

11.2.2 Parameterizing the Simple Avian Influenza Model

One of the main ingredients in developing models is determining reasonable parameter values for the model. We fix the time unit in years. The reason for that will become clear later. Determining parameters is typically done through fitting. Model (11.1)–(11.2) has eight parameters and four unknown initial conditions. The main data source is the cumulative number of H5N1 human cases given by the World Health Organization [166]. We give the data in Table 11.1.

Table 11.1 Number of cumulative human cases of H5N1 in units 10^5

Year	Time	Cases	Year	Time	Cases
2005	0	0.00047	2010	5	0.00467
	0.5	0.00108		5.5	0.005
2006	1	0.00148	2011	6	0.00516
	1.5	0.00229		6.5	0.00562
2007	2	0.00263	2012	7	0.00576
	2.5	0.00318		7.5	0.00607
2008	3	0.00351	2013	8	0.00610
	3.5	0.00387		8.5	0.00633
2009	4	0.00395	2014	9	0.0065
	4.5	0.00436			

If the data are taken at half-year intervals, that will give 19 data points, potentially not enough to fit all parameters and initial conditions. A better approach is to predetermine some of the parameters. The Food and Agriculture Organization of the United Nations (FAO) publishes statistics on livestock [58]. FAO gives that in 2012, there were 24 billion units of poultry worldwide [58]. We set the world's poultry population at 2400×10^7. Iwami [81] gives the mean lifespan of poultry to be two years. That translates into $\mu_d = 0.5$. Since the entire population Λ_d/μ_d is equal to 2400 we have $\Lambda_d = 1200$ in units of 10^7 per year. Iwami [81] also uses mean infectious period for domestic birds of 10 days, that is, $\nu_d = 36.5$ years^{-1}.

The natural lifespan of humans throughout the world varies significantly from country to country. We take an average value of human lifespan to be 65 years. Therefore, $\mu = 1/65$. The world human population has been on average approximately 6.5 billion over these 10 years. That gives a value of $\Lambda = 1000$ births per

year in units of 10^5 individuals. Finally, we preestimate some of the initial conditions: $S_d(0) = 2400$, $I(0) = 0.0007$ and $S(0) = 65000$.

Alternatively, one can get the information about the poultry and the human population of only the affected countries, of which there are 16. Data exist in the same data sources. This is left as an exercise (see Problem 11.1).

To determine β_d, β, $I_d(0)$, we fit the model (11.1)–(11.2) to the data, and we estimate $I_d(0) = 0.3936$, $\beta = 0.000000035327684$, and $\beta_d = 0.015489231377$. The fit is given in Fig. 11.1. The Matlab code that executes the fitting is given in the appendix.

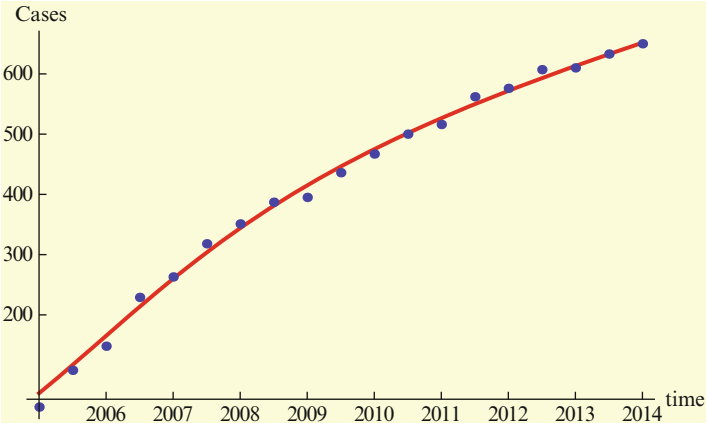

Fig. 11.1 Fit of model (11.1)–(11.2) to cumulative number of human cases of H5N1 given in Table 11.1

The estimated reproduction number with the fitted data is $\mathscr{R}_d = 1.00471$.

11.2.3 Evaluating Avian Influenza Control Strategies

Control strategies that are currently in place have the goal of delaying or preventing the emergence of a pandemic H5N1 strain. These measures currently involve the following [108]:

- Vaccination of poultry;
- Culling/destroying infected and potentially exposed poultry;
- Reducing contact with poultry by wearing protective gear;
- Isolation of humans infected with H5N1 and tracing the source of infection of the isolated individuals;
- Increasing biosecurity of poultry rearing;
- Education of poultry workers and health personnel.

Multiple control measures are applied differently in different countries. Evaluating their overall effectiveness is not a trivial task. Typically, this is done by collecting the opinions of experts. Here we present a more objective approach. Suppose the goal of the control measures is to reduce the number of cases of H5N1 in humans, that is, the goal is to minimize I^*. Each of the control measures impacts certain parameter values. We evaluate the change that a 1% change in a parameter p makes on I^* through the concept of elasticity. Recall that the elasticity of I^* with respect to the parameter p is given by

$$\varepsilon_p = \frac{\partial I^*}{\partial p}\frac{p}{I^*}, \tag{11.3}$$

where

$$I^* = \frac{\Lambda}{\mu+\nu}\frac{\beta\mu_d(\mathscr{R}_d-1)}{\beta\mu_d(\mathscr{R}_d-1)+\mu\beta_d}.$$

We use Mathematica to compute the elasticities with the above-evaluated parameters. The elasticities are listed in Table 11.2.

Table 11.2 Table of elasticities of I^*

Parameter	Elasticity	Parameter	Elasticity
β_d	212.454	β	1
μ_d	−215.338	μ	− 1.00042
ν_d	−210.569	ν	− 0.999579
Λ_d	213.454	Λ	1

From this table, we see directly that control measures that are applied to poultry and affect poultry parameters are much more effective in influencing the prevalence in humans than control measures applied to humans. This result seems robust and independent of the model [108]. To compare the control measures, we determine which parameters each control measure would affect. For instance, culling affects μ_d and ν_d. Culling with repopulation affects μ_d, ν_d, and Λ_d. Vaccination affects β_d and ν_d. Wearing protective gear affects β. We define the overall effect of the control measure to be the sum of efficacies of the effect of the measure on each affected parameter. For instance, culling with repopulation increases Λ_d and increases μ_d and ν_d. Hence, the overall efficacy is $213.454 - 210.569 - 215.338 = -212.453$. We will take this number as an absolute value. We summarize the overall effects of each control measure in Table 11.3. The affected parameters in the educational control measure are hard to pinpoint and are omitted.

Table 11.3 suggests that culling without repopulation is the most effective strategy, but it is rarely applied. Without it, culling with repopulation and biosecurity are the two most efficient strategies, followed by vaccination. The low rank of vaccination comes from the fact that vaccination leads to asymptomatic diseases and increases the lifespan of infected poultry. At low levels, vaccination effectively

Table 11.3 List of control measures and their efficacies

Control measure	Affected parameters	Overall efficacy	Rank
Culling w/o repopulation	μ_d, ν_d	425.907	1
Culling with repopulation	μ_d,ν_d,Λ_d	212.453	3
Vaccination	β_d, ν_d, β	2.885	4
Biosecurity	β_d	212.454	2
Protective gear	β	1	5
Isolation	ν	0.999579	6

supports higher prevalence. One caveat that we should mention is that the actual control measures do not necessarily affect all parameters with 1% change. Pinpointing the exact change is not trivial.

11.3 Seasonality in Avian Influenza Modeling

It has been known for a long time that human influenza exhibits seasonality in the temperate zones. In more tropical climates, human flu shows more complex patterns. The reasons for the human flu seasonality remain unknown.

Avian influenza H5N1 affects many countries with different climates, and yet it exhibits seasonality similar to the human flu in temperate climates. This can be easily seen from the monthly human cases in Fig. 11.2.

Figure 11.2 shows that most of the cases occur in the period from December through March, and there are very few cases in the summer months. Moreover, in humans, seasonality can also be observed in H5N1 poultry outbreaks. To capture seasonality, we have to measure t in days or in months. Our preference will be to measure t in days.

11.3.1 An Avian Influenza Model with Seasonality

The cause of seasonality in H5N1 is completely unknown. Some authors have hypothesized that seasonality is intrinsic and should be modeled with autonomous models whose endemic equilibria can be destabilized and exhibit oscillations [107]. A more likely scenario is that seasonality is extrinsic. Perhaps the transmission rate β_d is not a constant, but a periodic function of t, or the survivability of H5N1 in the environment is periodic. A recent study considered a number of potential extrinsic mechanisms and their combinations as possible drivers of seasonality in H5N1 [157]. The study performed model selection on the resulting seven models and found out that Iwami's model with periodic transmission rate is the best fit to the cumulative number of human cases. We introduce that model here:

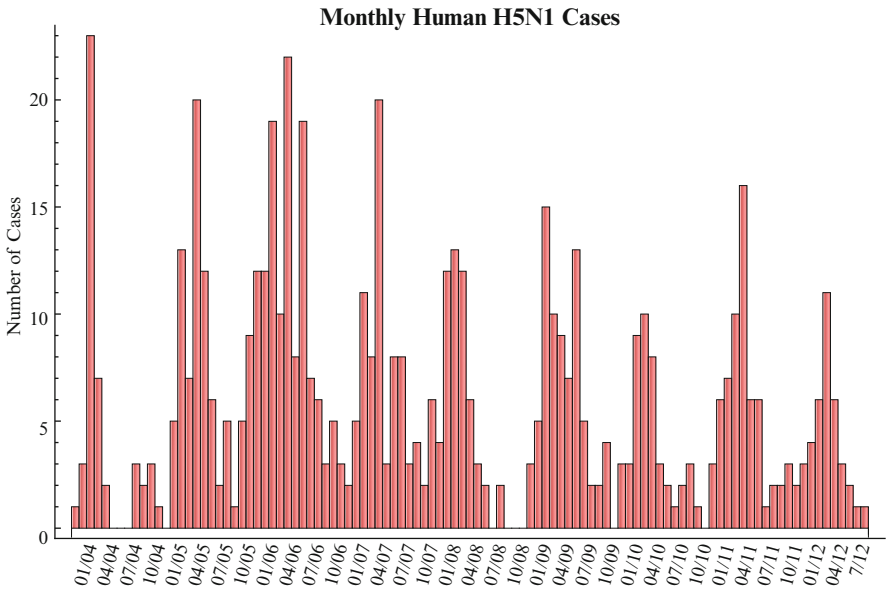

Fig. 11.2 Number of human cases of H5N1 by month. Data taken from [166]

$$\textbf{Domestic Birds:} \begin{cases} S_d'(t) = \Lambda_d - \beta_d(t) S_d I_d - \mu_d S_d, \\ I_d'(t) = \beta_d(t) S_d I_d - (\mu_d + \nu_d) I_d, \end{cases} \tag{11.4}$$

where the parameters have the same meanings as above. Seasonality is captured through a periodic transmission rate given by the periodic function

$$\beta_d(t) = \kappa_1 \sin\left(\frac{2\pi}{365}(t+\omega)\right) + \kappa_2. \tag{11.5}$$

Here κ_1 is the amplitude, ω is the horizontal shift, and κ_2 is the vertical shift. We want $\kappa_1 < \kappa_2$, so that $\beta_d(t) > 0$. The spillover model for humans takes the form

$$\textbf{Humans:} \begin{cases} S'(t) = \Lambda - \beta S I_d - \mu S, \\ I'(t) = \beta S I_d - (\mu + \nu) I. \end{cases} \tag{11.6}$$

Model (11.4)–(11.6) is well justified. As before, it can be fitted to the cumulative number of human cases of H5N1. We show the fit in Fig. 11.3, where we fitted the data through December 2009. These data are called calibration data. Then we extended the solution and plotted it alongside the incoming new data, called test data. It can be seen that the model describes well the incoming new data.

Model (11.4)–(11.6) is a model whose parameter β_d is an explicit function of the independent variable t. We have not considered models of this type before. Recall, however, that models in which one or more parameters are given functions of the independent variable are called *nonautonomous*.

An important class of nonautonomous models consists of those in which the parameters are periodic functions of the independent variable.

Definition 11.1. Models in which one or more parameters are periodic functions of the independent variable are called *periodic or seasonally forced* models.

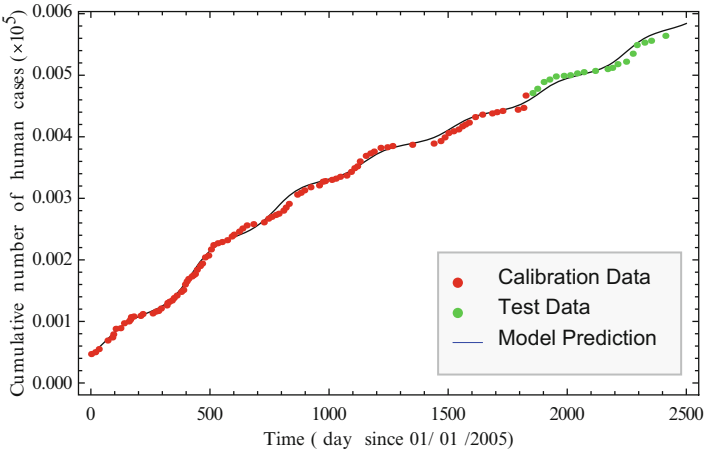

Fig. 11.3 Cumulative number of human cases of H5N1 in days

Methods for nonautonomous models include the Poincaré map and Floquet theory. We direct the reader to the many excellent books that cover this topic [155].

11.3.2 Tools For Nonautonomous Models

There are many tools that are designed to facilitate the study of nonautonomous periodic dynamical systems. Nonautonomous periodic dynamical systems are to a large extent analogous to autonomous dynamical systems. Here we explore two such tools.

11.3.2.1 The Poincaré Map

The Poincaré map was developed to study the intersection of the solution flow of a periodic orbit with the transversal cross section S. It is a tool for investigation of the n-dimensional dynamical system

$$x' = f(t,x). \tag{11.7}$$

The Poincaré map is defined as the point of return of the periodic orbit to S.

Suppose the flow ϕ generated by (11.7) is T-periodic, that is, $\phi(t+T,x_0)=\phi(t,x_0)$, and the cross section S of dimension $n-1$ is transversal to the vector field. Then the *Poincaré map* $\mathscr{P}(x):V\subset S\to S$ associates point x_0 in V with its point $\mathscr{P}(x_0)$ of the first return of the flow to S (see Fig. 11.4).

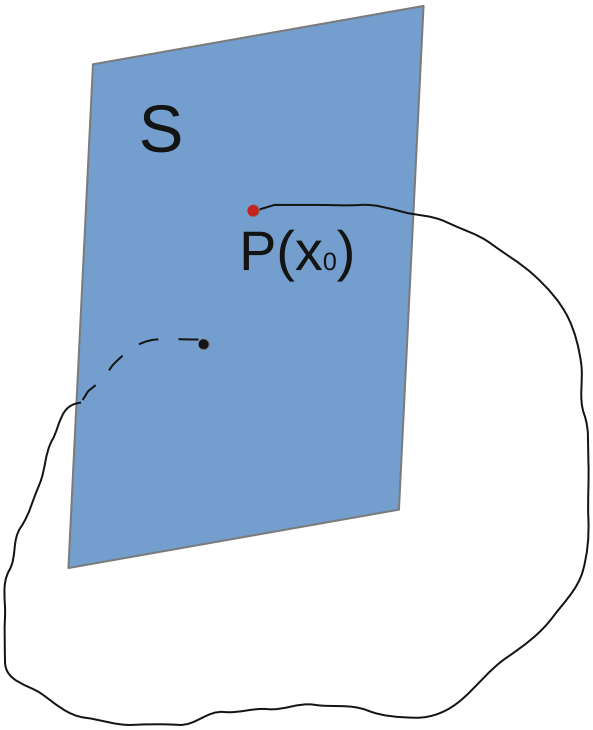

Fig. 11.4 A schematic description of the Poincaré map

The Poincaré map is relatively simple to study, but on the other hand, many of its properties are correlated to the properties of the flow. For instance, the stability of x of the map $\mathscr{P}(x)$ corresponds to the stability of the solution flow $\phi(t,x)$. If the solution flow has n (m) eigenvalues with negative (positive) real part, then the linearized map $D\mathscr{P}(x)$ has n (m) eigenvalues with modulus smaller (bigger) than one [155]. We utilize this property of the Poincaré map below.

11.3.2.2 Floquet Theory

Floquet theory provides another tool for investigating the local stability of solutions of periodic dynamical systems. Let

$$x' = A(t)x, \qquad x(0) = x_0, \tag{11.8}$$

be a linear nonautonomous periodic system, that is, $A(t+T) = A(t)$.

Definition 11.2. A matrix

$$\Phi(t) = [x^1(t), \ldots, x^n(t)],$$

where each column vector $x^j(t)$ is an independent solution of $x' = A(t)x$, is called a *fundamental matrix* of $x' = A(t)x$.

By definition,

$$\Phi'(t) = A(t)\Phi(t).$$

Theorem 11.1 (Floquet). *Each fundamental matrix of the T-periodic system $x' = A(t)x$ can be written as*

$$\Phi(t) = P(t)e^{Bt},$$

where $P(t)$ is T-periodic, $P(t+T) = P(t)$, and B is an $n \times n$ constant matrix.

The proof of this theorem can be found in [155] and is omitted. The following corollary shows the connection between a nonautonomous periodic linear system and the corresponding autonomous linear system.

Corollary 11.1. *The periodic system $x' = A(t)x$ is equivalent to the constant-coefficient system $y' = By$.*

Definition 11.3. The matrix $C = e^{BT}$ is called a *monodromy matrix*. The eigenvalues λ of the matrix B are called *Floquet exponents*. The eigenvalues $\rho = e^{\lambda T}$ of the matrix C are called *characteristic multipliers*.

11.3.2.3 Overview of Methods for Computing $\mathscr{R}_0$ in Periodic Models

When a nonautonomous system of differential equations is large, the analytical form of the reproduction number is difficult to compute. In this case, approximate methods must be used. Approximate methods compute an approximate value for the reproduction number. There are two types of approximate methods: analytical and computational. Among the analytical approximate methods is a method developed by Bacaër [19]. With this method, one calculates successive approximations of the reproduction number. Of course, analytically, one can compute perhaps two or three approximations, but these seem to be good enough. The advantage of the method is that one obtains an explicit formula for the approximate $\mathscr{R}_0$.

The reproduction number for periodic models has been more carefully defined in [164], in which the author proves its threshold properties and also gives an algorithm for its numerical computation. We notice that these results hold only when the coefficients of the nonautonomous model are periodic.

Another approach to computing the reproduction number is discussed in [45]. This approach is both approximate and exact for simpler models. The exact approach is very reminiscent of the approach we use below in this chapter for the computation of the reproduction number. Further discussion of the reproduction numbers including examples in which these computed reproduction numbers fail to provide a threshold for seasonally forced models can be found in [103].

11.3.3 Analyzing the Avian Influenza Model with Seasonality

Analyzing nonautonomous models is harder than analyzing autonomous models. For most of the nonautonomous models, even computing the reproduction number is a nontrivial task, and numerical methods must be used.

The model for domestic birds (11.4) can be separated from the full systems and investigated independently. The model is simple enough so that we can compute the reproduction number. The disease-free equilibrium of the model is time-independent and is given by $\mathscr{E}_0 = (\frac{\Lambda_d}{\mu_d}, 0)$. We linearize around the disease-free equilibrium. Let $x(t)$ be the perturbation of S_d, and $y(t)$ the perturbation in I_d. After dropping the quadratic terms and using the equations for the disease-free equilibrium to simplify, the system for the perturbations becomes

$$\begin{cases} x'(t) = -\beta_d(t)\dfrac{\Lambda_d}{\mu_d}y(t) - \mu_d x(t), \\ y'(t) = \beta_d(t)\dfrac{\Lambda_d}{\mu_d}y(t) - (\mu_d + \nu_d)y(t). \end{cases} \tag{11.9}$$

The second equation separates from the first. It is a linear equation with nonconstant coefficients. It can be solved explicitly. The solution is given by

$$y(t) = y(0)e^{\int_0^t \left(\beta_d(s)\frac{\Lambda_d}{\mu_d} - (\mu_d + \nu_d)\right) ds}. \tag{11.10}$$

To define the basic reproduction number, we first introduce the average of a periodic function over its period.

Definition 11.4. If $f(t)$ is a periodic function of period T, then the *average of* f is given by

$$< f >= \frac{1}{T}\int_0^T f(s)ds$$

Proposition 11.1. *If $f(t)$ is a periodic function of period T, then*

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t f(s)ds = < f > .$$

Proof. Let $t \in [nT,(n+1)T)$. Then $t = nT+\varepsilon$, where $\varepsilon \in [0,T)$. Furthermore,

$$\int_0^t f(s)ds = \int_0^{nT} f(s)ds + \int_{nT}^{nT+\varepsilon} f(s)ds = n\int_0^T f(s)ds + \int_0^{\varepsilon} f(s)ds.$$

Dividing by $t = nT+\varepsilon$ and taking the limit $n\to\infty$, we have

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t f(s)ds = \frac{1}{T}\int_0^T f(s)ds.$$

This completes the proof. □

Returning to Eq. (11.10), we see that

$$y(t) = y(0)e^{\left(\frac{1}{t}\int_0^t\left(\beta_d(s)\frac{\Lambda_d}{\mu_d}-(\mu_d+\nu_d)\right)ds\right)t} \approx y(0)e^{(<\beta_d>\frac{\Lambda_d}{\mu_d}-(\mu_d+\nu_d))t} \tag{11.11}$$

for t large enough. The expression on the right-hand side goes to $\pm\infty$ if and only if

$$< \beta_d > \frac{\Lambda_d}{\mu_d} - (\mu_d+\nu_d) > 0.$$

That prompts us to define the following reproduction number:

$$\mathscr{R}_0 = \frac{< \beta_d > \Lambda_d}{\mu_d(\mu_d+\nu_d)}. \tag{11.12}$$

We note that $\beta_d(t)$ is periodic with period 365 days. Hence $T = 365$. Clearly, we have the following traditional result:

Proposition 11.2. *If $\mathscr{R}_0 < 1$, then $y(t)\to 0$, and the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, then $|y(t)|\to\infty$, and the disease-free equilibrium is unstable.*

Furthermore, we can show that the disease-free equilibrium is globally stable. Indeed, we have the following result.

Proposition 11.3. *If $\mathscr{R}_0 < 1$, then $I_d(t)\to 0$ as $t\to\infty$.*

Proof. Adding the two equations, we have $N_d'(t) = \Lambda_d - \mu_d N_d - \nu_d I_d$, where $N_d = S_d + I_d$. This implies that $N_d'(t) \le \Lambda_d - \mu_d N_d$. We have shown before that in this case, $\limsup_t N_d(t) \le \frac{\Lambda_d}{\mu_d}$. Given $\varepsilon > 0$ such that

$$\mathscr{R}_0(\varepsilon) = \frac{< \beta_d > \left(\frac{\Lambda_d}{\mu_d} + \varepsilon\right)}{\mu_d + \nu_d} < 1,$$

there exists t_0 such that for every $t > t_0$, we have $N_d(t) < \frac{\Lambda_d}{\mu_d} + \varepsilon$. Since $\beta_d(t)$ is periodic, we can take $t_0 = nT$ and move the dynamical system so that $N_d(t) < \frac{\Lambda_d}{\mu_d} + \varepsilon$ is valid for all $t > 0$. Considering the equation for I_d, we have

$$I_d'(t) \leq [\beta(t)\left(\frac{\Lambda_d}{\mu_d} + \varepsilon\right) - (\mu_d + \nu_d)]I_d(t).$$

Solving this linear inequality, we have

$$I_d(t) \leq I_d(0)e^{\int_0^t [\beta(s)\left(\frac{\Lambda_d}{\mu_d} + \varepsilon\right) - (\mu_d + \nu_d)]ds}.$$

Thus, if $\mathscr{R}_0(\varepsilon) < 1$, we have $I_d(t) \to 0$ as $t \to \infty$. □

11.3.4 The Nonautonomous Avian Influenza Model with $\nu_d = 0$

The nonautonomous model (11.4) is very simple but capable of very complex behavior. To remove some of that complexity, in this subsection we analyze the model with $\nu_d = 0$. In this special case, the model can be reduced from a system of two equations to a single equation. In particular, we consider the system

$$\textbf{Domestic Birds:} \begin{cases} S_d'(t) = \Lambda_d - \beta_d(t)S_dI_d - \mu_dS_d, \\ I_d'(t) = \beta_d(t)S_dI_d - \mu_dI_d. \end{cases} \tag{11.13}$$

Adding the two equations, we obtain the equation of the total population size:

$$N_d'(t) = \Lambda_d - \mu_dN_d.$$

We know that the solution satisfies $N_d(t) \to \frac{\Lambda_d}{\mu_d}$. Let us assume for simplicity that $S_d(0) + I_d(0) = \frac{\Lambda_d}{\mu_d}$. Then $N_d(t) = \frac{\Lambda_d}{\mu_d}$ for all t. In this case,

$$S_d(t) = \frac{\Lambda_d}{\mu_d} - I_d(t).$$

From the second equation in (11.13) we obtain the following single equation in I_d:

$$I_d'(t) = \beta_d(t)(N_d - I_d(t))I_d(t) - \mu_dI_d(t), \tag{11.14}$$

where to simplify notation, we have set $N_d = \frac{\Lambda_d}{\mu_d}$. It can be shown as before that the reproduction number of this model is (see Problem 11.3)

$$\mathscr{R}_0 = \frac{< \beta_d > \Lambda_d}{\mu_d^2},$$

and the disease-free equilibrium is $\mathscr{E}_0 = (0)$. Equation (11.14) is a periodic Bernoulli equation [151] and has been studied before.

The main result for that equation is that it has a unique periodic solution $\xi(t)$ if $\mathscr{R}_0 > 1$. This periodic solution is globally asymptotically stable. We state these results in the following theorems:

Theorem 11.2. *Let $\beta_d(t)$ be periodic of period T. Assume also $\mathscr{R}_0 > 1$. Then equation (11.14) has a unique periodic solution $\xi(t)$.*

Proof. We consider Eq. (11.14) on the domain $\Omega = \{I_d : I_d \in [0, N_d]\}$. To show the existence of a periodic solution, we use the Poincaré map $\mathscr{P}$. The Poincaré map $\mathscr{P}$ maps the interval $[0, N_d]$ into itself. The Poincaré map is defined as follows. Let $I_d(0) = I_0$. Then

$$\mathscr{P}(I_0) = I_d(T, I_0),$$

where $I_d(t, I_0)$ is the solution of Eq. (11.14) that starts at I_0. In other words, $\mathscr{P}$ corresponds to the initial value I_0, the value of the solution at time $t = T$. Because of the properties of solutions to ODEs, the Poincaré map is one-to-one. Furthermore, it can be shown that it is continuously differentiable. It is not hard to show that $\mathscr{P}(0) = 0$ and $\mathscr{P}(N_d) < N_d$. The number $I_p \in [0, N_d]$ is an initial value of a periodic solution if and only if $\mathscr{P}(I_p) = I_p$, that is, if and only if I_p is a fixed point of the Poincaré map. Therefore, in order to show existence of a positive periodic solution of Eq. (11.14), we have to show that the Poincaré map has a fixed point. Define

$$v(t) = \frac{\partial I_d}{\partial I_0}(t, I_0).$$

Then the derivative of the Poincaré map is given as follows:

$$\mathscr{P}'(I_0) = \frac{\partial I_d}{\partial I_0}(T, I_0) = v(T).$$

To obtain the derivative of the Poincaré map, we differentiate equation (11.14) with respect to I_0. In this case, we obtain a differential equation in v:

$$v'(t) = v(t)[\beta_d(t)(N_d - I_d(t, I_0)) - \mu_d - \beta_d(t) I_d(t, I_0)]. \tag{11.15}$$

Differentiating the initial condition $I_d(0) = I_0$ with respect to I_0, we obtain that $v(0) = 1$. The differential equation for v can be solved, which gives the following expression for the derivative of the Poincaré map:

$$\mathscr{P}'(I_0) = e^{\int_0^T [\beta_d(t)(N_d - I_d(t, I_0)) - \mu_d - \beta_d(t) I_d(t, I_0)] dt}.$$

Clearly $\mathscr{P}'(I_0) > 0$, and hence the Poincaré map is increasing. Thus, if I_1 and I_2 are two initial conditions satisfying $I_1 < I_2$, then we have $\mathscr{P}(I_1) < \mathscr{P}(I_2)$. Furthermore,

$$\mathscr{P}'(0) = e^{T(<\beta_d> N_d - \mu_d)}.$$

Since $\mathscr{R}_0 > 1$, this means that the exponent is positive. Therefore, $\mathscr{P}'(0) > 1$. Hence for I_0 small enough,

$$\frac{\mathscr{P}(I_0) - \mathscr{P}(0)}{I_0} \approx \mathscr{P}'(0) > 1.$$

That implies that for I_0 small enough, $\mathscr{P}(I_0) > I_0$. Since, $\mathscr{P}(N_d) < N_d$, this means that the function $\mathscr{P}(I_0) - I_0$ changes sign in the interval $(0, N_d)$. Hence, there must exist I_p such that it becomes zero, that is, $\mathscr{P}(I_p) = I_p$.

To show uniqueness, we assume there are two distinct periodic solutions I_{p_1} and I_{p_2}. Without loss of generality, we may assume $I_{p_1} < I_{p_2}$. First, we note that if I_p is a periodic solution that satisfies model (11.14), then (see Problem 11.5)

$$\int_0^T [\beta_d(t)(N_d - I_d(t, I_p)) - \mu_d] dt = 0. \tag{11.16}$$

Second, for I_{p_1} and I_{p_2}, we have

$$|I_{p_1} - I_{p_2}| = |\mathscr{P}(I_{p_1}) - \mathscr{P}(I_{p_2})| = |\mathscr{P}'(I_m)||I_{p_1} - I_{p_2}|, \tag{11.17}$$

where I_m satisfies $I_{p_1} < I_m < I_{p_2}$. Furthermore, we have

$$\begin{aligned} \mathscr{P}'(I_m) &= e^{\int_0^T [\beta_d(t)(N_d - I_d(t, I_m)) - \mu_d - \beta_d(t) I(t, I_m)] dt} \\ &< e^{\int_0^T [\beta_d(t)(N_d - I_d(t, I_{p_1})) - \mu_d - \beta_d(t) I_d(t, I_m)] dt} \\ &< e^{-\int_0^T [\beta_d(t) I_d(t, I_m)] dt} < 1. \end{aligned} \tag{11.18}$$

Thus, we obtain a contradiction with (11.17). The contradiction is a result of the assumption that we have two distinct positive periodic solutions. □

Theorem 11.3. *Let $\beta_d(t)$ be periodic of period T. Assume also $\mathscr{R}_0 > 1$. Then the unique periodic solution $\xi(t)$ of Eq. (11.14) is globally stable, that is, if $I_d(t, I_0)$ is any solution starting from $I_d(0) = I_0$, then*

$$\lim_{t \to \infty} |I_d(t, I_0) - \xi(t)| = 0. \tag{11.19}$$

Proof. To complete the proof of the theorem, we have to establish the convergence to the periodic solution. We again assume $\mathscr{R}_0 > 1$, and we consider the solutions of Eq. (11.14). Let $I_d(t)$ be an arbitrary solution starting from the initial condition $I_d(0) = I_0$. We recall that I_p is the initial condition for the periodic solution. We assume that $I_p \neq I_0$. We have two choices, $\mathscr{P}(I_0) > I_0$ and $\mathscr{P}(I_0) < I_0$. We assume $\mathscr{P}(I_0) < I_0$. The other case can be addressed in a similar way. Since the Poincaré

map is increasing, we have $\mathscr{P}^n(I_0) < \mathscr{P}^{n-1}(I_0)$. Hence the sequence $\mathscr{P}^n(I_0)$ is a decreasing sequence. Since it is bounded from below, it must converge to a limit:

$$\lim_{n\to\infty} \mathscr{P}^n(I_0) = I_\infty.$$

It is not hard to see that the number I_∞ is a fixed point of the Poincaré map $\mathscr{P}(I_\infty) = I_\infty$. But the Poincaré map of model (11.14) has only two fixed points, $I_\infty = 0$ and $I_\infty = I_p$. Assume that $I_\infty = 0$. Then for some N, the number $\mathscr{P}^N(I_0)$ is small enough that from the properties of the Poincaré map, we have $\mathscr{P}^{N+1}(I_0) > \mathscr{P}^N(I_0)$, which contradicts the fact that the sequence is decreasing. Therefore, $I_\infty = I_p$. Consequently, the limit (11.19) holds. This completes the proof of the theorem. □

11.3.5 The Full Nonautonomous Avian Influenza Model

The nonautonomous model with $\nu_d \neq 0$ is a two-dimensional system and cannot be reduced to a single equation. Unlike autonomous two-dimensional models, which can exhibit only oscillations, two-dimensional nonautonomous models are capable of exhibiting chaotic behavior. This is the case with model (11.4). For small ν_d, the unique oscillatory solution is still stable, but as ν_d increases, period-doubling occurs, and the solution transitions to a chaotic solution. This can be seen in the bifurcation diagram in Fig. 11.5.

The chaotic solution exhibits the pattern typical for H5N1 outbreak. We show this in Fig. 11.6

Acknowledgements The author thanks Necibe Tuncer for help with fittings.

Appendix

In this appendix, we include Matlab code that fits model (11.1)–(11.2) to the data in Table 11.1.

```
function Ch11fitting_model1

clear all
close all
clc

load AFluDatCumHalf14.txt %Imports the data file

tdata =  AFluDatCumHalf14(:,1);
qdata =  AFluDatCumHalf14(:,2);

```

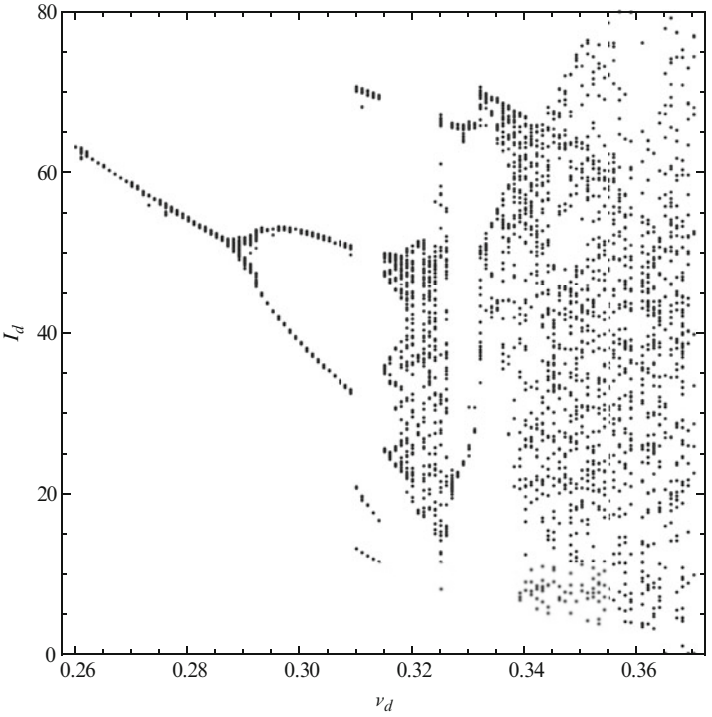

Fig. 11.5 Period-doubling and transition to chaos as ν_d increases. Parameters for the Figure are $\Lambda_D = 1020$, $\mu_d = 1/(2*356)$, $\kappa_1 = 0.00005111486$, $\kappa_2 = 0.00032621758$, $\omega = 127$

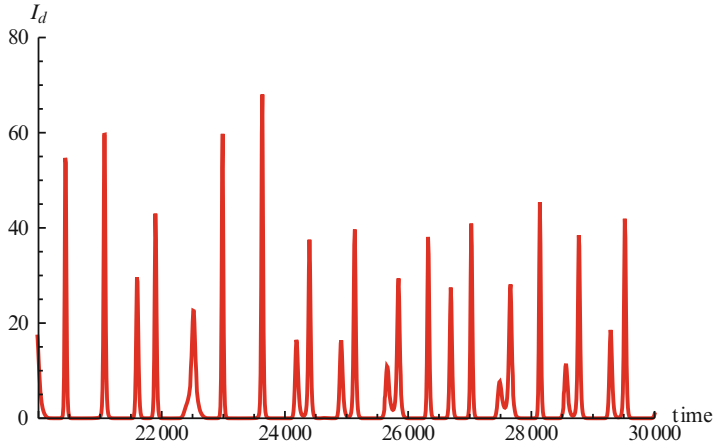

Fig. 11.6 Chaotic solution exhibits outbreak pattern. Parameters as in Fig. 11.5. In addition, $\nu_d = 0.35$

```
tforward = (0:0.01:9)';
tmeasure = [1:50:901]';

format long

    function dy = model_1(t,y,k)

    Lb = 1200;
    L  = 1000;
    mb = 1/2;
    nb = 36.5;
    mu = 1/65;
    nu = 36.5;

    dy = zeros(5,1);

    dy(1) = Lb -k(1)*y(1)*y(2)-mb*y(1);
    dy(2) = k(1)*y(1)*y(2)-(nb+mb)*y(2);
    dy(3) = L - k(2)*y(3)*y(2)-mu*y(3);
    dy(4) = k(2)*y(3)*y(2)-(mu+nu)*y(4);
    dy(5) = k(2)*y(3)*y(2);

    end

    function q = model1(k,tdata)

    [T,Y] = ode23s(@(t,y)(model_1(t,y,k)),tforward,
                           [2400  k(3)  65000  0.0007  ...
                               .0007]);

    q = Y(tmeasure(:),5);

    end

k = [0.0158  0.00000001063 0.9]; % Initial values for ...
    parameters

lb = [0.0           0.0              0.0];



for i = 1:5

  [k,resnorm] = lsqcurvefit(@model1,k,tdata,qdata,lb,[],...
   optimset('Disp','iter','TolX',10^(-20),'TolFun',10^(-20)))
 end


 [T,Y] = ode15s(@(t,y)(model_1(t,y,k)),tforward,[2400 ...
        k(3)  65000  .0007  .0007]);



figure(1)
```

```
64
65    plot(tdata,qdata,'r.');
66    hold on
67    plot(tforward,Y(:,5),'b-');
68
69
70  end
71
72  end
```

Problems

11.1. For model (11.1)–(11.2), obtain a list of the countries that have human cases from the main source of data [166]. Use [58] to obtain the number of poultry units for these countries. Use [167] to determine the human population of the affected countries. Parameterize model (11.1)–(11.2) with these data.

11.2. Consider the AI model with pandemic strain [80]

$$\textbf{Domestic Birds:} \begin{cases} S_d'(t) = \Lambda_d - \beta_d S_d I_d - \mu_d S_d, \\ I_d'(t) = \beta_b S_d I_d - (\mu_d + \nu_d) I_d. \end{cases} \tag{11.20}$$

The spillover model for humans with pandemic strain takes the form

$$\textbf{Humans:} \begin{cases} S'(t) = \Lambda - \beta S I_d - \beta_Z S Z - \mu S, \\ I'(t) = \beta S I_d - (\mu + \nu + \rho) I, \\ Z'(t) = \rho I + \beta_Z S Z - (\mu + \nu_Z) Z. \end{cases} \tag{11.21}$$

where Z is the number of individuals infected by the pandemic strain.

(a) Compute the reproduction numbers of the avian and the pandemic strains.
(b) Fit the model to the data in Table 11.1. Take $\nu_z = 36.5$. Estimate the reproduction numbers from the fit. The reproduction number of the pandemic strain should be between 1.5 and 3.
(c) Compute the invasion number of the pandemic strain.
(d) Compute the elasticity of the pandemic invasion number with respect to I_d^*. Culling facilitates invasion, but how pronounced is that effect?

11.3. Consider the reproduction number

$$\mathscr{R}_0 = \frac{< \beta_d > \Lambda_d}{\mu_d^2}.$$

Using the periodicity properties of sin, simplify it as much as possible.

11.4. Show that model (11.14) has as a reproduction number

$$\mathscr{R}_0 = \frac{< \beta_d > \Lambda_d}{\mu_d^2}.$$

11.5. Prove equality (11.16).

11.6. Let $\beta(t)$, $\mu(t)$, and $\gamma(t)$ be periodic with period T. Consider the model

$$I'(t) = \beta(t)(1 - I(t))I(t) - (\mu(t) + \gamma(t))I(t).$$

(a) Compute the reproduction number of this model.
(b) Show that $\mathscr{R}_0$ computed in (a) gives a threshold, that is, the DFE is locally asymptotically stable if $\mathscr{R}_0 < 1$ and unstable if $\mathscr{R}_0 > 1$.
(c) Sow that the DFE is globally asymptotically stable if $\mathscr{R}_0 < 1$.

11.7. Let $\beta(t)$, $\mu(t)$, and $\gamma(t)$ be periodic with period T. Consider the model (see Problem 11.6)

$$I'(t) = \beta(t)(1 - I(t))I(t) - (\mu(t) + \gamma(t))I(t).$$

(a) Show that the model has a unique periodic solution if $\mathscr{R}_0 > 1$.
(b) Show that the periodic solution is globally asymptotically stable.

11.8. Let $\beta(t)$, $\mu(t)$, and $\gamma(t)$ be periodic with period T. Consider the two-strain model

$$\begin{aligned} I_1'(t) &= \beta_1(t)(1 - I_1(t) - I_2(t))I_1(t) - (\mu_1(t) + \gamma_1(t))I_1(t), \\ I_2'(t) &= \beta_2(t)(1 - I_1(t) - I_2(t))I_2(t) - (\mu_2(t) + \gamma_2(t))I_2(t). \end{aligned} \tag{11.22}$$

(a) Define the reproduction number of each strain.
(b) Let $\xi_1(t)$ and $\xi_2(t)$ be the periodic solutions of strain one alone and strain two alone. Define the invasion numbers of strain one and strain two.

Hint: You have to look at the stability of the solution $(\xi_1(t), 0)$. Define the Floquet exponent for strain two and read off the invasion number.

Chapter 12
Age-Structured Epidemic Models

12.1 Introduction

Chronological age is perhaps one of the most important factors distinguishing individuals in a population that needs to be incorporated in population and epidemic models. Undoubtedly, vital characteristics such as birth and death rates differ markedly among the individuals of various ages. Age is also a key to capturing important mixing patterns in epidemic models. For instance, in childhood diseases, children predominantly mix with other children in similar age groups as well as with the individuals of the age groups of their parents and grandparents. Children are at greatest risk for contracting malaria and exhibiting strongest symptoms and highest death rate, yet malaria affects all age groups. Incidence of HIV is highest in the age groups from age twenty to age forty-five. Endemic models that incorporate births and deaths should also preferably incorporate age structure, since with time, the age profile of the population may change, and that may affect the dynamics of the disease.

Age-structured epidemic models are built on the basis of age-structured population models. There are several excellent introductory texts for age-structured population models [78, 46]. For completeness, we will introduce here first the linear age-structured population model. Most epidemic models use the linear model as a baseline population model.

12.2 Linear Age-Structured Population Model

The Malthusian model, which is also linear, considers a homogeneous population in which individuals are not distinguished by age [104]. The linear age-structured model is a strict analogue of the Malthusian model but allows for variability in age. As in the Malthusian model, the age-structured model considers a single population, not stratified by sex, which is isolated (no emigration and immigration) and

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_12

lives in an invariant habitat, that is, the birth and death rates can be assumed time-independent. The only characteristics by which individuals in the population differ is age.

Models that structure the population by age can be continuous or discrete. In this chapter, we will consider the continuous age-structured model. Because age and time are considered two independent variables, the resulting continuous models are cast as partial differential equation models. It was the physician A.G. McKendrick who first considered an age-structured PDE model for the dynamics of a one-sex population [115]. Earlier age-structured models were developed by Sharpe and Lotka [144, 3] but they were formulated as integral equations of Volterra type. We will first introduce McKendrick's age-structured PDE model, and then derive Lotka's model.

12.2.1 Derivation of the Age-Structured Model

McKendrick model introduces two independent variables: age a and time t. The distribution of the female population can be described by a function $u(a,t)$ that denotes the age-density of individuals at age a at time t. We assume $a \in [0, a_\dagger]$, $t \geq 0$, where $a_\dagger$ denotes the maximal age, which may be assumed infinite. If $a_\dagger = \infty$, then we assume that $u(a,t) = 0$ for all sufficiently large values of a. The function gives "density" rather than numbers, because the number of individuals of ages between a and $a + \Delta A$, where Δa is a small increment, at time t is approximately $u(a,t)\Delta a$. Thus, the total number of individuals of ages in the interval $[a_1, a_2]$ at time t is given by the integral

$$\int_{a_1}^{a_2} u(a,t)da.$$

Respectively, the total number of individuals in the population or the *total population size* at time t is given by

$$P(t) = \int_0^{a_\dagger} u(a,t)da.$$

Consider a *cohort* of individuals, that is, a group of individuals of age in an interval of length Δa. The number of individuals in that cohort is $u(a,t)\Delta a$. If a small interval of time Δt elapses, then those individuals who where of age a at time t will be of age $a + \Delta t$, and time is $t + \Delta t$. The number of these individuals is $u(a+\Delta t, t+\Delta t)\Delta a$. This is the same cohort. Their new number, however, is smaller than their original number, since some of these individuals will have died. We assume that the members of the population leave the population only by death. We denote the *age-specific per capita mortality rate* by $\mu(a)$. Then the number of individuals that die at age $[a, a+\Delta a]$ at time t is $\mu(a)u(a,t)\Delta a$. For the whole interval of time Δt, that number would be $\mu(a)u(a,t)\Delta a\Delta t$. The balance law can be written as

$$u(a+\Delta t,t+\Delta t)\Delta a - u(a,t)\Delta a = -\mu(a)\Delta t u(a,t)\Delta a. \tag{12.1}$$

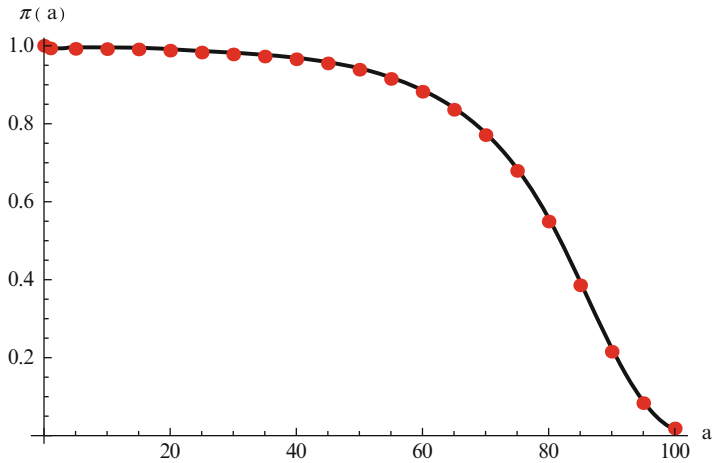

Fig. 12.1 Data on age-specific probability of survival in the United States in 2007. The data are taken from the life tables in the U.S. Vital Statistics Reports [15]. Data are interpolated to show the *continuous curve*

Dividing both sides by $\Delta a \Delta t$, we obtain a difference quotient on the left-hand side:

$$\frac{u(a+\Delta t,t+\Delta t)-u(a,t)}{\Delta t} = -\mu(a)u(a,t). \tag{12.2}$$

We take the limit as $t \to 0$. If $u(a,t)$ is a differentiable function with respect to each variable, we have

$$\begin{aligned}
&\lim_{\Delta t\to 0} \frac{u(a+\Delta t,t+\Delta t)-u(a,t)}{\Delta t} \\
&\quad = \lim_{\Delta t\to 0} \frac{u(a+\Delta t,t+\Delta t)-u(a,t+\Delta t)}{\Delta t} \\
&\qquad + \lim_{\Delta t\to 0} \frac{u(a,t+\Delta t)-u(a,t)}{\Delta t} \\
&\quad = u_a(a,t)+u_t(a,t).
\end{aligned} \tag{12.3}$$

We obtain McKendrick's equation, also referred to as the *McKendrick–von Foerster* equation, since it was later derived by von Foerster (1959) in the context of cell biology [63]:

$$u_a(a,t)+u_t(a,t) = -\mu(a)u(a,t). \tag{12.4}$$

The mortality rate is connected to the *probability of survival* (Fig. 12.1). Let $\pi(a)$ be the probability of survival from birth to age a. Suppose we start from a cohort of newborns of size P. Then $\pi(a)P$ gives the number of individuals from the cohort

who are of age a. Similarly, $\pi(a+\Delta a)P$ gives the number of individuals from the cohort who are of age $a+\Delta a$. Then $\pi(a+\Delta a)P-\pi(a)P$ is the number of individuals who have died in the age interval Δa. We have

$$\pi(a+\Delta a)P-\pi(a)P=-\mu(a)\pi(a)P\Delta a.$$

Canceling P and dividing by Δa, we have

$$\frac{\pi(a+\Delta a)-\pi(a)}{\Delta a}=-\mu(a)\pi(a).$$

Taking the limit as $\Delta a\to 0$ and assuming that the probability of survival $\pi(a)$ is a differentiable function, we obtain the following ordinary differential equation for $\pi(a)$:

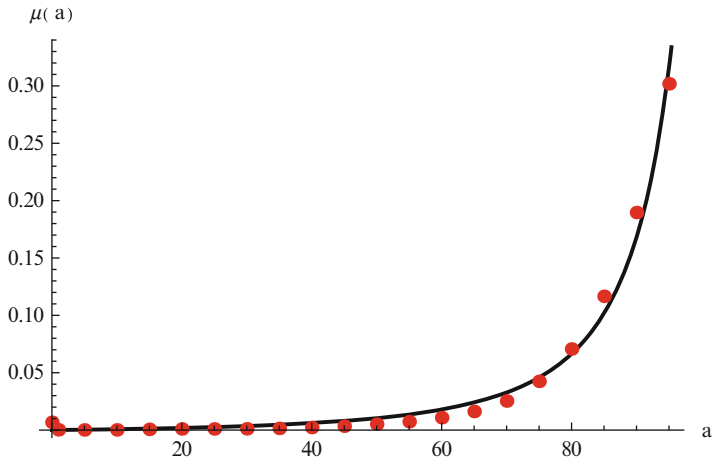

Fig. 12.2 Data on age-specific mortality in the United States in 2007. The data are taken from life tables of the the U.S. Vital Statistics Reports [15] as explained in the text. The age-specific function fitted to the data is $\mu(a)=\dfrac{0.748123a}{(110-a)^2}$

$$\pi'(a)=-\mu(a)\pi(a).$$

Since $\pi(0)$ is the probability of survival until age 0, we may assume that $\pi(0)=1$. Solving the initial value problem, we have the following constitutive form for the probability of survival until age a:

$$\pi(a)=e^{-\int_0^a \mu(s)ds}.$$

We further assume that the probability of survival until age a_1 is independent of the probability of survival until age a_2. If $a_2>a_1$, we have

$$\pi(a_2) = \pi(a_1)e^{-\int_{a_1}^{a_2}\mu(s)ds}. \tag{12.5}$$

Thus, $e^{-\int_{a_1}^{a_2}\mu(s)ds}$ is the probability that an individual will survive from age a_1 to age a_2. Age-dependent mortality $\mu(a)$ can be found in the national vital statistics files [121], but it is advisable that it be estimated from the life tables and the probability of survival. From the formula (12.5), we have

$$\int_{a_1}^{a_2}\mu(s)ds = \ln\left(\frac{\pi(a_1)}{\pi(a_2)}\right).$$

Hence approximately [12],

$$\mu(a_1) = \frac{1}{a_2 - a_1}\ln\left(\frac{\pi(a_1)}{\pi(a_2)}\right).$$

Computing the discrete values of $\mu(a)$, we may obtain a typical form of the age-specific mortality function. The age-specific mortality data for the United States in 2007 and a function $\mu(a)$ that fits them are given in Fig. 12.2.

Remark 12.1. If $a_\dagger < \infty$, then we **must** have

$$\lim_{a\to a_\dagger}\pi(a) = 0,$$

which means that no one survives until the maximal age. That, in turn, implies that

$$\lim_{a\to a_\dagger}\mu(a) = \infty.$$

Equation (12.4) is a first-order linear partial differential equation. It is defined on the domain

$$\mathscr{D} = \{(a,t) : a \geq 0, t \geq 0\},$$

that is, in the first quadrant. We will need conditions on the boundary of the domain to complete the model. In particular, we need to specify the age density for the initial population distribution at time $t = 0$:

$$u(a,0) = u_0(a),$$

where $u_0(a)$ is a given function. The function $u_0(a)$ is called the *initial population density*. It is assumed that $u_0(a) \geq 0$. Furthermore,

$$P_0 = \int_0^{a_\dagger} u_0(a)da < \infty,$$

where P_0 is the initial total population size.

The value $u(0,t)$ is the number of newborns at time t. To model the birth process, we introduce the *age-specific per capita birth rate* $\beta(a)$. Since there are $u(a,t)\Delta a$ females in the population with ages in the interval $[a, a+\Delta a]$, it follows that $\beta(a)u(a,t)\Delta a$ gives the number of births to females of age $[a, a+\Delta a]$ at time t.

Therefore, the total number of births at time t is the sum of all births at all ages: $\sum_i \beta(a_i)u(a_i,t)\Delta a$. Taking the limit as $\Delta a \to 0$, we obtain an integral in place of the sum. Thus, the **total birth rate** is given by

$$B(t) = \int_0^{a_\dagger} \beta(a)u(a,t)da,$$

which also gives the total number of newborns at time t. That is, we have

$$u(0,t) = \int_0^{a_\dagger} \beta(a)u(a,t)da.$$

Data on fertility for the United States in 2010 and a typical form of birth rate are given in Fig. 12.3.

The full McKendrick–von Foerster age-structured population model takes the form

$$\begin{cases} u_a(a,t) + u_t(a,t) = -\mu(a)u(a,t), \\ u(0,t) = \int_0^{a_\dagger} \beta(a)u(a,t)da, \\ u(a,0) = u_0(a). \end{cases} \tag{12.6}$$

The condition specified on the boundary $a = 0$, that is, $u(0,t)$, is called a **boundary condition**. Since the boundary condition is not specified through a given function

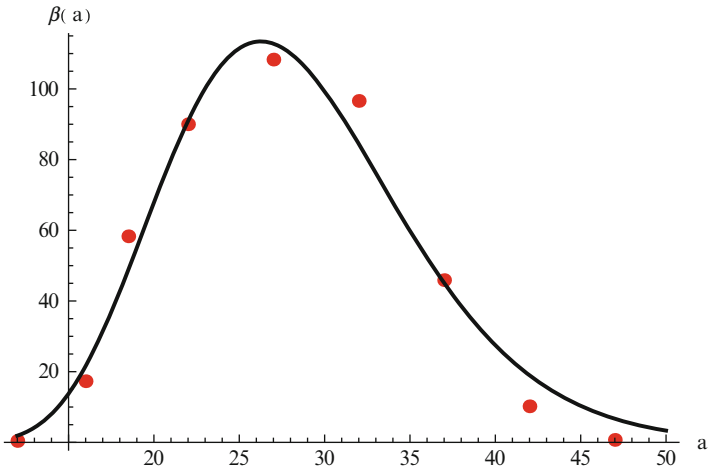

Fig. 12.3 Data on age-specific fertility rate in the United States in 2010. The data are taken from the U.S. Vital Statistics Reports [70]. The age-specific function fitted to the data is $\beta(a) = 8.151 * 10^{-7}(a-4.48268)^{10}e^{-0.459a}$

but through an equation that depends on the entire unknown function $u(a,t)$, this boundary condition is referred to as a **nonlocal boundary condition**. The condition giving a value to $u(a,0)$ is called an *initial condition*.

12.2.2 Reformulation of the Model Through the Method of the Characteristics. The Renewal Equation

Model (12.6) can be recast in an integral equation form, a form first derived by Sharpe and Lotka [3, 144]. To obtain the integral equation form, we must integrate the partial differential equation in model (12.6). That can be done through the **method of characteristics**. The method of characteristics typically applies to first-order hyperbolic partial differential equations. The method identifies curves, called characteristics, along which the partial differential equation reduces to an ordinary differential equation. Then the ordinary differential equation can be integrated from some initial data given on a suitable curve. Finally, the solution of the ordinary differential equation can be transformed into a solution for the original PDE.

An important step in the method of characteristics is identifying the characteristics of the PDE. This is relatively simple for *almost* linear PDEs. For an almost linear PDE of the form

$$c_1(x,y)\frac{\partial u}{\partial x} + c_2(x,y)\frac{\partial u}{\partial y} = c_0(x,y,u)$$

in parameterized form, the characteristics are given by

$$\begin{aligned} \frac{dx}{ds} &= c_1(x,y), \\ \frac{dy}{ds} &= c_2(x,y). \end{aligned} \tag{12.7}$$

Therefore, in our case we have

$$\frac{da}{ds} = 1, \qquad \frac{dt}{ds} = 1.$$

Dividing the one of the equations by the other, we have $dt/da = 1$. Hence $t = a + c$, where c is an arbitrary constant. This implies that the characteristics are lines of slope 1, called *characteristic lines* of the PDE. What is the practical significance of the characteristic lines? The value of $u(a,t)$ along the characteristic lines models one cohort of individuals. Thus the value of $u(a,t)$ is determined by previous values of $u(a,t)$ along the characteristics, and in particular by the value of u where the characteristic line crosses the boundary of $\mathscr{D}$.

The integral formulation can be derived from the partial differential equation through a rigorous procedure, called *integration along the characteristic lines*. To apply the procedure, we fix a point (a_0,t_0) in the first quadrant. We parameterize the characteristic line that goes through that point. If we denote by s the parameter for the fixed point (a_0,t_0) and a variable s, then $u(a_0+s,t_0+s)$ gives the value of u along the characteristic line. Define $v(s) = u(a_0+s,t_0+s)$. Then the derivative along the characteristic line can be written as

$$\frac{dv}{ds} = u_a + u_t.$$

We let also $\bar{\mu}(s) = \mu(a_0 + s)$. Then the partial differential equation in model (12.6) can be written as the following ordinary differential equation along the characteristic line:

$$\frac{dv}{ds} = -\bar{\mu}(s)v(s).$$

This ODE can easily be solved to give

$$v(s) = v(0)e^{-\int_0^s \bar{\mu}(\tau)d\tau}. \tag{12.8}$$

Now we have to interpret this solution in terms of u and two variables a and t. We consider two cases:

1. $a_0 \geq t_0$. The solution (12.8) gives, in the original variables,

$$u(a_0 + s, t_0 + s) = u(a_0, t_0)e^{-\int_0^s \mu(a_0+\tau)d\tau}. \tag{12.9}$$

   Now we specify our original point to have been fixed on one of the boundaries. Since we are in the case $a_0 \geq t_0$, we must take $t_0 = 0$. Then $s = t$ and $a_0 = a - t$. The choice of s and a_0 is made in such a way that $a_0 + s = a$ and $t_0 + s = t$. With these specifications, from (12.9) we obtain the solution in the case $a_0 \geq t_0$, that is, the solution when $a \geq t$:

$$\begin{aligned} u(a,t) &= u(a-t,0)e^{-\int_0^t \mu(a-t+\tau)d\tau}, \\ &= u(a-t,0)e^{-\int_{a-t}^a \mu(\sigma)d\sigma}, \\ &= u_0(a-t)\frac{\pi(a)}{\pi(a-t)}. \end{aligned} \tag{12.10}$$

2. $a_0 < t_0$. Again

$$u(a_0 + s, t_0 + s) = u(a_0, t_0)e^{-\int_0^s \mu(a_0+\tau)d\tau}. \tag{12.11}$$

   Now we set $a_0 = 0$. Then $s = a$, $t_0 = t - a$. Substituting above, we get for $a < t$,

$$u(a,t) = u(0,t-a)e^{-\int_0^a \mu(\tau)d\tau} = B(t-a)\pi(a).$$

So finally, the solution of the PDE can be written in the form

$$u(a,t) = \begin{cases} u_0(a-t)\dfrac{\pi(a)}{\pi(a-t)} & a \geq t, \\ B(t-a)\pi(a) & a < t. \end{cases} \tag{12.12}$$

This would have been an explicit solution of the partial differential equation if $B(t)$ were a given function. However, $B(t)$ depends on $u(a,t)$. Thus, this represen-

tation is another equation for $u(a,t)$. This representation is not easy to work with, since it is an equation for a function of two variables. It is customary to rewrite this representation in the form of an equation for the function $B(t)$. To do that, we substitute the expression for $u(a,t)$ from (12.12) into the formula for $B(t)$. We have two cases:

Case $a_\dagger < \infty$: In this case, we have

$$B(t) = \begin{cases} \displaystyle\int_0^t \beta(a)\pi(a)B(t-a)da + \int_t^{a_\dagger} \beta(a)\frac{\pi(a)}{\pi(a-t)}u_0(a-t)da & t \le a_\dagger, \\ \displaystyle\int_0^{a_\dagger} \beta(a)\pi(a)B(t-a)da & t > a_\dagger. \end{cases} \tag{12.13}$$

Case $a_\dagger = \infty$: In this case, we have only $t < \infty$. Therefore, we have the following equation for $B(t)$:

$$B(t) = \int_0^t \beta(a)\pi(a)B(t-a)da + F(t), \tag{12.14}$$

where $F(t)$ is a given function:

$$F(t) = \int_t^\infty \beta(a)\frac{\pi(a)}{\pi(a-t)}u_0(a-t)da = \int_0^\infty \beta(a+t)\frac{\pi(a+t)}{\pi(a)}u_0(a)da.$$

Equation (12.14) is a linear Volterra integral equation of convolution type with kernel $K(a) = \beta(a)\pi(a)$. The function $K(a)$ is sometimes referred to as a *maternity function* [78]. Equation (12.14) is called the **renewal equation** or **Lotka equation**. If we can solve (12.14) and determine $B(t)$ from it, then we can obtain $u(a,t)$ from (12.12).

12.2.3 Separable Solutions. Asymptotic Behavior

The McKendrick–von Foerster model (12.6) is a linear model. Thus its solutions are not necessarily bounded but grow or decay to zero nearly exponentially in time. It can be shown that all solutions approach in time a *separable solution*. Separable solutions for partial differential equation are solutions that can be written as a product of two functions: a function of age and a function of time, that is, the solution can be written in the form $u(a,t) = T(t)\phi(a)$. It can be shown (see Problem 12.4) that the time-dependent function $T(t)$ is actually an exponential function. Hence, the separable solutions of the model (12.6) are of the form

$$u(a,t) = e^{\lambda t}\phi(a).$$

To find the separable solutions, we have to find the number λ and the function $\phi(a)$. To find λ and $\phi(a)$, we substitute in the differential equation of model (12.6). Therefore, we have

$$e^{\lambda t}\phi'(a) + \lambda e^{\lambda t}\phi(a) = -\mu(a)e^{\lambda t}\phi(a),$$

where the prime denotes a derivative with respect to a. After canceling $e^{\lambda t}$, we obtain the differential equation for the function $\phi(a)$:

$$\phi'(a) + \lambda\phi(a) = -\mu(a)\phi(a).$$

This is a first-order linear ordinary differential equation, which can be explicitly solved. The solution is given by

$$\phi(a) = \phi_0 e^{-\lambda a}\pi(a),$$

where the constant $\phi_0 = \phi(0)$ is unknown and will be specified later. Hence, we have the following form for the solution $u(a,t)$:

$$u(a,t) = \phi_0 e^{\lambda t}e^{-\lambda a}\pi(a).$$

We seek to satisfy the boundary condition in model (12.6):

$$\phi_0 e^{\lambda t} = \phi_0 e^{\lambda t}\int_0^{a_\dagger} \beta(a)e^{-\lambda a}\pi(a)da.$$

Therefore, λ should be identified in such a way that the equation below is satisfied.

The characteristic equation of the McKendrick–von Foerster model is

$$\int_0^{a_\dagger} \beta(a)e^{-\lambda a}\pi(a)da = 1. \tag{12.15}$$

This is a transcendental equation, and it can have multiple solutions for λ that are real and complex. It turns out that the Eq. (12.15) has a unique real solution if $\beta(a)$ is positive on some positive interval:

Lemma 12.1. *Assume $\beta(a) \geq \hat{\beta} > 0$ for $a \in [a_1, a_2]$. Then there is a unique real solution λ^* such that Eq. (12.15) holds.*

The unique real solution λ^* of Eq. (12.15) gives the growth rate of the population for the age-structured model (12.6). It is similar to the population growth rate derived from the Malthusian model.

Definition 12.1. The parameter λ^* is called a *Malthusian parameter* or *growth rate of the population*.

Proof. Define

$$f(\lambda) = \int_0^{a_\dagger} \beta(a) e^{-\lambda a} \pi(a) da.$$

Because $\beta(a) \geqslant \hat{\beta} > 0$ on a nonzero interval, it follows that $f(\lambda)$ is a strictly decreasing function of λ. The function $f(\lambda)$ may not be defined for all values of λ, particularly if $a_\dagger = \infty$. Assume that $f(\lambda)$ is defined and continuous on the interval $(-L, \infty)$. We have

$$\lim_{\lambda \to -L} f(\lambda) = \infty \qquad \lim_{\lambda \to \infty} f(\lambda) = 0.$$

Hence, there is a unique real value λ^* that satisfies $f(\lambda^*) = 1$. □

The *net reproduction rate* of the population is defined as

$$\mathscr{R} = \int_0^{a_\dagger} \beta(a) \pi(a) da.$$

The following relationship between the net reproduction rate and the growth rate of the population exists:

$$\begin{aligned} \mathscr{R} > 1 \quad &\Longleftrightarrow \quad \lambda^* > 0, \\ \mathscr{R} = 1 \quad &\Longleftrightarrow \quad \lambda^* = 0, \\ \mathscr{R} < 1 \quad &\Longleftrightarrow \quad \lambda^* < 0, \end{aligned} \tag{12.16}$$

where $\mathscr{R}$ gives the number of progeny that one female in the population produces during her lifespan. If $\mathscr{R} > 1$, that is, one individual produces more than one individual, we have $\lambda^* > 0$, and the population is growing.

Finally, we can identify ϕ_0 from the initial condition. Recall that $P_0 = \int_0^{a_\dagger} u_0(a) da$ and $u(a,0) = u_0(a)$. On the other hand, $u(a,t) = \phi_0 e^{\lambda^* t} e^{-\lambda^* a} \pi(a)$. Therefore,

$$\int_0^{a_\dagger} u_0(a) da = \phi_0 \int_0^{a_\dagger} e^{-\lambda^* a} \pi(a) da.$$

Hence,

$$\phi_0 = \frac{\int_0^{a_\dagger} u_0(a) da}{\int_0^{a_\dagger} e^{-\lambda^* a} \pi(a) da}. \tag{12.17}$$

In the following example, we compute λ^*.

Example 12.1. Suppose

$$\begin{aligned} \beta(a) &= \bar{\beta} a e^{-ca}, \\ \mu(a) &= \mu, \end{aligned} \tag{12.18}$$

where $\bar{\beta}$, c, and μ are given constants. Take $a_\dagger = \infty$. Compute the growth rate of the population λ^* and the net reproduction rate $\mathscr{R}$.

Solution: The growth rate λ^* is a solution to the following equation:

$$\int_0^\infty \beta(a) e^{-\mu a} e^{-\lambda a} da = 1.$$

Replacing $\beta(a)$, we have

$$\bar{\beta} \int_0^\infty a e^{-ca} e^{-\mu a} e^{-\lambda a} da = 1.$$

Collecting terms, we obtain

$$\bar{\beta} \int_0^\infty a e^{-(c+\mu+\lambda)a} da = 1.$$

Computing the integral (think under what conditions this is possible), we have

$$\frac{\bar{\beta}}{(\lambda+\mu+c)^2} = 1. \tag{12.19}$$

The function

$$f(\lambda) = \frac{\bar{\beta}}{(\lambda+\mu+c)^2}$$

is defined and continuous on the interval $(-(\mu+c), \infty)$. Thus $L = \mu + c$. Furthermore, we have

$$\lim_{\lambda \to \infty} f(\lambda) = 0 \qquad \lim_{\lambda \to -(\mu+c)} f(\lambda) = \infty.$$

Then there is a unique real solution in the interval $(-(\mu+c), \infty)$. Solving Eq. (12.19), we obtain

$$\lambda^* = \sqrt{\beta} - (\mu + c).$$

Define

$$\mathscr{R} = f(0) = \frac{\bar{\beta}}{(\mu+c)^2}.$$

From the expressions for $\mathscr{R}$ and λ^*, we can see that $\mathscr{R} > 1$ if and only if $\lambda^* > 0$.

12.3 Age-Structured SIS Epidemic Models

Age-structured epidemic models are often formulated around the basic linear McKendrick–von Foerster model. Because the total population in that model has three modes of change, exponential growth, exponential decay, and constant, we

have to be careful in incorporating disease dynamics. The mass action incidence is intrinsically nonlinear and suggests that the model should exhibit convergence to an equilibrium. For consistency, we take the total population size to be constant. This means that we take the age-specific birth and death rate in such a way that $\mathscr{R} = 1$ or the growth rate of the population is zero. In particular, this assumption leads to the condition

$$\int_0^{a_\dagger} \beta(a)\pi(a)da = 1. \tag{12.20}$$

That is an appropriate assumption for age-dependent models with no disease-induced mortality.

12.3.1 Introduction of the SIR Age-Structured Epidemic Model

We begin with one of the simplest age-structured epidemic models, the SIR model. Because chronological age is a heterogeneity characteristic of all individuals in the population, all epidemic classes have to be structured by age. Hence, $s(a,t)$ denotes the density of susceptible individuals of age a at time t, and $i(a,t)$ denotes the density of infected/infectious individuals of age a at time t, and $r(a,t)$ denotes the density of recovered individuals of age a at time t. The age-dependent density of the total population $u(a,t)$ is given by

$$u(a,t) = s(a,t) + i(a,t) + r(a,t).$$

As before, the densities of the susceptible and infective individuals imply that $s(a,t)\Delta a$ denotes the number of susceptible individuals of age in the interval $[a, a+\Delta a]$ at time t, and $i(a,t)\Delta a$ denotes the number of infected individuals of age in the interval $[a, a+\Delta a]$ at time t, and similarly for r. Hence, the total numbers of susceptible, infective, and recovered individuals are

$$S(t) = \int_0^{a_\dagger} s(a,t)da, \qquad I(t) = \int_0^{a_\dagger} i(a,t)da, \qquad R(t) = \int_0^{a_\dagger} r(a,t)da.$$

The total population size is defined as

$$P(t) = S(t) + I(t) + R(t) = \int_0^{a_\dagger} u(a,t)da.$$

The change of each epidemic class is given in the same way as the change of the total population density $u(a,t)$ in the McKendrick–von Foerster equation, that is, as a sum of the age derivative and the time derivative: The right-hand side of an age-structured epidemic model is similar to the right-hand side of the corresponding ODE epidemic model, but all rates are age-dependent, and recruitment is incorporated as a boundary condition at age zero. Following these guidelines, we can write the age-specific SIS epidemic model as

$$\begin{aligned} s_a + s_t &= -\lambda(a,t)s(a,t) - \mu(a)s(a,t), \\ i_a + i_t &= \lambda(a,t)s(a,t) - (\mu(a)+\gamma(a))i(a,t), \\ r_a + r_t &= \gamma(a)i(a,t) - \mu(a)r(a,t), \end{aligned} \tag{12.21}$$

where $\lambda(a,t)$ is the age-specific force of infection, which depends on the infected individuals, $\mu(a)$ is the age-specific natural mortality rate, and $\gamma(a)$ is the recovery rate.

The general form of the force of infection is given by the expression

$$\lambda(a,t) = \rho(a)i(a,t) + \int_0^{a_\dagger} k(a,\tau)i(\tau,t)d\tau. \tag{12.22}$$

Several special cases of the force of infection have been considered in the literature. The *intracohort mixing* case assumes that individuals of a certain age infect individuals of the same age only. It is obtained from the above formula with $k(a,\tau) = 0$. In this case,

$$\lambda(a,t) = \rho(a)i(a,t).$$

Purely intracohort mixing may be appropriate for some childhood diseases. Purely *intercohort mixing* occurs when the mixing occurs among all age classes. It is obtained from the formula above with the assumption that $\rho(a) = 0$. In this case, the force of infection is given by the formula

$$\lambda(a,t) = \int_0^{a_\dagger} k(a,\tau)i(\tau,t)d\tau.$$

In this general case, the mixing kernel $k(a,\tau)$, which gives the rate of transmission between a susceptible of age a and infective of age τ, does not allow for the computation of an explicit reproduction number. Hence, often in practice, it is assumed that the mixing kernel is of *separable form*, that is, that it can be written as

$$k(a,\tau) = k_1(a)k_2(\tau).$$

Like the model for the total population size, the age-structured epidemic model (12.21) is also defined in the first quadrant, and initial and boundary conditions should be specified for the given model to be completely defined. The boundary conditions are specified at age $a = 0$. We may assume that all newborns are susceptible, in which case we specify

$$\begin{aligned} s(0,t) &= \int_0^{a_\dagger} \beta(a)u(a,t)da, \\ i(0,t) &= 0, \\ r(0,t) &= 0, \end{aligned} \tag{12.23}$$

where $\beta(a)$ is the age-specific birth rate. On the other hand, we may assume that some of the newborns are born infected, that is, that the disease exhibits *vertical transmission*, in which case we specify

$$\begin{aligned} s(0,t) &= \int_0^{a_\dagger} \beta(a)[s(a,t)+r(a,t)+(1-q)i(a,t)]da, \\ i(0,t) &= q\int_0^{a_\dagger} \beta(a)i(a,t)da, \\ r(0,t) &= 0, \end{aligned} \tag{12.24}$$

where $0 \le q \le 1$ is called a *vertical transmission parameter.* Finally, we have to specify initial conditions. These are usually given age-specific functions that specify the distribution of the susceptible and infected individuals at the starting time:

$$s(a,0) = s_0(a) \qquad i(a,0) = i_0(a) \qquad r(a,0) = r_0(a).$$

Model (12.21), like its age-independent counterpart, also has equilibria. In the next section, we discuss the equilibria and the reproduction number in the purely intercohort case without vertical transmission.

12.3.2 Equilibria and Reproduction Number

As before, equilibria are time-independent solutions of model (12.21) that also satisfy the boundary conditions. In the current example, we will take $a_\dagger = \infty$. In particular, these are age-specific densities $s(a)$ and $i(a)$ such that

$$\begin{aligned} s_a &= -\lambda(a)s(a) - \mu(a)s(a), \\ i_a &= \lambda(a)s(a) - (\mu(a)+\gamma(a))i(a), \\ r_a &= \gamma(a)i(a) - \mu(a)r(a), \\ s(0) &= \int_0^{\infty} \beta(a)u(a)da, \\ i(0) &= 0, \\ r(0) &= 0, \end{aligned} \tag{12.25}$$

where the force of infection is given by

$$\lambda(a) = k_1(a)\int_0^{\infty} k_2(\tau)i(\tau)d\tau = \hat{\lambda}k_1(a).$$

We note that the number $\hat{\lambda}$ denotes the integral. We should keep in mind that the number $\hat{\lambda}$ is not given, it depends rather on the number of infected individuals, which is to be determined.

System (12.25) clearly has a disease-free equilibrium as a solution: $i(a) = 0$, $r(a) = 0$, $\hat{\lambda} = 0$, and $s(a) = s(0)\pi(a)$, where $s(0)$ is an arbitrary constant and

$$\pi(a) = e^{-\int_0^a \mu(\tau)d\tau}.$$

To find the endemic equilibria, we may assume that $\hat{\lambda} \neq 0$. The equation for the density of susceptible individuals $s(a)$ can be solved to give

$$s(a) = s(0)e^{-\int_0^a (\hat{\lambda}k_1(s)+\mu(s))ds}. \tag{12.26}$$

If we add the three equations in (12.25), we obtain

$$u_a = -\mu(a)u(a).$$

We may solve this equation as before to get $u(a) = u(0)\pi(a)$, where $u(0)$ is an arbitrary constant. It can be shown that $u(0) = \phi_0$ from (12.17), where $\lambda^* = 0$. Consequently (see (12.20),

$$s(0) = \phi_0 \int_0^\infty \beta(a)\pi(a)da = \phi_0.$$

Hence, $s(0)$ is also the constant ϕ_0. Since $s(a)$ is given, to obtain the nonzero density of the infective individuals, we have to solve the equation for the infectives. To do that, we move the outflow terms from the right-hand side to the left-hand side and multiply by the appropriate integration factor:

$$e^{\int_0^a (\mu(\sigma)+\gamma(\sigma))d\sigma}[i_a + (\mu(a)+\gamma(a))i(a)] = \hat{\lambda}k_1(a)s(a)e^{\int_0^a (\mu(\sigma)+\gamma(\sigma))d\sigma}.$$

Then the left-hand side can be written as a complete derivative:

$$\frac{d}{da}\left[e^{\int_0^a (\mu(\sigma)+\gamma(\sigma))d\sigma}i(a)\right] = \hat{\lambda}k_1(a)s(a)e^{\int_0^a (\mu(\sigma)+\gamma(\sigma))d\sigma}.$$

We can rename the variable in the above equation from a to η. We then integrate from 0 to a. On the left-hand side, we have an integral after a derivative. Evaluating the integral, we have

$$e^{\int_0^a (\mu(\sigma)+\gamma(\sigma))d\sigma}i(a) - i(0) = \hat{\lambda}\int_0^a k_1(\eta)s(\eta)e^{\int_0^\eta (\mu(\sigma)+\gamma(\sigma))d\sigma}d\eta.$$

From the boundary conditions, we conclude that $i(0) = 0$. We multiply by the reciprocal of the integration factor to get

$$i(a) = \hat{\lambda}\int_0^a k_1(\eta)s(\eta)e^{-\int_\eta^a (\mu(\sigma)+\gamma(\sigma))d\sigma}d\eta.$$

Replacing $s(\eta)$ with its value, we obtain the final form of the infected individuals $i(a)$:

$$i(a) = \phi_0\hat{\lambda}\pi(a)\int_0^a k_1(\eta)e^{-\int_0^\eta \hat{\lambda}k_1(s)ds}e^{-\int_\eta^a \gamma(\sigma)d\sigma}d\eta. \tag{12.27}$$

We notice that this form is not explicit, because it depends on $\hat{\lambda}$, which is still to be determined. To compute $\hat{\lambda}$, we replace $i(a)$ in the definition of $\hat{\lambda}$. Hence,

$$\hat{\lambda} = \phi_0 \hat{\lambda} \int_0^\infty k_2(a)\pi(a) \int_0^a k_1(\eta) e^{-\int_0^\eta \hat{\lambda} k_1(s)ds} e^{-\int_\eta^a \gamma(\sigma)d\sigma} d\eta da.$$

Since $\hat{\lambda} \neq 0$, we may cancel to obtain the following equation for $\hat{\lambda}$:

$$1 = \phi_0 \int_0^\infty k_2(a)\pi(a) \int_0^a k_1(\eta) e^{-\int_0^\eta \hat{\lambda} k_1(s)ds} e^{-\int_\eta^a \gamma(\sigma)d\sigma} d\eta da.$$

Define the right-hand side of this equation as a function of $\hat{\lambda}$:

$$\mathscr{G}(\hat{\lambda}) = \phi_0 \int_0^\infty k_2(a)\pi(a) \int_0^a k_1(\eta) e^{-\int_0^\eta \hat{\lambda} k_1(s)ds} e^{-\int_\eta^a \gamma(\sigma)d\sigma} d\eta da.$$

If the function $k_1(a)$ is nonzero on some interval, then $\mathscr{G}(\hat{\lambda})$ is a strictly decreasing function of $\hat{\lambda}$. Hence, if the equation $\mathscr{G}(\hat{\lambda}) = 1$ has a positive solution, that solution must be unique. Furthermore,

$$\lim_{\hat{\lambda} \to \infty} \mathscr{G}(\hat{\lambda}) = 0.$$

Therefore, the existence of a positive solution of the equation $\mathscr{G}(\hat{\lambda}) = 1$ depends on the value of $\mathscr{G}(0)$. We define that value as the basic reproduction number of the disease:

$$\mathscr{R}_0 = \phi_0 \int_0^\infty k_2(a)\pi(a) \int_0^a k_1(\eta) e^{-\int_\eta^a \gamma(\sigma)d\sigma} d\eta da. \tag{12.28}$$

We conclude that if $\mathscr{R}_0 > 1$, then there is a unique positive solution $\hat{\lambda}^* > 0$ of the equation $\mathscr{G}(\hat{\lambda}) = 1$. To that solution $\hat{\lambda}^*$, there correspond unique nonzero values $s^*(a)$ and $i^*(a)$. The value of $r(a)$ can be computed from $r^*(a) = u(a) - s^*(a) - i^*(a)$. We have the following result:

Proposition 12.1. *Model (12.21) has a unique disease-free equilibrium* $\mathscr{E}_0 = (\phi_0 \pi(a), 0, 0)$. *If* $\mathscr{R}_0 > 1$, *the model also has a unique endemic equilibrium* $\mathscr{E}^* = (s^*(a), i^*(a), r^*(a))$.

12.3.3 Local Stability of the Disease-Free Equilibrium

In this section, we study the local stability of the equilibria. The approach is similar to the ODE models. We need to linearize around an equilibrium. To the system (12.21) we will associate the linear system for the perturbations around an equilibrium. We will consider only the first two equations of system (12.21). We will replace the equation for r with the equation for the total population size. Let $s(a,t) = s^*(a) + x(a,t)$, $i(a,t) = i^*(a) + y(a,t)$ and $u(a,t) = u^*(a) + p(a,t)$, where $x(a,t)$, $y(a,t)$, and $p(a,t)$ are the perturbation of the equilibrium. For now, s^* and i^* denote a generic equilibrium that could be the disease-free or the endemic equilibrium. Plugging in the expressions for $s(a,t)$ and $i(a,t)$, we obtain

$$\begin{aligned}
s_a^* + x_a + x_t &= -(\lambda^*(a) + \tilde{\lambda}(a,t))(s^*(a) + x(a,t)) \\
&\quad -\mu(a)(s^*(a) + x(a,t)), \\
i_a^* + y_a + y_t &= (\lambda^*(a) + \tilde{\lambda}(a,t))(s^*(a) + x(a,t)) \\
&\quad -(\mu(a) + \gamma(a))(i^*(a) + y(a,t)), \\
u_a^* + p_a + p_t &= -\mu(a)(u^*(a) + p(a,t)), \\
s^*(0) + x(0,t) &= \int_0^\infty \beta(a)(u^*(a) + p(a,t))da, \\
i^*(0) + y(0,t) &= 0, \\
u^*(0) + p(0,t) &= \int_0^\infty \beta(a)(u^*(a) + p(a,t))da,
\end{aligned} \tag{12.29}$$

where

$$\begin{aligned}
\lambda^*(a) &= k_1(a) \int_0^\infty k_2(a) i^*(a) da \\
\tilde{\lambda}(a,t) &= k_1(a) \int_0^\infty k_2(a) y(a,t) da.
\end{aligned} \tag{12.30}$$

Using the equations for the equilibria (12.25), we can simplify the above equations for the perturbations to the following system:

$$\begin{aligned}
x_a + x_t &= -\lambda^*(a)x(a,t) - \tilde{\lambda}(a,t)s^*(a), \\
&\quad -\tilde{\lambda}(a,t)x(a,t) - \mu(a)x(a,t), \\
y_a + y_t &= \lambda^*(a)x(a,t) + \tilde{\lambda}(a,t)s^*(a) + \tilde{\lambda}(a,t)x(a,t), \\
&\quad -(\mu(a) + \gamma(a))y(a,t), \\
p_a + p_t &= -\mu(a)p(a,t), \\
x(0,t) &= \int_0^\infty \beta(a)p(a,t)da, \\
y(0,t) &= 0, \\
u(0,t) &= \int_0^\infty \beta(a)p(a,t)da.
\end{aligned} \tag{12.31}$$

Finally, to obtain a linear system of the above system, we drop the nonlinear quadratic terms:

$$\begin{aligned}
x_a + x_t &= -\lambda^*(a)x(a,t) - \tilde{\lambda}(a,t)s^*(a) - \mu(a)x(a,t), \\
y_a + y_t &= \lambda^*(a)x(a,t) + \tilde{\lambda}(a,t)s^*(a) - (\mu(a) + \gamma(a))y(a,t), \\
p_a + p_t &= -\mu(a)p(a,t), \\
x(0,t) &= \int_0^\infty \beta(a)p(a,t)da, \\
y(0,t) &= 0, \\
p(0,t) &= \int_0^\infty \beta(a)p(a,t)da.
\end{aligned} \tag{12.32}$$

This is the linearized system for the perturbation of an equilibrium. Because this system is linear, it has separable solutions in which the time-dependent function is an exponential, just like the Mckendrick–von Foerster model. We look for those solutions in the form $x(a,t) = e^{\rho t}x(a)$, $y(a,t) = e^{\rho t}y(a)$, and $p(a,t) = e^{\rho t}p(a)$. The a-dependent functions $x(a)$, $y(a)$, and $p(a)$ are different from their a- and t-dependent counterparts, but we adopt that notation for convenience. Substituting the time-dependent solutions with the separable ones, and after canceling $e^{\rho t}$, we arrive at the following system:

$$\begin{aligned}
x_a + \rho x &= -\lambda^*(a)x(a) - \tilde{\lambda}(a)s^*(a) - \mu(a)x(a),\\
y_a + \rho y &= \lambda^*(a)x(a) + \tilde{\lambda}(a)s^*(a) - (\mu(a)+\gamma(a))y(a),\\
p_a + \rho p &= -\mu(a)p(a),\\
x(0) &= \int_0^\infty \beta(a)p(a)da,\\
y(0) &= 0,\\
p(0) &= \int_0^\infty \beta(a)p(a)da.
\end{aligned} \tag{12.33}$$

This is a linear eigenvalue problem, the same one that we obtain for ODEs. We want to find a number ρ and $x(a)$, $y(a)$, $p(a)$ that are nonzero and solve the above problem. In the ODE case, we take the characteristic equation, which is an appropriate determinant set equal to zero. Here we can obtain the characteristic equation by gradually eliminating $x(a)$, $y(a)$, and $p(a)$.

We derive the characteristic equation of the disease-free equilibrium: system (12.33) for the disease-free equilibrium takes the form

$$\begin{aligned}
x_a + \rho x &= -\tilde{\lambda}(a)\phi_0\pi(a) - \mu(a)x(a),\\
y_a + \rho y &= \tilde{\lambda}(a)\phi_0\pi(a) - (\mu(a)+\gamma(a))y(a),\\
p_a + \rho p &= -\mu(a)p(a),\\
x(0) &= \int_0^\infty \beta(a)p(a)da,\\
y(0) &= 0,\\
u(0) &= \int_0^\infty \beta(a)p(a)da.
\end{aligned} \tag{12.34}$$

We recall that

$$\tilde{\lambda}(a) = k_1(a)\int_0^\infty k_2(a)y(a)da = \tilde{\lambda}k_1(a),$$

where $\tilde{\lambda}$ denotes the integral. We note again that $\tilde{\lambda}$ is a number that is not known. We notice that in system (12.34), the equation for y separates from all others. We may solve it using similar techniques as before.

$$\begin{aligned}
y(a) &= \tilde{\lambda}\phi_0 \int_0^a k_1(\eta)\pi(\eta)e^{-\int_\eta^a(\mu(s)+\gamma(s)+\rho)ds}d\eta\\
&= \tilde{\lambda}\phi_0\pi(a)\int_0^a k_1(\eta)e^{-\int_\eta^a(\gamma(s)+\rho)ds}d\eta.
\end{aligned} \tag{12.35}$$

Substituting in the expression for $\tilde{\lambda}$, we obtain

$$\tilde{\lambda} = \tilde{\lambda}\phi_0 \int_0^\infty k_2(a)\pi(a) \int_0^a k_1(\eta) e^{-\int_\eta^a (\gamma(s)+\rho)ds} d\eta da.$$

Next, since $\tilde{\lambda}$ is nonzero, we may cancel it to obtain the following *characteristic equation* for ρ:

$$1 = \phi_0 \int_0^\infty k_2(a)\pi(a) \int_0^a k_1(\eta) e^{-\int_\eta^a (\gamma(s)+\rho)ds} d\eta da. \tag{12.36}$$

This equation has both real and complex solutions for ρ. The right-hand side is a function of ρ, say

$$\mathscr{H}(\rho) = \phi_0 \int_0^\infty k_2(a)\pi(a) \int_0^a k_1(\eta) e^{-\int_\eta^a (\gamma(s)+\rho)ds} d\eta da.$$

We notice that

$$\mathscr{H}(0) = \mathscr{R}_0.$$

For ρ real, $\mathscr{H}(\rho)$ is a decreasing function of ρ if $k_1(a) \geq 0$ is not identically zero. Hence, if we had $\mathscr{R}_0 > 1$, that would imply $\mathscr{H}(0) > 1$. Since

$$\lim_{\rho \to \infty} \mathscr{H}(\rho) = 0,$$

the equation $\mathscr{H}(\rho) = 1$ has a unique real solution $\rho^* > 0$. In this case, the disease-free equilibrium is unstable.

Consider now the case $\mathscr{R}_0 < 1$. In this case, the equation $\mathscr{H}(\rho) = 1$ also has a unique real solution, but $\rho^* < 0$. We show that all remaining complex solutions have real part that is negative. Assume that there is a complex solution $\rho = \xi_1 + i\xi_2$ such that $\xi_1 \geq 0$. Then, using Euler's formula for complex numbers, we have

$$\begin{aligned}
|\mathscr{H}(\rho)| &\leq \phi_0 \int_0^\infty k_2(a)\pi(a) \int_0^a k_1(\eta) e^{-\int_\eta^a \gamma(s)ds} |e^{-\rho(a-\eta)}| d\eta da \\
&= \phi_0 \int_0^\infty k_2(a)\pi(a) \int_0^a k_1(\eta) e^{-\int_\eta^a \gamma(s)ds} e^{-\xi_1(a-\eta)} \\
&\quad \times |\cos(\xi_2(a-\eta)) - i\sin(\xi_2(a-\eta))| d\eta da \\
&\leq \mathscr{H}(\xi_1) \leq \mathscr{H}(0) = \mathscr{R}_0 < 1.
\end{aligned} \tag{12.37}$$

In particular, this means that for every $\rho = \xi_1 + i\xi_2$ with $\xi_1 \geq 0$, we have $|\mathscr{H}(\rho)| < 1$. Therefore, such a ρ cannot be a solution of the equation $\mathscr{H}(\rho) = 1$. Consequently, that equation has complex solutions with negative real part only. This implies that the disease-free equilibrium is locally asymptotically stable in this case. We have established the following theorem:

Theorem 12.1. *If $\mathscr{R}_0 < 1$, the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, the disease-free equilibrium is unstable.*

Once we find a ρ that is a characteristic value, to that ρ corresponds a nonzero $y(a)$ given by (12.35). The values for $x(a)$ and $p(a)$ can also be computed.

Investigating the stability of the endemic equilibrium is rather cumbersome, and we omit it.

12.4 Numerical Methods for Age-Structured Models

Partial differential equation models, such as the ones considered in this section, cannot be solved analytically or even directly by a computer algebra system. One has to use numerical methods to obtain a discrete form of the model, which can then be coded in some programming language. Numerical methods solve a PDE *approximately*, not exactly. The methods used in solving age-structured models can be separated into several classes:

- finite-difference methods for the PDE model;
- the method of lines, in which the age variable is discretize, while the time variable is not, so the PDE system is converted to a large system of ODEs;
- discretization methods applied to the Lotka–Volterra integral form of the model.

12.4.1 A Numerical Method for the McKendrick–von Foerster Model

In this section, we introduce a finite difference method for solving age-structured PDE models. We will demonstrate the method on the McKendrick–von Foerster model (12.6). Finite difference methods approximate the solution at a number of points in the domain. Typical of these methods is that the derivative is approximated by a finite difference quotient. To introduce the method, we begin by discretizing the domain

$$\mathscr{D} = \{(a,t) : 0 \le a \le a_\dagger,\ t \ge 0\}.$$

We cannot compute to infinite time and age, so if $a_\dagger = \infty$, we take a number $A < \infty$ as a maximal age and a number $T < \infty$ as a maximal time. We divide the age direction and the time direction into equally spaced points. Since age and time progress together, it is reasonable to choose the step, that is, the distance between two points, to be the same in both directions. Thus, we take

$$\Delta a = \Delta t.$$

Then all the points in the age and time direction can be computed as

$$a_i = i\Delta t \qquad\qquad t_n = n\Delta t.$$

We set

$$M = \left[\frac{A}{\Delta t}\right], \qquad N = \left[\frac{T}{\Delta t}\right],$$

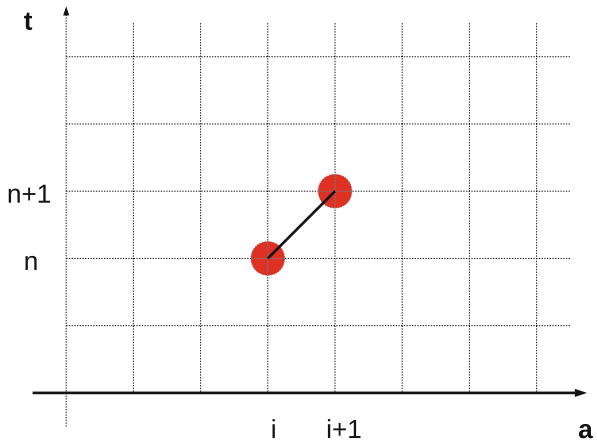

Fig. 12.4 Mesh for the finite difference numerical method. The derivative is computed as a difference of the values of the function along the characteristic line

where $[\cdot]$ denotes the integer part of the expression. We may modify A and T a little so that without loss of generality, we may assume that

$$A = M\Delta t \qquad T = N\Delta t.$$

This discretization generates a square mesh in the domain $\mathscr{D}$. The solution is computed at the points of that mesh $u(a_i, t_n)$. In fact, the numerical method computes approximations to the solution at the mesh points:

$$u_i^n \approx u(a_i, t_n).$$

The numerical method is generated by discretizing the differential equation, the boundary condition, and the initial condition. We evaluate the differential equation at $a = a_{i+1}$ and $t = t_n$. We have

$$\frac{\partial u}{\partial a}(a_{i+1}, t_n) + \frac{\partial u}{\partial a}(a_{i+1}, t_n) = -\mu(a_{i+1})u(a_{i+1}, t_n).$$

We replace the partial derivatives with difference quotients. The derivative in a is replaced by a backward difference quotient, while the derivative in t is replaced by a forward difference quotient:

$$\frac{u_{i+1}^n - u_i^n}{\Delta t} + \frac{u_{i+1}^{n+1} - u_{i+1}^n}{\Delta t} = -\mu_{i+1} u_{i+1}^n.$$

Since when we replace the partial derivatives with the difference quotient we make errors, we use the approximate solution notation u_i^n. We notice that the left-hand side of this equation can be significantly simplified. We obtain (see Fig. 12.4)

$$\frac{u_{i+1}^{n+1} - u_i^n}{\Delta t} = -\mu_{i+1} u_{i+1}^n. \tag{12.38}$$

This gives a discretization of the partial derivatives along the characteristic lines. Thus, the sum of the two partial derivatives is approximated by a difference quotient along the characteristic line, much like the one we used to derive the McKendrick–von Foerster model [see (12.1)]. This is a viable numerical method for the partial differential equation. However, this method is *explicit*, that is, the values of u at time level n are used to compute the values at time level $n+1$, but the right-hand side of the equation is computed at time level n. Explicit methods have various shortcomings, one of which being that the method approximates the exact solution for only very small values of Δt. This is somewhat inconvenient. For that reason, expression (12.38) is reformulated as an implicit method. That requires us to evaluate the right-hand side at level $n+1$. Thus, we obtain the following *finite difference method along the characteristic lines*:

$$\frac{u_{i+1}^{n+1} - u_i^n}{\Delta t} = -\mu_{i+1} u_{i+1}^{n+1}. \tag{12.39}$$

The above equation is a simple linear equation that can be solved for u_{i+1}^{n+1}:

$$u_{i+1}^{n+1} = \frac{u_i^n}{1 + \mu_{i+1}\Delta t}.$$

Hence, if we know all values at time level n, we can compute all values at time level $n+1$ except the first one. To start the process, we must know all values at time level zero. We can compute all values at time level zero from the initial condition. The values of u at time level zero are initialized by

$$u_i^0 = u_0(a_i), \qquad i = 0, \ldots, M.$$

Finally, to approximate the newborns, we discretize the boundary condition. The boundary condition is given by an integral. There are many methods to discretize integrals. However, since the derivative is discretized with a method that is not very accurate, using a highly accurate method (like Simpson's method) for the integral is not necessary. We will approximate the integral with the right endpoint rule. After we have computed all other values at time level n, we compute the newborns at time level n:

$$u_0^n = \sum_{i=1}^{n} \beta_i u_i^n \Delta t,$$

where $\beta(a_i) = \beta_i$. We obtain the following:

Numerical method for the McKendrick–von Foerster model:

$$\begin{cases} u_{i+1}^{n+1} = \dfrac{u_i^n}{1+\mu_{i+1}\Delta t}, & i = 0,\dots,M-1; n = 0,\dots N-1; \\ u_i^0 = u_0(a_i), & i = 0,\dots,M; \\ u_0^n = \sum_{i=1}^{n} \beta_i u_i^n \Delta t & n = 1,\dots,N. \end{cases} \tag{12.40}$$

The method we discuss here has numerous strengths: it preserves the positivity of the solution, it is implicit and therefore approximates the solution for all time steps, it can be solved explicitly (no linear or nonlinear solvers are necessary), and it very easy to code.

12.4.2 Numerical Method for the Age-Structured SIR Model

A numerical method for the age-structured SIR epidemic model (12.21) can be obtained that is similar to the method for the McKendrick–von Forster model. However, there is one significant difference. The right-hand side of the model is nonlinear. Thus, if we evaluate the entire right-hand side at the next time to obtain an implicit method, the resulting equation for the next time level will become nonlinear and very hard to solve. Hence, to obtain an implicit method and a linear equation for the next time level, we partially linearize the nonlinear term. We illustrate these concepts below. As before, we begin by discretizing the domain

$$\mathscr{D} = \{(a,t) : 0 \le a \le a_\dagger,\ t \ge 0\}.$$

We take a number $A < \infty$ as a maximal age (if $a_\dagger = \infty$) and a number $T < \infty$ as a maximal time. We divide the age direction and the time direction into equally spaced points. We discretize the space and time direction with the same step,

$$\Delta a = \Delta t.$$

Then all the points in the age and time direction can be computed as

$$a_i = i\Delta t \qquad t_n = n\Delta t.$$

We set

$$M = \left[\frac{A}{\Delta t}\right], \qquad N = \left[\frac{T}{\Delta t}\right],$$

where $[\cdot]$ denotes the integer part of the expression. We may modify A and T a little so that without loss of generality, we may assume that

$$A = M\Delta t \qquad T = N\Delta t.$$

This discretization generates a square mesh in the domain $\mathscr{D}$.

As before, we replace the sum of partial derivatives by a difference quotient along the characteristic lines. In particular,

$$\begin{aligned} s_a + s_t &\approx \frac{s_{i+1}^{n+1} - s_i^n}{\Delta t}, \\ i_a + i_t &\approx \frac{i_{i+1}^{n+1} - i_i^n}{\Delta t}, \\ r_a + r_t &\approx \frac{r_{i+1}^{n+1} - r_i^n}{\Delta t}. \end{aligned} \tag{12.41}$$

We evaluate the right-hand side at a_{i+1} and t_n and approximate. As a first step, we have

$$\begin{aligned} \frac{s_{i+1}^{n+1} - s_i^n}{\Delta t} &= -\lambda_{i+1}^n s_{i+1}^n - \mu_{i+1} s_{i+1}^n, \\ \frac{i_{i+1}^{n+1} - i_i^n}{\Delta t} &= \lambda_{i+1}^n s_{i+1}^n - (\mu_{i+1} + \gamma_{i+1}) i_{i+1}^n, \\ \frac{r_{i+1}^{n+1} - r_i^n}{\Delta t} &= \gamma_{i+1} i_{i+1}^n - \mu_{i+1} r_{i+1}^n, \end{aligned} \tag{12.42}$$

where λ_{i+1}^n is the force of infection evaluated at time level n. Here we can easily solve for level $n+1$, but the solution will be nonnegative and will approximate the continuous solution only for a very small step Δt. Thus, to make the method implicit, we replace

$$s_{i+1}^n \approx s_{i+1}^{n+1}, \qquad i_{i+1}^n \approx i_{i+1}^{n+1}, \qquad r_{i+1}^n \approx r_{i+1}^{n+1},$$

while keeping λ_{i+1}^n evaluated at level n. We obtain the following system approximating the differential equations:

$$\begin{cases} \dfrac{s_{i+1}^{n+1} - s_i^n}{\Delta t} = -\lambda_{i+1}^n s_{i+1}^{n+1} - \mu_{i+1} s_{i+1}^{n+1}, \\ \dfrac{i_{i+1}^{n+1} - i_i^n}{\Delta t} = \lambda_{i+1}^n s_{i+1}^{n+1} - (\mu_{i+1} + \gamma_{i+1}) i_{i+1}^{n+1} \quad i = 0, \ldots, M-1, \quad n = 0, \ldots, N-1, \\ \dfrac{r_{i+1}^{n+1} - r_i^n}{\Delta t} = \gamma_{i+1} i_{i+1}^{n+1} - \mu_{i+1} r_{i+1}^{n+1}. \end{cases} \tag{12.43}$$

This system is linear in s_{i+1}^{n+1}, i_{i+1}^{n+1}, r_{i+1}^{n+1} and can be easily solved. It is easy to verify that the solutions are nonnegative, and if all equation are added, the equation for the total population size can be obtained. At the same time, the method still possesses the good properties of the implicit schemes. We approximate the integral in the force of infection with a right-hand sum

$$\lambda_{i+1}^n = k_1^{i+1} \sum_{j=1}^{M} k_2^j i_j^n \Delta t.$$

The initial and boundary conditions can be similarly discretized:

$$\begin{cases} s_0^{n+1} = \sum_{j=1}^{M} \beta_j i_j^{n+1} \Delta t & n = 0, \ldots, N-1, \\ i_0^{n+1} = 0, \\ r_0^{n+1} = 0, \\ s_j^0 = s_0(a_j), \quad i_j^0 = i_0(a_j), \quad r_j^0 = r_0(a_j), & j = 0, \ldots, M. \end{cases} \tag{12.44}$$

In conclusion, we would like to point out that this method is convergent, that is, approximations of the solution converge to the true solution as $C\Delta t \to 0$ as $t \to 0$, where C is an appropriate constant. In numerical analysis, in this case often it is said that the numerical method is convergent of order $\mathcal{O}(\Delta t)$. Changing the method for the approximation of the derivative, the method for the approximation of the integrals, or the way we approximated the right-hand side can produce variations of the above method. Any numerical method is acceptable if it is convergent at least of order $\mathcal{O}(\Delta t)$. Methods for proving that a numerical method is convergent can be found in a number of papers that discuss numerical methods for age-structured population models [14, 118].

Problems

12.1. For the age-structured mortality function fitted in Fig. 12.2,

$$\mu(a) = \frac{0.748123a}{(110-a)^2},$$

compute the probability of survival $\pi(a)$.

12.2. Growth Rate of the Population
Consider the McKendrick–von Foerster model (12.6). Assume $\mu(a) = \mu$, a constant. Furthermore assume $\beta(a) = \bar{\beta} e^{-\alpha a}$, where $\bar{\beta}$ and α are constants. Assume $a_\dagger = \infty$.

(a) Compute the growth rate of the population λ^* in terms of μ, $\bar{\beta}$, and α.
(b) Compute the net reproduction rate of the population $\mathcal{R}$ in terms of μ, $\bar{\beta}$, and α.

12.3. Growth Rate of the Population
Consider the McKendrick–von Foerster model (12.6). Assume $\mu(a) = \mu$, a constant. Furthermore, assume

$$\beta(a) = \begin{cases} 0 & 0 \le a < A, \\ \bar{\beta} e^{-\alpha a} & a > A, \end{cases} \tag{12.45}$$

where $\bar{\beta}$ and α are also constants. Assume also $a_\dagger = \infty$.

(a) Compute the growth rate of the population λ^* in terms of μ, $\bar{\beta}$, and α if $\bar{\beta}A = 1$.
(b) Compute the net reproduction rate of the population $\mathcal{R}$ in terms of μ, $\bar{\beta}$, and α.

(c) What is the biological meaning of A? How does A affect the growth rate λ^* and the reproduction rate $\mathscr{R}$ of the population (are they increasing or decreasing functions of A)? Why does that make sense biologically?

12.4. SIR Model with Vertical Transmission

For the SIR model with vertical transmission, assume $a_\dagger = \infty$,

$$\begin{aligned} s_a + s_t &= -\lambda(a,t)s(a,t) - \mu(a)s(a,t), \\ i_a + i_t &= \lambda(a,t)s(a,t) - (\mu(a)+\gamma(a))i(a,t), \\ r_a + r_t &= \gamma(a)i(a,t) - \mu(a)r(a,t), \end{aligned} \tag{12.46}$$

with initial and boundary conditions:

$$\begin{aligned} s(0,t) &= \int_0^{a_\dagger} \beta(a)[s(a,t)+r(a,t)+(1-q)i(a,t)]da, \\ i(0,t) &= q\int_0^{a_\dagger} \beta(a)i(a,t)da, \\ r(0,t) &= 0. \end{aligned} \tag{12.47}$$

(a) Compute the disease-free equilibrium and the disease reproduction number of the model.
(b) Linearize around the disease-free equilibrium and show that the disease-free equilibrium is locally asymptotically stable if $\mathscr{R}_0 < 1$ and unstable if $\mathscr{R}_0 > 1$.

12.5. SIS Model

For the SIS model, assume $a_\dagger = \infty$,

$$\begin{aligned} s_a + s_t &= -\lambda(a,t)s(a,t) - \mu(a)s(a,t) + \gamma(a)i(a,t), \\ i_a + i_t &= \lambda(a,t)s(a,t) - (\mu(a)+\gamma(a))i(a,t), \end{aligned} \tag{12.48}$$

with initial and boundary conditions

$$\begin{aligned} s(0,t) &= \int_0^{a_\dagger} \beta(a)[s(a,t)+i(a,t)]da, \\ i(0,t) &= 0. \end{aligned} \tag{12.49}$$

(a) Compute the disease-free equilibrium and the disease reproduction number of the model.
(b) Linearize around the disease-free equilibrium and show that the disease-free equilibrium is locally asymptotically stable if $\mathscr{R}_0 < 1$ and unstable if $\mathscr{R}_0 > 1$.

12.6. Age-Structured SIR Vector-Borne Model

For the SIR model of a vector-borne disease, assume $a_\dagger = \infty$. The dynamics of the vector population are given by

$$\begin{aligned} V_S'(t) &= \Lambda_v - \lambda_H(t)V_S(t) - dV_S, \\ V_I'(t) &= \lambda_H(t)V_S(t) - dV_I(t), \end{aligned} \tag{12.50}$$

where $\lambda_H(t)$ is the force of infection generated by humans and is given by

$$\lambda_H = \int_0^\infty k_2(a)i(a,t)da.$$

The equations for the humans are given by

$$\begin{aligned} s_a + s_t &= -\lambda_v(a,t)s(a,t) - \mu(a)s(a,t), \\ i_a + i_t &= \lambda_v(a,t)s(a,t) - (\mu(a)+\gamma(a))i(a,t), \\ r_a + r_t &= \gamma(a)i(a,t) - \mu(a)r(a,t), \end{aligned} \tag{12.51}$$

where

$$\lambda_v(a,t) = k_1(a)V_I(t)$$

with initial and boundary conditions

$$\begin{aligned} s(0,t) &= \int_0^{a_\dagger} \beta(a)[s(a,t)+r(a,t)+i(a,t)]da, \\ i(0,t) &= 0, \\ r(0,t) &= 0. \end{aligned} \tag{12.52}$$

(a) Compute the disease-free equilibrium and the disease reproduction number of the model.
(b) Linearize around the disease-free equilibrium and show that the disease-free equilibrium is locally asymptotically stable if $\mathscr{R}_0 < 1$ and unstable if $\mathscr{R}_0 > 1$.

12.7. For the age-structured model of a vector-borne disease in Problem 12.6:

(a) Devise a numerical method to solve the problem. Use a Runge–Kutta (Runge–Kutta methods can be found in standard numerical analysis textbooks) method to discretize the equations for the vector.
(b) Show that the solution of your numerical scheme is nonnegative.
(c) Write a code for your numerical method and simulate the solution.

12.8. SIR Model with Two Strains

For the SIR model with two strains, assume $a_\dagger = \infty$. The model is introduced as follows:

$$\begin{aligned} s_a + s_t &= -\lambda_1(t)s(a,t) - \lambda_2(t)s(a,t) - \mu s(a,t), \\ i_a^1 + i_t^1 &= \lambda_1(t)s(a,t) - (\mu+\gamma_1)i^1(a,t), \\ i_a^2 + i_t^2 &= \lambda_2(t)s(a,t) - (\mu+\gamma_2(a))i^2(a,t), \\ r_a + r_t &= \gamma_1 i^1(a,t) + \gamma_2(a)i^2(a,t) - \mu r(a,t), \end{aligned} \tag{12.53}$$

where

$$\gamma_2(a) = \begin{cases} 0 & a < A \\ \gamma_2 & a > A \end{cases} \tag{12.54}$$

and the remaining parameters are assumed constant. The force of infections are given by

$$\lambda_1(t) = k_1 \int_0^\infty i^1(a,t)da, \qquad \lambda_1(t) = k_2 \int_0^\infty i^2(a,t)da.$$

Furthermore, the model is equipped with initial and the following boundary conditions:

$$\begin{aligned} s(0,t) &= \int_0^{a_\dagger} \beta(a)[s(a,t) + r(a,t) + i^1(a,t) + i^2(a,t)]da, \\ i^1(0,t) &= 0, \\ i^2(0,t) &= 0, \\ r(0,t) &= 0. \end{aligned} \tag{12.55}$$

(a) Compute the disease-free equilibrium and the disease reproduction number of the model.
(b) Write a numerical method to approximate the solution of the problem.
(c) Show that your solution of the numerical scheme is nonnegative.
(d) Write a computer program to simulate the model.

Chapter 13
Class-Age Structured Epidemic Models

13.1 Variability of Infectivity with Time-Since-Infection

Many diseases progress quickly. Once infected, an individual goes through a short incubation period and becomes infectious. In a matter of days, this individual has either recovered or is dead. That is the case with influenza, which many have experienced. Other diseases with short span are SARS, meningitis, plague, and many of the childhood diseases. We will call such diseases *quickly progressive diseases*. See Table 13.1 for a more extensive list of quickly progressive diseases. In modeling such diseases, it is acceptable to ignore host vital dynamics and to assume that the infectivity of infectious individuals is constant throughout their infectious period.

Other diseases infect their hosts for a long time, sometimes for the duration of the lifespan of the host. Examples of such diseases include HIV/AIDS, tuberculosis, and hepatitis C. These diseases necessarily include a long-term latent or chronic stage. Such diseases are called *slowly progressive diseases*. Table 13.1 contains a list of slowly progressive diseases. Models of slowly progressive diseases should include host demography.

Evidence exists that the infectivity, that is, the probability of infection, given a contact, may vary in time since the moment at which the infectious individual has become infected. This variability exists with fast diseases but is far more important with slow diseases. Infectivity for several common diseases is plotted in Fig. 13.1. The problem of variability of infectivity with time-since-infection has been studied most extensively in HIV. The *California Partners' Study* examined 212 females having regular sexual contacts with their HIV-infected male partners. Couples were followed for different durations (duration of exposure) up to 100 months. All partners were already infected before the contact began. Only about 20% of the females were eventually infected. Shiboski and Jewell [145] use the data to estimate a time-since-infection-dependent infectivity. No explicit form of the function is given. Generally, it is accepted that the viral load in HIV-infected patients is correlated with their infectivity. Since the viral load is high right after infection, and then during the time when AIDS develops, the infectivity is assumed to be higher for those two periods,

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_13

Table 13.1 Slow and fast diseases

Slow diseases	Length of infection	Fast diseases	Length of infection
HIV/AIDS	Lifelong	Influenza	2–10 days
Hepatitis C	Lifelong[a]	Measles	10–12 days
Tuberculosis	Lifelong[a]	Mumps	12–25 days
Genital herpes	Lifelong	Rubella	3–4 weeks
Hepatitis B	Lifelong	Chicken pox	17–30 days
Cervical cancer (HPV)	Lifelong	Dengue fever	10–30 days
Malaria	200[a] days	Ebola	3–6 weeks

[a]If not treated

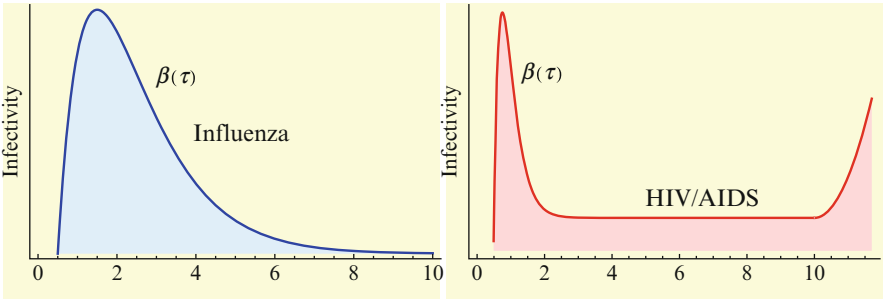

Fig. 13.1 Infectivity for the fast disease influenza [35] and the slow disease HIV [42]

and generally lower during the latent stage of the infection. Shiboski and Jewell's functions do not possess these properties, presumably because infected individuals were already past their acute stage when they were enrolled in the study. Shiboski and Jewell's functions first increase from 0 to 40 months after infection, and then rapidly decrease, so that they are nearly zero at 90 months after infection.

In fact, the very first epidemic model developed by Kermack and McKendrick [84] structures the infected individuals by the time-since-infection (also called *age of infection*). Kermack and McKendrick's motivation for inclusion of infection-age was not only that infectivity may change with infection-age but also that the possibility of recovery or death may depend on the time elapsed since infection. In modeling with ODEs, it is implicitly assumed that the time to recovery or death is exponentially distributed. This assumption may be weakened if infection-age is incorporated into the model. Although Kermack-McKendrick's age-since-infection model did not include birth and natural death in the population, more recent age-since-infection models of slowly progressive diseases include demography.

13.2 Time-Since-Infection Structured SIR Model

In this section, we consider a continuous version of the Kermack–McKendrick time-since-infection structured model. Because mass action incidence is used, the model can be used to describe diseases such as influenza and childhood diseases, but it is not suitable for HIV.

13.2.1 Derivation of the Time-Since-Infection Structured Model

Since infectivity for infectious individuals varies with the time since infection (see Fig. 13.1), we must keep track of the time that has elapsed since infection for each infected individual. Let τ denote the time-since-infection. The time-since-infection begins when an individual becomes infected and progresses with the chronological time t. Let $i(0,t)$ denote the number of individuals who have just become infected at time t. All individuals who become simultaneously infected make up one disease "cohort," that is, they have experienced the same life event together, namely getting infected with the disease. As time progresses, this group of people has the same time-since-infection τ. Let $i(\tau,t)$ be the density of infected individuals with time-since-infection τ at time t. The fact that $i(\tau,t)$ is a density means that $i(\tau,t)\Delta\tau$ is the number of individuals with time-since-infection in the interval $(\tau,\tau+\Delta\tau)$. Suppose that time Δt elapses. Then the same group of individuals who at time t had time-since-infection in the interval $(\tau,\tau+\Delta\tau)$, now at time $t+\Delta t$ have time-since infection in the interval $(\tau+\Delta t,\tau+\Delta\tau+\Delta t)$. The number of those individuals is given by $i(\tau+\Delta t,t+\Delta t)\Delta\tau$. Adding all individuals in all infection-age classes gives the total infected population:

$$I(t)=\int_0^\infty i(\tau,t)\,d\tau.$$

Since this is the same group of individuals, their numbers might have changed in the interval $(t,t+\Delta t)$ as a result of two possible events: some of them might have recovered, and other might have left the system due to natural causes (e.g., death). We assume that the lifespan of individuals in the system is exponentially distributed, and equal for susceptible, infected, and recovered individuals. Thus, susceptible, infected, and recovered individuals leave the system at a constant rate. Denote by μ the per capita rate at which individuals leave the system. The number of infected individuals who leave the system in the time interval $(t,t+\Delta t)$ is given by

$$\mu\Delta t i(\tau,t)\Delta\tau,$$

where $i(\tau,t)\Delta\tau$ is the number of people in the age interval $(\tau,\tau+\Delta\tau)$, and $\mu i(\tau,t)\Delta\tau$ is the number of people in that age interval who leave the system at time t. To model the number of individuals who recover in the time interval $(t,t+\Delta t)$, we denote the per capita recovery rate by γ. We will assume that the recovery rate

depends on the time-since-infection τ: $\gamma(\tau)$. Following a similar line of reasoning as in the case with the number of individuals who leave the system, the number of individuals who recover from this cohort of infecteds is given by

$$\gamma(\tau)\Delta t i(\tau,t)\Delta\tau.$$

The balance equation for the change in the number of infected individuals in this cohort is given by

$$i(\tau+\Delta t,t+\Delta t)\Delta\tau - i(\tau,t)\Delta\tau = -\gamma(\tau)\Delta t i(\tau,t)\Delta\tau - \mu\Delta t i(\tau,t)\Delta\tau.$$

Dividing by $\Delta\tau\Delta t$, we obtain

$$\frac{i(\tau+\Delta t,t+\Delta t)-i(\tau,t)}{\Delta t} = -\gamma(\tau)i(\tau,t) - \mu i(\tau,t). \tag{13.1}$$

We rewrite the left-hand side above as

$$\frac{i(\tau+\Delta t,t+\Delta t)-i(\tau,t+\Delta t)}{\Delta t} + \frac{i(\tau,t+\Delta t)-i(\tau,t)}{\Delta t} = -\gamma(\tau)i(\tau,t) - \mu i(\tau,t). \tag{13.2}$$

We take the limit as $\Delta t \to 0$. If the partial derivatives of the function i exist and are continuous, we can rewrite the equation above in the form

$$i_\tau(\tau,t) + i_t(\tau,t) = -\gamma(\tau)i(\tau,t) - \mu i(\tau,t). \tag{13.3}$$

This is a first-order partial differential equation. It is linear. It is defined on the domain

$$\mathscr{D} = \{(\tau,t) : \tau \geq 0, t \geq 0\}.$$

To complete the partial differential equation, we must derive a boundary condition along the boundary $\tau = 0$ and an initial condition.

To derive the boundary condition, let $S(t)$ be the number of susceptible individuals at time t, $R(t)$ the number of recovered individuals, and $N(t)$ the total population size:

$$N(t) = S(t) + \int_0^\infty i(\tau,t)\,d\tau + R(t).$$

The newly infected individuals have time-since-infection equal to zero and their number is given by $i(0,t)$. To derive the expression for newly infected individuals, we let $\beta(\tau)N$ be the infectivity of the infectious individuals, where N is the total population. The infectivity depends on the time-since-infection τ that has elapsed for the infecting individual. It is assumed that infectious individuals have different infectivities at different times-since-infection. This is the case with most infectious diseases. The probability that an infectious individual with time-since-infection equal to τ will come in a contact with a susceptible individual, given that the individual makes a contact, is $\frac{S}{N}$. Thus, this infectious individual with time-since-infection equal to τ will transmit the disease to

$$\beta(\tau)N\frac{S}{N} = \beta(\tau)S$$

individuals. Since there are $i(\tau,t)\Delta\tau$ infectious individuals with time-since-infection in the interval $(\tau, \tau+\Delta\tau)$, the total number of infections generated by such infectious individuals will be

$$\beta(\tau)Si(\tau,t)\Delta\tau.$$

Adding all newly infected individuals generated by all infected individuals in all time-since-infection classes, we get

$$i(0,t) = S\int_0^\infty \beta(\tau)i(\tau,t)\,d\tau.$$

This incidence is the equivalent of mass action incidence in the ODE case. This equation gives the *boundary condition* of the partial differential equation. To derive the equation that gives the dynamics of the susceptible individuals, we assume that the recruitment into the population occurs at a constant rate Λ. Thus, the equation becomes

$$S'(t) = \Lambda - S\int_0^\infty \beta(\tau)i(\tau,t)\,d\tau - \mu S(t).$$

Finally, the equation for the recovered individuals has as an inflow the total number of recovered individuals summed by all age-since-infection classes:

$$R'(t) = \int_0^\infty \gamma(\tau)i(\tau,t)d\tau - \mu R(t).$$

The susceptible, infected, and recovered populations make up the total population. The equations for the susceptible, infected, and recovered populations define a closed system of equations, which we will consider in itself:

$$\begin{cases} S'(t) = \Lambda - S(t)\displaystyle\int_0^\infty \beta(\tau)i(\tau,t)\,d\tau - \mu S(t), \\ i_\tau(\tau,t) + i_t(\tau,t) = -\gamma(\tau)i(\tau,t) - \mu i(\tau,t), \\ i(0,t) = S(t)\displaystyle\int_0^\infty \beta(\tau)i(\tau,t)\,d\tau, \\ R'(t) = \displaystyle\int_0^\infty \gamma(\tau)i(\tau,t)d\tau - \mu R(t). \end{cases} \tag{13.4}$$

The model is equipped with the following initial conditions:

$$\begin{aligned} S(0) &= S_0, \\ i(\tau,0) &= i_0(\tau), \\ R(0) &= R_0, \end{aligned} \tag{13.5}$$

where S_0 and R_0 are given numbers, and $i_0(\tau)$ is a given function that is assumed integrable. We note that

$$N_0 = S_0 + \int_0^\infty i_0(\tau)d\tau + R_0.$$

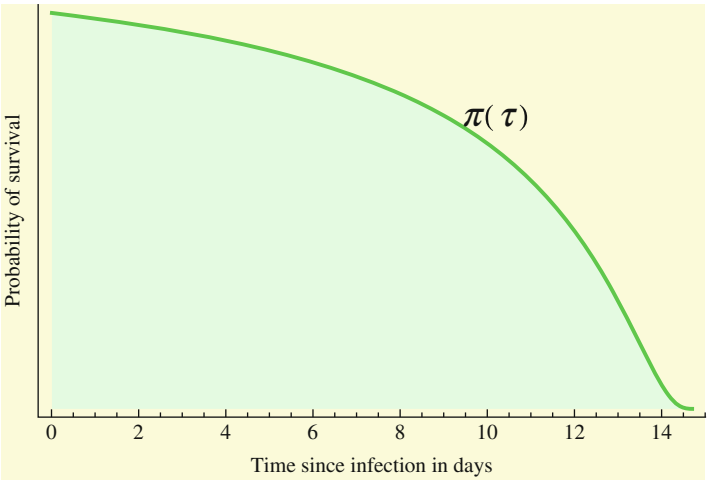

Fig. 13.2 A typical probability of survival in a class

Model (13.4) together with the initial conditions (13.5) is the time-since-infection structured Kermack–McKendrick SIR epidemic model.

Remark 13.1. Typically, the infectivity $\beta(\tau)$ is assumed to be a bounded function:

$$\bar{\beta} = \sup_\tau \beta(\tau).$$

A key quantity related to the survival of infectious individuals in a given class is $\pi(\tau)$, the probability of still being infectious τ time units after becoming infected. Then, if $\hat{I}$ individuals become infected at some moment of time, the number of those who are still infectious after τ time units is $\hat{I}\pi(\tau)$. Those numbers change in a small interval of time-since-infection $\Delta\tau$ by those who have stopped being infectious or those who have left the system:

$$\hat{I}\pi(\tau+\Delta\tau) - \hat{I}\pi(\tau) = -\mu\hat{I}\pi(\tau)\Delta\tau - \gamma(\tau)\hat{I}\pi(\tau)\Delta\tau.$$

The probability of still being infectious τ time units after becoming infected/infectious, $\pi(\tau)$, satisfies the following differential equation:

$$\pi'(\tau) = -\mu\pi(\tau) - \gamma(\tau)\pi(\tau),$$

whose solution is

$$\pi(\tau) = e^{-\mu\tau - \int_0^\tau \gamma(s)\,ds}.$$

Different assumptions on $\gamma(\tau)$ correspond to different real-life scenarios. If all infected individuals are assumed to recover or leave the infectious class by certain age-since-infection $\bar{\tau}$, their probability $\pi(\tau)$ of still being infectious, that is, of being in the class i, τ time units after becoming infected/infectious must tend to zero as $\tau \to \bar{\tau}^-$. This will occur if the function $\gamma(\tau)$ tends to ∞ as $\tau \to \bar{\tau}^-$. Thus, we may assume that

$$\gamma(\tau) \to \infty \qquad \text{as} \qquad \tau \to \bar{\tau}^-; \qquad\qquad \gamma(\tau) = 0 \qquad \text{for} \qquad \tau > \bar{\tau}.$$

A simple example of a possible function $\gamma(\tau)$ that tends to infinity is given by

$$\gamma(\tau) = \frac{0.2\bar{\tau}}{(\bar{\tau} - \tau)^2}.$$

The corresponding probability of survival in the infectious class when $\mu = 0$ is given by

$$\pi(\tau) = e^{-\dfrac{0.2\tau}{\bar{\tau} - \tau}}.$$

This probability of survival is graphed in Fig. 13.2 with $\bar{\tau} = 15$. The coefficient 0.2 is used to give the typical shape of the graph characterized by slow decrease for small τ and fast decrease for $\tau \approx \bar{\tau}$.

13.2.2 Equilibria and Reproduction Number of the Time-Since-Infection SIR Model

The model (13.4) is a first-order integrodifferential equation model. We would like to be able to say something about the solutions. Could a reproduction number $\mathscr{R}_0$ be defined such that the disease dies out if $\mathscr{R}_0 < 1$ and persists otherwise? First, one has to show that for each nonnegative and integrable initial condition (13.5), the model has a unique nonnegative solution. This result is not obvious, but the derivation is somewhat technical and will not be included. Next, we would like to see that the solutions are bounded. To see this, we must obtain the differential equation satisfied by the total population size. Integrating with respect to τ the PDE in system (13.4), we obtain

$$i(\tau,t)|_0^\infty + I' = -\int_0^\infty \gamma(\tau) i(\tau,t) d\tau - \mu I(t),$$

where I' above is the derivative of the total infected population size I with respect to t. If we assume $\lim_{\tau\to\infty} i(\tau,t) = 0$, the above equality leads to

$$I'(t) = S(t)\int_0^\infty \beta(\tau) i(\tau,t)\,d\tau - \int_0^\infty \gamma(\tau) i(\tau,t) d\tau - \mu I(t).$$

Adding the equation above to the equations for S' and R' from (13.4), we obtain

$$S' + I' + R' = \Lambda - \mu(S + I + R).$$

Thus, the total population size satisfies the usual equation, whose solution we know. In particular, we know that

$$\max N(t) \leq \max\left\{N_0, \frac{\Lambda}{\mu}\right\}.$$

Now we consider equilibria of the model. As before, to find the equilibria, we look for time-independent solutions $(S, i(\tau))$ that satisfy the system (13.4) with the time derivatives equal to zero. The system for the equilibria takes the form

$$\begin{cases} \Lambda - S\displaystyle\int_0^\infty \beta(\tau) i(\tau)\, d\tau - \mu S = 0, \\ i_\tau(\tau) = -\gamma(\tau) i(\tau) - \mu i(\tau), \\ i(0) = S\displaystyle\int_0^\infty \beta(\tau) i(\tau)\, d\tau, \\ \displaystyle\int_0^\infty \gamma(\tau) i(\tau) d\tau - \mu R = 0. \end{cases} \tag{13.6}$$

This system consists of one first-order ODE with initial condition that depends on the solution, and two algebraic equations. Clearly, $\mathscr{E}_0 = (\frac{\Lambda}{\mu}, 0, 0)$ is one solution of that system. This solution gives the *disease-free equilibrium*, where the age-since-infection distribution of infectious individuals is identically zero. The disease-free equilibrium always exists. An *endemic equilibrium* will be given by a nontrivial solution $\mathscr{E}^* = (S, i(\tau), R)$.

There is a typical approach for solving such systems. We first solve the differential equation whose solution is

$$i(\tau) = i(0)\pi(\tau). \tag{13.7}$$

This is not an explicit solution, since $i(0)$ depends on $i(\tau)$. The following notation is useful:

$$P = \int_0^\infty \pi(\theta)\, d\theta.$$

This notation occurs when we compute the total infectious population:

$$I = i(0)P.$$

The usual approach to solving the system (13.6) is to substitute the expression for $i(\tau)$ from (13.7) in the boundary condition and the total population size. We typically obtain a system for $i(0)$ and S. However, in this case, when we substitute $i(\tau)$ in the boundary condition, we obtain an explicit expression for the susceptible individuals in the endemic equilibrium:

$$S = \frac{1}{\int_0^\infty \beta(\tau)\pi(\tau)\,d\tau}. \tag{13.8}$$

From the third equation in (13.6), we can express R in terms of $i(0)$:

$$R = \frac{i(0)}{\mu}\int_0^\infty \gamma(\tau)\pi(\tau)d\tau = \frac{i(0)}{\mu}\Gamma. \tag{13.9}$$

We note that Γ is a given number. To find $i(0)$, we use the first equation in (13.6), which becomes

$$\Lambda - i(0) - \mu S = 0.$$

Substituting S, we obtain

$$i(0) = \Lambda\left(1 - \frac{1}{\mathscr{R}_0}\right), \tag{13.10}$$

where we have defined the basic reproduction number as

$$\mathscr{R}_0 = \frac{\Lambda}{\mu}\int_0^\infty \beta(\tau)\pi(\tau)\,d\tau.$$

From the above expressions, we see that the endemic equilibrium is unique and exists if and only if $\mathscr{R}_0 > 1$.

Remark 13.2. We notice that integration by parts gives the following identity:

$$\int_0^\infty \gamma(\tau)\pi(\tau)\,d\tau + \mu\int_0^\infty \pi(\tau)\,d\tau = 1,$$

which makes each term on the left-hand side less than one. The equation above says that the probability of leaving the infectious class i through leaving the system (dying), $\mu\int_0^\infty \pi(\tau)\,d\tau$, or through recovery, $\int_0^\infty \gamma(\tau)\pi(\tau)\,d\tau$, is equal to 1. Indeed, all individuals leave the infectious class through one of those two routes.

13.2.3 Local Stability of Equilibria

To investigate the local stability of the equilibria, we need to linearize the system. For a PDE model, that is done directly following the underlying linearization procedure. In particular, let $S(t) = S^* + x(t)$, $i(\tau,t) = i^*(\tau) + y(\tau,t)$ and $R(t) = R^* + z(t)$, where $x(t)$, $y(\tau,t)$, and $z(t)$ are the perturbations, and $(S^*, i^*(\tau), R^*)$ denotes a generic equilibrium. We substitute the expressions for S, $i(\tau,t)$, and R in the system (13.4):

$$\begin{cases} (S^*+x(t))' = \Lambda - (S^*+x(t))\displaystyle\int_0^\infty \beta(\tau)(i^*(\tau)+y(\tau,t))\,d\tau - \mu(S^*+x(t)), \\ (i^*(\tau)+y(\tau,t))_\tau + (i^*(\tau)+y(\tau,t))_t = -\gamma(\tau)(i^*(\tau)+y(\tau,t)) - \mu(i^*(\tau)+y(\tau,t)), \\ i^*(0)+y(0,t) = (S^*+x(t))\displaystyle\int_0^\infty \beta(\tau)(i^*(\tau)+y(\tau,t))\,d\tau, \\ (R^*+z(t))' = \displaystyle\int_0^\infty \gamma(\tau)(i^*(\tau)+y(\tau,t))d\tau - \mu(R^*+z(t)). \end{cases} \tag{13.11}$$

Multiplying out the expressions, we have

$$\begin{cases} x'(t) = \Lambda - S^*\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)d\tau - \mu S^* - x(t)\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau \\ \qquad\qquad -S^*\displaystyle\int_0^\infty \beta(\tau)y(\tau,t)\,d\tau - x(t)\int_0^\infty \beta(\tau)y(\tau,t)\,d\tau - \mu x(t), \\ i^*_\tau(\tau) + y_\tau(\tau,t) + y_t(\tau,t) = -\gamma(\tau)i^*(\tau) - \gamma(\tau)y(\tau,t) - \mu i^*(\tau) - \mu y(\tau,t), \\ i^*(0)+y(0,t) = S^*\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau + x(t)\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau \\ \qquad\qquad +S^*\displaystyle\int_0^\infty \beta(\tau)y(\tau,t))\,d\tau + x(t)\int_0^\infty \beta(\tau)y(\tau,t)\,d\tau. \\ z'(t) = \displaystyle\int_0^\infty \gamma(\tau)i^*(\tau)d\tau + \int_0^\infty \gamma(\tau)y(\tau,t)d\tau - \mu R^* - \mu z(t). \end{cases} \tag{13.12}$$

This system can be simplified further by the use of two techniques. First, we use the equations for the equilibria (13.6). This approach simplifies the system to

$$\begin{cases} x'(t) = -x(t)\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau \\ \qquad -S^*\displaystyle\int_0^\infty \beta(\tau)y(\tau,t)\,d\tau - x(t)\int_0^\infty \beta(\tau)y(\tau,t)\,d\tau - \mu x(t), \\ y_\tau(\tau,t) + y_t(\tau,t) = -\gamma(\tau)y(\tau,t) - \mu y(\tau,t), \\ y(0,t) = x(t)\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau \\ \qquad\qquad +S^*\displaystyle\int_0^\infty \beta(\tau)y(\tau,t))\,d\tau + x(t)\int_0^\infty \beta(\tau)y(\tau,t)\,d\tau, \\ z'(t) = \displaystyle\int_0^\infty \gamma(\tau)y(\tau,t)d\tau - \mu z(t). \end{cases} \tag{13.13}$$

Notice that after this transformation, system (13.13) contains only terms that include a perturbation. However, system (13.13) is not linear. Terms such as $x(t)\int_0^\infty \beta(\tau)$ $y(\tau,t)\,d\tau$ are quadratic in the perturbations. Since we assume that the perturbations are small, the quadratic terms must be much smaller. Therefore, the second technique that we use to simplify the system is to neglect the quadratic terms. After disregarding the quadratic terms, we obtain the following linear system in the perturbations:

$$\begin{cases} x'(t) = -x(t)\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau - S^*\int_0^\infty \beta(\tau)y(\tau,t)\,d\tau - \mu x(t), \\ y_\tau(\tau,t) + y_t(\tau,t) = -\gamma(\tau)y(\tau,t) - \mu y(\tau,t), \\ y(0,t) = x(t)\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau + S^*\int_0^\infty \beta(\tau)y(\tau,t))\,d\tau, \\ z'(t) = \displaystyle\int_0^\infty \gamma(\tau)y(\tau,t)d\tau - \mu z(t). \end{cases} \tag{13.14}$$

System (13.14) is a linear system for $x(t)$, $y(\tau,t)$, and $z(t)$. Just like linear systems of ODEs, the above system also has exponential solutions. Therefore, it is sensible to look for solutions of the form $x(t) = \bar{x}e^{\lambda t}$, $y(\tau,t) = \bar{y}(\tau)e^{\lambda t}$, $z(t) = \bar{z}e^{\lambda t}$, where $\bar{x}$, $\bar{y}(\tau)$, $\bar{z}$, and λ have to be determined in such a way that $\bar{x}$, $\bar{y}(\tau)$, $\bar{z}$ are not all zero. Substituting the constitutive form of the solutions in the system (13.14), we obtain the following problem for $\bar{x}$, $\bar{y}(\tau)$, $\bar{z}$, and λ (the bars have been omitted):

$$\begin{cases} \lambda x = -x\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau - S^*\int_0^\infty \beta(\tau)y(\tau)\,d\tau - \mu x, \\ y_\tau(\tau) + \lambda y(\tau) = -\gamma(\tau)y(\tau) - \mu y(\tau), \\ y(0) = x\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau + S^*\int_0^\infty \beta(\tau)y(\tau)\,d\tau, \\ \lambda z = \displaystyle\int_0^\infty \gamma(\tau)y(\tau)d\tau - \mu z. \end{cases} \tag{13.15}$$

Remark 13.3. Solutions of system (13.15) give the eigenvectors and eigenvalues λ of the differential operator. Eigenvalues are the only points in the spectrum of operators generated by ODEs. However, operators that originate from PDEs may have other points in the spectrum besides eigenvalues, which also contribute to the stability or instability of an equilibrium. It can be shown [112] that for the problems of type (13.4), knowing the distribution of the eigenvalues is sufficient to determine the stability of a given equilibrium. In other words, we have the same rules that are used in ODEs. In particular, if all eigenvalues have negative real parts, the corresponding equilibrium is locally stable; if there is an eigenvalue with a positive real part, then the equilibrium is unstable. Because of that, we will concentrate on investigating eigenvalues.

The next step will be to eliminate x, $y(\tau)$, and z so that an equation in λ is obtained. This process is different for the different equilibria, so we have to consider two cases. The first case is that of the disease-free equilibrium. Then $S^* = \frac{\Lambda}{\mu}$, $i^* = 0$, and $R^* = 0$. System (13.15) simplifies to the following system:

$$\begin{cases} \lambda x = -S^*\displaystyle\int_0^\infty \beta(\tau)y(\tau)\,d\tau - \mu x, \\ y_\tau(\tau) + \lambda y(\tau) = -\gamma(\tau)y(\tau) - \mu y(\tau), \\ y(0) = S^*\displaystyle\int_0^\infty \beta(\tau)y(\tau)\,d\tau, \\ \lambda z = \displaystyle\int_0^\infty \gamma(\tau)y(\tau)d\tau - \mu z. \end{cases} \tag{13.16}$$

It is easy to see that the equation for $y(\tau)$ is independent of x and z. Solving the differential equation, we have

$$y(\tau) = y(0)e^{-\lambda\tau}\pi(\tau).$$

Substituting this solution in the boundary condition and canceling $y(0)$ (assumed nonzero), we obtain the following *characteristic equation* for λ:

$$S^* \int_0^\infty \beta(\tau)e^{-\lambda\tau}\pi(\tau)\,d\tau = 1. \tag{13.17}$$

The above equation is a *transcendental equation*, and it may have many solutions. To show stability of the disease-free equilibrium, we need to show that all solutions λ of the above equation have negative real parts. If there is a solution λ with positive real part, then the disease-free equilibrium will be unstable. To investigate this, we define

$$\mathcal{G}(\lambda) = S^* \int_0^\infty \beta(\tau)e^{-\lambda\tau}\pi(\tau)\,d\tau.$$

We first notice that $\mathcal{G}(0) = \mathcal{R}_0$. Hence, if $\mathcal{R}_0 > 1$, and if $\beta(\tau)$ is strictly positive on a positive interval, then the function $\mathcal{G}(\lambda)$ as a function of the real variable λ is a decreasing function. Since $\mathcal{G}(0) > 1$ and $\lim_{\lambda\to\infty} G(\lambda) = 0$, there exists $\lambda^* > 0$ such that $\mathcal{G}(\lambda^*) = 1$. Consequently, the disease-free equilibrium is unstable. If, alternatively, $\mathcal{R}_0 < 1$, then for all $\lambda = a + bi$ with $a \geq 0$, we have

$$|\mathcal{G}(\lambda)| \leq S^* \int_0^\infty \beta(\tau)|e^{-\lambda\tau}|\pi(\tau)\,d\tau = S^* \int_0^\infty \beta(\tau)e^{-a\tau}\pi(\tau)\,d\tau \leq \mathcal{R}_0 < 1.$$

We conclude that those λ whose real part is nonnegative cannot satisfy the equation $\mathcal{G}(\lambda) = 1$. Therefore, the disease-free equilibrium is locally asymptotically stable in this case. We summarize these results in the following proposition:

Proposition 13.1. *If $\mathcal{R}_0 < 1$, the disease-free equilibrium is locally stable. If $\mathcal{R}_0 > 1$, then the disease-free equilibrium is unstable.*

We see that we obtain similar results as for ODE epidemiological models.

Now we turn to the stability of the endemic equilibrium. We consider system (13.15), where the equilibrium is the endemic equilibrium. The goal again is to eliminate x, $y(\tau)$, and z, but this time the first and the second equations are coupled. We can neglect the equation for z, since z does not participate in the first two equations. Because the differential equation depends on y and λ only, we can solve it and replace $y(\tau)$ by its expression in the boundary condition and in the equation for x. That will produce a linear system for the numbers $y(0)$ and x (assuming that λ is given). Namely, solving the differential equation, we get

$$y(\tau) = y(0)e^{-\lambda\tau}\pi(\tau).$$

From the first equation and the boundary equation, we obtain the system

$$\begin{cases} \lambda x = -x\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau - S^*y(0)\int_0^\infty \beta(\tau)e^{-\lambda\tau}\pi(\tau)\,d\tau - \mu x, \\ y(0) = x\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau + S^*y(0)\int_0^\infty \beta(\tau)e^{-\lambda\tau}\pi(\tau)\,d\tau. \end{cases} \tag{13.18}$$

There are many ways to solve this system in order to find a nontrivial solution. One way is to require that the determinant be zero:

$$\begin{vmatrix} \lambda+\mu+\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau & S^*\displaystyle\int_0^\infty \beta(\tau)e^{-\lambda\tau}\pi(\tau)\,d\tau \\ -\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau & 1-S^*\displaystyle\int_0^\infty \beta(\tau)e^{-\lambda\tau}\pi(\tau)\,d\tau \end{vmatrix} = 0. \tag{13.19}$$

Adding the second row to the first, we obtain

$$\begin{vmatrix} \lambda+\mu & 1 \\ -\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau & 1-S^*\displaystyle\int_0^\infty \beta(\tau)e^{-\lambda\tau}\pi(\tau)\,d\tau \end{vmatrix} = 0. \tag{13.20}$$

We notice that $\displaystyle\int_0^\infty \beta(\tau)i^*(\tau)\,d\tau$ is just a positive number. We denote that positive number by B. Expanding the determinant, we have

$$(\lambda+\mu)(1-S^*\int_0^\infty \beta(\tau)e^{-\lambda\tau}\pi(\tau)\,d\tau)+B=0.$$

The general idea is to rewrite this equation so that there are positive terms on both sides of the equation. In this case, a useful form for the characteristic equation of the endemic equilibrium is

$$\frac{\lambda+\mu+B}{\lambda+\mu} = S^*\int_0^\infty \beta(\tau)e^{-\lambda\tau}\pi(\tau)\,d\tau. \tag{13.21}$$

Now we will show that this equation cannot have solutions λ with positive real part. Let $\lambda = a+bi$, and assume $a \geq 0$. Taking the absolute value of both sides of the above equality, we have

$$\left|\frac{\lambda+\mu+B}{\lambda+\mu}\right| = \frac{\sqrt{(a+\mu+B)^2+b^2}}{\sqrt{(a+\mu)^2+b^2}} > 1.$$

On the other hand, for $a \geq 0$ we have

$$\begin{aligned} |S^*\int_0^\infty \beta(\tau)e^{-\lambda\tau}\pi(\tau)\,d\tau| &\leq S^*\int_0^\infty \beta(\tau)|e^{-\lambda\tau}|\pi(\tau)\,d\tau \\ &\leq S^*\int_0^\infty \beta(\tau)e^{-a\tau}\pi(\tau)\,d\tau \\ &\leq S^*\int_0^\infty \beta(\tau)\pi(\tau)\,d\tau = 1. \end{aligned} \tag{13.22}$$

This means that for λ with nonnegative real part, the left-hand side remains strictly greater than one, while the right-hand side is strictly less than one. Thus, such λ's cannot satisfy the characteristic equation (13.21). We conclude that the endemic equilibrium is locally asymptotically stable.

Stability for the endemic equilibrium in age-since-infection structured models is a rare event. It has been established that infection-age can destabilize the endemic equilibrium in a simple SI model of HIV [153]. The main differences of the HIV model from model (13.4) is that there is no recovery and that the incidence must be a standard incidence, since sexual contacts do not increase linearly with the population size.

13.3 Influenza Model Structured with Time-Since-Recovery

Besides the infectious class, other classes could be structured by the duration of residence in the class or *class-age*. This is necessary particularly when parameters describing the class may vary with the time individuals spend in the class.

13.3.1 Equilibria of the Time-Since-Recovery Model

To be more specific, let us consider again influenza. Instead of structuring the infectious class with time-since-infection, it may be more realistic to consider structuring the recovered class with time-since-recovery. Influenza strains are believed to impart permanent immunity to themselves and partial immunity to related influenza strains. Because the makeup of influenza strains that circulate in the population continuously changes, a recovered individual has an increasing probability of contracting influenza again. Our goal with this model is to present an example of a different class structure, as well as the fact that the endemic equilibrium does not necessarily needs to be locally stable.

To introduce the model, we have classes of susceptible individuals $S(t)$ who have never had influenza, infected individuals $I(t)$, and recovered individuals whose density is structured with time-since-recovery, $r(\tau,t)$. The model, first introduced in [154], is given below:

$$\begin{aligned} S'(t) &= \Lambda - \beta SI - \mu S, \\ I'(t) &= \beta SI + I\int_0^\infty \gamma(\tau) r(\tau,t)\, d\tau - (\mu+\alpha)I, \\ r_\tau + r_t &= -\gamma(\tau) I r(\tau,t) - \mu r(\tau,t), \\ r(0,t) &= \alpha I. \end{aligned} \tag{13.23}$$

This model is structured by time-since-recovery. The newly recovered individuals αI move into the recovery class with age-since-recovery equal to zero, that is, they

give the boundary condition of the PDE. Recovered individuals can become infected at a rate $\gamma(\tau)$. Realistically, we may expect that $\gamma(\tau)$ is an increasing function of τ. The model is augmented with the following initial conditions:

$$\begin{aligned} S(0) &= S_0, \\ I(0) &= I_0, \\ r(\tau,0) &= \phi(\tau). \end{aligned} \tag{13.24}$$

We define the probability of survival in the recovered class:

$$\pi(\tau) = e^{-I\int_0^\tau \gamma(\sigma)d\sigma} e^{-\mu\tau}.$$

To determine the equilibria, we set the time derivatives equal to zero. We have to solve the following system:

$$\begin{aligned} \Lambda - \beta SI - \mu S &= 0, \\ \beta SI + I\int_0^\infty \gamma(\tau) r(\tau)\, d\tau - (\mu+\alpha)I &= 0, \\ r_\tau &= -\gamma(\tau) I r(\tau) - \mu r(\tau), \\ r(0) &= \alpha I. \end{aligned} \tag{13.25}$$

Solving the differential equation, we obtain an expression for $r(\tau)$ in terms of the number of infected I:

$$r(\tau) = \alpha I \pi(\tau).$$

In addition, we express S in terms of I from the first equation:

$$S = \frac{\Lambda}{\beta I + \mu}.$$

Substituting in the second equation, we obtain an equation for I:

$$\frac{\beta \Lambda I}{\beta I + \mu} + \alpha I^2 \int_0^\infty \gamma(\tau)\pi(\tau)\, d\tau - (\mu+\alpha)I = 0. \tag{13.26}$$

This equation clearly has the disease-free equilibrium $\mathscr{E}_0 = (\frac{\Lambda}{\mu}, 0, 0)$. To find the endemic equilibria, we can cancel one I and obtain the following equation for $I \neq 0$:

$$\frac{\beta \Lambda}{\beta I + \mu} + \alpha I \int_0^\infty \gamma(\tau)\pi(\tau)\, d\tau = (\mu+\alpha). \tag{13.27}$$

We define the basic reproduction number

$$\mathscr{R}_0 = \frac{\beta \Lambda}{\mu(\alpha+\mu)}. \tag{13.28}$$

Proposition 13.2. *Assume $\mathscr{R}_0 > 1$. Then Eq. (13.27) has at least one positive solution.*

Proof. Integration by parts shows that

$$I\int_0^\infty \gamma(\tau)\pi(\tau)\,d\tau = 1-\mu\int_0^\infty \pi(\tau)\,d\tau.$$

Hence,

$$\lim_{I\to\infty} I\int_0^\infty \gamma(\tau)\pi(\tau)\,d\tau = 1.$$

We denote the left-hand side of (13.27) by $f(I)$. We have $f(0)=\frac{\beta\Lambda}{\mu}$. Hence, since $\mathscr{R}_0>1$, we have $f(0)>\mu+\alpha$. On the other hand, we have $\lim_{I\to\infty} f(I)=\alpha<\alpha+\mu$. This shows that the equation $f(I)=\alpha+\mu$ has at least one solution. □

It is biologically reasonable to assume that the reinfection rate $\gamma(\tau)$ is a bounded function and that its supremum does not exceed β:

$$\sup_\tau \gamma(\tau)\le\beta. \tag{13.29}$$

Under this condition, Thieme and Yang [154] showed that the endemic equilibrium $\mathscr{E}^*=(S^*,I^*,r^*(\tau))$ is unique. This equilibrium cannot be explicitly computed. If condition (13.29) is not satisfied, then backward bifurcation may occur, and multiple endemic equilibria are possible (see Problem 13.2).

13.3.2 Stability of Equilibria

In this subsection, we investigate the stability of the equilibria. We will concentrate primarily on the unique endemic equilibrium in the case that condition (13.29) holds. As before, we start from linearizing system (13.23). Let (S^*,I^*,r^*) denote a generic equilibrium. We set $S=S^*+x(t)$, $I=I^*+y(t)$, and $r(\tau,t)=r^*(\tau)+z(\tau,t)$. Substituting in the equations of (13.23), we obtain

$$\begin{aligned}
(S^*+x)'(t) &= \Lambda-\beta(S^*+x)(I^*+y)-\mu(S^*+x),\\
(I^*+y)'(t) &= \beta(S^*+x)(I^*+y)+(I^*+y)\int_0^\infty \gamma(\tau)(r^*+z(\tau,t))\,d\tau-(\mu+\alpha)(I^*+y),\\
(r^*+z)_\tau+z_t &= -\gamma(\tau)(I^*+y)(r^*(\tau)+z(\tau,t))-\mu(r^*+z(\tau,t)),\\
r^*(0)+z(0,t) &= \alpha(I^*+y).
\end{aligned} \tag{13.30}$$

Multiplying out and using the equations for the equilibria (13.25), we obtain the following system:

$$
\begin{aligned}
x'(t) &= -\beta S^* y - \beta I^* x - \beta xy - \mu x,\\
y'(t) &= \beta S^* y + \beta I^* x + \beta xy + I^* \int_0^\infty \gamma(\tau) z(\tau,t)\, d\tau + y \int_0^\infty \gamma(\tau) r^*(\tau)\, d\tau\\
&\quad + y \int_0^\infty \gamma(\tau) z(\tau,t)\, d\tau - (\mu+\alpha) y,\\
z_\tau + z_t &= -\gamma(\tau) y r^*(\tau) - \gamma(\tau) I^* z(\tau,t) - \gamma(\tau) y z(\tau,t), -\mu z(\tau,t),\\
z(0,t) &= \alpha y.
\end{aligned}
\tag{13.31}
$$

We also neglect the quadratic terms in the perturbations x, y, and z to obtain the following linear system in the perturbations:

$$
\begin{aligned}
x'(t) &= -\beta S^* y - \beta I^* x - \mu x,\\
y'(t) &= \beta S^* y + \beta I^* x + I^* \int_0^\infty \gamma(\tau) z(\tau,t)\, d\tau + y \int_0^\infty \gamma(\tau) r^*(\tau)\, d\tau - (\mu+\alpha) y,\\
z_\tau + z_t &= -\gamma(\tau) y r^*(\tau) - \gamma(\tau) I^* z(\tau,t) - \mu z(\tau,t),\\
z(0,t) &= \alpha y.
\end{aligned}
\tag{13.32}
$$

This is our linear system for the perturbations. To investigate the local stability of the equilibria, we have to study the solutions of this system. As before, we expect that the solutions are exponential. Therefore, we look for solutions of the form $x(t) = \bar{x}e^{\lambda t}$, $y(t) = \bar{y}e^{\lambda t}$, $z(\tau,t) = \bar{z}(\tau)e^{\lambda t}$. We obtain the following linear eigenvalue problem for $\bar{x}$, $\bar{y}$, and $\bar{z}(\tau)$, and the eigenvalue λ:

$$
\begin{aligned}
\lambda x &= -\beta S^* y - \beta I^* x - \mu x,\\
\lambda y &= \beta S^* y + \beta I^* x + I^* \int_0^\infty \gamma(\tau) z(\tau)\, d\tau + y \int_0^\infty \gamma(\tau) r^*(\tau)\, d\tau - (\mu+\alpha) y,\\
z_\tau + \lambda z &= -\gamma(\tau) y r^*(\tau) - \gamma(\tau) I^* z(\tau) - \mu z(\tau),\\
z(0) &= \alpha y,
\end{aligned}
\tag{13.33}
$$

where in the above, we have dropped the bars. To investigate the stability of the disease-free equilibrium, we have to write the above system for that equilibrium. This will simplify that system significantly:

$$
\begin{aligned}
\lambda x &= -\beta S^* y - \mu x,\\
\lambda y &= \beta S^* y - (\mu+\alpha) y,\\
z_\tau + \lambda z &= -\mu z(\tau,t),\\
z(0) &= \alpha y.
\end{aligned}
\tag{13.34}
$$

Problem 13.3 asks you to determine the stability of the disease-free equilibrium. Use the above system to answer Problem 13.3.

Here, we assume that condition (13.29) holds, and we continue with the investigation of the unique endemic equilibrium. The next step will be to solve system (13.33) and derive the characteristic equation. A typical way in which this can be done is to solve an ordinary differential equation and express z in terms of y. Then from the first equation, we can express x in terms of y. Then we may substitute the expressions for z and x in the second equation. We will obtain an equation in y only.

Since we are looking for a nontrivial eigenvector, we assume that y is not zero and we cancel it. We obtain an equation in λ that constitutes the characteristic equation.

To carry out this plan, we begin by solving the differential equation. The ODE is a first-order linear ODE with nonzero right-hand side. We move the terms that depend on z to the left-hand side and the terms that do not depend on z to the right-hand side. We multiply by the integrating factor $e^{(\lambda+\mu)\tau+I\int_0^\tau \gamma(\sigma)d\sigma}$. The differential equation becomes

$$\left[e^{(\lambda+\mu)\tau+I^*\int_0^\tau \gamma(\sigma)d\sigma}z\right]' = -\gamma(\tau)yr^*(\tau)e^{(\lambda+\mu)\tau+I^*\int_0^\tau \gamma(\sigma)d\sigma}.$$

Integrating both sides of this equation from 0 to τ and recalling that $z(0)=\alpha y$, we have

$$z(\tau) = \alpha y e^{-(\lambda+\mu)\tau-I^*\int_0^\tau \gamma(\sigma)d\sigma} - y\int_0^\tau \gamma(s)r^*(s)e^{-(\lambda+\mu)(\tau-s)-I^*\int_s^\tau \gamma(\sigma)d\sigma}ds.$$

This gives an expression for z in terms of y. This equation can be rewritten also in the form

$$z(\tau) = \alpha y e^{-(\lambda+\mu)\tau-I^*\int_0^\tau \gamma(\sigma)d\sigma} - yr^*(\tau)\int_0^\tau \gamma(s)e^{-\lambda(\tau-s)}ds.$$

From the first equation in system (13.33), we have

$$x = -\frac{\beta S^* y}{\lambda+\beta I^*+\mu}.$$

From the second equation in system (13.33), after substituting z and x and canceling y, we obtain the characteristic equation:

$$\begin{aligned}(\lambda+\mu+\alpha) &= \beta S^* - \frac{\beta I^*\beta S^*}{\lambda+\beta I^*+\mu} + \int_0^\infty \gamma(\tau)r^*(\tau)e^{-\lambda\tau}d\tau \\ &\quad - I\int_0^\infty \gamma(\tau)r^*(\tau)\int_0^\tau \gamma(s)e^{-\lambda(\tau-s)}dsd\tau + \int_0^\infty \gamma(\tau)r^*(\tau)d\tau.\end{aligned} \tag{13.35}$$

Integrating the double integral by parts, thinking of $I\gamma(\tau)e^{-I\int_0^\tau \gamma(\sigma)d\sigma}$ as u' and the rest as v, we can obtain the following simplified characteristic equation:

$$\begin{aligned}\lambda+\mu+\alpha = \frac{\beta S(\lambda+\mu)}{\lambda+\beta I^*+\mu} &+ \int_0^\infty \gamma(\tau)r^*(\tau)e^{-\lambda\tau}d\tau \\ &+(\lambda+\mu)\int_0^\infty r^*(\tau)\int_0^\tau \gamma(s)e^{-\lambda(\tau-s)}dsd\tau.\end{aligned} \tag{13.36}$$

This characteristic equation does not always have only roots with negative real parts. One can pose additional conditions that would imply stability. However, here we would like to show that instability and oscillations may occur in this model. To show this, we need to exhibit a specific example in which Hopf bifurcation can occur and sustained oscillations are possible. To demonstrate this, we consider the following special case:

Assumption: Assume

$$\gamma(\tau) = \begin{cases} 0 & 0 \le \tau \le A, \\ \beta & \tau > A, \end{cases} \tag{13.37}$$

where A is an arbitrary constant. This form of $\gamma(\tau)$ suggests that recovered individuals are completely protected for a period of time A and then completely susceptible again. This is a reasonable assumption in influenza modeling. With this $\gamma(\tau)$, the recovered individuals are given by the following expression:

$$r^*(\tau) = \begin{cases} \alpha I^* e^{-\mu\tau} & 0 \le \tau \le A, \\ \alpha I^* e^{-\mu\tau} e^{-\beta I^*(\tau - A)} & \tau > A. \end{cases} \tag{13.38}$$

When the class-age structured function is a step function, the class-age model becomes equivalent to a delay model. As a result, the characteristic equation (13.36) can be significantly simplified. We compute the integrals:

$$\begin{aligned} \int_0^\infty \gamma(\tau) r^*(\tau) e^{-\lambda\tau} d\tau &= \alpha\beta I^* \int_A^\infty e^{-\mu\tau} e^{-\beta I^*(\tau - A)} e^{-\lambda\tau} d\tau \\ &= \frac{\alpha\beta I^*}{\lambda + \mu + \beta I^*} e^{-(\lambda+\mu)A}. \end{aligned} \tag{13.39}$$

Given the value of the integral above with $\lambda = 0$, the equation for the equilibria becomes

$$\beta S^* + \frac{\alpha\beta I^*}{\mu + \beta I^*} e^{-\mu A} = (\mu + \alpha). \tag{13.40}$$

Recalling that $S^* = \Lambda/(\beta I^* + \mu)$, we can solve the resulting equation for I^* to obtain

$$I^* = \frac{\beta\Lambda - \mu(\mu + \alpha)}{\beta(\mu + \alpha + \alpha e^{-\mu A})}. \tag{13.41}$$

The double integral in (13.36) can also be computed:

$$\begin{aligned} \int_0^\infty r^*(\tau) \int_0^\tau \gamma(s) e^{-\lambda(\tau - s)} ds d\tau &= \int_A^\infty r^*(\tau) \int_A^\tau \gamma(s) e^{-\lambda(\tau - s)} ds d\tau \\ &= \beta\alpha I^* \int_A^\infty e^{-\mu\tau} e^{-\beta I^*(\tau - A)} \int_A^\tau e^{-\lambda(\tau - s)} ds d\tau \\ &= \frac{\alpha\beta I^* e^{\beta I^* A}}{\lambda} \int_A^\infty e^{-\mu\tau} e^{-\beta I^*\tau} e^{-\lambda\tau} (e^{\lambda\tau} - e^{\lambda A}) d\tau \\ &= \frac{\alpha\beta I^* e^{\beta I^* A}}{\lambda} \int_A^\infty e^{-\mu\tau} e^{-\beta I^*\tau} d\tau \\ &\quad - \frac{\alpha\beta I^* e^{\beta I^* A + \lambda A}}{\lambda} \int_A^\infty e^{-\mu\tau} e^{-\beta I^*\tau} e^{-\lambda\tau} d\tau \\ &= \frac{\alpha\beta I^* e^{-\mu A}}{\lambda(\mu + \beta I^*)} - \frac{\alpha\beta I^* e^{-\mu A}}{\lambda(\lambda + \mu + \beta I^*)} \\ &= \frac{\alpha\beta I^* e^{-\mu A}}{(\mu + \beta I^*)(\lambda + \mu + \beta I^*)}. \end{aligned} \tag{13.42}$$

Using the equation for the equilibria (13.40), the characteristic equation (13.36) can be simplified as follows:

$$\lambda+\mu+\alpha=\frac{\beta S^*(\lambda+\mu)}{\lambda+\beta I^*+\mu}+\frac{\alpha\beta I^* e^{-(\lambda+\mu)A}}{\lambda+\beta I^*+\mu}+\frac{\alpha\beta I^*(\lambda+\mu)e^{-\mu A}}{(\lambda+\beta I^*+\mu)(\beta I^*+\mu)}. \quad (13.43)$$

Collecting terms gives us

$$(\lambda+\mu+\alpha)(\lambda+\beta I^*+\mu)=(\lambda+\mu)\left[\beta S^*+\frac{\alpha\beta I^* e^{-\mu A}}{\mu+\beta I^*}\right]+\alpha\beta I^* e^{-(\lambda+\mu)A}. \quad (13.44)$$

Using (13.40) and simplifying, we obtain

$$(\lambda+\mu+\alpha)(\lambda+\beta I^*+\mu)-(\lambda+\mu)(\alpha+\mu)=\alpha\beta I^* e^{-(\lambda+\mu)A}. \quad (13.45)$$

Hence, the characteristic equation simplifies to the following transcendental equation:

$$\lambda^2+(\beta I^*+\mu)\lambda+\beta I^*(\mu+\alpha)=\alpha\beta I^* e^{-(\lambda+\mu)A}. \quad (13.46)$$

Let $\lambda=\xi+i\eta$. We can use the methodology first introduced in Chap. 4 to find eigenvalues with positive real part. We separate the real and the imaginary part in the equation above. We obtain the following system:

$$\begin{cases}\xi^2-\eta^2+(\beta I^*+\mu)\xi+\beta I^*(\mu+\alpha)=\alpha\beta I^* e^{-(\xi+\mu)A}\cos\eta A,\\ 2\xi\eta+(\beta I^*+\mu)\eta=-\alpha\beta I^* e^{-(\xi+\mu)A}\sin\eta A.\end{cases} \quad (13.47)$$

To find parameters that will give us oscillations, we proceed in the following way. We notice that the system above is linear in βI^* and $\alpha\beta I^*$. Hence, we can solve for these parameters. That cannot be done by hand, but a computer algebra system such as Mathematica can do it. The expressions we obtain are rather large, and we will not include them here. We view $\beta I^*=f(\eta)$ and $\alpha\beta I^*=g(\eta)$. We plot parametrically the points $(f(\eta),g(\eta))$ in the $(\beta I^*,\alpha\beta I^*)$-plane. We obtain the left figure in Fig. 13.3. Before plotting, we fix the other parameters as follows. Since the current worldwide lifespan of humans is 70 years, we define $\mu=1/(70*365)$ days^{-1}. The worldwide human population is 7 billion. So we compute $\Lambda=1000/365$ births per day (in units of 10^5). In this way, at equilibrium, where the population is Λ/μ, we would have 70,000 individuals (in units 10^5), that is, 7 billion individuals. We will compute the infectious period so that oscillations occur. We take $A=30$ days. That is, prior exposure to influenza protects a person completely for 30 days, after which one becomes completely susceptible again. That value for A will give too small a period of oscillations for influenza. If we want a more realistic period, we need to take $A=365$. We fix $\xi=0.01$. From the left figure of Fig. 13.3, we see that for some η's, the point in the $(\beta I^*,\alpha\beta I^*)$-plane is in the positive quadrant. Hence, it gives a viable point. We guess a value of η that gives positive $\beta I^*=f(\eta)$ and $\alpha\beta I^*=g(\eta)$. In the simulations, we took $\eta=0.19$. That gives $\beta I^*=0.172187$ and $\alpha\beta I^*=0.0896313$. Dividing the second of these numbers by the first, we get $\alpha=0.520548$, which gives an infectious period of less than two days. We determine β from the formula for I^*:

$$\beta I^* = \frac{\beta\Lambda - \mu(\mu+\alpha)}{(\mu+\alpha-\alpha e^{-\mu A})} = f(0.19) = 0.172187.$$

This gives $\beta = 0.0000482876$. To observe the principal eigenvalue, we plot the equations in system (13.47) as contour plots in the (ξ, η)-plane. The eigenvalues of Eq. (13.46) are given from the intersection of the two types of level curves. The right-hand picture of Figure 13.3 shows that Eq. (13.46) exhibits Hopf bifurcation and has a principal eigenvalue with positive real part.

To illustrate the oscillations, one needs to simulate the solution of system (13.23). As an integrodifferential PDE system with nonlocal boundary condition, the system cannot be automatically solved by a computer algebra system such as Matlab of Mathematica. A numerical method needs to be built for system (13.23) and coded in Matlab, Fortran, or C. In the next subsection, we discuss how to discretize the system.

13.3.3 Numerical Method for the Time-Since-Recovery Model

In this section, we build a numerical method and code model (13.23). The numerical method is a finite difference method that discretizes both the age and time variables and computes the solution at a number of points that form a mesh.

As a first step, we need to discretize the domain of the system (13.23). Recall that the domain is given by

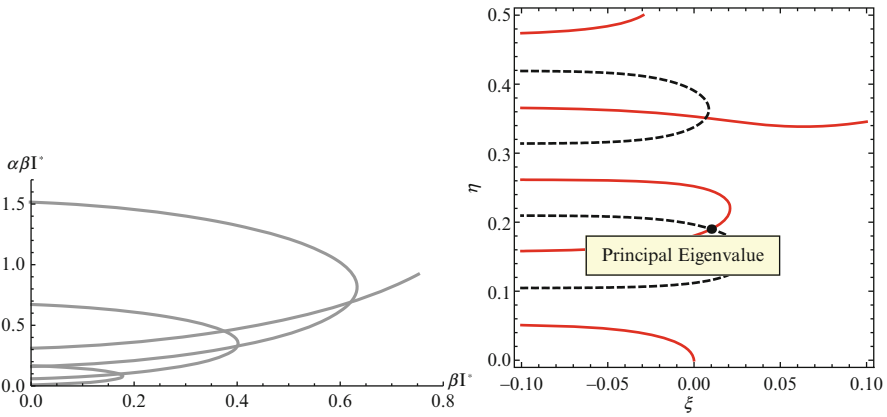

Fig. 13.3 The *left* figure gives the parametric plot in the $(\beta I^*, \alpha\beta I^*)$-plane. The *right* figure shows the level curves of system (13.47). The level curves of the first equation are given by a *continuous curve*, and the level curves of the second equation are given in *dashed curves*. The eigenvalues are given by the intersection of *continuous* and *dashed curves*. The principal eigenvalue is the one that is farthest to the *right*. The figure indicates it with a point. The principal eigenvalue has $\xi = 0.01$, as set in the computation (see text)

$$\mathscr{D} = \{(\tau,t) : \tau \geq 0, t \geq 0\}.$$

The domain $\mathscr{D}$ is an infinite domain. We cannot compute with infinite domains, so we have to truncate it in both the age and time directions. We will consider the finite domain

$$\bar{\mathscr{D}} = \{(\tau,t) : 0 \leq \tau \leq G, 0 \leq t \leq T\}.$$

What is a sensible way to choose G so that when we truncate the infinite integral in the boundary condition, we do not make too much of an error? One way is to choose G large enough that $e^{-\mu G}$ is almost zero. Since the solutions decay with $e^{-\mu\tau}$, computing in τ until this exponent becomes zero, at least in a computational sense, will guarantee that the error of replacing the infinite integral with a finite one will not modify the solution too much.

To discretize, we take the points along the age direction equally spaced with a step $\Delta\tau$: $\tau_k = k\Delta\tau$. Since both age and time progress simultaneously, we discretize the time with same step $\Delta t = \Delta\tau$. The points in the time direction are given by $t_n = n\Delta t$. These points discretize the domain $\bar{\mathscr{D}}$ with a discrete square mesh. We define the number of steps made in each direction as

$$K = \left[\frac{G}{\Delta t}\right], \qquad N = \left[\frac{T}{\Delta t}\right],$$

where $[\cdot]$ denotes the integer part of the expression. We may modify G and T a little so that without loss of generality, we may assume that

$$G = K\Delta t, \qquad T = N\Delta t.$$

To discretize the time-since-recovery model (13.23), we assume $S(t_n) \approx S^n$, $r(\tau_k, t_n) \approx r_k^n$, $I(t_n) \approx I^n$. We first discretize the equation for the susceptibles, which is a first-order ODE. We can use a backward difference to replace the time derivative. Thus, we evaluate the equation at time level t_{n+1} and apply a backward finite difference for the time derivative. We obtain

$$\frac{S^{n+1} - S^n}{\Delta t} = \Lambda - \beta S^{n+1} I^{n+1} - \mu S^{n+1}.$$

From here, we should be able to compute S^{n+1}, knowing the values at the nth time level. However, as the equation stands now, this is not possible, because we do not know I^{n+1}. To simplify the computations, we "linearize" the nonlinear term and evaluate I at time level n rather than time level $n+1$. This is legitimate in numerical methods, since I^n and I^{n+1} are close, and so their values should be close. The equation above becomes

$$\frac{S^{n+1} - S^n}{\Delta t} = \Lambda - \beta S^{n+1} I^n - \mu S^{n+1}.$$

Now we can use this equation to compute S^{n+1} from values at level n. Since the equation for I contains both S and r as well as I, we compute that last. Next we discretize the PDE. We evaluate it at t_{n+1} and τ_k. We have

$$r_\tau(\tau_k, t_{n+1}) + r_t(\tau_k, t_{n+1}) = -\gamma(\tau_k) I(t_{n+1}) r(\tau_k, t_{n+1}) - \mu r(\tau_k, t_{n+1}).$$

We replace the derivative in τ with a forward difference and the derivative in time with a backward difference. We obtain the following difference equation:

$$\frac{r_{k+1}^{n+1} - r_k^{n+1}}{\Delta t} + \frac{r_k^{n+1} - r_k^n}{\Delta t} = -\gamma_k I^{n+1} r_k^{n+1} - \mu r_k^{n+1}.$$

The left-hand side can be simplified by canceling the two r_k^{n+1} terms. The right-hand side contains r_k^{n+1}, while the left-hand side contains only r_{k+1}^{n+1} and r_k^n. It makes sense to replace r_k^{n+1} on the right-hand side with one of the two on the left. The better choice is to replace them with r_{k+1}^{n+1}. This makes the method "implicit," that is, the right-hand side depends on the time level that we are computing. We need to solve the equation for r_{k+1}^{n+1}, but that is not difficult, since the equation is linear in r. As with the S equation, we "linearize" the nonlinear term by computing I at time level n rather than time level $n+1$. The discretization of the PDE becomes

$$\frac{r_{k+1}^{n+1} - r_k^n}{\Delta t} = -\gamma_k I^n r_{k+1}^{n+1} - \mu r_{k+1}^{n+1}.$$

This equation can be solved easily for r_{k+1}^{n+1}. It gives a formula for the computation of the next value of r along the characteristic line:

$$r_{k+1}^{n+1} = \frac{r_k^n}{1 + \gamma_k I^n \Delta t + \mu \Delta t}.$$

From the boundary condition, we have

$$r_0^{n+1} = \alpha I^{n+1}.$$

Here we do not "linearize" at the previous level. This means that we cannot compute the boundary condition until we have computed I^{n+1}. The reason we do not linearize here is that the αI in the equation for I will be computed at level $n+1$, so that the method is implicit, but the two terms have to cancel each other if we are to obtain the equation for the total population size. We may use the right-endpoint rule to compute the integral in the equation for I. In this way, we may avoid using the boundary condition for r in the equation for I. We compute the equation for I at time level $n+1$, discretize the derivative with a backward difference, and the integral with a right-endpoint rule sum:

$$\frac{I^{n+1} - I^n}{\Delta t} = \beta S^{n+1} I^{n+1} + I^{n+1} \sum_{k=1}^{K} \gamma_k r_k^{n+1} \Delta t - (\mu + \alpha) I^{n+1}.$$

We replace I^{n+1} with I^n in the terms with S and r so that they agree with the corresponding terms in the equations for S and r. We obtain

$$\frac{I^{n+1}-I^n}{\Delta t} = \beta S^{n+1} I^n + I^n \sum_{k=1}^{K} \gamma_k r_k^{n+1} \Delta t - (\mu+\alpha) I^{n+1}.$$

In this way, we can compute I^{n+1} before we compute the boundary condition for r. In case we want to use a different rule for the integral, such as the trapezoidal rule, we have to solve a system of equations to find the solution at time level $n+1$. With this scheme, the computation is performed time level after time level. To begin the computation, we initialize all variables with the initial conditions that give the values at time level zero:

$$S^0 = S_0 \qquad I^0 = I_0, \qquad r_k^0 = \phi_k \quad , k = 0, \ldots, K.$$

We summarize the numerical method below:

$$\begin{cases} S^{n+1} = \dfrac{\Lambda \Delta t + S^n}{1 + \beta I^n \Delta t + \mu \Delta t} & n = 0, \ldots, N-1, \\ I^{n+1} = \dfrac{I^n + \beta \Delta t S^{n+1} I^n + I^n \Delta t \sum_{k=1}^{K} \gamma_k r_k^{n+1} \Delta t}{1 + \alpha \Delta t + \mu \Delta t} & n = 0, \ldots, N-1, \\ r_{k+1}^{n+1} = \dfrac{r_k^n}{1 + \gamma_k I^n \Delta t + \mu \Delta t}, & k = 0, \ldots, K-1, \\ & n = 0, \ldots, N-1, \\ S^0 = S_0 & \\ I^0 = I_0 & \\ r_k^0 = \phi_k & k = 0, \ldots, K. \end{cases} \tag{13.48}$$

The numerical method in (13.48) is given by a *difference scheme*. It can be shown (but we will not do so here) that the solutions of the difference scheme converge to the solution of the continuous problem (13.23) with the same speed as $C\Delta t$ converges to zero as Δt converges to zero. Here C is an appropriate constant. In this case, we say that the method has *convergence rate* $\mathcal{O}(\Delta t)$. The method (13.48) has other important strengths. In particular, its solutions are always nonnegative for every value of the step Δt. Finally, it is easy to code and has relatively low computational complexity (number of operations). An appropriate size of the step for running this method is $\Delta t = 0.01$. Smaller step sizes are also appropriate, but one has to keep in mind that as the step size decreases, the time needed to perform the computation increases. We ran the method with the parameters estimated in the previous subsections. The number of infected individuals, which exhibits sustained oscillations, is plotted in Fig. 13.4.

Acknowledgements The author thanks Necibe Tuncer for her help with the Matlab code and checking.

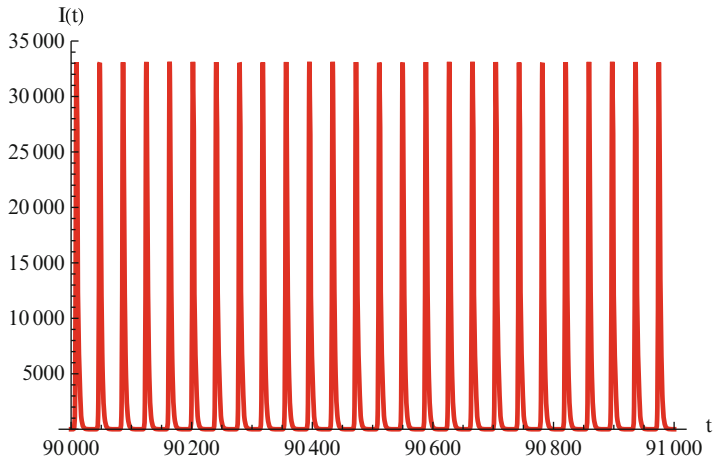

Fig. 13.4 Oscillations in the number of infected individuals $I(t)$

Appendix

In this appendix we include the Matlab code that executes the numerical method in the section.

```
function [S, I] = sir3(M,N,dt)



T = dt*N;
G = dt*M;

Lambda = 1000/365;
mu = 1/(70*365);
alpha = 0.520548;
beta = 0.0000482876;

S = zeros(N,1);
I = zeros(N,1);
rold = zeros(N,1);
rnew = zeros(N,1);

S(1) = 7000;
I(1) = 30000;

for i = 1:M
    rold(i) = 10;
end

t = 0:dt:T;
ttau = 0:dt:G;
```

```
for n = 1:N

    S(n+1) = (Lambda*dt+S(n))/(1+beta*I(n)*dt + mu*dt);

    Int = 0.0;

    for i = 1:M

        tau = i*dt;

        if tau < 30

            q = 0.0;

        elseif   tau ≥ 30

            q = 1;

        end

     rnew(i+1) = rold(i)/(1 + q*beta*I(n)*dt + mu*dt);

     Int = Int + q*rnew(i+1)*dt;

    end

    I(n+1) = (I(n) + beta*S(n+1)*I(n)*dt + ...
        I(n)*dt*beta*Int)/(1 + (mu + alpha)*dt);

    rnew(1) = alpha*I(n+1);


    for i = 1:M

        rold(i) = rnew(i);

    end

end

plot(t, I, '-r')
xlim([90000,91000])


end
```

Problems

13.1. Consider the SIR model with age-of-infection (13.4). Assume that the transmission rate and the recovery rate are given by the following functions:

$$\gamma(\tau) = \begin{cases} 0 & \tau \le A \\ \gamma & \tau > A \end{cases} \tag{13.49}$$

and $\beta(\tau) = \beta\tau e^{k\tau}$.

(a) Compute the probability of survival in the infectious class: $\pi(\tau) = e^{-\mu\tau}e^{-\int_0^\tau \gamma(s)ds}$.
(b) Compute the reproduction number in terms of k and A.
(c) Compute the endemic equilibrium in terms of k and A.

13.2. Backward Bifurcation in the Time-Since-Recovery Model
Consider Eq. (13.27) with $\gamma(\tau) = \gamma$, a constant.

(a) Show that if $\gamma \le \beta$ and $\mathscr{R}_0 > 1$, the equation (13.27) has a unique nonzero solution. Furthermore, show that if $\gamma \le \beta$ and $\mathscr{R}_0 < 1$, the equation (13.27) has no solutions.
(b) Show that if $\gamma > \beta$ and $\mathscr{R}_0 < 1$, the equation (13.27) may have two solutions.
(c) For $\alpha = 0.05$, $\mu = 1/(365*70)$, $\beta = 0.021$, and $\gamma = 0.025$, use a computer algebra system to draw the backward bifurcation diagram of I^* with respect to $\mathscr{R}_0$.

13.3. Consider the model with time-since-recovery (13.23). Show that if $\mathscr{R}_0 < 1$, the disease-free equilibrium is locally asymptotically stable. Furthermore, show that if $\mathscr{R}_0 > 1$, the disease-free equilibrium is unstable.

13.4. HIV/AIDS Model
Consider the following model of HIV:

$$\begin{cases} S'(t) = \Lambda - \dfrac{S(t)}{N(t)}\displaystyle\int_0^\infty \beta(\tau)i(\tau,t)\,d\tau - \mu S(t), \\ i_\tau(\tau,t) + i_t(\tau,t) = -\gamma(\tau)i(\tau,t) - \mu i(\tau,t), \\ i(0,t) = \dfrac{S(t)}{N(t)}\displaystyle\int_0^\infty \beta(\tau)i(\tau,t)\,d\tau, \end{cases} \tag{13.50}$$

where $S(t)$ are the susceptible individuals, $i(\tau,t)$ is the density of the infected individuals, $N(t)$ is the total population size. We have to use standard incidence in HIV models. (Why?)

(a) Compute $\mathscr{R}_0$ and the disease-free equilibrium. Show that if $\mathscr{R}_0 < 1$, the disease-free equilibrium is locally stable and that otherwise, it is unstable.
(b) Compute the endemic equilibrium.
(c) Derive the characteristic equation of the endemic equilibrium.
(d) Take $\beta(\tau) = \tau e^{-c\tau}$. Is the endemic equilibrium stable or unstable in this case?

13.5. HIV/AIDS Model
Consider the following model of HIV:

$$\begin{cases} S'(t) = \Lambda - \dfrac{S(t)}{N(t)} \displaystyle\int_0^\infty \beta(\tau) i(\tau,t)\, d\tau - \mu S(t), \\ i_\tau(\tau,t) + i_t(\tau,t) = -\gamma(\tau) i(\tau,t) - \mu i(\tau,t), \\ i(0,t) = \dfrac{S(t)}{N(t)} \displaystyle\int_0^\infty \beta(\tau) i(\tau,t)\, d\tau, \end{cases} \tag{13.51}$$

where $S(t)$ are the susceptible individuals, $i(\tau,t)$ is the density of the infected individuals, $N(t)$ is the total population size. Assume

$$\beta(\tau) = \tau e^{-c\tau}.$$

(a) Compute $\mathscr{R}_0$ and the disease-free equilibrium.
(b) Compute the endemic equilibrium.
(c) Derive the characteristic equation of the endemic equilibrium.
(d) Is the endemic equilibrium stable or can it become unstable in this case?

13.6. HIV/AIDS Model
Consider the following model of HIV:

$$\begin{cases} S'(t) = \Lambda - \dfrac{S(t)}{N(t)} \displaystyle\int_0^\infty \beta(\tau) i(\tau,t)\, d\tau - \mu S(t), \\ i_\tau(\tau,t) + i_t(\tau,t) = -\gamma(\tau) i(\tau,t) - \mu i(\tau,t), \\ i(0,t) = \dfrac{S(t)}{N(t)} \displaystyle\int_0^\infty \beta(\tau) i(\tau,t)\, d\tau, \end{cases} \tag{13.52}$$

where $S(t)$ are the susceptible individuals, $i(\tau,t)$ is the density of the infected individuals, $N(t)$ is the total population size.

(a) Derive a numerical method for model (13.52).
(b) Write a Matlab code to simulate the method. How do you know whether your code computes correctly? Compare the equilibrium computed by the code with the one that you computed in Problem 13.5.

13.7. Time-Since-Vaccination Model
Many vaccines wane, and the waning depends on the time elapsed since the individual was vaccinated. Consider the following model with time-since-vaccination τ:

$$\begin{cases} S'(t) = \Lambda - \beta S(t) I(t) - (\mu + \psi) S(t) + \displaystyle\int_0^\infty \omega(\tau) v(\tau,t)\, d\tau, \\ I'(t) = \beta S(t) I(t) - (\mu + \alpha) I, \\ v_\tau + v_t = -\omega(\tau) v(\tau,t) - \mu v(\tau,t), \\ v(0,t) = \alpha I + \psi S, \end{cases} \tag{13.53}$$

where $v(\tau,t)$ is the density of vaccinated individuals structured by the time-since-vaccination τ, and ψ is the vaccination rate.

(a) Interpret all terms in the model. What is the assumed efficacy of the vaccine in this model?
(b) Compute the reproduction number $\mathscr{R}_0(\psi)$.
(c) Compute the disease-free equilibrium. Show that if $\mathscr{R}_0(\psi) < 1$, the disease-free equilibrium is locally asymptotically stable; otherwise, the disease-free equilibrium is unstable.
(d) Compute the endemic equilibrium.

13.8. Time-Since-Vaccination Model
Many vaccines wane, and the waning depends on the time elapsed since the individual was vaccinated. Consider the following model with time-since-vaccination τ:

$$\begin{cases} S'(t) = \Lambda - \beta S(t)I(t) - (\mu+\psi)S(t) + \int_0^\infty \omega(\tau)v(\tau,t)\,d\tau, \\ I'(t) = \beta S(t)I(t) - (\mu+\alpha)I, \\ v_\tau + v_t = -\omega(\tau)v(\tau,t) - \mu v(\tau,t), \\ v(0,t) = \alpha I + \psi S, \end{cases} \tag{13.54}$$

where $v(\tau,t)$ is the density of vaccinated individuals structured by the time-since-vaccination τ, and ψ is the vaccination rate.

(a) Write a numerical method for the model above. Show that your numerical method preserves the positivity of solutions.
(b) Write a Matlab code to simulate the model. How do you know whether your code computes correctly? Compare the equilibrium computed by the code with the one that you computed in Problem 13.7.

13.9. Time-Since-Infection Model of Vector-Borne Disease
Consider the following model of a vector-borne disease, structured by time-since-infection τ:

$$\begin{cases} S_v' = \Lambda_v - S_v \int_0^\infty \beta_H(\tau)i(\tau,t)d\tau - \mu_v S_v, \\ I_v' = S_v \int_0^\infty \beta_H(\tau)i(\tau,t)d\tau - \mu_v I_v, \\ S_H' = \Lambda_H - \beta_v S_H I_v - \mu_H S_H, \\ i_\tau + i_t = -(\alpha_H(\tau)+\mu_H)i(\tau,t), \\ i(0,t) = \beta_v S_H I_v, \\ R_H' = \int_0^\infty \alpha_H(\tau)i(\tau,t)d\tau - \mu_H R_H, \end{cases} \tag{13.55}$$

where S_v, I_v are the susceptible and infected vectors, S_H, $i(\tau,t)$, and R_H are the susceptible, infected, and recovered humans.

(a) Compute the reproduction number $\mathscr{R}_0$.
(b) Compute the disease-free equilibrium. Show that if $\mathscr{R}_0 < 1$, the disease-free equilibrium is locally asymptotically stable; otherwise, it is unstable.
(c) Compute the endemic equilibrium.
(d) Derive the characteristic equation of the endemic equilibrium. Can you show local stability of the endemic equilibrium?

Chapter 14
Immuno-Epidemiological Modeling

14.1 Introduction to Immuno-Epidemiological Modeling

To better understand the spread of infectious diseases in populations, we need to realize that each infectious person harbors the pathogen, and that pathogen is in dynamic interplay with the host immune system. The host's propensity to transmit the pathogen or die from it depends on the amount of the pathogen in the system as well as the intensity of the immune response. The spread of diseases on a population level depends on these within-host disease characteristics of infectious individuals. Further understanding of epidemiological processes relies on our knowledge of within-host (immunological) processes and the links between the two scales. Just as the epidemiological processes have been extensively modeled within their own scale, immunological processes have also been modeled widely. A variety of within-host dynamical models exists for most human diseases such as HIV, HCV, influenza, and malaria. The within-host and between-host models are the two basic building blocks of a new class of models, called immuno-epidemiological models. The new subject of immuno-epidemiology merges individual and population-oriented approaches to examine how within-host pathogen dynamics affect the population dynamics of micro- and macro-parasites to produce the epidemiological patterns of infection observed in the host populations [18]. There are a number of mathematical approaches to modeling simultaneously the two scales; however, we will concentrate here on the nested modeling. This modeling approach uses a dynamical within-host model and embeds it into an age-since-infection structured epidemiological model (see Fig. 14.1). The two models are linked through the age-since-infection variable of the epidemiological model, which is the time variable of the immunological model. Furthermore, the two models are linked through the age-since-infection-dependent epidemiological parameters, which are taken as functions of the immunologically dependent variables, such as pathogen load and immune response. In the process of formulating age-since-infection structured epidemiological models, we argued that the infection transmission rate should not be constant but that it should depend on

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_14

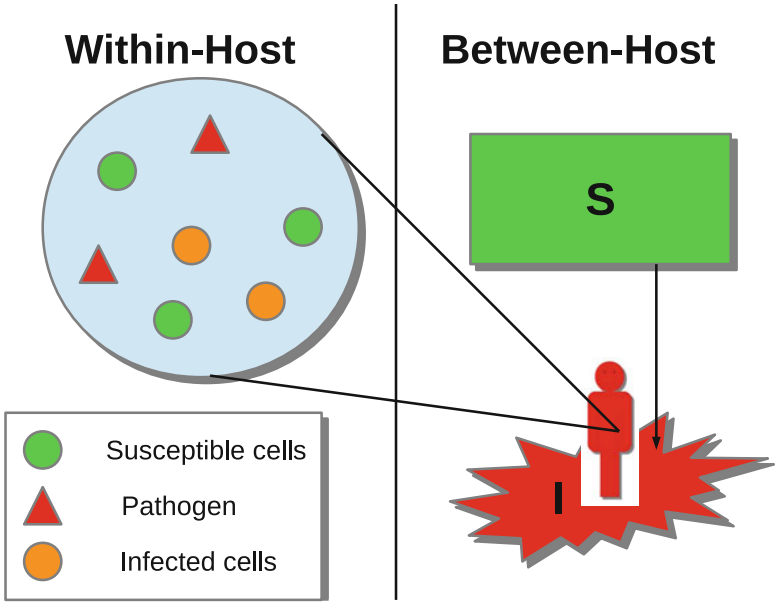

Fig. 14.1 Diagram of nesting within-host and between-host scales

the time-since-infection. The rationale for this is that particularly for HIV, infectivity is actually dependent on the within-host viral load, and since this viral load is changing as the infection progresses, so should the transmission rate. In this chapter, we will make this dependence of the transmission rate on the within-host viral load explicit, connecting the epidemiological transmission rate to the pathogen load. This new class of models, called *nested immuno-epidemiological models*, will help us address a number of questions that we could not address before. For instance, how do medications, which affect specific immunological parameters, affect the between-host distribution of disease? These models are also particularly suitable for addressing questions on the evolution of virulence.

To build the immuno-epidemiological models, we need to be able to build immunological models first, so in the next section, we briefly introduce within-host processes and their dynamical modeling. Several excellent books are devoted entirely to within-host modeling [105, 127, 168].

14.2 Within-Host Modeling

Recall that a *pathogen* or infectious agent is a microorganism that causes disease in its host. Pathogens can be of many types. Pathogens include viruses, bacteria, prions, and fungi. Pathogens use host resources, such as healthy cells, to replicate and interact with the immune system of the host.

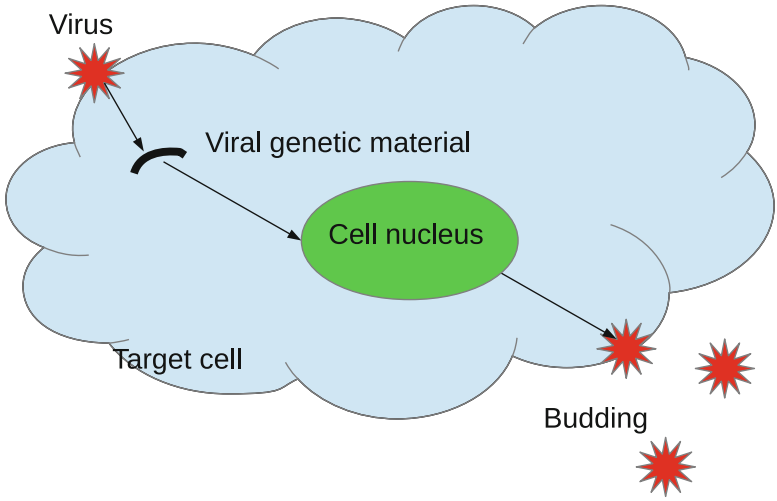

Fig. 14.2 Interaction of a virus with a target cell. The virus enters the cell and loses its envelope. The virus's genetic material interacts with the genetic material of the target cell to reproduce more copies of genetic material. New viruses are assembled and leave the cell

14.2.1 Modeling Replication of Intracellular Pathogens

Many pathogens are intracellular, which means that they need to enter a cell to reproduce. The cells that a pathogen uses to reproduce are a specific type of cells, called *target cells*. For instance, in HIV, the pathogen is the human immunodeficiency virus, and the target cells are the CD4 cells. In human tuberculosis, the pathogen is the bacterium *M. tuberculosis*, and the target cells are the macrophages. The target cells that do not contain a pathogen will be called susceptible cells. The target cells that contain pathogen are called infected cells. The newly produced pathogen can exit the host cell either through *budding* out of the cell membrane or through destroying the cell, a process called *lysis*. The pathogen can exist in a free state and interact with the host immune system (Fig. 14.2).

This process can be modeled with a system similar to the epidemiological models. If we denote the susceptible target cells by $T(t)$, the infected target cells by $I(t)$, and the free virus by $V(t)$, then the model becomes

$$\begin{aligned} T' &= r - \rho TV - dT, \\ I' &= \rho TV - \delta I, \\ V' &= pI - cV, \end{aligned} \tag{14.1}$$

where d is the uninfected target cells' natural death rate, δ is the infected target cells' death rate, c is the clearance rate of the virus, p is the virus production rate of infected cells, ρ is the infection rate of target cells by free virus, and r is the rate of production of target cells. This model captures well the within-host dynamics of HIV and HCV. Analysis of within-host models can be performed similarly to epidemiological models. We expect again two types of equilibria: infection-free

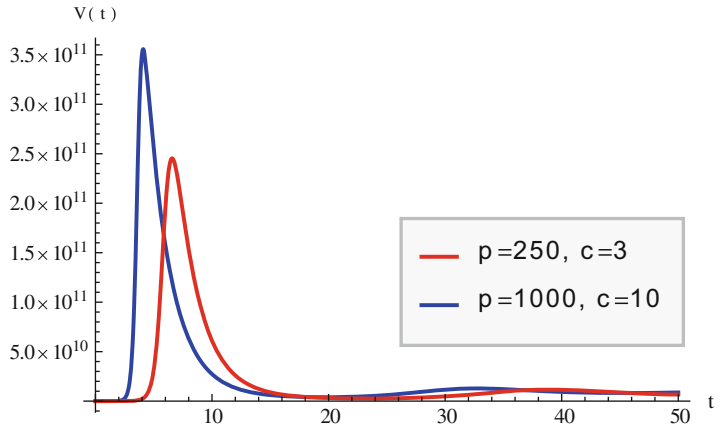

Fig. 14.3 The within-host viral load exhibits an early peak before stabilizing at equilibrium levels. Parameter values are $r = 50{,}000{,}000$, $\rho = 0.000000000015$, $d = 0.01$, $\delta = 0.5$; p and c are given in the legend. As p increases, the peak occurs sooner and is more pronounced

equilibria and infection equilibria. The infection-free equilibrium is adequate for diseases in which the virus is eventually cleared from the body, such as influenza. An infection equilibrium corresponds to chronic diseases, such as HIV and HCV. To find the equilibria, we set the right-hand side in system (14.1) to zero. Expressing I in terms of V from the last equation, $I = \frac{c}{p}V$, and replacing it in the second equation, we obtain $T = \frac{\delta c}{\rho p}$. Then, using the value of T, we obtain V from the first equation:

$$V = \frac{rp}{\delta c} - \frac{d}{\rho}.$$

In analogy with the epidemiological reproduction number, here we can define the *immunological* reproduction number:

$$\Re_0 = \frac{rp\rho}{\delta dc}.$$

The immunological reproduction number gives the number of secondary viral particles that one viral particle will produce in the entirely susceptible target cell population. The solution of the system for the equilibria is then given by

$$T^* = \frac{\delta c}{\rho p} \qquad I^* = \frac{cd}{p\rho}(\Re_0 - 1), \qquad V^* = \frac{d}{\rho}(\Re_0 - 1). \tag{14.2}$$

Model (14.1) has been completely analyzed [49]. The analysis shows that if $\Re_0 < 1$, the infection-free equilibrium is globally asymptotically stable. If $\Re_0 > 1$, the infection-free equilibrium is unstable, and the infection equilibrium is globally asymptotically stable. The model exhibits the typical peak at early infection that many pathogens display after becoming adapted to the host and beginning to replicate (see Fig. 14.3).

14.2.2 Modeling the Interaction of the Pathogen with the Immune System

The immune system is a biological structure within the host that has evolved to protect the host from disease. The immune system is very complicated and mounts a response to a foreign invader that is both general and specific to the invader. When a pathogen enters a host, it first encounters the *innate immune system*. The innate immune system triggers the *adaptive immune system*, which mounts an antigen-specific response.

The pathogen or some part of it, called an antigen, is engulfed by APCs *antigen presenting cells* and "presented" to the T-helper cells. T-helper cells further activate the killer T-cells, which are a part of the *cellular immune response*. The killer T-cells destroy the cells already infected by the pathogen. B-cells are also activated further in the immune response. B cells, which are a part of the *humoral immune response*, recognize the whole pathogen without presentation. The B cells engulf the pathogen and decompose it, expressing certain parts of it on its surface. As the B cells begin to divide, they secrete antibodies that circulate in the blood and mark the pathogen and the infected target cells for destruction. Schematically, the main players in the immune response to a pathogen are described in Fig. 14.4. APC, such as macrophages and dendritic cells, help initiate the adaptive immune response. They have a dual role. On the one hand, they engulf and then digest cellular parts and pathogens. On the other, they present the pathogen to the T-helper cells, thus activating the corresponding T-helper cells.

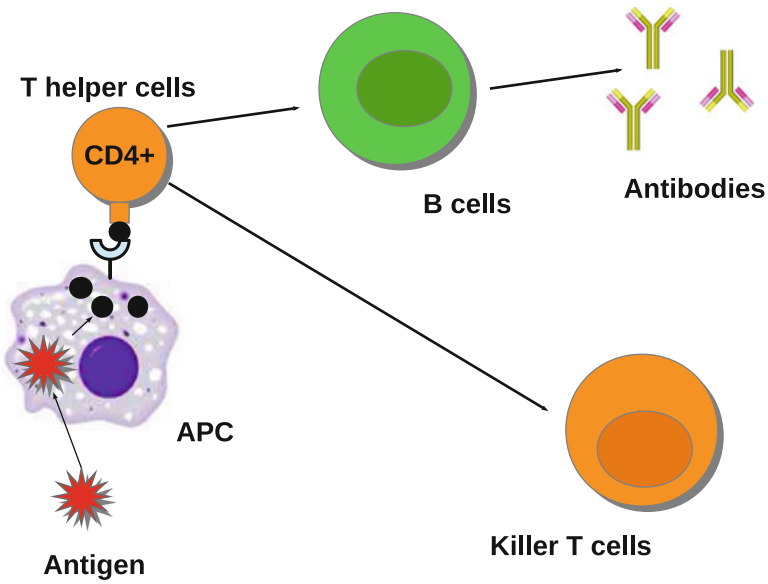

Fig. 14.4 The activation of the immune system by a pathogen

Mathematical models of the interplay of pathogens with the immune system vary from very simple to quite complicated. One of the simplest models was proposed by Gilchrist and Sasaki [65], which includes only the pathogen $P(t)$ and the immune cells $B(t)$:

$$\begin{aligned} P' &= rP - cBP, \\ B' &= aBP, \end{aligned} \tag{14.3}$$

where r is the pathogen reproduction rate, c is the pathogen clearance rate by the B-cells, and a is the stimulation of B-cell production by the pathogen. This model is very simple, and it can be solved, although an implicit solution is obtained. Problem 14.1 asks you to do that. The model has a set of equilibria with $P^* = 0$ and B^* arbitrary: $\mathscr{E}_0 = (0, B^*)$. Over the long term, the pathogen is always cleared, so the model mimics only acute infection and recovery. This is illustrated in Fig. 14.5. The outcome of the dynamics is sensitive to the initial conditions.

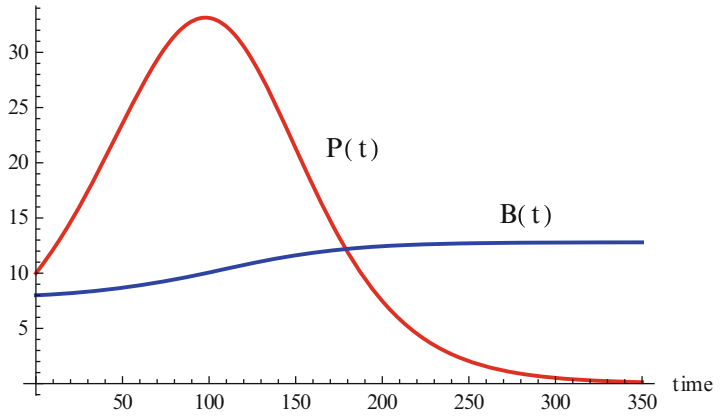

Fig. 14.5 The activation of the immune system by a pathogen given by system 14.3. Parameters are $r = 0.1$, $c = 0.01$, $a = 0.0001$, $P(0) = 10$, $B(0) = 8$

Mohtashemi and Levins [95] considered two other features of the immune system: the spontaneous production of specific cells and their decay. In this case, the model can capture both acute and chronic infection:

$$\begin{aligned} P' &= rP - cBP, \\ B' &= kP - \delta B + h, \end{aligned} \tag{14.4}$$

where h is the constant production of specific cells, and δ is their clearance rate. The production of B cells is assumed proportional to the pathogen with constant of proportionality k. Model (14.4) has two equilibria: a pathogen-free equilibrium $\mathscr{E}_0 = (0, h/\delta)$ and a coexistence equilibrium. The coexistence equilibrium is present if and only if the immunological reproduction number

$$\mathfrak{R}_0 = \frac{\delta r}{hc}$$

is greater than one: $\Re_0 > 1$. The coexistence equilibrium is given by $\mathscr{E}^* = (P^*, B^*)$, where

$$P^* = \frac{h}{k}(\Re_0 - 1) \qquad B^* = \frac{r}{c}.$$

It can be shown that the coexistence equilibrium is globally stable. More immunological models can be found in the problems section.

14.2.3 Combining Intracellular Pathogen Replication and Immune Response

More complex within-host models combine the intracellular replication of the pathogen and the immune response. The model is based on model (14.1) but includes the killer T-cells, which destroy the infected target cells, and antibodies that help destroy the free pathogen. Both immune responses are stimulated by the pathogen, so they function in competition. The model below was presented in [168]. We denote again by T the target cells, by I the infected target cells, and by V the virus. Killer T-cells are denoted by Z, and antibodies are denoted by A:

$$\begin{aligned} T' &= r - \rho TV - dT, \\ I' &= \rho TV - \delta I - \psi IZ, \\ V' &= pI - cV - qVA, \\ Z' &= aIZ - \mu Z, \\ A' &= bVA - \nu A. \end{aligned} \tag{14.5}$$

Infected target cells are killed at a rate ψIZ, and pathogen is destroyed at a rate qVA. The killer T-cells' response is stimulated at a rate aIZ, and the antibodies are stimulated at a rate bVA.

Equilibria of model (14.5) are solutions of the system obtained by setting the right-hand side of system (14.5) to zero. We have a number of solutions of this system. First, we have the infection-free, immune-response-free equilibrium $\mathscr{E}_0 = (\frac{r}{d}, 0, 0, 0, 0)$. Infection develops when the immune reproduction number is greater than one:

$$\Re_0 = \frac{rp\rho}{\delta dc} > 1.$$

The infection- and immune-response-free equilibrium is given by $\mathscr{E}_1 = (T_1, I_1, V_1, 0, 0)$, where

$$T_1 = \frac{\delta c}{\rho p} \qquad I_1 = \frac{cd}{p\rho}(\Re_0 - 1) \qquad V_1 = \frac{d}{\rho}(\Re_0 - 1).$$

The system has three more solutions: one in which the killer T-cells are present but the antibody response is not, another in which the antibody response is present but the killer T-cells are not, and finally a third one, in which both immune responses are present. To define the killer T-cells' immune response, we define a killer T-cell's *invasion number* (see Chap. 8)

$$\Re_1 = \frac{p\rho}{c\delta}\frac{rca}{\rho p\mu + dca}.$$

If $\Re_1 > 1$, the infection, killer T-cell immune response equilibrium is present: $\mathscr{E}_2 = (T_2, I_2, V_2, Z_2, 0)$, where

$$T_2 = \frac{r}{\rho V_2 + d} = \frac{rca}{\rho p\mu + dca}, \qquad I_2 = \frac{\mu}{a} \qquad V_2 = \frac{p\mu}{ca} \qquad Z_2 = \frac{\delta}{\psi}(\Re_1 - 1).$$

When the antibody response is present but the killer T-cell response is not, we obtain another equilibrium. The infection, antibody immune response equilibrium exists if the antibody invasion number $\Re_2$ is greater than 1, where

$$\Re_2 = \frac{p\rho}{c\delta}\frac{rb}{\rho\nu + bd}.$$

In this case, we obtain the equilibrium $\mathscr{E}_3 = (T_3, I_3, V_3, 0, A_3)$, where

$$T_3 = \frac{r}{\rho V_3 + d} = \frac{rb}{\rho\nu + bd}, \qquad I_3 = \frac{\rho\nu}{\delta b}\frac{rb}{\rho\nu + bd}, \qquad V_3 = \frac{\nu}{b}, \qquad A_3 = \frac{c}{q}(\Re_2 - 1).$$

The final equilibrium is an infection, killer T-cells, and antibody response equilibrium with all components present. This equilibrium exists if the killer T-cells' invasion number in the presence of antibodies $\Re_3$ and the antibody invasion number in the presence of killer T-cells $\Re_4$ are both greater than one. The two invasion numbers are defined as follows:

$$\Re_3 = \frac{\rho}{\delta}\frac{\nu a}{b\mu}\frac{rb}{\rho\nu + bd}, \qquad \Re_4 = \frac{p}{c}\frac{\mu b}{a\nu}.$$

We note that $\Re_3\Re_4 = \Re_2$. The infection, killer T-cells, and antibody response equilibrium $\mathscr{E}_4 = (T_4, I_4, V_4, Z_4, A_4)$ is given by

$$\begin{aligned} T_4 &= \frac{rb}{\rho\nu + bd}, \qquad I_4 = \frac{\mu}{a}, \qquad V_4 = \frac{\nu}{b}, \\ Z_4 &= \frac{\delta}{\psi}(\Re_3 - 1), \qquad A_4 = \frac{c}{q}(\Re_4 - 1). \end{aligned} \tag{14.6}$$

The complete analysis of model (14.5) is complex. No oscillations have been found in this model [169].

14.3 Nested Immuno-Epidemiological Models

Each of the models (14.1), (14.3), (14.4), and (14.5) as well as any other immunological model can be coupled with an appropriate epidemiological model to form a two-scale immuno-epidemiological model. The model that will be formulated and used depends on the disease we want to study as well as the question we would like to address. If we would like to study the impact of medications, we need to use model (14.1), since most medications impact the production of the virus. If we

would like to study coevolution, we need to use model (14.3) or model (14.4), since they involve the two variables that evolve in the parasites and the host: the reproduction rate of the parasite r and the immune response of the host a. Coupling model (14.5) with an appropriate epidemiological model may allow us to study both the impact of drugs on the epidemiology of the disease and the coevolution of parasites and hosts. In the next subsection, we will compose a model that will allow us to study the epidemiology of the disease as well as the evolution of the parasite.

14.3.1 Building a Nested Immuno-Epidemiological Model

The first nested immuno-epidemiological model was proposed by Gilchrist and Sasaki [65]. Their goal was to study the coevolution of pathogens and hosts. The idea is simple. The infected individuals in the population are structured by time and time-since-infection $i(\tau,t)$, where τ, the time-since-infection, is the independent variable in the immunological model. We take model (14.1) with τ as independent variable. Then $T(\tau)$, $I(\tau)$, $V(\tau)$ are functions of τ. We embed (14.1) into an epidemiological model, say of HIV. We take a simple SI with time-since-infection:

$$\begin{aligned} S' &= \Lambda - \frac{S}{N}\int_0^\infty \beta(V(\tau))i(\tau,t)d\tau - m_0 S, \\ i_\tau + i_t &= -m(V(\tau))i, \\ i(0,t) &= \frac{S}{N}\int_0^\infty \beta(V(\tau))i(\tau,t)d\tau, \end{aligned} \tag{14.7}$$

where $\beta(\tau)$ is the transmission rate and $m(\tau)$ is the death rate; N is the total population size, and m_0 is the natural death rate:

$$N = S + \int_0^\infty i(\tau,t)d\tau.$$

The second step in linking the immunological and epidemiological models is to link the epidemiological parameters to the immunological variables. The simplest scenario will be to assume the epidemiological parameters proportional to the viral load. In particular,

$$\beta(V(\tau)) = bV(\tau) \qquad m(V(\tau)) = m_0 + m_1 V(\tau). \tag{14.8}$$

The immuno-epidemiological model (14.1)–(14.7) is perhaps one of the simplest models of this type. It has multiple drawbacks:

1. It assumes that all individuals in the population experience the same within-host dynamics. This problem can be remedied if multiple groups of infected individuals are included in the population.
2. The immunological model does not experience the growth of the viral load associated with AIDS. This problem can be remedied if an AIDS compartment is included in the epidemiological model:

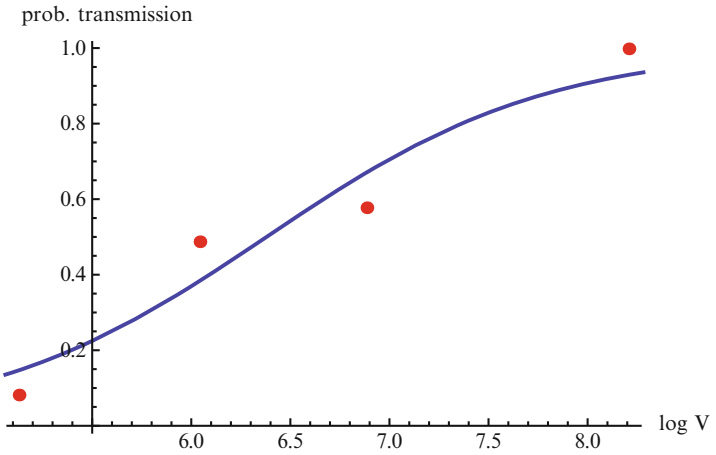

Fig. 14.6 Data and fit of the function $1/(1+Be^{-rx\ln 10})$ to the data, where $x = \log V$. Data taken from [92]

$$\begin{aligned} S' &= \Lambda - \frac{S}{N}\int_0^\infty \beta(V(\tau))i(\tau,t)d\tau - m_0 S, \\ i_\tau + i_t &= -m(V(\tau))i - \gamma(\tau)i, \\ i(0,t) &= \frac{S}{N}\int_0^\infty \beta(V(\tau))i(\tau,t)d\tau, \\ A' &= \int_0^\infty \gamma(\tau)i(\tau,t)d\tau - \nu A, \end{aligned} \tag{14.9}$$

where $A(t)$ gives the number of people with AIDS. The dependence of the rate of transition to AIDS $\gamma(\tau)$ on the immunological parameters is not clear.

3. The dependence of $\beta(\tau)$ and $m(\tau)$ on the viral load as specified in Eq. (14.8) can be criticized:

   - The linear dependence of $\beta(\tau)$ is not ideal. In particular, we know that $\beta(\tau) = C(\tau)q(\tau)$, where $C(\tau)$ is the contact rate and $q(\tau)$ is the probability of transmission. In HIV, the contact rate either remains constant or decreases with the viral load, as infected individuals become progressively sicker. The probability of transmission increases with the viral load, but it is bounded by 1. Lange and Ferguson [92] found that the dependence of the risk of infection on $\log(V(\tau))$ is best fitted by an S-shaped function. We used the data in [92] to fit the function $1/(1+Be^{-rx\ln 10})$ to the data (see Fig. 14.6). The best-fitted parameters were $B = 7952.52$ and $r = 0.6$. Recalling that $x = \log_{10} V(\tau)$, we obtain the following function of the viral load:

$$q(\tau) = \frac{1}{1+BV^{-r}(\tau)} = \frac{V^r(\tau)}{V^r(\tau)+B}.$$

If the contact rate is assumed constant, the following dependence of β on the viral load is a reasonable approximation of the above function, although [92] suggests a different relationship:

$$\beta(V(\tau)) = \frac{bV(\tau)}{B+V(\tau)}, \tag{14.10}$$

where b is an appropriate constant.

- The equation given in (14.8) for $m(\tau)$ suggests that given that the viral load is maximal during the acute stage, the maximal death rate is also maximal during the acute stage, which is not the case in HIV. This problem may be remedied by the introduction of an AIDS class.

Despite its shortcomings, the immuno-epidemiological model (14.1)–(14.7) is an adequate tool to study the impact of the within-host dynamics and HIV medications on the epidemiology of the disease.

14.3.2 Analysis of the Immuno-Epidemiological Model

To connect the immunological model to the epidemiology, we compute key epidemiological quantities such as the reproduction number and the prevalence in terms of the immunological quantities. To find the equilibria of the epidemiological model, we consider the system

$$\begin{aligned} \Lambda - \frac{S}{N}\int_0^\infty \beta(V(\tau))i(\tau)d\tau - m_0 S = 0,& \\ \frac{\partial i}{\partial \tau} = -m(V(\tau))i,& \\ i(0) = \frac{S}{N}\int_0^\infty \beta(V(\tau))i(\tau)d\tau.& \end{aligned} \tag{14.11}$$

This system has the disease-free equilibrium $\mathscr{E}_0 = (\frac{\Lambda}{m_0}, 0)$. To compute the endemic equilibrium, we define

$$\pi(\tau) = e^{-\int_0^\tau m(V(\sigma))d\sigma}.$$

We recall that $\pi(\tau)$ is the probability of survival in the infectious class. We use $\pi(\tau)$ to solve the differential equation: $i(\tau) = i(0)\pi(\tau)$. We replace the solution in the third equation in (14.11). After canceling $i(0)$, we obtain the following expression for the fraction of susceptible individuals:

$$\frac{S}{N} = \frac{1}{\int_0^\infty \beta(V(\tau))\pi(\tau)d\tau}.$$

The fraction of susceptible individuals is less than one. We need the fraction on the right to be less than one. This prompts us to define the following expression as the *epidemic basic reproduction number*:

$$\mathscr{R}_0 = \int_0^\infty \beta(V(\tau))\pi(\tau)d\tau. \tag{14.12}$$

We notice that the epidemic reproduction number depends on the immunological parameters through the viral load $V(\tau)$. The epidemic basic reproduction number $\mathscr{R}_0$ also depends on the immune basic reproduction number $\mathfrak{R}_0$. It is easy to see that the endemic equilibrium $\mathscr{E}^* = (S^*, i^*(\tau))$ exists if and only if $\mathscr{R}_0 > 1$. To compute the components of the equilibrium, we first notice that

$$\frac{S^*}{N^*} + \frac{\int_0^\infty i^*(\tau)}{N^*} = 1.$$

Furthermore,

$$\frac{\int_0^\infty i^*(\tau)}{N^*} = \frac{i^*(0)\Pi}{N^*} = 1 - \frac{1}{\mathscr{R}_0},$$

where

$$\Pi = \int_0^\infty \pi(\tau)d\tau.$$

Using the first and the third equations in (14.11), we obtain $\Lambda - i^*(0) - m_0 S^* = 0$. Dividing by N^* and replacing the fractions of S^*/N^* and $i^*(0)/N^*$, we obtain the following equation for N:

$$\frac{\Lambda}{N^*} = \frac{1}{\Pi}\left(1 - \frac{1}{\mathscr{R}_0}\right) + m_0 \frac{1}{\mathscr{R}_0}.$$

Solving for N, we have

$$N^* = \frac{\Lambda \Pi \mathscr{R}_0}{\mathscr{R}_0 - 1 + \Pi m_0}.$$

The components of the equilibrium are given by

$$\begin{aligned} S^* &= \frac{\Lambda \Pi}{\mathscr{R}_0 - 1 + \Pi m_0}, \\ i^*(\tau) &= i^*(0)\pi(\tau), \\ i^*(0) &= \frac{\Lambda(\mathscr{R}_0 - 1)}{\mathscr{R}_0 - 1 + \Pi m_0}. \end{aligned} \tag{14.13}$$

Analysis of the stability of equilibria is fairly similar to that in age-since-infection models. To investigate the stability of the disease-free equilibrium, we linearize the model around the disease-free equilibrium $\mathscr{E}_0 = (S_0, 0)$. If $S(t) = S_0 + x(t)$, where $S_0 = \Lambda/m_0$, $i(\tau,t) = y(\tau,t)$, and $N(t) = S_0 + n(t)$, then we note that

$$\frac{S}{N} = \frac{S_0 + x(t)}{S_0 + n(t)} = \frac{S_0 + x(t)}{S_0(1 + n(t)/S_0)} \approx \frac{S_0 + x(t)}{S_0}\left(1 - \frac{n(t)}{S_0}\right) \approx 1 + \frac{x(t)}{S_0} - \frac{n(t)}{S_0}.$$

The linearized system becomes

$$\begin{aligned} x' &= -\int_0^\infty \beta(V(\tau))y(\tau,t)d\tau - m_0 x, \\ y_\tau + y_t &= -m(V(\tau)y(\tau,t), \\ y(0,t) &= \int_0^\infty \beta(V(\tau))y(\tau,t)d\tau. \end{aligned} \tag{14.14}$$

Since this is a linear model, we look for solutions in exponential form $x(t) = xe^{\lambda t}$ and $y(\tau,t) = y(\tau)e^{\lambda t}$. The exponential solutions satisfy the following linear eigenvalue problem:

$$\begin{aligned} \lambda x &= -\int_0^\infty \beta(V(\tau))y(\tau)d\tau - m_0 x, \\ y_\tau + \lambda y &= -m(V(\tau)y(\tau), \\ y(0) &= \int_0^\infty \beta(V(\tau))y(\tau)d\tau. \end{aligned} \tag{14.15}$$

Solving the differential equation, we have

$$y(\tau) = y(0)e^{-\lambda\tau}\pi(\tau).$$

We replace $y(\tau)$ in the initial condition; we cancel $y(0)$, which is assumed to be nonzero; and we obtain the following characteristic equation:

$$1 = \int_0^\infty \beta(V(\tau))e^{-\lambda\tau}\pi(\tau)d\tau. \tag{14.16}$$

If we denote the right-hand side of this equality by $\mathcal{G}(\lambda)$ and we assume that the real part of λ is nonnegative ($\Re\lambda \geq 0$), we see that

$$|\mathcal{G}(\lambda)| \leq \mathcal{G}(\Re\lambda) \leq \mathcal{G}(0) = \mathcal{R}_0.$$

Therefore, if $\mathcal{R}_0 < 1$ for $\Re\lambda \geq 0$, then $|\mathcal{G}(\lambda)| < 1$, so the equality cannot have a solution of this type. We conclude that all solutions of equality (14.16) have negative real parts, and the disease-free equilibrium is locally asymptotically stable. If $\mathcal{R}_0 > 1$, then $\mathcal{G}(0) > 1$. For λ real, $\mathcal{G}(\lambda)$ is a decreasing function approaching zero as $\lambda \to \infty$. We conclude that there is a unique positive λ^* such that $\mathcal{G}(\lambda^*) = 1$. This implies that the disease-free equilibrium is unstable. So the epidemiological reproduction number plays the usual role of a threshold for the stability of the disease-free equilibrium.

14.3.3 Dependence of $\mathscr{R}_0$ and Prevalence on Immunological Parameters

The reproduction number and the prevalence $\int_0^\infty i(\tau)d\tau$ depend on the immunological parameters. Since an explicit solution to model (14.1) cannot be found, this dependence is implicit. To obtain some intuition about this dependence, we will consider the very special case that the immunological reproduction number is greater than one, $\mathfrak{R}_0 > 1$, and the initial conditions of (14.1) are exactly at the infection equilibrium. Then the solution of model (14.1) is known, and it is given by the infection equilibrium (14.2). In this case, the transmission $\beta(V^*)$ and the death rate $m(V^*)$ are constant, and an explicit expression can be obtained. With $\beta(V)$ given by (14.10), we have the following epidemiological reproduction number:

$$\mathscr{R}_0 = \frac{bV^*}{B+V^*}\frac{1}{m_0+m_1V^*}, \tag{14.17}$$

where V^* is given in (14.2). The epidemiological reproduction number depends on the immunological parameters only through the equilibrial viral load V^*. To get a general idea of how the reproduction number behaves with respect to the immunological parameters, we take the derivative of $\mathscr{R}_0$ with respect to an arbitrary immunological parameter p_i:

$$\frac{\partial \mathscr{R}_0}{\partial p_i} = \frac{\partial \mathscr{R}_0}{\partial V^*}\frac{\partial V^*}{\partial p_i}.$$

Taking the derivative of $\mathscr{R}_0$ with respect to V^*, we have

$$\frac{\partial \mathscr{R}_0}{\partial V^*} = \frac{b(Bm_0 - m_1(V^*)^2)}{(B+V^*)^2(m_0+m_1V^*)^2}.$$

The derivative shows that $\mathscr{R}_0$ has a unique critical point at

$$V^*_{\text{crit}} = \sqrt{\frac{Bm_0}{m_1}}.$$

Since $\frac{\partial V^*}{\partial p_i}$ is of definite sign, positive or negative, the sign of the derivative $\frac{\partial \mathscr{R}_0}{\partial p_i}$ changes only once. The value p_i^{crit} corresponds to the critical value of V^*. When $p_i < p_i^{\text{crit}}$ and $\frac{\partial V^*}{\partial p_i} > 0$, then $V^* < V^*_{\text{crit}}$ and $\frac{\partial \mathscr{R}_0}{\partial p_i} > 0$. Similarly, if $p_i > p_i^{\text{crit}}$ and $\frac{\partial V^*}{\partial p_i} > 0$, then $V^* > V^*_{\text{crit}}$ and $\frac{\partial \mathscr{R}_0}{\partial p_i} < 0$. The same reasoning implies the same result when $\frac{\partial V^*}{\partial p_i} < 0$. This says that $\mathscr{R}_0$ has a unique maximum with respect to every immunological parameter. We note that this dependence of the epidemiological reproduction number on immunological parameters is determined by our choice of $\beta(V)$.

Table 14.1 Table of parameter values

Param.	Value	Units	Param.	Value	Units
r	6	10^7 cells/day	d	0.01	day^{-1}
ρ	0.00015	1/{virion*day}	δ	1	day^{-1}
p	1,000	virions	c	23	day^{-1}
b	1.3	day^{-1}	B[a]	8,000	virions
m_0	1/(365*70)	day^{-1}	m_1	0.00002	1/{virions*day}
Λ	275	10^3 people/day			

[a]As estimated from fitting in Fig. 14.6.

The maximum of the reproduction number is given by

$$\mathcal{R}_0^{\max} = \frac{b\sqrt{\frac{Bm_0}{m_1}}}{(B+\sqrt{\frac{Bm_0}{m_1}})(m_0+m_1\sqrt{\frac{Bm_0}{m_1}})}.$$

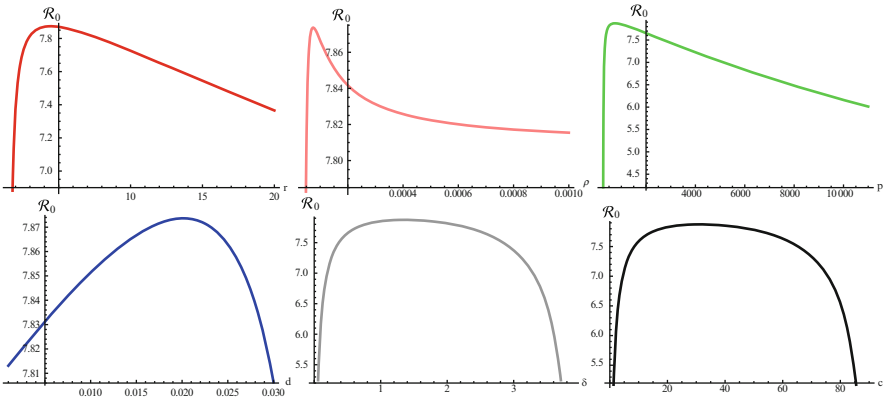

Fig. 14.7 This figure gives the dependence of the epidemiological reproduction number on the within-host parameters. The *first row* gives dependence with respect to which the immunological reproduction number is increasing. The *second row* gives the dependence with respect to which the immunological reproduction number is decreasing. Other parameters are fixed as in Table 14.1

One nonobvious conclusion of this expression is that the maximum of the reproduction number depends only on the epidemiological parameters and not on the within-host parameters. To plot the dependence of the epidemiological reproduction number on the immunological parameters, we determine reasonable parameter values for HIV. These are taken from [110] and are given in Table 14.1.

The dependence of the reproduction number on the immunological parameters is given graphically in Fig. 14.7. All graphs exhibit approximately the same behavior. They grow, reach a maximal value, and then decline. Because of this dependence of

the reproduction number on immunological parameters, it is said that the pathogen evolves to maximize the epidemiological reproduction number.

The prevalence is given by

$$\int_0^\infty i^*(\tau)d\tau = \frac{\Lambda(\mathscr{R}_0 - 1)}{\frac{bV^*}{B+V^*} - m_1 V^*}, \tag{14.18}$$

where $\mathscr{R}_0$ is given in (14.17). The prevalence is graphed in Fig. 14.8. The prevalence is in general a decreasing function with respect to immune parameters for which the immunological reproduction number is an increasing function and an increasing function with respect to immune parameters for which the immunological reproduction number is decreasing. The prevalence experiences a sudden rise close to the critical value at which the immunological reproduction number becomes equal to one. That may be an artifact of the fact that the dependence on the immunological parameters is defined only for $\mathfrak{R}_0 > 1$, so when $\mathfrak{R}_0$ approaches one from the right-hand side, the prevalence collapses and becomes negative.

Protease inhibitors are medications used in treatment of HIV. They reduce the number of infectious virions produced. In other words, they reduce p. As p decreases, within-host viral load decreases, but that increases the prevalence in the population, as Fig. 14.8 suggests. The phenomenon that HIV medications *increase*

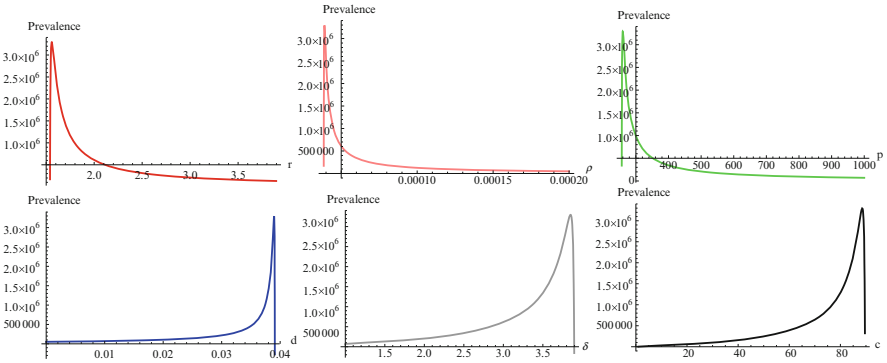

Fig. 14.8 This figure gives the dependence of the prevalence on the within-host parameters. The *first row* gives dependence on parameters with respect to which the immunological reproduction number is increasing. The *second row* gives the dependence on parameters with respect to which the immunological reproduction number is decreasing. Other parameters are fixed as in Table 14.1

population prevalence, as counterintuitive as it may seem, is not without a basis in reality. Table 6.2 gives the world prevalence of HIV. The numbers clearly show that prevalence has been increasing ever since HIV became a public health problem.

Similar results about the dependence of $\mathscr{R}_0$ and prevalence on the within-host parameters can be obtained if the full immuno-epidemiological model is used with τ-dependent viral load. In this case, the dependence of $\mathscr{R}_0$ and prevalence on the within-host parameters cannot be explicitly obtained, but it must be simulated. In addition, dependence of these key epidemiological quantities on the initial parameters of the immune model, $T(0)$, $V(0)$, and $I(0)$, can be obtained.

14.3.4 Sensitivity and Elasticity of $\mathscr{R}_0$ and Prevalence with Respect to Immunological Parameters

Recall that the *sensitivity* of a quantity Q with respect to a parameter p_i is defined as $\partial Q/\partial p_i$. The sensitivity has the drawback that it is not rescaled and does not give the change of the quantity Q relative to the size of the quantity. To address this issue, *elasticity* is defined as

$$\frac{\partial Q}{\partial p_i}\frac{p_i}{Q}.$$

When Q is a simple function of the parameter p_i, computing elasticities is not difficult. The elasticity of the epidemiological reproduction number $\mathscr{R}_0$ given by expression (14.17) is given by

$$\frac{\partial \mathscr{R}_0}{\partial p_i}\frac{p_i}{\mathscr{R}_0} = \frac{b(Bm_0 - m_1(V^*)^2)}{(B+V^*)^2(m_0+m_1V^*)^2}\frac{\partial V^*}{\partial p_i}\frac{p_i}{\mathscr{R}_0}.$$

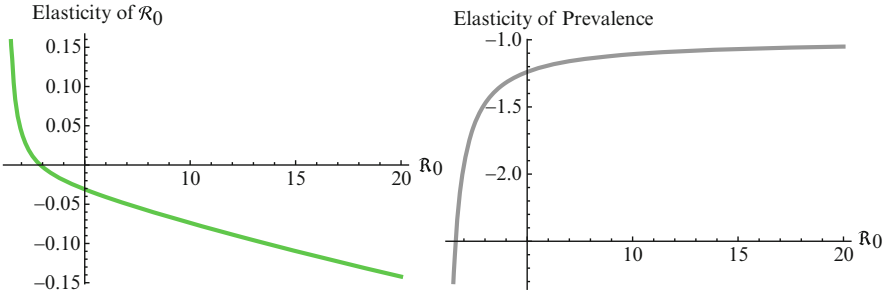

Fig. 14.9 This figure gives the elasticities of $\mathscr{R}_0$ (*left figure*) and the prevalence (*right figure*) with respect to the immunological reproduction number $\Re_0$ in the case that those are given by formulas (14.17) and (14.18). Parameters are fixed as in Table 14.1

Computing the explicit elasticity of the prevalence is more complicated, but it could be done in the same way. The elasticities of the epidemiological reproduction number and the prevalence with respect to the immunological reproduction number are given in Fig. 14.9. Both $\mathscr{R}_0$ and the prevalence are most elastic with respect to the immunological reproduction number when the immunological reproduction number $\Re_0$ is close to the threshold value 1. As $\Re_0$ becomes larger, the magnitude of the elasticity declines. The epidemiological reproduction number remains less elastic to $\Re_0$ than the prevalence. The higher elasticity of $\mathscr{R}_0$ and the prevalence when $\Re_0 \to 1_+$ suggests that medications, which work to decrease the within-host reproduction number, also increase the elasticity of the epidemiological reproduction number and the prevalence. The smaller the value of $\Re_0$, the greater the effect on the epidemiological reproduction number and the prevalence.

Computing the elasticities with respect to the equilibrium of the immune system is valuable and illuminating, but it is not sufficient for understanding the full dependence on the age-since-infection and the impact of the initial viral load that is transferred, $V(0)$. So how do we compute the elasticities of $\mathscr{R}_0$ and the prevalence if they are not functions, but functionals, of the immunological parameters? The main difficulty lies in computing $\partial \mathscr{R}_0/\partial p_i$, where $\mathscr{R}_0$ is given by formula (14.12). Replacing $\beta(V)$ with its value from (14.10), we have the following expression:

$$\mathscr{R}_0 = \int_0^\infty \frac{bV(\tau)}{B+V(\tau)} e^{-m_0\tau - m_1\int_0^\tau V(\sigma)d\sigma} d\tau. \tag{14.19}$$

We will illustrate the computation of $\partial R_0/\partial p$, where p is the immunological parameter in the equation for V. In the formula for $\mathscr{R}_0$, only $V(\tau)$ depends on p, so we differentiate with respect to V and then with respect to p:

$$\begin{aligned}\frac{\partial \mathscr{R}_0}{\partial p} = &\int_0^\infty \frac{bB}{(B+V(\tau))^2}\frac{\partial V(\tau)}{\partial p} e^{-m_0\tau - m_1\int_0^\tau V(\sigma)d\sigma} d\tau \\ &-m_1\int_0^\infty \frac{bV(\tau)}{B+V(\tau)}\int_0^\tau \frac{\partial V(\sigma)}{\partial p} d\sigma e^{-m_0\tau - m_1\int_0^\tau V(\sigma)d\sigma} d\tau.\end{aligned} \tag{14.20}$$

To compute this derivative of the epidemiological reproduction number, we need $\frac{\partial V(\tau)}{\partial p}$. This last derivative can be obtained by differentiating model (14.1) with respect to p. In this case, we will have the derivatives of T, I, and V with respect to p. We will use the following notation:

$$X(\tau) = \frac{\partial T(\tau)}{\partial p}, \qquad Y(\tau) = \frac{\partial I(\tau)}{\partial p}, \qquad Z(\tau) = \frac{\partial V(\tau)}{\partial p}. \tag{14.21}$$

Differentiation (14.1) with respect to p and coupling the original equations with the equations for X, Y, Z, we obtain the following system:

$$\begin{aligned} T' &= r - \rho TV - dT, \\ I' &= \rho TV - \delta I, \\ V' &= pI - cV, \\ X' &= -\rho XV - \rho TZ - dX, \\ Y' &= \rho XV + \rho TZ - \delta Y, \\ Z' &= I + pY - cZ, \end{aligned} \tag{14.22}$$

where $'$ denotes the derivative with respect to τ. Solving this system will give us $Z(\tau)$, or $\frac{\partial V(\tau)}{\partial p}$, which we may use to compute the elasticity of $\mathscr{R}_0$ or the prevalence.

14.4 A Nested Immuno-Epidemiological Model with Immune Response

In this section, we consider the role of the immune response in the dependence of the epidemiological reproduction number on immune parameters, and particularly on those that are affected by medication. We will consider a simplified version of model (14.1). In particular, we will assume that the free virus equilibrates much faster and that $\frac{dV}{d\tau} \approx 0$, which would imply that the free virus is proportional to the number of infected cells:

$$V \approx \frac{p}{c} I.$$

Substituting this approximation of the free virus into the remaining equations, we may eliminate the free virus compartment. We further modify model (14.1) by including the number of killer T cells $Z(\tau)$ that attack and destroy the infected target cells. Since in HIV, the target cells are CD4 cells, which are also helper T cells, the new killer T-cells depend on the quantity of infected cells $I(\tau)$ and helper T-cells $T(\tau)$. The within-host model becomes

$$\begin{aligned} T' &= r - \rho TI - dT, \\ I' &= \rho TI - \delta I - \psi IZ, \\ Z' &= aTI - \mu Z, \end{aligned} \tag{14.23}$$

where μ is the natural death rate of the killer T-cells, ψ is the killing rate, and all other parameters have the same meaning as before. To account for the immune response in the epidemiological model, we assume that the onset of immune response causes adverse effects in the host, which increases the virulence of the disease. We modify the epidemiological model accordingly:

$$\begin{aligned} S' &= \Lambda - \frac{S}{N} \int_0^\infty \beta(I(\tau)) i(\tau,t) d\tau - m_0 S, \\ i_\tau + i_t &= -m(I(\tau))i - \eta(Z(\tau))i, \\ i(0,t) &= \frac{S}{N} \int_0^\infty \beta(I(\tau)) i(\tau,t) d\tau, \end{aligned} \tag{14.24}$$

where $\eta(Z(\tau))$ is the additional host mortality incurred by the onset of the immune response. The epidemiological parameters are linked to the within-host variables as before: $m(I(\tau)) = m_0 + m_1 I(\tau)$, $\beta(I(\tau)) = bI(\tau)/(B + I(\tau))$. The new epidemiological parameter η is linked as follows:

$$\eta(Z(\tau)) = m_2 \max\left\{\frac{dZ}{d\tau}, 0\right\}.$$

First, we investigate the within-host model. The equilibria of the model satisfy the system

$$\begin{aligned} r - \rho TI - dT &= 0, \\ \rho TI - \delta I - \psi IZ &= 0, \\ aTI - \mu Z &= 0. \end{aligned} \tag{14.25}$$

The system has an infection-free equilibrium $\mathscr{E}_0 = (\frac{r}{d}, 0, 0)$. The system also has an infection equilibrium $\mathscr{E}^* = (T^*, I^*, Z^*)$. To find the values, we notice that we can cancel I in the second equation and use the third equation to eliminate I in the first equation to obtain the following reduced system:

$$\begin{aligned} r - \frac{\rho\mu}{a} Z - dT &= 0, \\ \rho T - \delta - \psi Z &= 0. \end{aligned} \tag{14.26}$$

Solving this system for T^* and Z^*, we obtain

$$T^* = \frac{\delta\rho\mu + \psi ar}{\rho^2\mu + \psi ad} \qquad Z^* = \frac{a(r\rho - d\delta)}{\rho^2\mu + \psi ad}. \tag{14.27}$$

The expression for Z^* is not always positive. This prompts us to define the immunological reproduction number as

$$\Re_0 = \frac{r\rho}{d\delta}.$$

So $Z^* > 0$ if and only if $\Re_0 > 1$. To obtain I^*, we use the third equation in (14.25) and replace T^* and Z^*. We obtain

$$I^* = \frac{\mu(r\rho - d\delta)}{\delta\rho\mu + \psi ar}.$$

To see the dependence of Z^* on the infection rate ρ, we plot Z^* against ρ for several values of a in Fig. 14.10. Figure 14.10 shows that for small infection rate ρ, the immune response becomes more and more effective as the infection rate increases, so a better reproduction of the pathogen stimulates a better immune response. However, as the infection rate increases past a critical size, the immune response is less and

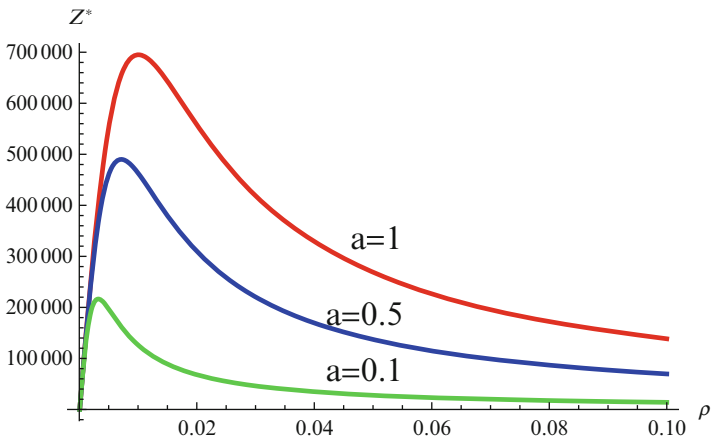

Fig. 14.10 Dependence of the equilibrial killer T-cell population on ρ for different values of a. Parameters as given in Table 14.2

Table 14.2 Table of parameter values

Param.	Value	Units	Param.	Value	Units
r	140	10^5 cells/day	d	0.01	day^{-1}
ρ	0.00015	1/(virion*day)	δ	1	day^{-1}
a	0.005	1/(cells*day)	μ	0.01	day^{-1}
b	1.3	day^{-1}	B[a]	8,000	virions
m_0	1/(365*70)	day^{-1}	m_1	0.00002	1/{virions*day}
Λ	275	10^3 people/day	ψ	0.0001	1/(cell*day)
Z_0	$0.1*10^{-10}$	10^5 cells			

[a] As estimated from fitting in Fig. 14.6.

less effective. This scenario suggests that the immune system may become "exhausted" if the pathogen is reproducing too rapidly. To compute the epidemiological reproduction number and its dependence on the within-host parameters, we further simplify system (14.23) by assuming that the target cell population has reached an equilibrium and the only dynamic variable that remains is the immune response. Hence, $dT/d\tau \approx 0$ and $dI/d\tau \approx 0$. These are simplifying assumptions, and the resulting new model does not necessarily have the dynamic properties of the original one. From the first two equations, we have

$$\begin{aligned} r - \rho TI - dT &= 0, \\ \rho TI - \delta I - \psi IZ &= 0. \end{aligned} \tag{14.28}$$

Using these two equations, we express T and I in terms of Z. From the second equation, we have

$$T = \frac{\delta + \psi Z}{\rho}.$$

From the first equation, we have $TI = (r - dT)/\rho$. We can obtain TI in terms of Z:

$$TI = \frac{1}{\rho}(r - dT) = \frac{1}{\rho}\left(r - \frac{d}{\rho}(\delta + \psi Z)\right) = \frac{1}{\rho^2}(r\rho - \delta d - d\psi Z).$$

Replacing this expression in the dynamic equation for Z, we obtain

$$Z' = \frac{a}{\rho^2}(r\rho - d\delta - d\psi Z) - \mu Z = \frac{a}{\rho^2}(r\rho - d\delta) - \frac{ad\psi + \mu\rho^2}{\rho^2}Z = \nu(Z^* - Z), \tag{14.29}$$

where $\nu = \frac{ad\psi + \mu\rho^2}{\rho^2}$. We will assume that the initial condition Z_0 for this equation satisfies $Z_0 < Z^*$. In this case, $Z(\tau) < Z^*$. The solution to Eq. (14.29) is given by

$$Z(\tau) = Z_0 e^{-\nu\tau} + Z^*(1 - e^{-\nu\tau}).$$

Hence,

$$Z(\tau) - Z_0 = (Z^* - Z_0)(1 - e^{-\nu\tau}). \tag{14.30}$$

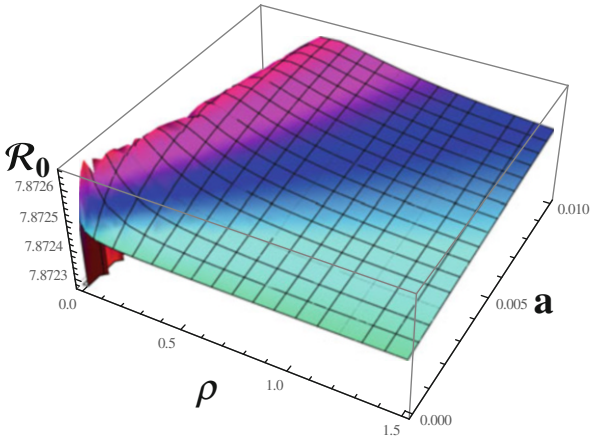

Fig. 14.11 Dependence of the epidemiological reproduction number $\mathscr{R}_0$ on ρ and a. Parameters as given in Table 14.2

We will need this expression in the epidemiological reproduction number. In analogy with the epidemiological reproduction number (14.19), the present epidemiological reproduction number is given by

$$\mathscr{R}_0 = \int_0^\infty \frac{bI(\tau)}{B+I(\tau)} e^{-m_0\tau} e^{-m_1\int_0^\tau I(\sigma)d\sigma} e^{-m_2\int_0^\tau Z'(\sigma)d\sigma} d\tau. \tag{14.31}$$

To gain insight into the dependence of the epidemiological reproduction number on the immunological parameters, we consider the approximation $I(\tau) \approx I^*$ and $Z(\tau)$ as given by (14.30). Integrating the derivative of Z, we have

$$\mathscr{R}_0 = \int_0^\infty \frac{bI^*}{B+I^*} e^{-m_0\tau} e^{-m_1 I^*\tau} e^{-m_2(Z^*-Z_0)(1-e^{-\nu\tau})} d\tau. \tag{14.32}$$

We go one step further, and we approximate $(1-e^{-\nu\tau}) \approx \nu\tau$. In this case, the expression under the main integral can be integrated, and we obtain the following explicit dependence of $\mathscr{R}_0$ on the within-host parameters:

$$\mathscr{R}_0 \approx \frac{bI^*}{B+I^*} \frac{1}{m_0 + m_1 I^* + m_2(Z^*-Z_0)\nu}. \tag{14.33}$$

The dependence of $\mathscr{R}_0$ on both ρ and a is given in Fig. 14.11. The dependence of $\mathscr{R}_0$ on ρ for several values of a is given in Fig. 14.12. It is easy to see that $\mathscr{R}_0$ is first increasing with ρ and then decreasing. As the value of a increases, that is, as the immune response becomes more vigorous, the maximum $\mathscr{R}_0^{\max}$ becomes smaller and shifts to the right. This means that the maximum $\mathscr{R}_0^{\max}$ is achieved for a larger value of ρ. So the more intensive the immune response, the more vigorously pathogens have to reproduce to maximize their epidemiological reproduction number. Even more surprising is the dependence of $\mathscr{R}_0$ on the immune response. In principle, one

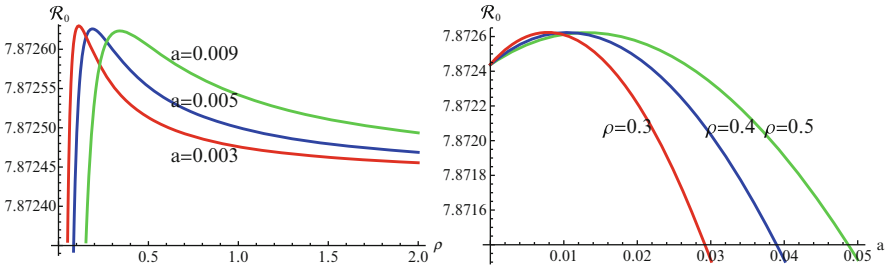

Fig. 14.12 Dependence of the epidemiological reproduction number $\mathscr{R}_0$ on ρ for several values of a (*left*) and dependence of the epidemiological reproduction number $\mathscr{R}_0$ on a for several values of ρ (*right*). Parameters as given in Table 14.2

expects that $\mathscr{R}_0$ is a decreasing function of a; however, in this case, a week immune response may benefit the pathogen and increase the epidemiological reproduction number.

Problems

14.1. Acute Infection Immune Model

(a) Find an implicit solution of model (14.3).
(b) Plot the solutions as a function of time with parameters $r = 1, c = 0.01, a = 0.001$.

14.2. Immune Model with Oscillations

The following simple immune model with specific and aspecific immunity was proposed in [135]:

$$\begin{aligned} P' &= rP - \frac{cBP}{1+k_cP} - \frac{mP}{1+k_mP}, \\ B' &= \frac{aPB}{1+k_aP} - \delta B + h, \end{aligned} \tag{14.34}$$

where m is the level of aspecific immunity, k_c and k_m modulate the functional response, k_a allows for different rules of immune response. The other parameters are as in the text.

(a) Compute the equilibria and the immune reproduction number.
(b) Investigate the stability of the pathogen-free equilibrium.
(c) Investigate the stability of the coexistence equilibrium.
(d) Simulate the model and show the sustained oscillations.

14.3. Immune Model
Consider the following simple immune model with infected cells I and killer T cells Z:

$$\begin{aligned} I' &= \frac{rI}{1+aZ} - \mu I - cIZ, \\ Z' &= \frac{bIZ}{1+kI} - \delta Z, \end{aligned} \tag{14.35}$$

where a and k modulate the functional response, c is the killing of infected cells by the immune response, b is the stimulation of the immune response by the infected cells, δ and μ are natural clearance rates.

(a) Compute the equilibria and the conditions for their existence.
(b) Investigate the stability of the pathogen-free equilibrium. Compute the immune reproduction number.
(c) Investigate the stability of the coexistence equilibrium.

14.4. Immune Model
Consider the following simple immune model with infected cells I and killer T cells Z:

$$\begin{aligned} I' &= \frac{rI}{1+aZ} - \mu I - \frac{cIZ}{1+mI}, \\ Z' &= \frac{bIZ}{1+kI} - \delta Z, \end{aligned} \tag{14.36}$$

where a, k, and m modulate the functional response, c is the killing of infected cells by the immune response, b is the stimulation of the immune response by the infected cells, δ and μ are natural clearance rates.

(a) Compute the equilibria and the conditions for their existence.
(b) Investigate the stability of the pathogen-free equilibrium. Compute the immune reproduction number.
(c) Investigate the stability of the coexistence equilibrium.

14.5. Pathogen Replication Model
Consider the following simple model that has been proposed to study HCV [48]:

$$\begin{aligned} T' &= s + r_1 T\left(1 - \frac{T+I}{T_{\max}}\right) - dT - \beta TV, \\ I' &= \beta TV + r_2 I\left(1 - \frac{T+I}{T_{\max}}\right) - \delta I, \\ V' &= pI - cV, \end{aligned} \tag{14.37}$$

where s, r_1, and r_2 are replication rates, c is the clearing of the virus, p is the virus production rate, δ is the death rate of infected cells, $T_{\max}$ is the maximum number of target cells, and β is the infection rate.

(a) Compute the equilibria and the conditions for their existence.

(b) Investigate the stability of the pathogen-free equilibrium. Compute the immune reproduction number.
(c) Show that backward bifurcation occurs in this model. Compute the stabilities of the infection equilibria.

14.6. Nested Model of Malaria
Consider the following simple model that has been proposed to study within-host transmission of malaria with antibody response:

$$\begin{aligned} R' &= r - dR - \rho RP, \\ I' &= \rho RP - \delta I, \\ P' &= pI - cV - \eta PA, \\ A' &= kPA - \mu A, \end{aligned} \tag{14.38}$$

where R is the number of red blood cells, P is the malaria parasite, I gives the number of infected red blood cells, and A is the antibody response. This model is nested into the between-host model of malaria:

$$\begin{aligned} S_H'(t) &= \Lambda_H - \beta_1 S_H(t) I_V(t) - m_0 S_H(t), \\ \frac{\partial i_H}{\partial \tau} + \frac{\partial i_H}{\partial t} &= -m(P) i_H, \\ i_H(0,t) &= \beta_1 S_H(t) I_V(t), \\ S_V'(t) &= \Lambda_V - S_V(t) \int_0^\infty \beta(P(\tau)) i_H(\tau,t) d\tau - m_V S_V(t), \\ I_V' &= S_v(t) \int_0^\infty \beta(P(\tau)) i_H(\tau,t) d\tau - m_V I_V(t), \end{aligned} \tag{14.39}$$

where we have assumed that humans are chronically infected and do not recover (that is, $\Re_0 > 1$). The parameters are linked as follows: $\beta(P(\tau)) = bP/(B+P)$ and $m(P(\tau)) = m_0 + m_1 P$.

(a) Compute the equilibria of the within-host model and the conditions for their existence.
(b) Compute the equilibria of the between-host model and the conditions for their existence.
(c) Compute the within-host and the between-host reproduction numbers.
(d) How does the epidemiological reproduction number depend on the within-host parameters?

14.7. Nested Immuno-Eco-Epidemiological Model
Consider a predator–prey model with disease in prey. Assume that we account for the within-host immunological dynamics of the prey with the following within-host model:

$$V' = rV\left(1-\frac{V}{K}\right) - \eta V z,$$
$$z' = \frac{\rho V z}{A+V} - \mu z, \tag{14.40}$$

where V is the viral load and z is the immune response; K is the carrying capacity of the virus, η is the killing rate of the virus by the immune response, ρ is the immune response rate, μ is the natural death rate of the immune cells. The epidemiological model is a predator–prey model:

$$\begin{aligned} S'(t) &= \Lambda - S(t)\int_0^\infty \beta(V)i(\tau,t)d\tau - \frac{a_1 S(t)P(t)}{1+cS(t)} - m_0 S(t),\\ \frac{\partial i}{\partial \tau} + \frac{\partial i}{\partial t} &= -m(V)i - \alpha P(t) i(\tau,t),\\ i(0,t) &= S(t)\int_0^\infty \beta(V)i(\tau,t)d\tau,\\ P' &= \frac{a_2 S(t)P(t)}{1+cS(t)} - dP(t) + k\alpha P(t)\int_0^\infty i(\tau,t)d\tau. \end{aligned} \tag{14.41}$$

The parameters are linked as follows: $\beta(V(\tau)) = bV/(B+V)$ and $m(V(\tau)) = m_0 + m_1 V$.

(a) Compute the equilibria of the within-host model and the conditions for their existence.
(b) Compute the equilibria of the between-host model and the conditions for their existence.
(c) Compute the within-host and the between-host reproduction numbers.
(d) How does the epidemiological reproduction number depend on the within-host parameters?

Chapter 15
Spatial Heterogeneity in Epidemiological Models

15.1 Introduction

Classical epidemic models assume that the entire population lives in one area and is well mixed. That assumption is not necessarily satisfied. For instance, the population may be living on isolated islands or in different countries while traveling from one location to another. This spatial heterogeneity affects the transmission of the disease. To understand the precise impact of spatial heterogeneity on the dynamics of a disease, we have to build models that account for that heterogeneity. Spatially explicit models are also more effective in evaluating control strategies, particularly those applied to movement of individuals.

Multiple modeling approaches have been used to account for space and movement of individuals in epidemic models. Space can be incorporated as a variable in discrete or continuous form. One of the most popular discrete-space modeling approaches is the *metapopulation approach* [16, 52, 102]. We will introduce the metapopulation approach in the next section. Other discrete space modeling approaches include epidemic spatial networks [20, 21, 43, 123], cellular automata [101, 147], and lattice epidemic models [137, 141]. Continuous space models include diffusion epidemic models [34, 82, 156], integrodifferential equation epidemic models [116, 136] and integrodifference equation epidemic models [8].

In this chapter, we will focus on metapopulation models and diffusion models as two types of modeling techniques incorporating movement of individuals. Each of these model types requires a different mathematical tool. Metapopulation models typically represent a large system of ordinary differential equations. Diffusion epidemic models, on the other hand, are small systems of partial differential equations.

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_15

15.2 Metapopulation Modeling of Epidemic Spread

The concept of metapopulation does not originate in epidemiology, but in ecology. A *metapopulation* [95] is a group of populations of the same species that leave in spatially isolated areas but interact on some level. Metapopulations occur when different populations live in fragmented habitats but are connected through migration. The isolated areas that are occupied by each population are called *patches*. In epidemiology, patches may be cities, countries, islands, or other geographically autonomous regions. A necessary requirement of a metapopulation is that patches be connected through migration. *Migration* is defined as physical movement of individuals from one area to another. Movement can be short-term or long-term. In short-term movement, individuals visit another location for a period of time and return to the home patch. Even though the movement is short-term, it still allows an infected individual to transmit the pathogen to a susceptible individual of a different patch, thus spreading the disease to other locations. Long-term migration arises when individuals move to another location and settle there. The two types of movement are modeled differently. Short-term movement has been called *Lagrangian movement*, and the corresponding models are called *Lagrangian metapopulation models*. Long-term movement has been called *Eulerian movement*, and the corresponding models are called *Eulerian metapopulation models* [44].

Metapopulation epidemic models consist of n patches. The population of each patch is assumed to be homogeneously mixing. It is divided into the typical epidemiological classes of susceptibles, infectives, and other classes. The sizes of each of these classes are different on different patches. Individuals of some or all of the classes travel between the patches, which leads to the movement of the disease (Fig. 15.1).

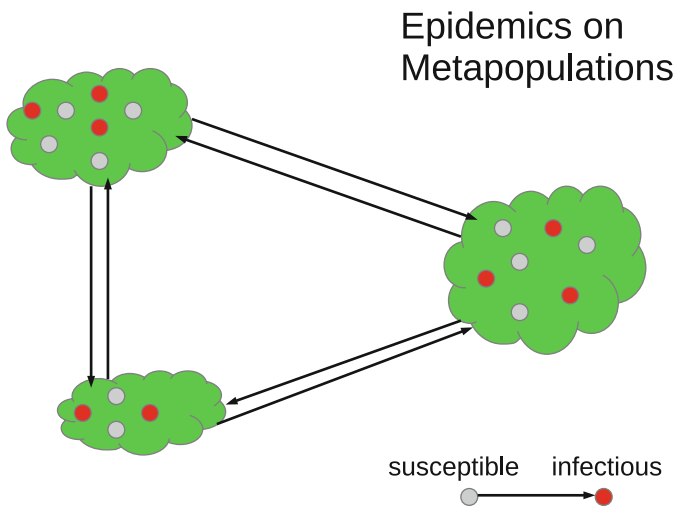

Fig. 15.1 Schematic representation of a metapopulation epidemic model

Historically, early metapopulation epidemic models incorporated short-term movement between the patches. Correspondingly, we begin by describing the Lagrangian modeling framework in the next subsection. Later, Eulerian models were developed, in which the movement is explicit and occurs at certain rates. We will consider an Eulerian modeling framework in the second subsection of this section.

15.2.1 Lagrangian Movement Epidemic Models

We begin by assuming that the total population occupies n regions or patches. The population of patch i is denoted by N_i, and the total population size is $N = N_1 + \cdots + N_n$. The population of the ith region is infected by a pathogen and consists of S_i susceptible individuals, I_i infected individuals, and R_i recovered/immune individuals. We have $N_i = S_i + I_i + R_i$. In terms of movement, we assume that the members of each region make short visits to at least some of the other regions and return to their home patch. Furthermore, we shall make the simplifying assumption that all members of each region spend the same amount of time visiting other regions, but that time depends on the visited region. While commuting to other regions, infected visitors can transmit the disease to susceptibles in the visited region, while susceptible visitors can acquire the disease from infected members of the visited region. Long-term migration of the members of this community will not be incorporated in this model. Furthermore, we assume for simplicity that the population size of each region remains constant, that is, births $\mu_i N_i$ are balanced by deaths, where μ_i is the death rate in region i. With these assumptions, the model takes the form

$$\begin{aligned} S_i' &= \mu_i N_i - S_i \sum_{j=1}^{n} \beta_{ij} I_j - \mu_i S_i, \\ I_i' &= S_i \sum_{j=1}^{n} \beta_{ij} I_j - (\mu_i + \gamma_i) I_i, \qquad i = 1, \ldots, n, \end{aligned} \tag{15.1}$$

where we have omitted the equations for the recovered. Here, γ_i represents recovery in the ith patch, and β_{ij} are the transmission rates of infected individuals from patch j to susceptible individuals in patch i. We note again that since we are assuming constant population size in each region, the birth rate equals the death rate. In addition, if we add the equations for S_i, I_i, and R_i, we obtain the following differential equation for the population size of patch i: $N_i' = 0$.

We obtain the disease-free equilibrium: $\mathscr{E}_0 = (N_1, 0, N_2, 0, \ldots, N_n, 0)$. If we imagine that commuting between patches does not occur, that is, if $\beta_{ij} = 0$ for $i \neq j$, then we can define a reproduction number for each patch:

$$\mathscr{R}_i = \frac{\beta_{ii} N_i}{\mu_i + \gamma_i}.$$

If $\mathscr{R}_i < 1$, then the disease will disappear in patch i if it is isolated from the metapopulation. If $\mathscr{R}_i > 1$, then the disease will persist in patch i if it is isolated from the metapopulation.

Definition 15.1. We call patch i a *sink* if $\mathscr{R}_i < 1$. We call patch i a *source* if $\mathscr{R}_i > 1$.

We can apply the next-generation approach to compute the reproduction number of the system, but that value is implicit and does not give any insights as to when the disease persists and when it dies out in the metapopulation. Following the next-generation approach, we have

$$F = \begin{pmatrix} \beta_{11}N_1 & \beta_{12}N_1 & \dots & \beta_{1n}N_1 \\ & & \vdots & \\ \beta_{n1}N_n & \beta_{n2}N_n & \dots & \beta_{nn}N_n \end{pmatrix}, \qquad V = \begin{pmatrix} (\mu_1+\gamma_1) & 0 & \dots & 0 \\ & & \vdots & \\ 0 & 0 & \dots & (\mu_n+\gamma_n) \end{pmatrix}. \tag{15.2}$$

Hence we define the basic reproduction number as $\mathscr{R}_0 = \rho(FV^{-1})$, where

$$FV^{-1} = \begin{pmatrix} \dfrac{\beta_{11}N_1}{\mu_1+\gamma_1} & \dfrac{\beta_{12}N_1}{\mu_2+\gamma_2} & \cdots & \dfrac{\beta_{1n}N_1}{\mu_n+\gamma_n} \\ & & \vdots & \\ \dfrac{\beta_{n1}N_n}{\mu_1+\gamma_1} & \dfrac{\beta_{n2}N_n}{\mu_2+\gamma_2} & \cdots & \dfrac{\beta_{nn}N_n}{\mu_n+\gamma_n} \end{pmatrix}. \tag{15.3}$$

To derive a condition for disease persistence or extinction, we follow [163]. We order the variables $(S_1,\dots,S_n,I_1,\dots,I_n)$, and we consider the Jacobian of system (15.1). It takes the form

$$J = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}. \tag{15.4}$$

The Jacobian is block-diagonal, since the matrix A_{21} is the zeroth matrix. Hence, its eigenvalues are precisely the eigenvalues of A_{11} and A_{22}. The matrix A_{11} is equal to $\operatorname{diag}(-\mu_1,\dots,-\mu_n)$. The matrix A_{22} is given by

$$A_{22} = \begin{pmatrix} \beta_{11}N_1 - (\mu_1+\gamma_1) & \beta_{12}N_1 & \dots & \beta_{1n}N_1 \\ & & \vdots & \\ \beta_{n1}N_n & \beta_{n2}N_n & \dots & \beta_{nn}N_n - (\mu_n+\gamma_n) \end{pmatrix}. \tag{15.5}$$

The eigenvalues of A_{11} are clearly negative. Consequently, the stability of the disease-free equilibrium depends on the eigenvalues of A_{22}. The matrix $-A_{22}$ has nonpositive off-diagonal entries, and thus it possesses the Z-pattern. Then A_{22} will have eigenvalues with negative real parts if and only if $-A_{22}$ is an M-matrix. For $-A_{22}$ to be an M-matrix, it is sufficient that a vector $v > 0$ exist such that $-A_{22}v > 0$. Let $v = (1,1,\dots,1)^T$. Then the condition $-A_{22}v > 0$ is equivalent to the following conditions being satisfied:

$$(\mu_i+\gamma_i) - N_i \sum_{j=1}^{n} \beta_{ij} > 0, \quad i = 1,\dots,n. \tag{15.6}$$

The above condition can also be rewritten in the form

$$(\mu_i+\gamma_i)(1-\mathscr{R}_i) > N_i \sum_{j\neq i}^{n} \beta_{ij}, \quad i=1,\ldots,n. \tag{15.7}$$

So if (15.7) holds, then the disease will be eliminated. However, (15.7) also implies that $\max\{\mathscr{R}_1,\ldots,\mathscr{R}_n\} < 1$, so in this case, all patches must be sinks. However, if even one patch is a source, that is, if there exists j_0 such that $\mathscr{R}_{j_0} > 1$, then condition (15.7) is violated, and the disease persists in the populations. One nonobvious conclusion of (15.7) is that the disease may persist even if all patches are sinks, that is, $\max\{\mathscr{R}_1,\ldots,\mathscr{R}_n\} < 1$.

15.2.2 Eulerian Movement Epidemic Models

To formulate the Eulerian movement epidemic model, we assume that the population lives on n patches. In Eulerian movement, it is assumed that individuals move to another patch and settle there, becoming a part of the population of the host patch. Further, we assume that within each patch, the population mixes homogeneously and the distribution of the disease is described by a classical SIR endemic model. We denote the migration rates from patch j to patch i by m_{ij}. We assume $m_{ii} = 0$. Furthermore, the migration rates of susceptible and infective may be different, so the m_{ij}^S denote the migration rates of susceptibles, and the m_{ij}^I denote the migration rates of infectives. A more complex model of this form was considered in [17]:

$$\begin{aligned} S_i' &= \mu_i N_i - \beta_i S_i I_i - \mu_i S_i - \sum_{j=1}^{n} m_{ji}^S S_i + \sum_{j=1}^{n} m_{ij}^S S_j, \\ I_i' &= \beta_i S_i I_i - (\mu_i+\gamma_i) I_i, - \sum_{j=1}^{n} m_{ji}^I I_i + \sum_{j=1}^{n} m_{ij}^I I_j, \qquad i=1,\ldots,n. \end{aligned} \tag{15.8}$$

In the above model, we have once again omitted the equation of the recovered. The total birth/recruitment rate into patch i is $\mu_i N_i$, the mortality rate in patch i is μ_i, the transmission rate in patch i is β_i, and the recovery rate in patch i is γ_i. The total population size is given by $N(t) = S_1(t) + I_1(t) + \cdots + S_n(t) + I_n(t)$, and it satisfies the differential equation

$$N'(t) = 0.$$

Hence, the total population size remains constant.

To find the disease-free equilibrium, let $I_i = 0$ for $i = 1,\ldots,n$. The disease-free equilibrium is given by $\mathscr{E}_0 = (S_1,\ldots,S_n,0,\ldots,0)$. Then the disease-free equilibrium satisfies the following system of equations:

$$\mu_i N_i - \mu_i S_i - \sum_{j=1}^{n} m_{ji}^S S_i + \sum_{j=1}^{n} m_{ij}^S S_j = 0, \qquad i=1,\ldots,n.$$

We rewrite this system in the form

$$\mu_i S_i + \sum_{j=1}^{n} m_{ji}^S S_i - \sum_{j=1}^{n} m_{ij}^S S_j = \mu_i N_i \qquad i = 1, \ldots, n.$$

The matrix of this system is

$$M = \begin{pmatrix} \mu_1 + \sum_{j=1}^{n} m_{j1}^S & -m_{12}^S \ \ldots & -m_{1n}^S \\ & \vdots & \\ -m_{n1}^S & -m_{n2}^S \ \ldots & \mu_n + \sum_{j=1}^{n} m_{jn}^S \end{pmatrix}. \tag{15.9}$$

The matrix M possesses the Z-pattern. Furthermore, for $v = (1, \ldots, 1)^T$, $M^T v > 0$. Hence, M is an M-matrix. As a result, it has a nonnegative inverse, and therefore the system has a nonnegative solution. We conclude that there is a unique nonnegative disease-free equilibrium.

To compute the reproduction number, we use the next-generation approach. The infective variables are $I_1, \ldots, I_n$. We compute the matrices F and V. The new infections are given by the incidence terms. Hence $F = \text{diag}(\beta_1 S_1, \ldots, \beta_n S_n)$. All remaining terms give V:

$$V = \begin{pmatrix} (\mu_1 + \gamma_1 + \sum_{j=1}^{n} m_{j1}^I) & -m_{12}^I \ \ldots & -m_{1n}^I \\ & \vdots & \\ -m_{n1}^I & -m_{n2}^I \ \ldots & (\mu_n + \gamma_n + \sum_{j=1}^{n} m_{jn}^I) \end{pmatrix}. \tag{15.10}$$

The reproduction number is then defined as $\mathscr{R}_0 = \rho(FV^{-1})$. It follows from the results in [159] that if $\mathscr{R}_0 < 1$, the disease-free equilibrium is locally asymptotically stable, and if $\mathscr{R}_0 > 1$, then the disease-free is unstable.

We define the reproduction number of each patch as

$$\mathscr{R}_i = \frac{\beta_i S_i}{\mu_i + \gamma_i}. \tag{15.11}$$

Exercise 15.3 asks you to show that for $n = 2$, the reproduction number $\mathscr{R}_0$ computed by the next-generation approach is bracketed between the two patch reproduction numbers.

15.3 Spatial Models with Diffusion

Reaction–diffusion equations are one of the most important tools for modeling spatial movement. They are continuous in both space and time and represent second-order partial differential equations. Reaction–diffusion equations offer the best modeling of spatial propagation that occurs as a result of diffusion. They consist of a diffusion component that is the second-order derivative of the unknown function(s) and a reaction part that resembles the right-hand side of an ordinary differential equation model.

15.3.1 Derivation of Reaction–Diffusion Equations

There are various approaches for the derivation of reaction–diffusion equations. We will use the approach that captures the movement discretely and then obtains a continuous equation by passing to the limit. To simplify matters, we will imagine that the movement occurs in one dimensional direction, or along a line. The description here follows that in [33]. Denote the probability that an individual is at location x at time t by $p(x,t)$. We discretize both space and time. Space is discretized with a step Δx, and time is discretized with a step Δt. The probability $p(x,t+\Delta t)$ that an individual will be at location x at time $t+\Delta t$ is the sum of the probabilities of being at location $x-\Delta x$ at time t and moving to the right with probability α or being in position $x+\Delta x$ at time t and moving to the left with probability $1-\alpha$, that is,

$$p(x,t+\Delta t) = \alpha p(x-\Delta x,t) + (1-\alpha)p(x+\Delta x,t).$$

Subtracting $p(x,t)$ from both sides and dividing by Δt, we have

$$\begin{aligned}\frac{p(x,t+\Delta t)-p(x,t)}{\Delta t} &= \frac{1}{2\Delta t}[p(x-\Delta x,t)-2p(x,t)+p(x+\Delta x,t)] \\ &\quad + \frac{1/2-\alpha}{\Delta t}[p(x+\Delta x,t)-p(x-\Delta x,t)]. \qquad (15.12)\end{aligned}$$

At this point, we have to make an assumption for the correlation between Δt and Δx. In particular, we assume that $D=(\Delta x)^2/(2\Delta t)$ and $v=2(1/2-\alpha)\Delta x/\Delta t$ are constants. Because of these definitions of D and v, they have units distance^2/time and distance/time respectively. The constant D is called the *diffusion constant*. Then the above expression becomes

$$\begin{aligned}\frac{p(x,t+\Delta t)-p(x,t)}{\Delta t} &= D\left[\frac{p(x-\Delta x,t)-2p(x,t)+p(x+\Delta x,t)}{(\Delta x)^2}\right] \\ &\quad + v\left[\frac{p(x+\Delta x,t)-p(x-\Delta x,t)}{2\Delta x}\right]. \qquad (15.13)\end{aligned}$$

Passing to the limit as $\Delta t\to 0$ and $\Delta x\to 0$, we obtain the following differential equation:

$$\frac{\partial p}{\partial t} = D\frac{\partial^2 p}{\partial x^2} + v\frac{\partial p}{\partial x}. \qquad (15.14)$$

This is the basic equation for directed diffusion, in which case the probabilities of moving left or right are not equal. If we assume that the probabilities of moving left or right are equal, that is, $\alpha=1/2$, then $v=0$, and we obtain the following simplified equation:

$$\frac{\partial p}{\partial t} = D\frac{\partial^2 p}{\partial x^2}. \qquad (15.15)$$

This is the basic equation that models *random diffusion*. We will work primarily with random diffusion models.

Equation (15.14) models only the diffusion part of a population model. To incorporate population growth and disease transmission, we have to have a *reaction part*, which comes from a regular ordinary differential equation epidemic model. For instance, consider an ordinary differential equation SI model

$$\begin{aligned} S' &= -\beta SI \\ I' &= \beta SI. \end{aligned} \tag{15.16}$$

Now we assume that susceptible and infected individuals are spatially distributed, $S(x,t)$ and $I(x,t)$, where x is a one-dimensional spatial variable. We add diffusion to this model to obtain the following epidemic reaction–diffusion model:

$$\begin{aligned} \frac{\partial S}{\partial t} &= -\beta S(x,t)I(x,t) + D\frac{\partial^2 S}{\partial x^2}, \\ \frac{\partial I}{\partial t} &= \beta S(x,t)I(x,t) + D\frac{\partial^2 I}{\partial x^2}, \end{aligned} \tag{15.17}$$

where β is still a constant independent of space. We denote by $N(x,t)$ the total population size. Hence, we have $N(x,t) = S(x,t) + I(x,t)$. The total population size satisfies an equation of the form (15.15):

Models of the form (15.14) and (15.17) will have a unique solution if two types of additional conditions are specified for them:

1. *Initial conditions.* These specify given spatial distributions at time $t = 0$. These are similar to the initial conditions specified for ordinary differential equation models.
2. *Boundary conditions.* These specify the values of the unknown function(s) at the boundaries of the spatial region.

Boundary conditions can be of different types and specify different types of problems when coupled with the same system of differential equations. If the domain is infinite, say the whole real axis, then typically a condition governing the growth of the functions at $\pm\infty$ are specified. In particular, with problem (15.17) we will have

$$S(x,t) \to 0, \qquad I(x,t) \to 0 \quad \text{as} \quad x \to \pm\infty.$$

If the spatial domain is finite, say $[0,L]$, then the following types of boundary conditions are often used:

- **Dirichlet Boundary Conditions.** For Dirichlet boundary conditions, the unknown function(s) are prescribed given known values at the boundary of the domain. In particular, for model (15.17), we have

  $$S(0,t) = S_0(t) \qquad S(L,t) = S_L(t), \qquad I(0,t) = I_0(t), \qquad I(L,t) = I_L(t),$$

  where $S_0(t), S_L(t), I_0(t), I_L(t)$ are given known functions. A commonly used set of Dirichlet boundary conditions prescribes the values at the boundary to be zero:

  $$S(0,t) = 0, \qquad S(L,t) = 0, \qquad I(0,t) = 0, \qquad I(L,t) = 0.$$

In this case, the Dirichlet boundary conditions are called *homogeneous Dirichlet boundary conditions*. These conditions are often called "absorbing" or "deadly," because the boundary "absorbs" all individuals who encounter it.

- **Neumann Boundary Conditions.** For the Neumann boundary conditions, the fluxes through the boundary of the unknown function(s) are prescribed given known values of the fluxes at the boundary of the domain. In particular, for model (15.17), we have

$$\frac{\partial S}{\partial x}(0,t) = \hat{S}_0(t), \qquad \frac{\partial S}{\partial x}(L,t) = \hat{S}_L(t), \qquad \frac{\partial I}{\partial x}(0,t) = \hat{I}_0(t), \qquad \frac{\partial I}{\partial x}(L,t) = \hat{I}_L(t),$$

where $\hat{S}_0(t), \hat{S}_L(t), \hat{I}_0(t), \hat{I}_L(t)$ are given known functions. Commonly used Neumann boundary conditions are those such that the prescribed flux values at the boundary are zero:

$$\frac{\partial S}{\partial x}(0,t) = 0, \qquad \frac{\partial S}{\partial x}(L,t) = 0, \qquad \frac{\partial I}{\partial x}(0,t) = 0, \qquad \frac{\partial I}{\partial x}(L,t) = 0.$$

In this case, the Neumann boundary conditions are called *homogeneous Neumann boundary conditions*. These conditions are often called "no-flux" conditions, because the boundary does not allow individuals to pass through it. These are the most commonly used in population models.

- **Robin Boundary Conditions.** For the Robin boundary conditions, a linear combination of the values and the fluxes through the boundary of the unknown function(s) are prescribed given known values at the boundary of the domain. In particular, for model (15.17), we have

$$\begin{aligned} \alpha_1 \frac{\partial S}{\partial x}(0,t) + \gamma_1 S(0,t) &= \hat{S}_0(t), & \alpha_2 \frac{\partial S}{\partial x}(L,t) + \gamma_2 S(L,0) &= \hat{S}_L(t), \\ \alpha_3 \frac{\partial I}{\partial x}(0,t) + \gamma_3 I(0,t) &= \hat{I}_0(t), & \alpha_4 \frac{\partial I}{\partial x}(L,t) + \gamma_4 I(L,t) &= \hat{I}_L(t), \end{aligned} \tag{15.18}$$

where $\hat{S}_0(t), \hat{S}_L(t), \hat{I}_0(t), \hat{I}_L(t)$ are given known functions. Commonly used Neumann boundary conditions are those such that the prescribed flux values at the boundary are zero:

$$\begin{aligned} \alpha_1 \frac{\partial S}{\partial x}(0,t) + \gamma_1 S(0,t) &= 0, & \alpha_2 \frac{\partial S}{\partial x}(L,t) + \gamma_2 S(L,0) &= 0, \\ \alpha_3 \frac{\partial I}{\partial x}(0,t) + \gamma_3 I(0,t) &= 0, & \alpha_4 \frac{\partial I}{\partial x}(L,t) + \gamma_4 I(L,t) &= 0. \end{aligned} \tag{15.19}$$

In this case, the Robin boundary conditions are called *homogeneous Robin boundary conditions*. These conditions are often called "mixed" conditions.

A combination of a differential equation and Dirichlet boundary conditions is called a *Dirichlet problem*. Similarly, a combination of a differential equation and Neumann boundary conditions is called a *Neumann problem*.

15.3.2 Equilibria and Their Local Stability

Reaction–diffusion equations, just like ordinary differential equations, have time-independent solutions, called *equilibria*. Equilibrium solutions may be constant in space, in which case they are called *spatially homogeneous*, or they may depend on the spatial variable explicitly, in which case they are called *spatially heterogeneous*.

To illustrate this concept, we consider model (15.17) with homogeneous Neumann boundary conditions. The total population size N satisfies the following problem:

$$\begin{aligned} \frac{\partial N}{\partial t} &= D\frac{\partial^2 N}{\partial x^2}, \\ \frac{\partial N}{\partial x}(0,t) &= \frac{\partial N}{\partial x}(L,t) = 0. \end{aligned} \tag{15.20}$$

If the initial conditions are constant in space, this system clearly has the constant N as a solution, where the value of N is given by the initial conditions. Expressing $S(x,t) = N - I(x,t)$ and eliminating S, we may reduce the system (15.17) to the single differential equation

$$\begin{aligned} \frac{\partial I}{\partial t} &= \beta I(x,t)(N - I(x,t)) + D\frac{\partial^2 I}{\partial x^2}, \\ \frac{\partial I}{\partial x}(0,t) &= \frac{\partial I}{\partial x}(L,t) = 0. \end{aligned} \tag{15.21}$$

Equation (15.21) was first proposed by Fisher [62] to model gene spread in a population and has been widely studied since then. Now this equation is referred to as the *Fisher–Kolmogorov equation.*

Equilibrium solutions of model (15.21) satisfy

$$\begin{aligned} 0 &= \beta I(x)(N - I(x)) + D\frac{\partial^2 I}{\partial x^2}, \\ \frac{\partial I}{\partial x}(0) &= \frac{\partial I}{\partial x}(L) = 0. \end{aligned} \tag{15.22}$$

This system clearly has two spatially homogeneous solutions: $I = 0$ and $I = N$. The system (15.21) does not have spatially heterogeneous solutions. To see this, integrate the first equation in system (15.22) to obtain

$$\frac{\partial I}{\partial x} = -\frac{\beta}{D}\int_0^x I(y)(N - I(y))dy.$$

The first boundary condition is clearly satisfied. However, for nonnegative I, smaller than N, $\dfrac{\partial I}{\partial x}$ is a function that is increasing in absolute value and negative. Thus, $\dfrac{\partial I}{\partial x}(L) < 0$, which contradicts the second boundary condition.

Local stability of these equilibria is determined by the following theorem:

Theorem 15.1. *The equilibrium $I = 0$ is unstable. The equilibrium $I = N$ is locally asymptotically stable.*

Proof. We begin with the equilibrium $I = 0$. We linearize around this equilibrium. In particular, we set $I(x,t) = y(x,t)$. We obtain the following linear equation for the perturbation:

$$\begin{aligned} \frac{\partial y}{\partial t} &= \beta N y(x,t) + D\frac{\partial^2 y}{\partial x^2}, \\ \frac{\partial y}{\partial x}(0,t) &= \frac{\partial y}{\partial x}(L,t) = 0. \end{aligned} \tag{15.23}$$

A traditional approach to solving Eq. (15.23) is to look for solutions in separable form, that is, we look for a solution of the form

$$y(x,t) = X(x)T(t),$$

where $X(x)$ and $T(t)$ are two unknown functions. Since $y(x,t) = X(x)T(t)$ is a solution, it should satisfy system (15.23). We substitute it into the differential equation to obtain

$$X(x)T'(t) = \beta N X(x)T(t) + DT(t)X''(x),$$

where the primes denote derivatives with respect to the variable of the function. We divide both sides by $DX(x)T(t)$, and we move the constant βN to the left-hand side of the equality:

$$\frac{T'(t)}{DT(t)} - \frac{\beta N}{D} = \frac{X''(x)}{X(x)}.$$

The function on the left is a function of t, while the function on the right is a function of x. The only way these two functions may be equal is if they are equal to the same constant. Hence,

$$\frac{T'(t)}{DT(t)} - \frac{\beta N}{D} = \frac{X''(x)}{X(x)} = -k.$$

This leads to two separate ordinary differential equations:

$$\begin{aligned} T'(t) + (kD - \beta N)T(t) &= 0, \\ X''(x) + kX(x) &= 0 \qquad \text{with} \qquad X'(0) = X'(L) = 0. \end{aligned} \tag{15.24}$$

The constant k is chosen to be positive. If $k = 0$, the above equation has an arbitrary constant as a solution. If $k < 0$, the second-order ordinary differential equation has only a trivial solution. The general solution of the second-order equation in (15.24) when $k > 0$ is given by

$$X(x) = C_1 \cos\sqrt{k}x + C_2 \sin\sqrt{k}x,$$

where C_1 and C_2 are to be determined from the boundary conditions. To satisfy the boundary conditions, we differentiate the expression above:

$$X'(x) = -C_1\sqrt{k}\sin\sqrt{k}x + C_2\sqrt{k}\cos\sqrt{k}x.$$

Hence,

$$X'(0) = C_2\sqrt{k} = 0.$$

This implies that $C_2 = 0$. Furthermore,

$$X'(L) = -C_1\sqrt{k}\sin\sqrt{k}L = 0.$$

This can be satisfied if $\sin\sqrt{k}L = 0$ or if $\sqrt{k}L = n\pi$ for $n = 1, 2, \ldots$. Consequently, we have a sequence of values

$$k_n = \frac{(n\pi)^2}{L^2}$$

and a sequence of functions corresponding to them:

$$X_n(x) = \cos\sqrt{k_n}x.$$

The functions $X_n(x)$ are called eigenfunctions, and the values k_n are called eigenvalues. The solution to the first-order ordinary differential equation in (15.24) is given by

$$T(t) = T_0 e^{-(kD-\beta N)t}.$$

We have a sequence of solutions corresponding to the different values of k:

$$T_n(t) = T_{0n} e^{-(k_n D-\beta N)t}.$$

Hence, one of the solutions is given by

$$y_n(x,t) = B_n\cos\sqrt{k_n}x e^{-(k_n D-\beta N)t},$$

where we have combined all constants in the constant B_n. Summing over all solutions, we obtain

$$y(x,t) = B_0 e^{\beta N t} + \sum_{n=1}^{\infty} B_n\cos\sqrt{k_n}x e^{-(k_n D-\beta N)t}.$$

It is not hard to see that the first term in the above sum always approaches infinity as $t \to \infty$. Hence, $y(x,t) \to \infty$ as $t \to \infty$. Therefore, $I = 0$ is unstable.

Remark 15.1. The coefficients B_n can be determined from the initial condition. If $y(x,0) = y_0(x)$, then we must have for $t = 0$,

$$y_0(x) = B_0 + \sum_{n=1}^{\infty} B_n\cos\sqrt{k_n}x.$$

Multiplying the above expression by $\cos\sqrt{k_{n0}}x$, where $n0$ is arbitrary fixed integer, we obtain

$$y_0(x)\cos\sqrt{k_{n0}}x = B_0\cos\sqrt{k_{n0}}x + \sum_{n=1}^{\infty} B_n\cos\sqrt{k_{n0}}x\cos\sqrt{k_n}x.$$

Integrating from 0 to L, we have

$$B_{n0} = \frac{2}{L}\int_0^L y_0(x)\cos\sqrt{k_{n0}}x\,dx.$$

Next, we consider the case $I = N$. We linearize around this equilibrium. In particular, we set $I(x,t) = y(x,t) + N$. We obtain the following linear equation for the perturbation:

$$\begin{aligned} \frac{\partial y}{\partial t} &= -\beta N y(x,t) + D\frac{\partial^2 y}{\partial x^2}, \\ \frac{\partial y}{\partial x}(0,t) &= \frac{\partial y}{\partial x}(L,t) = 0. \end{aligned} \tag{15.25}$$

Following the same approach as before with βN replaced by $-\beta N$, we arrive at the following solution of the linearized problem (15.25):

$$y(x,t) = B_0 e^{-\beta N t} + \sum_{n=1}^{\infty} B_n\cos\sqrt{k_n}x\,e^{-(k_n D+\beta N)t}.$$

Clearly, we have $y(x,t)\to 0$ as $t\to\infty$. Hence, the equilibrium $I = N$ is locally asymptotically stable.

Epidemiologically speaking, this model predicts that in the long term, the entire population will eventually become infected, a result that is not particularly realistic.

15.3.3 Traveling-Wave Solutions

Time-dependent diffusion equations have another important type of special solution called *traveling-wave solutions*. These are solutions that propagate without a change in shape and at a constant speed. Traveling-wave solutions are important in epidemiology because they model the geographic spread of a diseases in a wavelike manner. Traveling-wave solutions are typically considered on an infinite domain.

To illustrate the presence of traveling-wave solutions, we consider the model

$$\begin{aligned} \frac{\partial I}{\partial t} &= \beta I(x,t)(N - I(x,t)) + D\frac{\partial^2 I}{\partial x^2}, \\ \lim_{x\to-\infty}\frac{\partial I}{\partial x} &= \lim_{x\to+\infty}\frac{\partial I}{\partial x} = 0. \end{aligned} \tag{15.26}$$

We note that the boundary conditions are equivalent to assuming that

$$\lim_{x\to-\infty} I(x,t) = A, \qquad \lim_{x\to+\infty} I(x,t) = B,$$

where A and B are appropriate constants. We note that solutions of (15.26) are unique when the system is equipped with initial conditions.

Definition 15.2. A *traveling-wave solution* is a solution that can be expressed in a single variable $z = x - vt$:

$$I(x,t) = I(x-vt) = I(z).$$

The constant v is called the *wave speed*.

The wave speed is undetermined, but we may assume $v > 0$, that is, that the wave travels to the right. To obtain the differential equation for the traveling-wave solution, we substitute $I(z)$ in the differential equation of (15.26). From the partial differential equation, we will obtain an ordinary differential equation in the function of a single variable $I(z)$. To express the partial derivatives in the single variable, we use the chain rule. Thus,

$$\begin{aligned} \frac{\partial I}{\partial t} &= \frac{dI}{dz}\frac{\partial z}{\partial t} = -v\frac{dI}{dz}, \\ \frac{\partial I}{\partial x} &= \frac{dI}{dz}\frac{\partial z}{\partial x} = \frac{dI}{dz}. \end{aligned} \tag{15.27}$$

Hence,

$$\frac{\partial^2 I}{\partial x^2} = \frac{d^2 I}{dz^2}.$$

Then, substituting in the partial differential equation, we obtain the following ordinary differential equation in $I(z)$:

$$\begin{aligned} D\frac{d^2 I}{dz^2} + v\frac{dI}{dz} + \beta I(z)(N - I(z)) &= 0, \\ \lim_{z\to-\infty} \frac{\partial I}{\partial z} = \lim_{z\to+\infty} \frac{\partial I}{\partial z} &= 0. \end{aligned} \tag{15.28}$$

The solutions of the traveling-wave equations are typically not unique. Since v is unknown, one may have different solutions for different values of v as v varies. Moreover, if $I(z)$ solves (15.28) for some fixed value of v, then for that same value of v, $I(z+c)$, where c is any constant, also solves (15.28). For both v and c fixed, the solution of the traveling wave equations are normally unique. Furthermore, the initial conditions specified with problem (15.26) are not used with the traveling-wave solutions. The traveling-wave solution is a solution to the full partial differential equation only if the traveling-wave solution is consistent with the initial condition specified with (15.26). However, one may often observe that a given initial condition approaches a traveling-wave solution as $t \to \infty$. This is illustrated in Fig. 15.2.

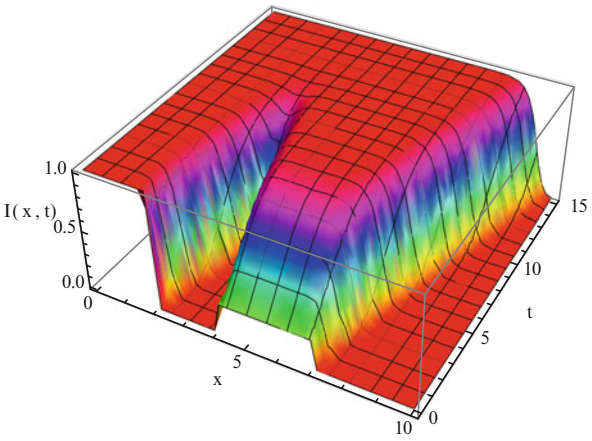

Fig. 15.2 Initial condition tends to a traveling-wave solution as $t \to \infty$ for model (15.26). The initial condition taken is a step function

The equation for the traveling-wave solution is a second-order nonlinear ordinary differential equation model. To prove the existence of a traveling wave, one rewrites the second-order equation into a system of first-order ordinary differential equations. In particular, we set $\frac{dI}{dz} = Y(z)$. Then (15.28) becomes the system

$$\begin{aligned} \frac{dI}{dz} &= Y, \\ \frac{dY}{dz} &= -\frac{v}{D}Y - \frac{\beta}{D}I(z)(N - I(z)). \end{aligned} \tag{15.29}$$

The traveling-wave solution must tend to the stationary points of the system (15.29) as $z \to \pm\infty$. This is necessary, since $I(z)$ tends to a constant as $z \to \pm\infty$. The investigation of the traveling-wave solution begins by classifying the equilibria of this planar system. In particular, if the variables are ordered (I, Y), then the system above has two equilibria: $(N, 0)$ and $(0, 0)$. To investigate the type of the equilibria and their stability, we look at the Jacobian:

$$J(I,Y) = \begin{pmatrix} 0 & 1 \\ -\frac{\beta}{D}(N - I) + \frac{\beta}{D}I & -\frac{v}{D} \end{pmatrix}. \tag{15.30}$$

We compute the Jacobian at the $(0,0)$ equilibrium:

$$J(0,0) = \begin{pmatrix} 0 & 1 \\ -\frac{\beta N}{D} & -\frac{v}{D} \end{pmatrix}. \tag{15.31}$$

The trace of this matrix is negative, since $v > 0$ and $D > 0$ and the determinant $\frac{\beta N}{D}$ is positive. Thus, $(0,0)$ is locally asymptotically stable. To find out what type it is, we compute the characteristic equation

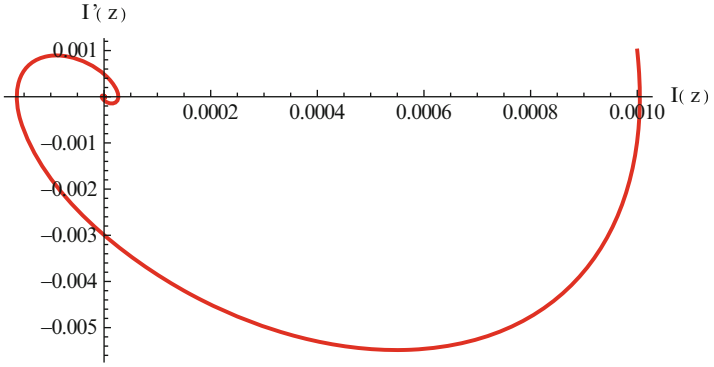

Fig. 15.3 Solution of model (15.29) in the case $v < \sqrt{\beta ND}$. One can see that near $(0,0)$, $I(z)$ can become negative. Parameters are taken as $\beta = 1$, $N = 1$, $v = 0.1$, $D = 0.01$

$$(-\lambda)(-\lambda - \frac{v}{D}) + \frac{\beta N}{D} = 0.$$

Thus, the characteristic equation becomes

$$\lambda^2 + \frac{v}{D}\lambda + \frac{\beta N}{D} = 0,$$

and the eigenvalues are given by

$$\lambda_{1,2} = \frac{-\frac{v}{D} \pm \sqrt{\left(\frac{v}{D}\right)^2 - 4\frac{\beta N}{D}}}{2}.$$

Hence, if $\left(\frac{v}{D}\right)^2 - 4\frac{\beta N}{D} > 0$, then $(0,0)$ is a stable node. Therefore, the origin is a stable node if the wave speed satisfies $v > 2\sqrt{\beta ND}$. If $\left(\frac{v}{D}\right)^2 - 4\frac{\beta N}{D} < 0$, then $(0,0)$ is a stable spiral. Therefore, the origin is a stable spiral if the wave speed satisfies $v < 2\sqrt{\beta ND}$. In this case, the solution near the equilibrium $(0,0)$ looks as shown in Fig. 15.3. In this case, clearly $I(z)$ can become negative for some values of z. That makes the solution biologically unrealistic, since the number of infected should be nonnegative. Thus, this case does not produce a traveling wave.

Next, we determine the stability and the type of the equilibrium $(N,0)$. Evaluating the Jacobian yields

$$J(N,0) = \begin{pmatrix} 0 & 1 \\ \frac{\beta N}{D} & -\frac{v}{D} \end{pmatrix}. \tag{15.32}$$

In this case, the trace is negative, but the determinant is also negative. Hence $(N,0)$ is unstable. Since the determinant is negative, the product of the eigenvalues is

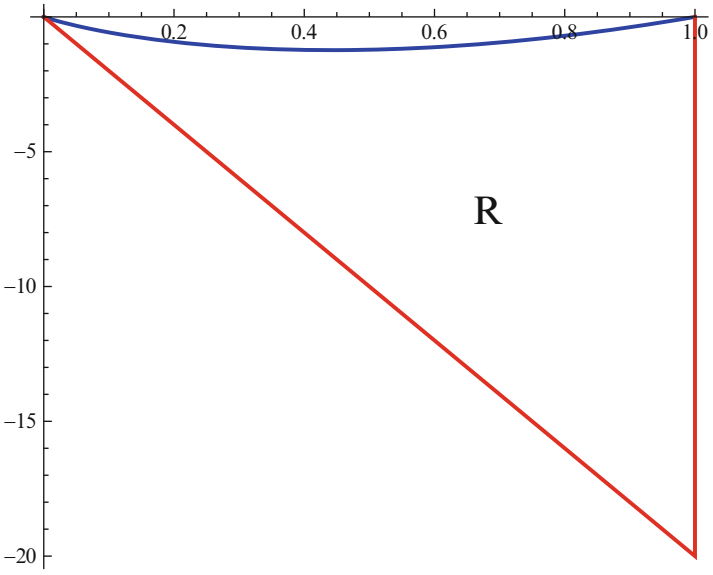

Fig. 15.4 The region **R** (in *red*) with the heteroclinic orbit (in *blue*). Parameter values are $\beta = 1$, $D = 0.01$, $N = 1$, $v = 2\sqrt{\beta DN}$

negative, so there is one positive and one negative eigenvalue. Hence, $(N,0)$ is a saddle. To find the eigenvalues, we derive the characteristic equation:

$$\lambda^2 + \frac{v}{D}\lambda - \frac{\beta N}{D} = 0.$$

Hence, the eigenvalues are

$$\lambda_{\pm} = \frac{-\frac{v}{D} \pm \sqrt{\left(\frac{v}{D}\right)^2 + 4\frac{\beta N}{D}}}{2}. \tag{15.33}$$

To show the existence of a traveling wave in the case $v > 2\sqrt{\beta ND}$, we have to find a connection between the equilibria $(0,0)$ and $(N,0)$ in the phase portrait of system (15.29). Such a connection is called a *heteroclinic orbit*. The heteroclinic orbit will leave the equilibrium $(N,0)$, which is a saddle, along its unstable manifold and will approach equilibrium $(0,0)$. First, we need to find the gradient of the unstable manifold at the equilibrium $(N,0)$. For that purpose, we require the eigenvectors of the Jacobian at $(N,0)$. Assume that the eigenvectors are $v_{\pm} = (1, q_{\pm})^T$, where $q_{\pm}$ are to be determined. Thus, we have

$$\begin{pmatrix} 0 & 1 \\ \frac{\beta N}{D} & -\frac{v}{D} \end{pmatrix} \begin{pmatrix} 1 \\ q_{\pm} \end{pmatrix} = \lambda_{\pm} \begin{pmatrix} 1 \\ q_{\pm} \end{pmatrix}. \tag{15.34}$$

From here, it is easy to see that $q_{\pm} = \lambda_{\pm}$. Hence the gradient of the unstable manifold is $v_+ = (1, \lambda_+)^T$, where $\lambda_{\pm}$ are given by (15.33). Next, we show that the unstable manifold leaving $(N,0)$ enters the region (see Fig. 15.4)

$$\mathbf{R} = \{(I,Y)|Y \le 0, I \in [0,N], Y \ge -\frac{\nu}{D}I\}$$

and can never leave that region. To see this, we consider system (15.29) and rewrite the system as the equation

$$\frac{dY}{dI} = -\frac{v}{D} - \frac{\beta}{D}\frac{I(N-I)}{Y}. \tag{15.35}$$

We consider the slope of the solutions of Eq. (15.35) along the boundaries of the region $\mathbf{R}$. Along the boundary

$$\mathbf{B}_1 = \{(I,Y)|Y = 0, I \in (0,N)\},$$

we have $\frac{dY}{dI} = -\infty$, so the trajectories point vertically into the region $\mathbf{R}$ (recall that the derivative gives $\tan\theta$, where θ is the angle of the tangent vector with the positive x-axis). Next, along the boundary

$$\mathbf{B}_2 = \{(I,Y)|Y \le 0, I = N\},$$

the slope of the trajectories is $\frac{dY}{dI} = -\frac{v}{D} < 0$. Hence, θ is in the forth quadrant and points into the region $\mathbf{R}$. Finally, along the boundary

$$\mathbf{B}_3 = \{(I,Y)|Y \le 0, I \in (0,N], Y = -\frac{\nu}{D}I\},$$

we show that $Y \ge -\frac{\nu}{D}I$. To see this, notice that

$$-\frac{\beta}{D}\frac{I(N-I)}{Y} > 0$$

in $\mathbf{R}$. Hence from (15.35), we have

$$\frac{dY}{dI} \ge -\frac{\nu}{D}. \tag{15.36}$$

Integrating this inequality, we have

$$Y(I) \ge -\frac{\nu}{D}I.$$

If $\frac{dY}{dI}$ is bigger than the right-hand side of the inequality (15.36), it will point inside the region $\mathbf{R}$. Hence, all solutions that enter region $\mathbf{R}$ stay there and approach the equilibrium $(0,0)$. This is true in particular for the solution that begins along the unstable manifold of equilibrium $(N,0)$. That solution is plotted in Fig. 15.5. In this case, we see that $I'(z) < 0$ but $I(z) \in [0,N]$, so this is a heteroclinic connection that gives a traveling wave.

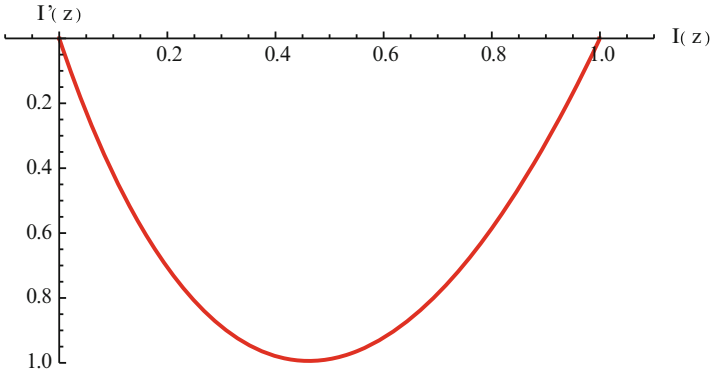

Fig. 15.5 Solution of model (15.29) in the case $v > 2\sqrt{\beta ND}$. One can see that $I(z) \in [0,N]$ and $I'(z) < 0$. Parameters are taken as $\beta = 1$, $N = 1$, $v = 0.25$, $D = 0.01$

Definition 15.3. The minimal speed $v_{\min}$ for which a traveling wave exists is called the *minimal wave speed*.

In this case, the minimal wave speed is $v_{\min} = 2\sqrt{\beta ND}$. The existence of a traveling-wave solution to the Fisher–Kolmogorov equation was first proved by Kolmogorov et al. [1].

15.3.4 Turing Instability

Turing instability refers to diffusion-driven destabilization of a spatially homogeneous steady state of the a reaction–diffusion system. This phenomenon was first observed by Alan Turing in 1952 [158]. We will follow Turing's 1952 paper and examine analytically the linear stability analysis of the simplest possible reaction–diffusion system that forms a pattern from a spatially homogeneous state. The analysis leads to several insights. The first is that at least two interacting classes are needed for pattern formation to occur, that is, pattern formation occurs in at least a 2×2 system. The second is Turing's most surprising insight, that diffusion in a 2×2 system can actually have a destabilizing influence. This observation is contrary to intuition, since diffusion by itself has a stabilizing effect on a system. A third insight is that pattern formation in a system will not occur unless the diffusion coefficients of the two classes differ substantially. In other words, Turing instability occurs only if the diffusion coefficients for the two classes are different.

The following subsection contains the linear stability analysis of a general 2×2 system. After that, we consider a simple epidemic model and show that diffusion-driven instability occurs.

15.3.4.1 Turing's Instability in a General 2×2 System

To illustrate the general approach, consider the following system:

$$\begin{aligned} u_t &= D_u u_{xx} + f(u,v), \\ v_t &= D_v v_{xx} + g(u,v), \end{aligned} \tag{15.37}$$

for $x \in [0,L]$ and $t \geq 0$. The problem may be posed with constant Dirichlet boundary conditions:

$$u(0,t) = u(L,t) = u^*, \qquad v(0,t) = v(L,t) = v^*, \tag{15.38}$$

where u^* and v^* are constants satisfying $f(u^*,v^*) = 0$ and $g(u^*,v^*) = 0$. The problem can also be posed with homogeneous Neumann boundary conditions:

$$u_x(0,t) = u_x(L,t) = v_x(0,t) = v_x(L,t) = 0. \tag{15.39}$$

We will consider the problem with the Neumann boundary condition. We rewrite model (15.37) in vector form:

$$\mathbf{u}_t = \mathbf{D}\,\mathbf{u}_{xx} + \mathbf{F}(\mathbf{u}),$$

where

$$\mathbf{u} = \begin{pmatrix} u \\ v \end{pmatrix}, \qquad \mathbf{F}(\mathbf{u}) = \begin{pmatrix} f(u,v) \\ g(u,v) \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} D_u & 0 \\ 0 & D_v \end{pmatrix}. \tag{15.40}$$

Let $\mathbf{u}^*$ be the solution of $\mathbf{F}(\mathbf{u}) = 0$. Then $\mathbf{u}^*$ is a spatially homogeneous equilibrium of model (15.37). It is also an equilibrium solution of the ODE model obtained from (15.37), where $\mathbf{D} = \mathbf{0}$. For Turing instability, we require that $\mathbf{u}^*$ be a locally stable solution of the ODE model but an unstable solution of the PDE model (15.37). To find conditions for this situation to occur, we let $\mathbf{w} = \mathbf{u} - \mathbf{u}^*$. Linearizing system (15.37), we obtain

$$\mathbf{w}_t = \mathbf{D}\,\mathbf{w}_{xx} + \mathbf{F}(\mathbf{u}^*) + \mathbf{J}\,\mathbf{w},$$

where $\mathbf{J}$ is the Jacobian of the system (15.37) evaluated at $\mathbf{u}^*$, given by

$$\mathbf{J} = \begin{pmatrix} f_u & f_v \\ g_u & g_v \end{pmatrix}\Big|_{\mathbf{u}=\mathbf{u}^*}.$$

We recall that $\mathbf{F}(\mathbf{u}^*) = 0$, so that the linearized system becomes

$$\mathbf{w}_t = \mathbf{D}\,\mathbf{w}_{xx} + \mathbf{J}\,\mathbf{w}. \tag{15.41}$$

This equation is equipped with Neumann boundary conditions:

$$\mathbf{w}_x(0,t) = \mathbf{w}_x(L,t) = \mathbf{0}.$$

System (15.41) is a linear system, and we can search for a solution in the form of a sum of separable solutions. We consider a general separable solution

$$\mathbf{w}(x,t) = T(t)\mathbf{X}(x),$$

where we assume that $T(t)$ is a scalar function of t. Substituting in system (15.41) and dividing by $T(t)$, we obtain

$$\frac{T'(t)}{T(t)}\mathbf{X}(x) = \mathbf{D}\mathbf{X}''(x) + \mathbf{J}\mathbf{X}(x). \tag{15.42}$$

For this equality to hold, $\dfrac{T'(t)}{T(t)}$ must be a quantity independent of t. Hence it must be a constant, say λ. So we have that

$$T(t) = T_0 e^{\lambda t},$$

where T_0 is an appropriate constant. From (15.42), we have the following eigenvalue problem:

$$\lambda\,\mathbf{X}(x) = \mathbf{D}\mathbf{X}''(x) + \mathbf{J}\mathbf{X}(x). \tag{15.43}$$

The key step here is to observe that we can find a nontrivial solution of (15.43) if we consider only $\mathbf{X}(x)$ that satisfy the system

$$\begin{aligned} \mathbf{X}''(x) + k^2\mathbf{X}(x) &= 0, \\ \mathbf{X}'(0) = \mathbf{X}'(L) &= 0. \end{aligned} \tag{15.44}$$

We investigated this problem before, and we saw that it has solutions in the form of trigonometric functions. In particular,

$$k = \frac{n\pi}{L} \qquad \mathbf{X}_k(x) = \mathbf{A}_k \cos(kx).$$

Therefore, we will obtain a solution in the form of a series of trigonometric functions. This process defines the k_n for $n = 1, 2 \ldots$. Now we return to problem (15.43), where we still have to specify the corresponding λ's. The λ's are a solution to the following system:

$$\lambda\,\mathbf{X}(x) = -\mathbf{D}k^2\mathbf{X}(x) + \mathbf{J}\mathbf{X}(x), \tag{15.45}$$

which has a nontrivial solution if and only if the determinant is zero. Hence, we want

$$|\mathbf{J} - \mathbf{D}k^2 - \lambda\mathbf{I}| = 0,$$

where $\mathbf{I}$ is the 2×2 identity matrix. In expanded form this determinant is

$$\begin{vmatrix} f_u - D_u k^2 - \lambda & f_v \\ g_u & g_v - D_v k^2 - \lambda \end{vmatrix} = 0. \tag{15.46}$$

This gives the following quadratic equation in λ:

$$\lambda^2 + [(D_u + D_v)k^2 - (f_u + g_v)]\lambda + [D_u D_v k^4 - (D_v f_u + D_u g_v)k^2 + (f_u g_v - g_u f_v)] = 0.$$

For each k_n, $n = 1, 2\ldots$, there are two solutions for λ from the above equation: $\lambda_{1n}(k_n)$ and $\lambda_{2n}(k_n)$. With $\mathbf{X}_n(x)$ solving (15.44), we have the following solution for each k_n:

$$(T_{01}e^{\lambda_{1n}(k_n)t} + T_{02}e^{\lambda_{2n}(k_n)t})\mathbf{X}_n(x).$$

Thus the complete solution of problem (15.41) is given by

$$\mathbf{w}(x,t) = \sum_n (T_{01}e^{\lambda_{1n}(k_n)t} + T_{02}e^{\lambda_{2n}(k_n)t})\mathbf{X}_n(x).$$

This solution will be stable if $\Re\lambda_{1n} < 0$ and $\Re\lambda_{2n} < 0$ for every n. On the other hand, it will be unstable if there exist n and λ_{1n} or λ_{2n} such that the real part is positive.

Recall that for Turing instability to occur, we want the ODE model obtained from $\mathbf{D} = \mathbf{0}$ to have a locally stable equilibrium. This means that the determinant (15.46) with $D_u = D_v = k^2 = 0$ should have eigenvalues that have negative real parts. That is the case if the $\mathrm{Tr}(\boldsymbol{J}) < 0$ and $\mathrm{Det}(\boldsymbol{J}) > 0$, that is, if we have

$$\begin{aligned} f_u + g_v &< 0, \\ f_u g_v - f_v g_u &> 0. \end{aligned} \tag{15.47}$$

When the diffusion coefficients are not zero, the conditions for stability of the determinant (15.46) for each k_n are

$$\begin{aligned} f_u + g_v - D_u k^2 - D_v k^2 &< 0, \\ (f_u - D_u k^2)(g_v - D_v k^2) - f_v g_u &> 0. \end{aligned} \tag{15.48}$$

Because of assumptions (15.47), the first inequality above is always satisfied. Thus, the only way we may destabilize the diffusion equilibrium with the diffusion is if we find that

$$(f_u - D_u k^2)(g_v - D_v k^2) - f_v g_u < 0.$$

The left-hand side of that inequality is a quadratic function of k^2: $H(k^2) = Ak^4 - Bk^2 + C$, where $A = D_u D_v$, $B = D_v f_u + D_u g_v$, and $C = f_u g_v - g_u f_v$. We conclude that instability occurs if the roots of $H(k^2) = 0$ are real, say k_1^2 and k_2^2, $k^2 \in [k_1^2, k_2^2]$, *and* there exists a solution of problem (15.44) for k^2 in that range. Thus, Turing instability occurs if

$$\begin{aligned} B > 0, &\qquad D_v f_u + D_u g_v > 0, \\ B^2 > 4AC, &\qquad D_v f_u + D_u g_v > 2\sqrt{D_u D_v (f_u g_v - g_u f_v)}. \end{aligned} \tag{15.49}$$

Notice that the first condition cannot hold if $D_u = D_v$. That is why for Turing instability we always need very distinct diffusion rates. Satisfying the conditions above is not a trivial matter, and we illustrate this in the example below.

15.3.4.2 Turing Instability in an SI Epidemic Model

In this subsection, we consider a specific example and show that the conditions for Turing instability can hold. Consider the following SI model, which comes from [164]:

$$\begin{aligned}\frac{\partial S}{\partial t} &= rN\left(1-\frac{N}{K}\right)-\beta\frac{SI}{N}-\mu S+D_S\frac{\partial^2 S}{\partial x^2},\\ \frac{\partial I}{\partial t} &= \beta\frac{SI}{N}-(\mu+\gamma)I+D_I\frac{\partial^2 I}{\partial x^2},\end{aligned} \tag{15.50}$$

where $N(x,t)=S(x,t)+I(x,t)$. This model is considered for $x\in[0,L]$ and $t\in[0,\infty)$. The model is equipped with Neumann boundary conditions:

$$S_x(0,t)=S_x(L,t)=I_x(0,t)=I_x(L,t)=0.$$

First we look for a spatially homogeneous solution (S^*,I^*), where S^* and I^* satisfy the system

$$\begin{aligned}rN\left(1-\frac{N}{K}\right)-\beta\frac{SI}{N}-\mu S &= 0,\\ \beta\frac{SI}{N}-(\mu+\gamma)I &= 0.\end{aligned} \tag{15.51}$$

We define the reproduction number

$$\mathscr{R}_0=\frac{\beta}{\mu+\gamma}.$$

The model has the disease-free spatially homogeneous equilibrium $\mathscr{E}_0=(K(1-\frac{\mu}{r}),0)$, but we will be more interested in the endemic equilibrium $\mathscr{E}^*=(S^*,I^*)$. To find the values of the endemic equilibrium, we assume $\mathscr{R}_0>1$. We note that $N^*=S^*+I^*$. Correspondingly, $1=s^*+i^*$, where $s^*=S^*/N^*$ and $i^*=I^*/N^*$. From the second equation in (15.51), we have $s^*=1/\mathscr{R}_0$. Hence, $i^*=1-1/\mathscr{R}_0$. To find N^*, we add the two equations, and we have

$$rN^*\left(1-\frac{N^*}{K}\right)=\gamma I^*+\mu N^*.$$

Dividing by N^*, replacing i^* with its value, and solving for N^*, we obtain

$$N^*=K\left(1-\frac{1}{\mathscr{R}_d}\right),$$

where we have defined the following demographic reproduction number:

$$\mathscr{R}_d = \frac{r\mathscr{R}_0}{\mu\mathscr{R}_0 + \gamma(\mathscr{R}_0 - 1)}.$$

Finally, we have

$$\begin{aligned} S^* &= \frac{1}{\mathscr{R}_0} K\left(1 - \frac{1}{\mathscr{R}_d}\right), \\ I^* &= \left(1 - \frac{1}{\mathscr{R}_0}\right) K\left(1 - \frac{1}{\mathscr{R}_d}\right). \end{aligned} \tag{15.52}$$

Proposition 15.1. *If $\mathscr{R}_0 > 1$ and $\mathscr{R}_d > 1$, then there exists a unique spatially homogeneous endemic equilibrium.*

To search for Turing instability, we derive the characteristic equation (15.46):

$$\begin{vmatrix} f_S - D_S k^2 - \lambda & f_I \\ g_S & g_I - D_I k^2 - \lambda \end{vmatrix} = 0, \tag{15.53}$$

where

$$\begin{aligned} f_S &= r\left(1 - \frac{N^*}{K}\right) - \frac{r}{K}N^* - \beta i^* + \beta s^* i^* - \mu = -r\left(1 - \frac{2}{\mathscr{R}_d}\right) - \beta\left(1 - \frac{1}{\mathscr{R}_0}\right)^2 - \mu, \\ f_I &= r\left(1 - \frac{N^*}{K}\right) - \frac{r}{K}N^* - \beta s^* + \beta s^* i^* = -r\left(1 - \frac{2}{\mathscr{R}_d}\right) - \beta\frac{1}{\mathscr{R}_0^2}, \\ g_S &= \beta i^* - \beta s^* i^* = \beta\left(1 - \frac{1}{\mathscr{R}_0}\right)^2, \\ g_I &= -\beta s^* i^* = -\beta\frac{1}{\mathscr{R}_0}\left(1 - \frac{1}{\mathscr{R}_0}\right), \end{aligned} \tag{15.54}$$

where in the last expression, we have taken the value of s^* into account to simplify.

Denote the matrix whose characteristic equation is given in (15.53) by $\mathscr{M}$. Then $\mathrm{Tr}(\mathscr{M}) < 0$, since we have assumed that the spatially homogeneous steady state is locally stable (the spatially homogeneous steady state is the steady state of the ODE system obtained from assuming $D_S = D_I = 0$). On the other hand, $\mathrm{Det}(\mathscr{M})$ might be positive or negative, depending on the diffusion rates and k^2. Define

$$H(k^2) = (f_S - D_S k^2)(g_I - D_I k^2) - g_S f_I.$$

We plot $H(k^2)$ in Fig. 15.6.

Figure 15.6 shows that $H(k^2)$ becomes negative for a range of values of k. If we pick $L = \pi$, then we can take $k = 1$. So Turing instability occurs. To simulate the Turing instability, we use Mathematica, in which we code equations (15.50). We take the initial conditions to be a perturbation of the spatially homogeneous equilibrium: $S_0(x) = 0.101453 + 0.001\cos(10x), I_0(x) = 0.071 + 0.001\cos(10x)$. We simulate for a long time, and we look for a spatially heterogeneous solution in time. In Fig. 15.7, we plot $I(x,t)$, where the solution becomes heterogeneous in space in later times.

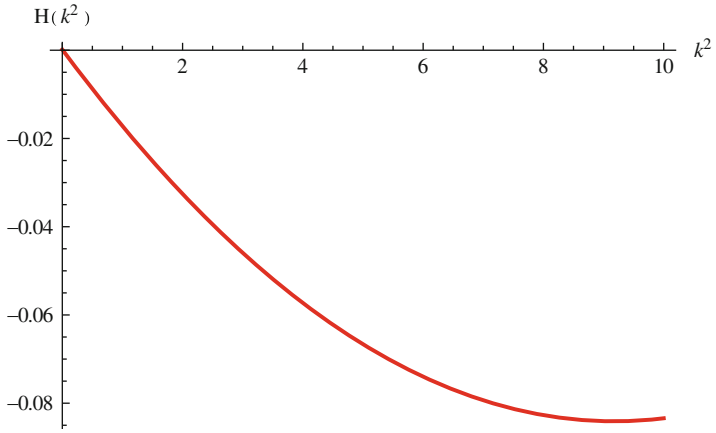

Fig. 15.6 Plot of $H(k^2)$ for $r = 0.25$, $\gamma = 0.05$, $\mu = 0.0001$, $D_S = 0.0001$, $D_I = 10$, $K = 1$, $\beta = 0.08517$. In this case, $\mathscr{R}_0 = 1.7$ and $\mathscr{R}_d = 1.20842$. We have $H(0) = 0.000151214$, which is positive. Clearly, $H(k^2)$ becomes negative for a range of values of k. Thus, Turing instability occurs

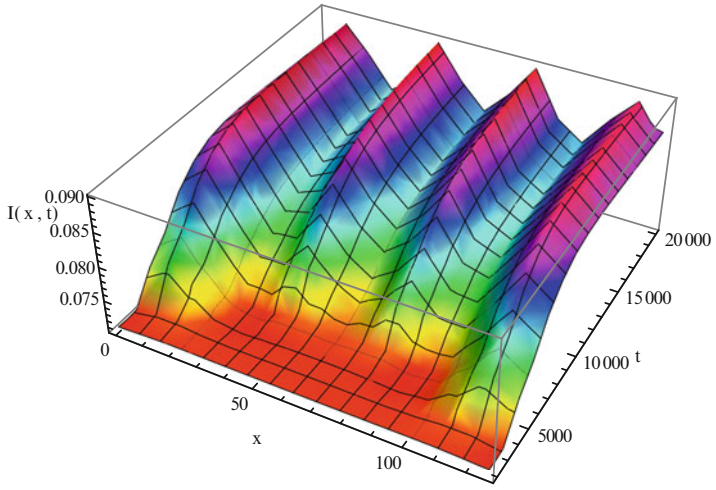

Fig. 15.7 Plot of $I(x,t)$. Parameters are chosen as $r = 0.25$, $\gamma = 0.05$, $\mu = 0.0001$, $D_S = 0.0001$, $D_I = 10$, $K = 1$, $\beta = 0.08517$, and $L = 40\pi$. In this case, $\mathscr{R}_0 = 1.7$ and $\mathscr{R}_d = 1.20842$. The solution exhibits spatial heterogeneity for $T = 20{,}000$. Thus, Turing instability occurs

Problems

15.1. Two-Patch Lagrange Movement Model

Consider model (15.1) with $n = 2$.

(a) Use the next-generation approach to compute an explicit formula for $\mathscr{R}_0$.
(b) Use a computer algebra system to simulate the case in which $\mathscr{R}_1 < 1$ and $\mathscr{R}_2 < 1$ but $\mathscr{R}_0 > 1$ and the disease persists in the metapopulation.

15.2. Two-Patch Lagrange Movement Model
Consider model (15.1) with $n = 2$. Show that the model has a unique endemic equilibrium if $\mathscr{R}_0 > 1$.

15.3. Two-Patch Eulerian Movement Model
Consider model (15.8) with $n = 2$.

(a) Compute the disease-free equilibrium. Use the next-generation approach to compute an explicit formula for $\mathscr{R}_0$.
(b) Compute the elasticities of $\mathscr{R}_0$ with respect to the migration rates. How do the migration rates affect the reproduction number?
(c) The patch reproduction numbers are given in (15.11). Show that $\min\{\mathscr{R}_1, \mathscr{R}_2\} < \mathscr{R}_0 < \max\{\mathscr{R}_1, \mathscr{R}_2\}$.

Hint: Assume without loss of generality that $\mathscr{R}_1 > \mathscr{R}_2$. Use the next-generation approach and derive the equation whose principal solution gives $\mathscr{R}_0$. Plot this quadratic polynomial and show that its value at $\mathscr{R}_1$ is positive, while at $\mathscr{R}_2$, it is negative. What does that mean?

15.4. Two-Patch Eulerian Movement Model
Consider model (15.8) with $n = 2$. Show that the model has a unique endemic equilibrium if $\mathscr{R}_0 > 1$.

15.5. SIS Model with Diffusion
Consider the following SIS model with diffusion:

$$\begin{aligned}\frac{\partial S}{\partial t} &= -\beta S(x,t)I(x,t) + \gamma I(x,t) + D\frac{\partial^2 S}{\partial x^2},\\ \frac{\partial I}{\partial t} &= \beta S(x,t)I(x,t) - \gamma I(x,t) + D\frac{\partial^2 I}{\partial x^2},\end{aligned} \tag{15.55}$$

with Neumann boundary conditions

$$\lim_{x\to-\infty}\frac{\partial S}{\partial x} = \lim_{x\to\infty}\frac{\partial S}{\partial x} = \lim_{x\to-\infty}\frac{\partial I}{\partial x} = \lim_{x\to\infty}\frac{\partial I}{\partial x} = 0.$$

(a) Argue that the total population size is constant, and eliminate the equation for S.
(b) Show that the diffusion epidemic model that consists of a single equation for I has a traveling-wave solution.
(c) Simulate a traveling-wave solution for system (15.55).

15.6. SIR Model with Diffusion in Infectives
The following SIR model with diffusion in infectives was proposed in [82] to model rabies in foxes:

$$\begin{aligned}\frac{\partial S}{\partial t} &= -\beta S(x,t)I(x,t),\\ \frac{\partial I}{\partial t} &= \beta S(x,t)I(x,t) - \gamma I(x,t) + D\frac{\partial^2 I}{\partial x^2},\end{aligned} \tag{15.56}$$

where the equation for the recovered has been omitted. Consider model (15.56) with Neumann boundary conditions

$$\lim_{x\to-\infty}\frac{\partial S}{\partial x}=\lim_{x\to\infty}\frac{\partial S}{\partial x}=0,\qquad \lim_{x\to-\infty}\frac{\partial I}{\partial x}=\lim_{x\to\infty}\frac{\partial I}{\partial x}=0.$$

(a) Look for a traveling-wave solution and reduce the system to two equations of a single variable $z=x-vt$:

$$\begin{aligned} vS' &= \beta SI,\\ DI''+vI'+\beta SI-\gamma I &= 0. \end{aligned} \tag{15.57}$$

(b) Express I from the first equation $I=\frac{vS'}{\beta S}$ and substitute it in the second, to obtain

$$DI''+vI'+vS'-\gamma\frac{vS'}{\beta S}=0.$$

Integrate the above equation from z to ∞ and obtain the following planar system for the traveling-wave solution:

$$\begin{aligned} S' &= \frac{\beta}{v}SI,\\ I' &= -\frac{v}{D}I-\frac{v}{D}S+\frac{\gamma v}{\beta D}\ln S+\frac{v}{D}S_\infty-\frac{\gamma v}{\beta D}\ln S_\infty, \end{aligned} \tag{15.58}$$

where $\lim_{z\to\infty}S=S_\infty$.

(c) Show that if

$$\frac{\beta S_\infty}{\gamma}<1, \tag{15.59}$$

then system (15.58) has two equilibria. Otherwise, it has one.

(d) Show that if (15.59) holds, then one of the equilibria is a saddle, and the other one is stable. Argue that there are a homoclinic orbit and a traveling-wave solution.

(d) Simulate a traveling-wave solution for system (15.56).

15.7. SI Model with Diffusion: Turing Instability

Consider the following model:

$$\begin{aligned} S_t &= rS\left(1-\tfrac{S}{K}\right)-\tfrac{\beta SI}{1+\alpha S}-\mu S+D_S S_{xx},\\ I_t &= \tfrac{\beta SI}{1+\alpha S}-(\mu+\gamma)I+D_I I_{xx}, \end{aligned} \tag{15.60}$$

with Neumann boundary conditions.

(a) Compute the disease-free equilibrium, the spatially homogeneous endemic equilibrium, and the reproduction number.

(b) Derive conditions for the stability of the endemic equilibrium in the case $D_S=D_I=0$.

(c) Show that Turing instability may occur.
(d) Simulate the spatially heterogenous solutions for a given time.

15.8. SI Model with Diffusion: Turing Instability
Consider the following model, in which the transmission rate is assumed not constant but linear in I:

$$\begin{aligned} S_t &= \Lambda - \beta(1+\nu I)IS - \mu S + D_S S_{xx}, \\ I_t &= \beta(1+\nu I)IS - (\alpha+\mu)I + D_I I_{xx}, \end{aligned} \tag{15.61}$$

with Neumann boundary conditions.

(a) Compute the disease-free equilibrium, the spatially homogeneous endemic equilibria, and the reproduction number.
(b) Derive conditions for the stability of the endemic equilibrium when $\mathscr{R}_0 > 1$ in the case $D_S = D_I = 0$.
(c) Show that Turing instability may occur.
(d) Simulate the spatially heterogenous solutions for given times.

Chapter 16
Discrete Epidemic Models

16.1 Single-Species Discrete Population Models

The continuous population models that we have considered in previous chapters model population and epidemic processes that occur continuously in time. In particular, they assume that births and deaths in the population occur continuously. This assumption is true for the human population, but many insect and plant populations have discrete, nonoverlapping generations. Such populations reproduce during specific time intervals of the year. Consequently, population censuses are taken at those specific times. As a result, modeling such populations and the distribution of disease in them should happen at discrete times. In this chapter we introduce discrete single-species population and epidemic models.

16.1.1 Simple Discrete Population Models

We assume that we measure the population at discrete, equally spaced, moments of time: $t_0, t_1, \ldots, t_n, \ldots$, and we find that the population numbers at these moments of time are N_t, where t takes the values of $t_0, t_1, \ldots, t_n, \ldots$. For simplicity, we will set $N_{t_n} = N_n$. Thus, the population size is described by a sequence: $N_1, N_2, \ldots, N_n, \ldots$. A discrete population model can be written in the following general form:

$$N_{n+1} = \mathscr{F}(N_n), \tag{16.1}$$

where $\mathscr{F}$ is a specified function of N_n. That is, if we know the population size at time t_n, the model tells us what the populations size at time t_{n+1} should be. Such a model is equipped with a given initial condition: the population size N_0 at time t_0 is given. Another way to rewrite Eq. (16.1) is

$$N_{n+1} = N_n f(N_n). \tag{16.2}$$

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3_16

The function $f(N_n)$ is called a *fitness function* or *per capita rate of population growth* or *net reproduction rate*.

Definition 16.1. Equations of the form (16.1) are called *difference equations*.

Such difference equations are of *first order*, because they contain only one time step. They are also *autonomous*, because $\mathscr{F}$ does not depend explicitly on the time t_n. The simplest discrete population model is derived under the assumption that individuals die with constant probability d. Furthermore, we assume that b individuals are born per individual in the population. The model then becomes

$$N_{n+1} = N_n + bN_n - dN_n,$$

that is, the number of individuals at the time step t_{n+1} is the number from the time step t_n plus those who have been born, minus those who have died. Defining $\mathscr{R} = 1 + b - d$, we obtain the following *linear* discrete equation of population growth:

$$N_{n+1} = \mathscr{R} N_n. \tag{16.3}$$

The parameter $\mathscr{R}$ is called the *net reproduction number*. We note that $\mathscr{R} > 0$, since b and d are probabilities and are less than one. Model (16.3) is a discrete analogue of the Malthusian equation. Equation (16.3) is a special case of Eq. (16.2) with $f(N_n) = \mathscr{R}$. Model (16.3) can be solved. Given initial population size N_0, we have

$$\begin{aligned} N_1 &= \mathscr{R} N_0, \\ N_2 &= \mathscr{R} N_1 = \mathscr{R}^2 N_0, \\ &\vdots \\ N_n &= \mathscr{R} N_{n-1} = \mathscr{R}^n N_0. \end{aligned} \tag{16.4}$$

If $\mathscr{R} > 1$, then each individual on average leaves more than one descendant, and the population grows geometrically. If $\mathscr{R} < 1$, then each individual leaves fewer than one descendant, and the population declines geometrically. If $\mathscr{R} = 1$, the population remains constant. These model predictions are valid under the assumption that the resources are unlimited.

In practice, populations do not experience unlimited growth, so models that predict asymptotically bounded growth are more realistic. One such model is the discrete analogue of the logistic equation. To derive such an analogue, we approximate the continuous time derivative with $N_{n+1} - N_n$, assuming that the time step is equal to one. Thus the discrete logistic equation takes the form

$$N_{n+1} = N_n + rN_n \left(1 - \frac{N_n}{K}\right). \tag{16.5}$$

First we factor N_n and $r+1$. Furthermore, we make the following changes in dependent variables and parameters:

$$y_n = \frac{r}{r+1}\frac{N_n}{K} \qquad a = r + 1.$$

We obtain a classical form for the discrete logistic equation:

$$y_{n+1} = a y_n (1 - y_n).$$

This method for producing discrete equations is not foolproof, however. The discrete logistic equation above is not well posed, in the sense that its solutions can become negative. This is not hard to see. Suppose we start from $y_0 = 0.5$ and $a = 6$. Then $y_1 = 1.5$. Consequently, $y_2 < 0$. Thus, the logistic equation is not a very good discrete population model.

We can derive a discrete version of the **simplified logistic model**. Suppose the population increases in each time interval by a constant amount Λ, and that $\gamma \leq 1$ is the probability for survival of individuals to the next time period. Then the **simplified logistic model** takes the form

$$N_{n+1} = \Lambda + \gamma N_n. \tag{16.6}$$

This model can also be solved explicitly:

$$\begin{aligned} &N_1 = \Lambda + \gamma N_0, \\ &N_2 = \Lambda + \gamma(\Lambda + \gamma N_0) = \gamma^2 N_0 + (1+\gamma)\Lambda, \\ &\vdots \\ &N_n = \gamma^n N_0 + (1 + \gamma + \cdots + \gamma^{n-1})\Lambda. \end{aligned} \tag{16.7}$$

Hence,

$$N_n = \begin{cases} N_0 + \Lambda n, & \gamma = 1, \\ \gamma^n \left(N_0 - \dfrac{\Lambda}{1-\gamma} \right) + \dfrac{\Lambda}{1-\gamma}, & \gamma < 1. \end{cases} \tag{16.8}$$

Other discrete population models have been proposed that guarantee that the population remains positive for all times. One such model, proposed by Bill Ricker [138], is the **Ricker model**:

$$N_{n+1} = N_n e^{r\left(1 - \frac{N_n}{K}\right)}. \tag{16.9}$$

Another model also widely used is the **Beverton–Holt model** [23], also called the Verhulst equation:

$$N_{n+1} = \frac{r N_n}{A + N_n}. \tag{16.10}$$

A generalization of the Beverton–Holt model can be made that is known as the **Hassell equation** [72]:

$$N_{n+1} = \frac{r N_n}{(A + N_n)^b}, \tag{16.11}$$

where $b > 0$ is a positive parameter.

16.1.2 Analysis of Single-Species Discrete Models

Difference equations also have solutions that do not depend on time, called *equilibria*. Since the solution does not depend on time, all members of the sequence have the same value, that is, we have

$$N_n = N^* \qquad \text{for} \quad \text{all} \quad n \geq 0.$$

Consequently, equilibria of the difference equation (16.1) must satisfy $N^* = \mathscr{F}(N^*)$.

Definition 16.2. A value N^* that satisfies

$$N^* = \mathscr{F}(N^*)$$

is called a *fixed point* of the function $\mathscr{F}$.

Example 16.1. Consider the equilibria of the logistic equation

$$N^* = N^* + rN^*\left(1 - \frac{N^*}{K}\right).$$

The solutions of this equation are $N_1^* = 0$ and $N_2^* = K$, that is, the equilibria in the discrete case are exactly the same as in the continuous case. The equilibrium $N_1^* = 0$ is called a *trivial equilibrium*, while the equilibrium $N_2^* = K$ is called a *nontrivial equilibrium*.

To describe the behavior of the solutions near an equilibrium, we use again a process called *linearization*. Let N^* be the equilibrium, and u_n the perturbation of the solution from the equilibrium, that is,

$$N_n = N^* + u_n.$$

Substituting this equation into Eq. (16.1), we have $u_{n+1} + N^* = \mathscr{F}(u_n + N^*)$. Expanding $\mathscr{F}$ in a Taylor series and neglecting all terms containing powers of u_n greater than one, we obtain

$$u_{n+1} + N^* = \mathscr{F}(N^*) + \mathscr{F}'(N^*)u_n.$$

Recall that since N^* is an equilibrium, we have $N^* = \mathscr{F}(N^*)$. Hence, we obtain the following linearized equation:

$$u_{n+1} = \mathscr{F}'(N^*)u_n. \tag{16.12}$$

We note that $\mathscr{F}'(N^*)$ is a fixed number, which may be positive or negative. If we consider

$$u_{n+1} = |\mathscr{F}'(N^*)|u_n, \tag{16.13}$$

then Eq. (16.13) is exactly the discrete Malthus equation. Consequently, we have the following:

1. If $|\mathscr{F}'(N^*)| < 1$, then $u_n \to 0$. Hence, $N_n - N^* \to 0$ and $N_n \to N^*$. This is the case if N_0 is close enough to N^*, that is, this result is local. In this case, we call N^* *locally asymptotically stable*.
2. If $|\mathscr{F}'(N^*)| > 1$, then $u_n \to \infty$. Hence $N_n - N^* \to \infty$, and N_n diverges from N^*. This is the case if N_0 is close enough to N^*. In this case, we call N^* *unstable*.

We note that if $|\mathscr{F}'(N^*)| = 1$, we cannot draw conclusions from the local analysis.

We summarize the above discussion in the following theorem:

Theorem 16.1. *The equilibrium N^* of the discrete equation (16.1) is locally asymptotically stable if and only if $|\mathscr{F}'(N^*)| < 1$. The equilibrium N^* of the discrete equation (16.1) is unstable if and only if $|\mathscr{F}'(N^*)| > 1$.*

To illustrate the use of the theorem above, we consider the local stability of the equilibria of the logistic equation.

Example 16.2. In the case of the logistic equation (16.5), the function $\mathscr{F}$ is given by

$$\mathscr{F}(N) = N + rN\left(1 - \frac{N}{K}\right).$$

The derivative is given by

$$\mathscr{F}'(N) = 1 + r\left(1 - \frac{N}{K}\right) - \frac{r}{K}N.$$

In the case of the trivial equilibrium $N^* = 0$, we have

$$\mathscr{F}'(0) = 1 + r > 1.$$

Consequently, the trivial equilibrium is always unstable. Now we consider the nontrivial equilibrium $N^* = K$. We have

$$\mathscr{F}'(K) = 1 - r.$$

So if $|1 - r| < 1$, or equivalently, if $0 < r < 2$, then the nontrivial equilibrium is locally asymptotically stable.

When $r > 2$, simulations suggests that the logistic equation can experience very complex behavior. To investigate this behavior through simulations, we will study the nondimensionalized version of the logistic equation:

$$y_{n+1} = \rho y_n(1 - y_n). \tag{16.14}$$

Recall that $\rho = 1 + r$, so we can expect complex behavior for $\rho > 3$. We notice that the corresponding equilibria of the nondimensional logistic model are $y^* = 0$ and $y^* = 1$. The first complexity that appears is a 2-cycle.

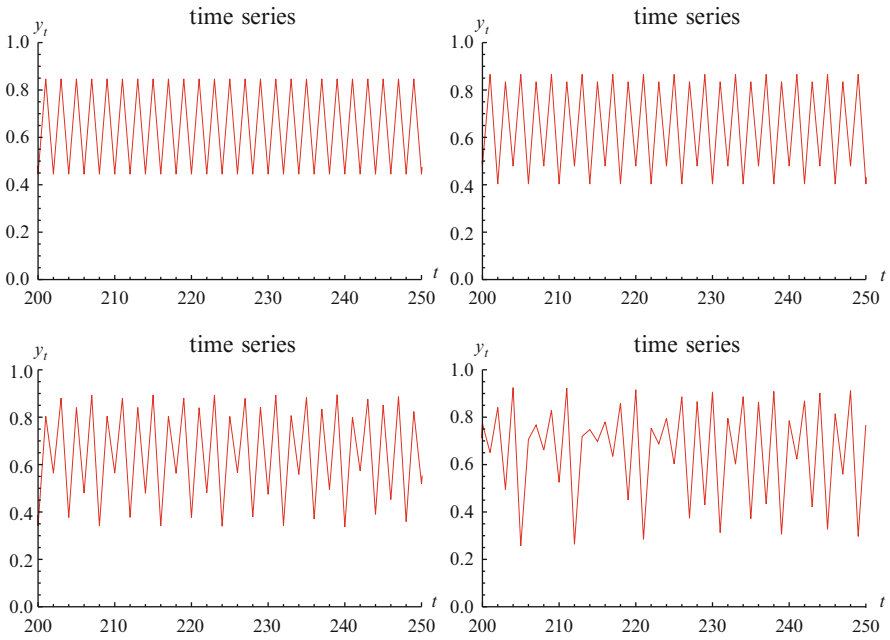

Fig. 16.1 The figure shows time series of the model $y_{n+1} = \rho y_n(1-y_n)$ for different values of ρ. The first figure shows a two-cycle with $\rho = 3.43$. The second figure on the *top line* shows a 4-cycle with $\rho = 3.47$. The *left* figure on the *bottom row* shows an 8-cycle with $\rho = 3.58$. The *right* figure on the *bottom row* shows chaos with $\rho = 3.7$

Definition 16.3. A *2-cycle* of model (16.1) is a system of two solutions y_1 and y_2 such that

$$\begin{aligned} y_1 &= \mathscr{F}(y_2), \\ y_2 &= \mathscr{F}(y_1). \end{aligned} \tag{16.15}$$

In model (16.14), $\mathscr{F}(y) = \rho y(1-y)$. As ρ increases, the system experiences a process, called *period-doubling*, to a 4-cycle. Similarly, a *4-cycle* of model (16.14) is a system of four solutions y_1, y_2, y_2, y_4 such that

$$\begin{aligned} y_1 &= \mathscr{F}(y_4), \\ y_2 &= \mathscr{F}(y_1), \\ y_3 &= \mathscr{F}(y_2), \\ y_4 &= \mathscr{F}(y_3). \end{aligned} \tag{16.16}$$

Further period-doubling occurs to an 8-cycle. The period-doubling continues until the system begins to exhibit chaos. We illustrate period-doubling and chaos in Fig. 16.1.

We need single-species discrete population models to capture the demographic processes in epidemic models. Many books focus on single-species discrete models and provide an excellent introduction to these models (for instance, see [27, 90]).

16.2 Discrete Epidemic Models

Just like single-species population models, discrete epidemic models can also be obtained from a discretization of the continuous epidemic models. However, this approach results in models that have issues like those of the discrete logistic equation. To avoid these problems, a modeling approach specific to discrete models should be taken. We follow here the approach of Castillo-Chavez and Yakubu [39].

16.2.1 A Discrete SIS Epidemic Model

We begin with a general population model

$$N_{n+1} = f(N_n) + \gamma N_n, \tag{16.17}$$

where $\gamma < 1$ is the probability of survival to the next time period, and $f(N_n)$ is a recruitment function. We assume that the disease does not affect the population dynamics, that is, we assume that the disease is nonfatal and does not affect the birth process. We will build an SIS epidemic process on top of the demographic process. We denote by S_n and I_n the susceptible and infected individuals at time t_n. Individuals survive with probability $\gamma < 1$ (die with probability $1-\gamma$) in each generation. Infected individuals recover with probability $1-\sigma$ (do not recover with probability $\sigma < 1$) in each generation. In each generation, susceptible individuals become infected with probability $1-G$ (remain susceptible with probability G). The function G is a function of the prevalence I_n/N_n, which is weighted with coefficient α. The model assumes a sequential process: at each generation, γS_n susceptibles survive, and the surviving susceptibles become infected with probability $1-G$. Similarly, γI_n infected individuals survive, and the surviving ones recover with probability $(1-\sigma)$:

$$\begin{aligned} S_{n+1} &= f(N_n) + \gamma S_n G\left(\frac{\alpha I_n}{N_n}\right) + \gamma(1-\sigma)I_n, \\ I_{n+1} &= \gamma S_n \left(1 - G\left(\frac{\alpha I_n}{N_n}\right)\right) + \gamma\sigma I_n. \end{aligned} \tag{16.18}$$

The function G must satisfy the following conditions:

1. $G : [0,\infty) \to [0,1]$.
2. $G(0) = 1$.
3. G is a monotone decreasing function with $G'(x) < 0$ and $G''(x) \geq 0$.

An example of such a function that we will use is $G(x) = e^{-x}$. Another example is $G(x) = A/(x+A)$. Adding the two equations in system (16.18) gives Eq. (16.17).

16.2.2 Analysis of Multidimensional Discrete Models

In this subsection, we introduce the techniques that help us analyze systems of discrete equations. Suppose that we are given the following system:

$$\mathbf{x}_{n+1} = \mathscr{F}(\mathbf{x}_n), \tag{16.19}$$

where $\mathbf{x}$ is an M-dimensional vector of variables. As before, an *equilibrium* of system (16.19) is the solution of the problem

$$\mathbf{x}^* = \mathscr{F}(\mathbf{x}^*).$$

To find the behavior of the solutions near an equilibrium, we use linearization. We set $\mathbf{x}_n = \mathbf{u}_n + \mathbf{x}^*$. We obtain the following linear system:

$$\mathbf{u}_{n+1} = \mathbf{J}(\mathbf{x}^*)\mathbf{u}_n, \tag{16.20}$$

where $\mathbf{J}$ is the Jacobian of the system, that is,

$$\mathbf{J}(\mathbf{x}^*) = \left(\begin{array}{ccc} \dfrac{\partial \mathscr{F}_1}{\partial x_1} & \cdots & \dfrac{\partial \mathscr{F}_1}{\partial x_M} \\ & \vdots & \\ \dfrac{\partial \mathscr{F}_M}{\partial x_1} & \cdots & \dfrac{\partial \mathscr{F}_M}{\partial x_M} \end{array} \right) |_{\mathbf{x}=\mathbf{x}^*}. \tag{16.21}$$

Definition 16.4. An equilibrium point $\mathbf{x}^*$ is said to be *locally asymptotically stable* if there exists a neighborhood U of x^* such that for each starting value $\mathbf{x}_0 \in U$, we get

$$\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}^*. \tag{16.22}$$

The equilibrium point $\mathbf{x}^*$ is called *unstable* if x^* is not (locally asymptotically) stable.

The limit (16.22) holds if for system (16.20), we have $\lim_{n\to\infty} \mathbf{u}_n = 0$. The following theorem gives the conditions for convergence of solutions of the linear system (16.20) to zero:

Theorem 16.2. *Let* $\mathbf{J}$ *be an* $M \times M$ *matrix with* $\rho(\mathbf{J}) < 1$*, where*

$$\rho(\mathbf{J}) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } \mathbf{J}\}.$$

Then every solution of (16.20) satisfies

$$\lim_{n\to\infty} \mathbf{u}_n = 0.$$

If $\rho(\mathbf{J}) > 1$*, then there are solutions that go to infinity.*

This implies the following criterion for stability of an equilibrium $\mathbf{x}^*$ of system (16.19).

Theorem 16.3. *Consider the nonlinear autonomous system (16.19). Suppose* $\mathscr{F} : \mathscr{D} \to \mathscr{D}$, *where* $\mathscr{D} \subset \mathbf{R}^M$ *and* $\mathscr{D}$ *is an open set. Suppose* $\mathscr{F}$ *is twice continuously differentiable in some neighborhood of a fixed point* $\mathbf{x}^* \in \mathscr{D}$. *Let* $\mathbf{J}(\mathbf{x}^*)$ *be the Jacobian matrix of* $\mathscr{F}$ *evaluated at* $\mathbf{x}^*$. *Then the following hold:*

1. $\mathbf{x}^*$ *is locally asymptotically stable if all eigenvalues of* $\mathbf{J}(\mathbf{x}^*)$ *have magnitude less than one.*
2. $\mathbf{x}^*$ *is unstable if at least one eigenvalue of* $\mathbf{J}(\mathbf{x}^*)$ *has magnitude greater than one.*

The Routh–Hurwitz criterion will not be helpful here in determining which matrices are stable, since Routh–Hurwitz identifies matrices whose eigenvalues lie in the left half of the complex plane. However, there is an analogous criterion that can help determine whether the spectral radius of a matrix is smaller than one. This criterion is called the **Jury conditions**. Let

$$p(\lambda) = |\mathbf{J} - \lambda I| = a_M \lambda^M + \cdots + a_1 \lambda + a_0,$$

where $a_M = 1$. To introduce the Jury conditions, we first have to introduce the **Jury array**. The Jury array is composed as follows: First we write out a row of the coefficients, and then we write out another row with the same coefficients in reverse order. The first two rows of the Jury array are composed of the coefficients of the polynomial $p(\lambda)$ above. Once we have the first two rows of the **a** coefficients, the next two rows are of the **b** coefficients, and so on. We obtain the array of Table 16.1, where the **b** coefficients, **c** coefficients, etc., are composed as follows:

Table 16.1 Jury array

Number	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
(1)	a_0	a_1	$\ldots$	a_{M-1}	a_M
(2)	a_M	a_{M-1}	$\ldots$	a_1	a_0
(3)	b_0	b_1	$\ldots$	b_{M-1}	
(4)	b_{M-1}	b_{M-2}	$\ldots$	b_0	
$\vdots$	$\vdots$	$\vdots$	$\vdots$		
(2M-3)	v_0	v_1	v_2		

$$b_k = \begin{vmatrix} a_0 & a_{M-k} \\ a_M & a_k \end{vmatrix} \qquad c_k = \begin{vmatrix} b_0 & b_{M-1-k} \\ b_{M-1} & b_k \end{vmatrix} \qquad d_k = \begin{vmatrix} c_0 & c_{M-2-k} \\ c_{M-2} & c_k \end{vmatrix}.$$

Jury Conditions
The Jury conditions require *all* of the following conditions to be met. If all the conditions are satisfied, then the spectral radius of the matrix is less than one, and the matrix is stable:

1. $p(1) > 0$.
2. $(-1)^M p(-1) > 0$.
3. $|a_0| < a_M$.
4. Once the Jury array has been composed, the Jury conditions also require

$$\begin{aligned} |b_0| &> |b_{M-1}|, \\ |c_0| &> |c_{M-2}|, \\ |d_0| &> |d_{M-3}|. \\ &\vdots \end{aligned} \tag{16.23}$$

In the case $M = 1$, the Jury conditions do not apply, but in this case, the eigenvalue is known explicitly, and its magnitude can be compared with one. In the cases $M = 2, 3, 4$, we write the Jury conditions in Table 16.2.

Table 16.2 Jury Conditions

Degree	Condition	Condition	Condition	Condition	Condition
$M=2$	$p(1)>0$	$p(-1)>0$	$\|a_0\|<1$		
$M=3$	$p(1)>0$	$p(-1)<0$	$\|a_0\|<1$	$\|a_0^2-1\|>\|a_0a_2-a_1\|$	
$M=4$	$p(1)>0$	$p(-1)>0$	$\|a_0\|<1$	$\|a_0^2-1\|>\|a_0a_3-a_1\|$	$\|b_0^2-b_3^2\|>\|b_0b_2-b_3b_1\|$

16.2.3 Analysis of the SIS Epidemic Model

In this section, we analyze model (16.18) with a specific fertility function. In particular, we choose the discrete simplified logistic model, where we know that the population tends to a constant size as $n \to \infty$. We will study the following epidemic model with a general force of infection G:

$$\begin{aligned} S_{n+1} &= \Lambda + \gamma S_n G\left(\frac{\alpha I_n}{N_n}\right) + \gamma(1-\sigma)I_n, \\ I_{n+1} &= \gamma S_n \left(1 - G\left(\frac{\alpha I_n}{N_n}\right)\right) + \gamma\sigma I_n. \end{aligned} \tag{16.24}$$

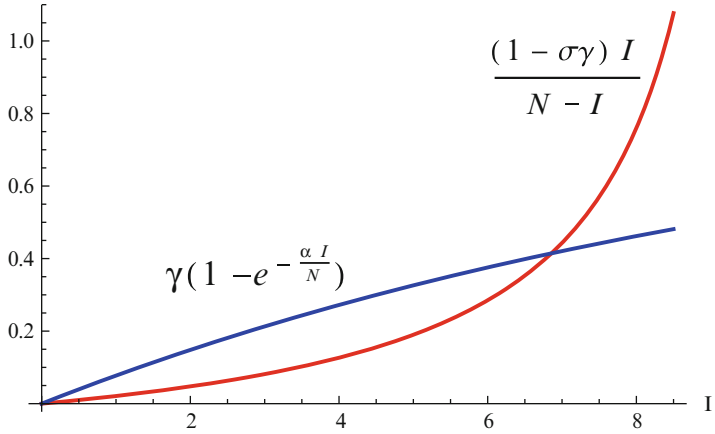

Fig. 16.2 The figure shows the two functions on the two sides of Eq. (16.27). Here $\alpha = 0.9$, $\sigma = 0.9$, $\gamma = 0.9$, and $N = 10$

The equilibria of the system above satisfy

$$\begin{aligned} S &= \Lambda + \gamma S G\left(\frac{\alpha I}{N}\right) + \gamma(1-\sigma)I, \\ I &= \gamma S\left(1 - G\left(\frac{\alpha I}{N}\right)\right) + \gamma\sigma I. \end{aligned} \tag{16.25}$$

Adding the equations, we have $N = \Lambda + \gamma N$. Hence $N = \Lambda/(1-\gamma)$. The system clearly has the disease-free equilibrium $\mathscr{E}_0 = (N, 0)$. To find the endemic equilibria, we write $S = N - I$ and substitute in the equation for I:

$$(1-\sigma\gamma)I = \gamma(N-I)\left(1 - G\left(\frac{\alpha I}{N}\right)\right). \tag{16.26}$$

This is a nonlinear equation for I. It has $I = 0$ as a solution. We need to find a condition under which this equation has a nonzero solution. The equation can be rewritten also as

$$(1-\sigma\gamma)\frac{I}{N-I} = \gamma\left(1 - G\left(\frac{\alpha I}{N}\right)\right). \tag{16.27}$$

The function on the right is increasing and concave down. The function on the left is increasing and concave up, tending to infinity as $I \to N$. Besides the common point at zero, these functions have another unique common point if and only if the slope at zero of the function on the left is smaller than the slope at zero of the function on the right (see Fig. 16.2), that is, if

$$(1-\sigma\gamma) < -\alpha\gamma G'(0).$$

This condition gives the reproduction number. We define

$$\mathscr{R}_0 = \frac{-\alpha\gamma G'(0)}{(1-\sigma\gamma)}. \tag{16.28}$$

We note that the reproduction number is positive, since $G'(0) < 0$. We summarize these results in the following proposition:

Proposition 16.1. *Assume $\mathscr{R}_0 < 1$. Then model (16.24) has only the disease-free equilibrium $\mathscr{E}_0 = (N,0)$. If $\mathscr{R}_0 > 1$, then model (16.24) has the disease-free equilibrium and a unique endemic equilibrium $\mathscr{E}^* = (S^*, I^*)$, where $I^* > 0$ is the unique positive solution of Eq. (16.27) and $S^* = N - I^*$.*

We use the theoretical results in the previous subsection to establish the local stability of equilibria. The following theorem summarizes the results:

Theorem 16.4. *The disease-free equilibrium is locally asymptotically stable if $\mathscr{R}_0 < 1$ and unstable if $\mathscr{R}_0 > 1$. The endemic equilibrium is locally asymptotically stable if $\mathscr{R}_0 > 1$.*

Proof. We begin by computing the generic form of the Jacobian:

$$J = \begin{pmatrix} \gamma G\left(\frac{\alpha I}{N}\right) - \gamma\alpha\frac{SI}{N^2}G'\left(\frac{\alpha I}{N}\right) & \gamma\alpha\left[\frac{S}{N} - \frac{SI}{N^2}\right]G'\left(\frac{\alpha I}{N}\right) + \gamma(1-\sigma) \\ \gamma\left(1 - G\left(\frac{\alpha I}{N}\right)\right) + \gamma\alpha\frac{SI}{N^2}G'\left(\frac{\alpha I}{N}\right) & -\gamma\alpha\left[\frac{S}{N} - \frac{SI}{N^2}\right]G'\left(\frac{\alpha I}{N}\right) + \gamma\sigma \end{pmatrix}, \tag{16.29}$$

where we recall that $N = S + I$. To find the stability of the disease-free equilibrium, we evaluate the Jacobian at the disease-free equilibrium:

$$J(\mathscr{E}_0) = \begin{pmatrix} \gamma G(0) & \gamma\alpha G'(0) + \gamma(1-\sigma) \\ \gamma(1 - G(0)) & -\gamma\alpha G'(0) + \gamma\sigma \end{pmatrix}. \tag{16.30}$$

The characteristic equation now becomes $|J(\mathscr{E}_0) - \lambda I| = 0$. Recall that $G(0) = 1$, so the characteristic determinant is upper triangular, and the eigenvalues are $\lambda_1 = \gamma$ and $\lambda_2 = -\gamma\alpha G'(0) + \gamma\sigma$. Both eigenvalues are positive, and λ_1 is by assumption less than one, while λ_2 is less than one if and only if $\mathscr{R}_0 < 1$.

To determine the stability of the endemic equilibrium, we first observe that from equality (16.26), we have the following inequality:

$$(1 - \gamma\sigma) > -\gamma(1 - G(\frac{\alpha I^*}{N^*})) - \alpha\gamma\frac{S^*}{N^*}G'(\frac{\alpha I^*}{N^*}). \tag{16.31}$$

This inequality simply says that at the point where the two curves intersect, the slope of the left one is larger than the slope of the right one. This is easy to see from their graphs. The characteristic polynomial is given by

$$|J-\lambda I| = \begin{vmatrix} \gamma G\left(\frac{\alpha I}{N}\right) - \gamma\alpha\frac{SI}{N^2}G'\left(\frac{\alpha I}{N}\right) - \lambda & \gamma\alpha\left[\frac{S}{N} - \frac{SI}{N^2}\right]G'\left(\frac{\alpha I}{N}\right) + \gamma(1-\sigma) \\ \gamma\left(1 - G\left(\frac{\alpha I}{N}\right)\right) + \gamma\alpha\frac{SI}{N^2}G'\left(\frac{\alpha I}{N}\right) & -\gamma\alpha\left[\frac{S}{N} - \frac{SI}{N^2}\right]G'\left(\frac{\alpha I}{N}\right) + \gamma\sigma - \lambda \end{vmatrix}. \tag{16.32}$$

We can manipulate the determinant to simplify the characteristic polynomial. In particular, adding the first line to the second, we have

$$|J-\lambda I| = \begin{vmatrix} \gamma G\left(\frac{\alpha I^*}{N^*}\right) - \gamma\alpha\frac{SI}{N^2}G'\left(\frac{\alpha I^*}{N^*}\right) - \lambda & \gamma\alpha\left[\frac{S^*}{N^*} - \frac{S^*I^*}{N^{*2}}\right]G'\left(\frac{\alpha I^*}{N^*}\right) + \gamma(1-\sigma) \\ \gamma - \lambda & \gamma - \lambda \end{vmatrix} = 0. \tag{16.33}$$

Factoring out $\gamma-\lambda$, we see that one of the eigenvalues is $\lambda_1 = \gamma$. This eigenvalue is positive and less than one. The second eigenvalue is obtained from the remaining determinant

$$\begin{vmatrix} \gamma G\left(\frac{\alpha I^*}{N^*}\right) - \gamma\alpha\frac{SI}{N^2}G'\left(\frac{\alpha I^*}{N^*}\right) - \lambda & \gamma\alpha\left[\frac{S^*}{N^*} - \frac{S^*I^*}{N^{*2}}\right]G'\left(\frac{\alpha I^*}{N^*}\right) + \gamma(1-\sigma) \\ 1 & 1 \end{vmatrix} = 0. \tag{16.34}$$

This gives, after some simplification,

$$\lambda_2 = -\gamma\left(1 - G\left(\frac{\alpha I^*}{N^*}\right)\right) - \alpha\gamma\frac{S^*}{N^*}G'\left(\frac{\alpha I^*}{N^*}\right) + \gamma\sigma.$$

Inequality (16.31) implies that $\lambda_2 < 1$. Furthermore, $\lambda_2 > -\gamma\left(1 - G\left(\frac{\alpha I^*}{N^*}\right)\right) > -1$. Hence $|\lambda_2| < 1$, and the endemic equilibrium is locally asymptotically stable. □

In this SIS example, we did not necessarily need the Jury conditions, because the two-equation model can be reduced to a single equation if we take into account the fact that the total population size is asymptotically constant.

16.3 Discrete SEIS Model

One can formulate discrete variants of all classical continuous epidemic models. In this section, we formulate a discrete version of an SEIS model that consists of three equations: one for the susceptible S_n, one for the exposed E_n, and one for the

infectious I_n individuals. We will use again an asymptotically constant population size and a general function for the force of infection. The model takes the form

$$\begin{aligned} S_{n+1} &= \Lambda + \gamma S_n G\left(\frac{\alpha I_n}{N_n}\right) + \gamma(1-\delta)I_n, \\ E_{n+1} &= \gamma S_n \left(1 - G\left(\frac{\alpha I_n}{N_n}\right)\right) + \gamma\sigma E_n, \\ I_{n+1} &= \gamma(1-\sigma)E_n + \gamma\delta I_n, \end{aligned} \tag{16.35}$$

where γ is the probability of survival to the next time period, $1-\sigma$ is the probability of progression to infectiousness, and $1-\delta$ is the probability of recovery. Again, the function G must satisfy the following conditions:

1. $G : [0,\infty) \to [0,1]$.
2. $G(0) = 1$.
3. G is a monotone decreasing function with $G'(x) < 0$ and $G''(x) \geq 0$.

Equilibria are solutions of the following system:

$$\begin{aligned} S &= \Lambda + \gamma S G\left(\frac{\alpha I}{N}\right) + \gamma(1-\delta)I, \\ E &= \gamma S \left(1 - G\left(\frac{\alpha I}{N}\right)\right) + \gamma\sigma E, \\ I &= \gamma(1-\sigma)E + \gamma\delta I. \end{aligned} \tag{16.36}$$

Adding the three equations, we have $N = \Lambda + \gamma N$. This gives the equilibrium total population size $N = \Lambda/(1-\gamma)$. The system has the disease-free equilibrium $\mathscr{E}_0 = (\frac{\Lambda}{1-\gamma}, 0, 0)$. Problem 16.4 asks you to compute the reproduction number, which is given by the following expression:

$$\mathscr{R}_0 = \frac{-\alpha\gamma^2(1-\sigma)G'(0)}{(1-\sigma\gamma)(1-\delta\gamma)}. \tag{16.37}$$

Problem 16.4 asks you to establish the following proposition:

Proposition 16.2. *If $\mathscr{R}_0 < 1$, then the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, the disease-free equilibrium is unstable, and there is a unique endemic equilibrium.*

To obtain the equation for the endemic equilibrium, we express E in terms of I from the last equation in system (16.36): $E = QI$, where

$$Q = \frac{1-\gamma\delta}{\gamma(1-\sigma)}.$$

We can express S in terms of I: $S = N - QI - I$. We replace these values in the second equation to obtain an equation for I:

$$(1-\gamma\sigma)QI = \gamma(N-(Q+1)I)\left(1-G\left(\frac{\alpha I}{N}\right)\right). \tag{16.38}$$

Every value of I that solves Eq. (16.38) gives an equilibrium $\mathscr{E} = (S^*, E^*, I^*)$. As before, it can be seen that the equation above has a unique nontrivial equilibrium $I^* > 0$. At the unique endemic equilibrium, the slopes of the two curves are related as follows:

$$(1-\gamma\sigma)Q > -\gamma(Q+1)\left(1-G\left(\frac{\alpha I^*}{N}\right)\right) - \alpha\gamma\frac{S^*}{N}G'\left(\frac{\alpha I^*}{N}\right). \tag{16.39}$$

Replacing the value of Q and taking a common denominator leads to the inequality

$$\begin{aligned}(1-\gamma\sigma)(1-\gamma\delta) \;>\; & -\gamma(1-\gamma\delta+\gamma(1-\sigma))\left(1-G\left(\frac{\alpha I^*}{N}\right)\right) \\ & -\alpha\gamma^2(1-\sigma)\frac{S^*}{N}G'\left(\frac{\alpha I^*}{N}\right).\end{aligned} \tag{16.40}$$

Now we are ready to establish a partial result on the stability of the endemic equilibrium:

Proposition 16.3. *Assume $\mathscr{R}_0 > 1$. If*

$$\sigma + \delta + G\left(\frac{\alpha I^*}{N}\right) - 1 > 0,$$

then the unique endemic equilibrium $\mathscr{E} = (S^*, E^*, I^*)$ is locally asymptotically stable.

Proof. The Jacobian at the endemic equilibrium is given by

$$J = \begin{pmatrix} \gamma G\left(\frac{\alpha I}{N}\right) - A & -A & \gamma\alpha\frac{S}{N}G'\left(\frac{\alpha I}{N}\right) - A + \gamma(1-\delta) \\ \gamma\left(1-G\left(\frac{\alpha I}{N}\right)\right) + A & \gamma\sigma + A & -\gamma\alpha\frac{S}{N}G'\left(\frac{\alpha I}{N}\right) + A \\ 0 & \gamma(1-\sigma) & \gamma\delta \end{pmatrix}, \tag{16.41}$$

where $A = \gamma\alpha\frac{SI}{N^2}G'\left(\frac{\alpha I}{N}\right)$. We consider the characteristic equation $|J - \lambda I| = 0$. Adding the first and the third rows in the determinant to the second row, we obtain

$$\begin{vmatrix} \gamma G\left(\frac{\alpha I}{N}\right) - A - \lambda & -A & \gamma\alpha\frac{S}{N}G'\left(\frac{\alpha I}{N}\right) - A + \gamma(1-\delta) \\ \gamma - \lambda & \gamma - \lambda & \gamma - \lambda \\ 0 & \gamma(1-\sigma) & \gamma\delta - \lambda \end{vmatrix} = 0. \tag{16.42}$$

Factoring out $\gamma-\lambda$, we see that the first eigenvalue is $\lambda_1=\gamma$. This eigenvalue is positive and less than one. The remaining eigenvalues are solutions of the characteristic equation

$$\begin{vmatrix} \gamma G\left(\frac{\alpha I}{N}\right)-A-\lambda & -A & \gamma\alpha\frac{S}{N}G'\left(\frac{\alpha I}{N}\right)-A+\gamma(1-\delta) \\ 1 & 1 & 1 \\ 0 & \gamma(1-\sigma) & \gamma\delta-\lambda \end{vmatrix}=0. \quad (16.43)$$

From here we obtain the quadratic polynomial

$$\begin{aligned} p(\lambda) &= \lambda^2-(\gamma\delta+\gamma G-\gamma(1-\sigma))\lambda \\ &\quad +[-\gamma^2(1-\sigma)G+\alpha\gamma^2(1-\sigma)\frac{S}{N}G' \\ &\quad +\gamma^2(1-\sigma)(1-\delta)+\gamma^2\delta G]=0, \end{aligned} \quad (16.44)$$

where G and G' have the usual argument. We can write the polynomial as $p(\lambda)=\lambda^2+a_1\lambda+a_0$. Rewriting inequality (16.40) as

$$\begin{aligned} \alpha\gamma^2(1-\sigma)\frac{S^*}{N}G'\left(\frac{\alpha I^*}{N}\right) > &-\gamma(1-\gamma\delta+\gamma(1-\sigma))\left(1-G\left(\frac{\alpha I^*}{N}\right)\right) \\ &-(1-\gamma\sigma)(1-\gamma\delta), \end{aligned} \quad (16.45)$$

we will use it to bound the polynomial from below. Applying this inequality to the constant term of the polynomial $p(\lambda)$, we have

$$\begin{aligned} p(\lambda) &> \lambda^2-(\gamma\delta+\gamma G-\gamma(1-\sigma))\lambda \\ &\quad +[-\gamma^2(1-\sigma-\delta)G-\gamma(1-\gamma\delta+\gamma(1-\sigma))(1-G) \\ &\quad -(1-\gamma\sigma)(1-\gamma\delta)+\gamma^2(1-\sigma)(1-\delta)] \\ &= \lambda^2-(\gamma\delta+\gamma G-\gamma(1-\sigma))\lambda \\ &\quad +[\gamma^2(1-\sigma-\delta)(1-G)-\gamma(1-\gamma\delta+\gamma(1-\sigma))(1-G) \\ &\quad -(1-\gamma\sigma)(1-\gamma\delta)+\gamma^2\sigma\delta]. \end{aligned} \quad (16.46)$$

On combining the coefficients of the two terms $(1-G)$, the above right-hand side simplifies to

$$p(\lambda)>\lambda^2-(\gamma\delta+\gamma G-\gamma(1-\sigma))\lambda-\gamma(1-G)-1+\gamma\delta+\gamma\sigma.$$

We need to check the Jury conditions. Clearly, $p(1)>0$. Furthermore, according to our assumption,

$$p(-1)>2\gamma(\delta+\sigma+G-1)>0.$$

Finally, we need to show that the constant term of the polynomial $p(\lambda)$ satisfies $|a_0| < 1$. We bound the constant term from above and from below:

$$a_0 > -\gamma(1-G) - 1 + \gamma\delta + \gamma\sigma = \gamma(\delta + \sigma + G - 1) - 1 > -1.$$

In addition, if $1 - \delta - \sigma > 0$, then

$$a_0 < \gamma^2(1-\delta)(1-\sigma) - \gamma(1-\delta-\sigma)G < 1.$$

If $1 - \delta - \sigma < 0$, we have

$$a_0 < \gamma^2(1-\delta-\sigma)(1-G) + \gamma^2\sigma\delta < \gamma^2\sigma\delta < 1.$$

We conclude that $|a_0| < 1$. The Jury conditions now imply that the endemic equilibrium is stable. □

In conclusion, discrete models look simpler and perhaps more natural, but their analysis is far more complicated than the analysis of continuous models. Furthermore, even very simple single-species discrete models are capable of exhibiting very complex, even chaotic, dynamics.

16.4 Next-Generation Approach for Discrete Models

As the discrete models become more and more realistic, computation of $\mathscr{R}_0$ becomes harder or impossible to do via the Jacobian approach. In analogy with the continuous case, a version of the next-generation approach for discrete models was developed [9].

16.4.1 Basic Theory

To introduce the next-generation approach for discrete models, let $\mathbf{x} = (x^1, \ldots, x^m)^T$ be the vector of dependent variables, and let

$$\mathbf{x}_{n+1} = \mathbf{F}(\mathbf{x}_n) \qquad n = 0, 1, \ldots$$

be the dynamical system over discrete time intervals with $F : \mathbb{R}_+^m \longrightarrow \mathbb{R}_+^m$ and $F \in C^1(\mathbb{R}_+^m)$. As in the continuous case, we order the variables so that the first $k < m$, denoted by $\mathbf{y} = (y^1, \ldots, y^k)^T$, are the infected states such as exposed, infectious, isolated, and the remaining $m-k$ states $\mathbf{z} = (z^{k+1}, \ldots, z^m)^T$ are the uninfected states, such as susceptible, recovered, vaccinated. In this case, the system can be written as

$$\begin{pmatrix} \mathbf{y}_{n+1} \\ \mathbf{z}_{n+1} \end{pmatrix} = \begin{pmatrix} \mathbf{F}_0(\mathbf{x}_n) \\ \mathbf{F}_1(\mathbf{x}_n) \end{pmatrix}. \tag{16.47}$$

We assume that there exists a unique disease-free equilibrium where $\mathbf{y} = 0$, and therefore the disease-free equilibrium is given by $(0, \mathbf{z}^*)^T$. Furthermore, linearizing the discrete system around the disease-free equilibrium gives

$$\xi_{n+1} = J\xi_n,$$

where ξ_n is the vector of perturbations, and J is the Jacobian evaluated at the disease-free equilibrium. The $m \times m$ Jacobian has the following form:

$$J = \begin{pmatrix} F+T & 0 \\ A & C \end{pmatrix}, \tag{16.48}$$

where $k \times k$ submatrices F and T are nonnegative, 0 is the zeroth matrix. Furthermore, we assume that $F+T$ is irreducible. Matrix F is a result of differentiation and evaluation at the disease-free equilibrium of the new infections, and matrix T is the result of differentiation and evaluation at the disease-free equilibrium of the transition states (recovery, death). The submatrix F is known as the fertility matrix, and T as the transition matrix. We assume that the disease-free equilibrium is locally asymptotically stable, that is $\rho(C) < 1$, where $\rho(C)$ is the spectral radius of C. In addition, we require $\rho(T) < 1$. Since J is block-triangular, the stability of the disease-free equilibrium depends on the eigenvalues of $F+T$. The next-generation matrix is

$$Q = F(I-T)^{-1},$$

where I is the $k \times k$ identity matrix. The basic reproduction number is defined as the spectral radius of the matrix Q, that is,

$$\mathscr{R}_0 = \rho(F(I-T)^{-1}).$$

16.4.2 Examples

In this subsection, we introduce several more complex discrete epidemic models and use the next-generation approach to compute the reproduction number.

16.4.2.1 SEIS Model

As a first example, we illustrate the theory on example (16.35). For this model, the infected vector is $\mathbf{y} = (E, I)^T$, and the uninfected vector is $\mathbf{z} = (S)$. Arranging the system so that the first equations are for the infected variables, we have

$$\begin{aligned} E_{n+1} &= \gamma S_n \left(1 - G\left(\frac{\alpha I_n}{N_n}\right)\right) + \gamma\sigma E_n, \\ I_{n+1} &= \gamma(1-\sigma)E_n + \gamma\delta I_n, \\ S_{n+1} &= \Lambda + \gamma S_n G\left(\frac{\alpha I_n}{N_n}\right) + \gamma(1-\delta)I_n. \end{aligned} \tag{16.49}$$

The disease-free equilibrium is given by $(0,0,\frac{\Lambda}{1-\gamma})$. The Jacobian is given by

$$J = \begin{pmatrix} \gamma\sigma & -\gamma\alpha G'(0) & 0 \\ \gamma(1-\sigma) & \gamma\delta & 0 \\ 0 & \gamma\alpha G'(0) + \gamma(1-\delta) & \gamma \end{pmatrix}. \tag{16.50}$$

First, $C = (\gamma)$ and $\rho(C) = \gamma < 1$. The Jacobian is block-triangular. The important step is to identify the matrices F and T. The new infections term is associated with the function G. Hence the matrix F is given by

$$F = \begin{pmatrix} 0 & -\gamma\alpha G'(0) \\ 0 & 0 \end{pmatrix}. \tag{16.51}$$

We notice that the entries in F are nonnegative, since $G'(0) < 0$. The transition matrix T is given by

$$T = \begin{pmatrix} \gamma\sigma & 0 \\ \gamma(1-\sigma) & \gamma\delta \end{pmatrix}. \tag{16.52}$$

Using Mathematica, we can invert $I - T$ to obtain

$$(I-T)^{-1} = \begin{pmatrix} \frac{1}{1-\gamma\sigma} & 0 \\ \frac{\gamma(1-\sigma)}{(1-\gamma\delta)(1-\gamma\sigma)} & \frac{1}{1-\gamma\delta} \end{pmatrix}. \tag{16.53}$$

Hence,

$$F(I-T)^{-1} = \begin{pmatrix} \frac{-\gamma^2\alpha G'(0)(1-\sigma)}{(1-\gamma\delta)(1-\gamma\sigma)} & \frac{-\gamma\alpha G'(0)}{1-\gamma\delta} \\ 0 & 0 \end{pmatrix}. \tag{16.54}$$

The spectral radius of the above matrix gives the reproduction number

$$\mathscr{R}_0 = \frac{-\gamma^2\alpha G'(0)(1-\sigma)}{(1-\gamma\delta)(1-\gamma\sigma)}.$$

16.4.2.2 A Two-Patch SIS Model

In this subsection we introduce a two-patch SIS model based on the one-patch SIS model (16.18). We assume that the movement occurs after the infection and recovery

process. Individuals move from patch one to patch two with probability d_1 and vice versa with probability d_2. We furthermore assume that the probability of survival of individuals in both patches is the same. This assumption can be easily relaxed.

The SIS model with movement takes the form

$$\begin{aligned}
S^1_{n+1} &= (1-d_1)[\Lambda_1 + \gamma S^1_n G_1\left(\frac{\alpha_1 I^1_n}{N^1_n}\right) + \gamma(1-\sigma_1)I^1_n] \\
&\quad + d_2[\Lambda_2 + \gamma S^2_n G_2\left(\frac{\alpha_2 I^2_n}{N^2_n}\right) + \gamma(1-\sigma_2)I^2_n], \\
I^1_{n+1} &= (1-d_1)[\gamma S^1_n\left(1 - G_1\left(\frac{\alpha_1 I^1_n}{N^1_n}\right)\right) + \gamma\sigma_1 I^1_n] \\
&\quad + d_2[\gamma S^2_n\left(1 - G_2\left(\frac{\alpha_2 I^2_n}{N^2_n}\right)\right) + \gamma\sigma_2 I^2_n], \\
S^2_{n+1} &= +d_1[\Lambda_1 + \gamma S^1_n G_1\left(\frac{\alpha_1 I^1_n}{N^1_n}\right) + \gamma(1-\sigma_1)I^1_n] \\
&\quad + (1-d_2)[\Lambda_2 + \gamma S^2_n G_2\left(\frac{\alpha_2 I^2_n}{N^2_n}\right) + \gamma(1-\sigma_2)I^2_n], \\
I^2_{n+1} &= d_1[\gamma S^1_n\left(1 - G_1\left(\frac{\alpha_1 I^1_n}{N^1_n}\right)\right) + \gamma\sigma_1 I^1_n] \\
&\quad + (1-d_2)[\gamma S^2_n\left(1 - G_2\left(\frac{\alpha_2 I^2_n}{N^2_n}\right)\right) + \gamma\sigma_2 I^2_n].
\end{aligned} \tag{16.55}$$

We begin by determining the disease-free equilibrium. It is given by $\mathscr{E}_0 = (S^1, 0, S^2, 0)$, where S^1 and S^2 are solutions of the following system:

$$\begin{aligned}
S^1 &= (1-d_1)[\Lambda_1 + \gamma S^1] + d_2[\Lambda_2 + \gamma S^2], \\
S^2 &= d_1[\Lambda_1 + \gamma S^1] + (1-d_2)[\Lambda_2 + \gamma S^2].
\end{aligned} \tag{16.56}$$

First, we see that

$$N = S^1 + S^2 = \frac{\Lambda_1 + \Lambda_2}{1-\gamma}.$$

Solving system (16.56), we obtain

$$\begin{aligned}
S^1 &= \frac{(1-(1-d_2)\gamma)[(1-d_1)\Lambda_1 + d_2\Lambda_2] + d_2\gamma[d_1\Lambda_1 + (1-d_2)\Lambda_2]}{\Delta}, \\
S^2 &= \frac{d_1\gamma[(1-d_1)\Lambda_1 + d_2\Lambda_2] + (1-(1-d_1)\gamma)[d_1\Lambda_1 + (1-d_2)\Lambda_2]}{\Delta},
\end{aligned} \tag{16.57}$$

where $\Delta = (1-(1-d_1)\gamma)(1-(1-d_2)\gamma) - d_1 d_2 \gamma^2$. The matrix C is given by

$$C = \begin{pmatrix} (1-d_1)\gamma & d_2\gamma \\ d_1\gamma & (1-d_2)\gamma \end{pmatrix}. \tag{16.58}$$

It is not hard to show that $\rho(C) = \gamma < 1$. Next, we construct the matrix $F + T$:

$$F + T = \begin{pmatrix} (1-d_1)[-\gamma\alpha_1 G_1'(0) + \gamma\sigma_1] & d_2[-\gamma\alpha_2 G_2'(0) + \gamma\sigma_2] \\ d_1[-\gamma\alpha_1 G_1'(0) + \gamma\sigma_1] & (1-d_2)[-\gamma\alpha_2 G_2'(0) + \gamma\sigma_2] \end{pmatrix}. \quad (16.59)$$

The matrix F consists of all terms that involve G'; the matrix T consists of all remaining terms. Therefore,

$$F = \begin{pmatrix} -(1-d_1)\gamma\alpha_1 G_1'(0) & -d_2\gamma\alpha_2 G_2'(0) \\ -d_1\gamma\alpha_1 G_1'(0) & -(1-d_2)\gamma\alpha_2 G_2'(0) \end{pmatrix}, \quad (16.60)$$

and the matrix $I - T$ is given by

$$I - T = \begin{pmatrix} 1-(1-d_1)\gamma\sigma_1 & -d_2\gamma\sigma_2 \\ -d_1\gamma\sigma_1 & 1-(1-d_2)\gamma\sigma_2 \end{pmatrix}. \quad (16.61)$$

To invert $I - T$, we compute the determinant $\Delta = (1-(1-d_1)\gamma\sigma_1)(1-(1-d_2)\gamma\sigma_2) - d_1 d_2 \gamma^2 \sigma_1 \sigma_2$. Hence,

$$(I-T)^{-1} = \frac{1}{\Delta}\begin{pmatrix} 1-(1-d_2)\gamma\sigma_2 & d_2\gamma\sigma_2 \\ d_1\gamma\sigma_1 & 1-(1-d_1)\gamma\sigma_1 \end{pmatrix}. \quad (16.62)$$

The next-generation matrix takes the form

$$F(I-T)^{-1} = \frac{1}{\Delta}\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad (16.63)$$

where

$$\begin{aligned} A &= -(1-d_1)\gamma\alpha_1 G_1'(0)[1-(1-d_2)\gamma\sigma_2] - d_1 d_2 \gamma^2 \sigma_1 \alpha_2 G_2'(0), \\ B &= -(1-d_1) d_2 \gamma^2 \sigma_2 \alpha_1 G_1'(0) - d_2\gamma\alpha_2 G_2'(0)[1-(1-d_1)\gamma\sigma_1], \\ C &= -d_1\gamma\alpha_1 G_1'(0)[1-(1-d_2)\gamma\sigma_2] - d_1(1-d_2)\gamma^2\sigma_1\alpha_2 G_2'(0), \\ D &= -d_1 d_2 \gamma^2 \sigma_2 \alpha_1 G_1'(0) - (1-d_2)\gamma\alpha_2 G_2'(0)[1-(1-d_1)\gamma\sigma_1]. \end{aligned} \quad (16.64)$$

The reproduction number is given by

$$\mathscr{R}_0 = \rho(F(I-T)^{-1}) = \frac{A + D + \sqrt{(A-D)^2 + 4BC}}{2\Delta}.$$

We note that in this example, it would have been impossible to compute $\mathscr{R}_0$ with the Jacobian approach.

16.4.2.3 A Discrete SARS Model

In this section, we consider a discrete SARS model with quarantine and isolation. Let S_n denote the susceptibles, E_n the exposed, I_n the individuals showing symptoms, Q_n the quarantined, J_n the isolated, and R_n the recovered individuals. In SARS,

the exposed individuals are infectious with reduced infectivity. The coefficient of reduction is q. The model takes the form

$$\begin{aligned}
S_{n+1} &= \Lambda + \gamma\alpha_1 S_n G\left(\frac{I_n + qE_n}{N_n - Q_n - J_n}\right) + \gamma\rho(1-\alpha_1)S_n + \gamma\alpha_4(1-\eta_1)Q_n,\\
E_{n+1} &= \gamma\alpha_1 S_n\left(1 - G\left(\frac{I_n + qE_n}{N_n - Q_n - J_n}\right)\right) + \gamma E_n(\alpha_2\sigma + (1-\alpha_2)\rho),\\
I_{n+1} &= \alpha_2\gamma(1-\sigma)E_n + \gamma I_n(\alpha_3\sigma + (1-\alpha_3)r_2),\\
Q_{n+1} &= \gamma(1-\rho)((1-\alpha_1)S_n + (1-\alpha_2)E_n) + \gamma Q_n(\alpha_4\eta_1 + (1-\alpha_4)\eta_2),\\
J_{n+1} &= \alpha_3\gamma(1-\sigma)I_n + (1-\alpha_4)\gamma(1-\eta_2)Q_n + \gamma r_1 J_n,\\
R_{n+1} &= \gamma(1-r_1)J_n + \gamma(1-\alpha_3)(1-r_2)I_n + \gamma R_n,
\end{aligned} \tag{16.65}$$

where the parameters are given in Table 16.3.

Table 16.3 Parameter meanings

Parameter	Meaning	Parameter	Meaning
Λ	Recruitment	γ	Probability of survival
q	Reduction in infectivity for exposed	$1-\rho$	Probability of quarantine
$1-\sigma$	Probability of isolation	α_i	Convex combination coefficients
$1-r_1$	Probability of recovery of isolated	$1-r_2$	Probability of recovery of infected
$1-\eta_1$	Probability of ending quarantine to susceptible class	$1-\eta_2$	Probability of ending the quarantine to isolated class

We apply the next-generation approach to compute the reproduction number. The disease-free equilibrium is given by $\mathscr{E}_0 = (S^*, 0, 0, 0, 0, 0)$, where

$$S^* = \frac{\Lambda}{1-\gamma}.$$

The vector of infected classes is (E, I, Q, J). Hence, the matrix $F+T$ is given by

$$F+T = \begin{pmatrix}
-\gamma\alpha_1 qG'(0) + \alpha_2\gamma\sigma + (1-\alpha_2)\gamma\rho & -\gamma\alpha_1 G'(0) & 0 & 0\\
\alpha_2\gamma(1-\sigma) & \gamma(\alpha_3\sigma + (1-\alpha_3)r_2) & 0 & 0\\
\gamma(1-\rho)(1-\alpha_2) & 0 & \gamma(\alpha_4\eta_1 + (1-\alpha_4)\eta_2) & 0\\
0 & \alpha_3\gamma(1-\sigma) & \gamma(1-\alpha_4)(1-\eta_2) & \gamma r_1
\end{pmatrix}. \tag{16.66}$$

The matrix F is written as $F = (f_{ij})$, where $f_{11} = -\gamma q\alpha_1 G'(0)$ and $f_{12} = -\gamma\alpha_1 G'(0)$, while the remaining entries are zero. The matrix $I-T$ is given by

$$I-T=\begin{pmatrix} 1-\alpha_2\gamma\sigma-(1-\alpha_2)\gamma\rho & 0 & 0 & 0 \\ -\alpha_2\gamma(1-\sigma) & 1-\gamma(\alpha_3\sigma+(1-\alpha_3)r_2) & 0 & 0 \\ -\gamma(1-\rho)(1-\alpha_2) & 0 & 1-\gamma(\alpha_4\eta_1+(1-\alpha_4)\eta_2) & 0 \\ 0 & -\alpha_3\gamma(1-\sigma) & -\gamma(1-\alpha_4)(1-\eta_2) & 1-\gamma r_1 \end{pmatrix}. \tag{16.67}$$

Because of the structure of F, only the first 2×2 block of $(I-T)^{-1}$ is important for the reproduction number. Because of the block-triangular form of $I-T$, that first 2×2 block of $(I-T)^{-1}$ is obtained from inverting the first 2×2 block of $(I-T)$. Thus we have

$$(I-T)^{-1}=\begin{pmatrix} \dfrac{1}{1-\alpha_2\gamma\sigma-(1-\alpha_2)\gamma\rho} & 0 & 0\,0 \\ \dfrac{\alpha_2\gamma(1-\sigma)}{\Delta} & \dfrac{1}{1-\gamma(\alpha_3\sigma+(1-\alpha_3)r_2)} & 0\,0 \\ * & * & *\,* \\ * & * & *\,* \end{pmatrix}, \tag{16.68}$$

where $\Delta=(1-\alpha_2\gamma\sigma-(1-\alpha_2)\gamma\rho)(1-\gamma(\alpha_3\sigma+(1-\alpha_3)r_2))$. The matrix $F(I-T)^{-1}$ has a very simple form, whose principal eigenvalue is not hard to determine. Hence, the reproduction number is given by

$$\mathscr{R}_0=\rho(F(I-T)^{-1})=\frac{-\gamma\alpha_1 qG'(0)}{1-\alpha_2\gamma\sigma-(1-\alpha_2)\gamma\rho}+\frac{-\alpha_1\gamma G'(0)\alpha_2\gamma(1-\sigma)}{\Delta}.$$

The first term of the reproduction number gives the number of secondary infections produced by an exposed individual; the second term gives the number of secondary infections produced by an infectious individual.

Problems

16.1. Ricker Model
Consider the Ricker model (16.9).

(a) Find the equilibria of the Ricker model.
(b) Determine the stability of the equilibria of the Ricker model.
(c) Does the Ricker model have 2-cycles?
(d) Does the Ricker model exhibit chaos?

16.2. Beverton–Holt Model
Consider the Beverton-Holt model (16.10).

(a) Find the equilibria of the Beverton–Holt model.
(b) Determine the stability of the equilibria of the Beverton–Holt model.
(c) Does the Beverton–Holt model have 2-cycles?
(d) Does the Beverton–Holt model exhibit chaos?

16.3. Hassell Model
Consider the Hassell model (16.11).

(a) Find the equilibria of the Hassell model.
(b) Determine the stability of the equilibria of the Hassell model.
(c) Does the Hassell model have 2-cycles?
(d) Does the Hassell model exhibit chaos?

16.4. SEIS Epidemic Model
Consider the discrete SEIS model (16.35).

(a) Derive the reproduction number $\mathscr{R}_0$.
(b) Use the Jury conditions to show that if $\mathscr{R}_0 < 1$, then the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, then the disease-free equilibrium is unstable.
(c) Show that if $\mathscr{R}_0 > 1$, there is a unique endemic equilibrium.

16.5. SI Epidemic Model
Consider the following SI epidemic model:

$$\begin{aligned} S_{n+1} &= \Lambda + \gamma S_n G\left(\frac{\alpha I_n}{N_n}\right), \\ I_{n+1} &= \gamma S_n \left(1 - G\left(\frac{\alpha I_n}{N_n}\right)\right) + \gamma\sigma I_n, \end{aligned} \tag{16.69}$$

where G has the same properties as in the text and $\sigma < 1$.

(a) Derive the reproduction number $\mathscr{R}_0$.
(b) Use the Jury conditions to show that if $\mathscr{R}_0 < 1$, then the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, then the disease-free equilibrium is unstable.
(c) Show that if $\mathscr{R}_0 > 1$, there is a unique endemic equilibrium.
(d) Consider the stability of the endemic equilibrium. When is it stable?

16.6. SIRS Epidemic Model
Consider the following SIRS epidemic model:

$$\begin{aligned} S_{n+1} &= \Lambda + \gamma S_n G\left(\frac{\alpha I_n}{N_n}\right) + \gamma(1-\delta)R_n, \\ I_{n+1} &= \gamma S_n \left(1 - G\left(\frac{\alpha I_n}{N_n}\right)\right) + \gamma\sigma I_n, \\ R_{n+1} &= \gamma(1-\sigma)I_n + \gamma\delta R_n, \end{aligned} \tag{16.70}$$

where G has the same properties as in the text.

(a) Derive the reproduction number $\mathscr{R}_0$.
(b) Use the Jury conditions to show that if $\mathscr{R}_0 < 1$, then the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, then the disease-free equilibrium is unstable.
(c) Show that if $\mathscr{R}_0 > 1$, then there is a unique endemic equilibrium.
(d) Consider the stability of the endemic equilibrium. When is it stable?

16.7. SIS Epidemic Model with Environmental Transmission
Consider the following SIS epidemic model with environmental transmission:

$$\begin{aligned} S_{n+1} &= \Lambda + \gamma S_n e^{-\frac{\alpha I_n}{N_n} - \beta P_n} + \gamma(1-\sigma) I_n, \\ I_{n+1} &= \gamma S_n \left(1 - e^{-\frac{\alpha I_n}{N_n} - \beta P_n}\right) + \gamma\sigma I_n, \\ P_{n+1} &= \rho I_n + \delta P_n, \end{aligned} \tag{16.71}$$

where P_n is the amount of the pathogen in the environment.

(a) Derive the reproduction number $\mathscr{R}_0$.
(b) Use the Jury conditions to show that if $\mathscr{R}_0 < 1$, then the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, then the disease-free equilibrium is unstable.
(c) Show that if $\mathscr{R}_0 > 1$, then there is a unique endemic equilibrium.
(d) Consider the stability of the endemic equilibrium. When is it stable?

16.8. SIRS Epidemic Model with Vaccination
Consider the following SIRS epidemic model:

$$\begin{aligned} S_{n+1} &= \Lambda + \rho\gamma S_n G\left(\frac{\alpha I_n}{N_n}\right) + \gamma(1-\delta) R_n + (1-\rho)\gamma\psi S_n, \\ I_{n+1} &= \rho\gamma S_n \left(1 - G\left(\frac{\alpha I_n}{N_n}\right)\right) + \gamma\sigma I_n, \\ R_{n+1} &= \gamma(1-\sigma) I_n + (1-\rho)\gamma(1-\psi) S_n + \gamma\delta R_n, \end{aligned} \tag{16.72}$$

where G has the same properties as in the text.

(a) Derive the reproduction number $\mathscr{R}_0$.
(b) Use the Jury conditions to show that if $\mathscr{R}_0 < 1$, then the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, then the disease-free equilibrium is unstable.
(c) Show that if $\mathscr{R}_0 > 1$, there is a unique endemic equilibrium.
(d) Consider the stability of the endemic equilibrium. When is it stable?

16.9. SIS Epidemic Model with Two Strains
Consider the following SIS epidemic model with two strains:

$$\begin{aligned} S_{n+1} &= \Lambda + \gamma S_n e^{-(\alpha_1 I_n + \alpha_2 J_n)} + \gamma(1-\sigma_1) I_n + \gamma(1-\sigma_2) J_n, \\ I_{n+1} &= \gamma \frac{\alpha_1 S_n I_n}{\alpha_1 I_n + \alpha_2 J_n} \left(1 - e^{-(\alpha_1 I_n + \alpha_2 J_n)}\right) + \gamma \sigma_1 I_n, \\ J_{n+1} &= \gamma \frac{\alpha_2 S_n J_n}{\alpha_1 I_n + \alpha_2 J_n} \left(1 - e^{-(\alpha_1 I_n + \alpha_2 J_n)}\right) + \gamma \sigma_2 J_n, \end{aligned} \tag{16.73}$$

where I_n denotes infection with strain one, and J_n denotes infection with strain two.

(a) Derive the reproduction numbers of strain one and strain two $\mathscr{R}_1$ and $\mathscr{R}_2$. Set $\mathscr{R}_0 = \max\{\mathscr{R}_1, \mathscr{R}_2\}$.
(b) Use the Jury conditions to show that if $\mathscr{R}_0 < 1$, then the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, then the disease-free equilibrium is unstable.
(c) Show that if $\mathscr{R}_1 > 1$, there is a unique endemic equilibrium corresponding to strain one. Show that if $\mathscr{R}_2 > 1$, there is a unique endemic equilibrium corresponding to strain two.
(d) Consider the stability of the endemic equilibrium corresponding to strain one. When is it stable?

16.10. SIS Epidemic Model with Two Strains and Mutation
Consider the following SIS epidemic model with two strains:

$$\begin{aligned} S_{n+1} &= \Lambda + \gamma S_n e^{-(\alpha_1 I_n + \alpha_2 J_n)} + \rho\gamma(1-\sigma_1) I_n + \gamma(1-\sigma_2) J_n, \\ I_{n+1} &= \gamma \frac{\alpha_1 S_n I_n}{\alpha_1 I_n + \alpha_2 J_n} \left(1 - e^{-(\alpha_1 I_n + \alpha_2 J_n)}\right) + \rho\gamma\sigma_1 I_n + (1-\rho)\mu\gamma I_n, \\ J_{n+1} &= \gamma \frac{\alpha_2 S_n J_n}{\alpha_1 I_n + \alpha_2 J_n} \left(1 - e^{-(\alpha_1 I_n + \alpha_2 J_n)}\right) + \gamma \sigma_2 J_n + (1-\rho)(1-\mu)\gamma I_n, \end{aligned} \tag{16.74}$$

where I_n denotes infection with strain one, and J_n denotes infection with strain two.

(a) Derive the reproduction numbers of strain one and strain two $\mathscr{R}_1$ and $\mathscr{R}_2$. Set $\mathscr{R}_0 = \max\{\mathscr{R}_1, \mathscr{R}_2\}$.
(b) Use the Jury conditions to show that if $\mathscr{R}_0 < 1$, then the disease-free equilibrium is locally asymptotically stable. If $\mathscr{R}_0 > 1$, then the disease-free equilibrium is unstable.
(c) Show that if $\mathscr{R}_1 > 1$, there is a unique endemic equilibrium corresponding to strain one. Show that there is a unique coexistence equilibrium.
(d) Consider the stability of the endemic equilibrium corresponding to strain one. When is it stable?

References

1. A. N. KOLMOGOROV, I. G. PETROVSKII, AND N. S. PISKUNOV, *A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem*, in selected works of A. N. Kolmogorov, V. M. Tikhomiror, ed., Kluwer Academic Publishers, 1991, pp. 242–270.
2. B. ADAMS, E. HOLMES, C.ZHANG, M. MAMMEN, S. NIMMANNITYA, S. KALAYANAROOJ, AND M. BOOTS, *Cross-protective immunity can account for the alternating epidemic pattern of dengue virus serotypes circulating in Bangkok*, PNAS, 103 (2006), pp. 14234–14239.
3. A.J. LOTKA, *The stability of the normal age distributions*, Proc. Natl. Acad. Sci. USA, 8 (1922), pp. 339–345.
4. H. AKAIKE, *Information theory and an extension of the maximum likelihood principle*, in Second International Symposium on Information Theory (Tsahkadsor, 1971), Akadémiai Kiadó, Budapest, 1973, pp. 267–281.
5. S. ALCHON, *A Pest in the land: New world epidemic in a global perspective*, University of New Mexico Press, 2003.
6. M. E. ALEXANDER AND S. M. MOGHADAS, *Bifurcation analysis of an SIRS epidemic model with generalized incidence*, SIAM J. Appl. Math., 65 (2001), pp. 1794–1916.
7. L. J. S. ALLEN, *An introduction to stochastic epidemic models*, in Mathematical epidemiology, vol. 1945 of Lecture Notes in Math., Springer, Berlin, 2008, pp. 81–130.
8. L. J. S. ALLEN AND R. K. ERNEST, *The impact of long-range dispersal on the rate of spread in population and epidemic models*, in Mathematical approaches for emerging and reemerging infectious diseases: an introduction (Minneapolis, MN, 1999), vol. 125 of IMA Vol. Math. Appl., Springer, New York, 2002, pp. 183–197.
9. L. J. S. ALLEN AND P. VAN DEN DRIESSCHE, *The basic reproduction number in some discrete-time epidemic models*, J. Difference Equ. Appl., 14 (2008), pp. 1127–1147.
10. R. ANDERSON AND R. MAY, *Directly transmitted infections diseases: control by vaccination*, Science, 215 (1982), pp. 1053–1050.
11. R. M. ANDERSON AND R. M. MAY, *Population biology of infectious disease: Part I*, Nature, 280 (1979), pp. 361–367.
12. O. ANGULO, J. C. LÓPEZ-MARCOS, AND F. A. MILNER, *The application of an age-structured model with unbounded mortality to demography*, Math. Biosci., 208 (2007), pp. 495–520.
13. S. ANIŢA, V. ARNĂUTU, AND V. CAPASSO, *An introduction to optimal control problems in life sciences and economics*, Modeling and Simulation in Science, Engineering and Technology, Birkhäuser/Springer, New York, 2011. From mathematical models to numerical simulation with MATLAB.

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3

14. T. ARBOGAST AND F. A. MILNER, *A finite difference method for a two-sex model of population dynamics*, SIAM J. Numer. Anal., 26 (1989), pp. 1474–1486.
15. E. ARIAS, *United States life tables, 2007*, National Vital Statistics Reports, 59 (2011).
16. J. ARINO AND P. VAN DEN DRIESSCHE, *A multi-city epidemic model*, Math. Popul. Stud., 10 (2003), pp. 175–193.
17. ———, *Metapopulations epidemic models. A survey*, Math. Popul. Stud., 48 (2006), pp. 1–12.
18. H. B., *Immunoepidemiology—bridging the gap between immunology and epidemiology*, Trends Parasitol., 17 (2001), pp. 102–106.
19. N. BACAËR AND R. OUIFKI, *Growth rate and basic reproduction number for population models with a simple periodic factor*, Math. Biosci., 210 (2007), pp. 647–658.
20. C. T. BAUCH AND A. P. GALVANI, *Using network models to approximate spatial point-process models*, Math. Biosci., 184 (2003), pp. 101–114.
21. C. T. BAUCH, A. P. GALVANI, AND D. J. D. EARN, *Group interest versus self-interest in smallpox vaccination policy*, Proc. Natl. Acad. Sci. USA, 100 (2003), pp. 10564–10567 (electronic).
22. D. BERNOULLI, *An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it, reprint*, Rev. Med. Virol., 14 (2004), pp. 275–288.
23. R. J. H. BEVERTON AND S. J. HOLT, *On the Dynamics of Exploited Fish Populations*, vol. XIX of Fishery Investigations Series II, Ministry of Agriculture, Fisheries and Food, 1957.
24. S. BHATT, P. W. GETHING, O. J. BRADY, J. P. MESSINA, A. W. FARLOW, C. L. MOYES, J. M. DRAKE, J. S. BROWNSTEIN, A. G. HOEN, O. SANKOH, M. F. MYERS, D. B. GEORGE, T. JAENISCH, G. R. W. WINT, C. P. SIMMONS, T. W. SCOTT, J. J. FARRAR, AND S. I. HAY, *The global distribution and burden of dengue*, Nature, 496 (2013), pp. 504–507.
25. S. BONHOEFFER AND M. NOWAK, *Mutation and the evolution of parasite virulence*, Proc. R. Soc. London, B, 258 (1994), pp. 133–140.
26. F. BRAUER, *Compartmental models in epidemiology*, in Mathematical epidemiology, vol. 1945 of Lecture Notes in Math., Springer, Berlin, 2008, pp. 19–80.
27. F. BRAUER AND C. CASTILLO-CHAVEZ, *Mathematical models in population biology and epidemiology*, vol. 40 of Texts in Applied Mathematics, Springer, New York, second ed., 2012.
28. F. BRAUER AND J. A. NOEL, *The Qualitative Theory of Ordinary Differential Equations: An Introduction*, Dover Publications, Inc., New York, 1989.
29. H.-J. BREMERMANN AND H. R. THIEME, *A competitive exclusion principle for pathogen virulence*, J. Math. Biol., 27 (1989), pp. 179–190.
30. C. BUCK, A. LLOPIS, E. NÁJERA, AND M. TERRIS, *The Challenge of Epidemiology: Issues and Selected Readings*, vol. 505 of Scientific Publication, Pan American Health Organization, 1998.
31. K. P. BURNHAM AND D. R. ANDERSON, *Model selection and multimodel inference: A practical information-theoretic approach*, Springer, New York, second ed., 2002.
32. K. P. BURNHAM AND D. R. ANDERSON, *Multimodel inference: understanding AIC and BIC in model selection*, Sociol. Methods Res., 33 (2004), pp. 261–304.
33. R. S. CANTRELL AND C. COSNER, *Spatial ecology via reaction–diffusion equations*, Wiley Series in Mathematical and Computational Biology, John Wiley & Sons Ltd., Chichester, 2003.
34. V. CAPASSO AND K. KUNISCH, *A reaction–diffusion system modelling man–environment epidemics*, Ann. Differential Equations, 1 (1985), pp. 1–12.
35. F. CARRAT, J. LUONG, H. LAO, A.-V. SALLÉ, C. LAJAUNIE, AND H. WACKERNAGEL, *A "small-world-like" model for comparing interventions aimed at preventing and controlling influenza pandemics*, BMC Med, 4 (2006).
36. C. CASTILLO-CHAVEZ, Z. FENG, AND W. HUANG, *On the computation of $\mathscr{R}_0$ and its role on global stability*, in Mathematical approaches for emerging and reemerging infectious diseases: an introduction (Minneapolis, MN, 1999), vol. 125 of IMA Vol. Math. Appl., Springer, New York, 2002, pp. 229–250.

37. C. Castillo-Chavez, H. W. Hethcote, V. Andreasen, S. A. Levin, and W. M. Liu, *Epidemiological models with age structure, proportionate mixing, and cross-immunity*, J. Math. Biol., 27 (1989), pp. 233–258.
38. C. Castillo-Chavez and B. Song, *Dynamical models of tuberculosis and their applications*, Math. Biosci. Eng., 1 (2004), pp. 361–404.
39. C. Castillo-Chavez and A.-A. Yakubu, *Discrete-time S-I-S models with complex dynamics*, in Proceedings of the Third World Congress of Nonlinear Analysts, Part 7 (Catania, 2000), vol. 47, 2001, pp. 4753–4762.
40. CDC, *Key facts about avian influenza (bird flu) and highly pathogenic avian influenza A (H5N1) virus*.
41. C. Chicone, *Ordinary differential equations with applications*, vol. 34 of Texts in Applied Mathematics, Springer, New York, second ed., 2006.
42. M. S. Cohen, *Preventing sexual transmission of HIV*, Clin. Infect. Dis., 45 (2007), pp. S287–S292.
43. V. Colizza and A. Vespignani, *Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: theory and simulations*, J. Theoret. Biol., 251 (2008), pp. 450–467.
44. C. Cosner, J. C. Beier, R. S. Cantrell, D. Impoinvil, L. Kapitanski, M. D. Potts, A. Troyo, and S. Ruan, *The effects of human movement on the persistence of vector-borne diseases*, J. Theoret. Biol., 258 (2009), pp. 550–560.
45. F. A. B. Coutinho, M. N. Burattini, L. F. Lopez, and E. Massad, *Threshold conditions for a non-autonomous epidemic system describing the population dynamics of dengue*, Bull. Math. Biol., 68 (2006), pp. 2263–2282.
46. J. M. Cushing, *An introduction to structured population dynamics*, vol. 71 of CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998.
47. J. M. Cushing, B. Dennis, R. F. Costantino, R. Desharnais, and S. M. Henson, *Chaos in ecology: experimental nonlinear dynamics*, Academic Press, New York, 2003.
48. H. Dahari, A. Lo, R. M. Ribeiro, and A. S. Perelson, *Modeling hepatitis C virus dynamics: liver regeneration and critical drug efficacy*, J. Theoret. Biol., 247 (2007), pp. 371–381.
49. P. De Leenheer and H. L. Smith, *Virus dynamics: a global analysis*, SIAM J. Appl. Math., 63 (2003), pp. 1313–1327.
50. C. V. de León, *Constructions of Lyapunov functions for classics SIS, SIR and SIRS epidemic model with variable population size*, Foro-Red-Mat: Revista electrónica de contenido matemático, 26 (2009).
51. C. V. de León and J. A. C. Hernández, *Local and global stability of host–vector disease models*, Foro-Red-Mat: Revista electrónica de contenido matemático, 25 (2008).
52. T. Dhirasakdanon, H. R. Thieme, and P. Van Den Driessche, *A sharp threshold for disease persistence in host metapopulations*, J. Biol. Dyn., 1 (2007), pp. 363–378.
53. R. P. Dickinson and R. J. Gelinas, *Sensitivity analysis of ordinary differential equation systems—a direct method*, J. Computational Phys., 21 (1976), pp. 123–143.
54. O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, *On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations*, J. Math. Biol., 28 (1990), pp. 365–382.
55. K. Diez and J. Heesterbeek, *Bernoulli's epidemiological model revisited*, Math. Biosci., 180 (2002), pp. 1–21.
56. K. E. and C. Williams, *Infectious disease epidemiology: theory and practice, 3rd ed.*, Jones and Bartlett Learning, LLC, Burlington, MA, 2014.
57. M. El Kahoui and A. Weber, *Deciding Hopf bifurcations by quantifier elimination in a software-component architecture*, J. Symbolic Comput., 30 (2000), pp. 161–179.
58. FAO, *Faostat*.
59. G. Fan, J. Liu, P. van den Driessche, J. Wu, and H. Zhu, *The Impact of maturation delay of mosquitoes on the transmission of West Nile virus*, Math. Biosci., 228 (2010), pp.119–126

60. Z. FENG, C. CASTILLO-CHAVEZ, AND A. CAPURRO, *A model for tuberculosis with exogenous reinfection*, Theor. Pop, Biol., 57 (2000), pp. 235–247.
61. Z. FENG AND H. R. THIEME, *Recurrent outbreaks of childhood diseases revisited: the impact of isolation*, Math. Biosci., 128 (1995), pp. 93–130.
62. R. A. FISHER, *The genetical theory of natural selection*, Oxford,UK, 1930.
63. H. V. FOERSTER, *Some remarks on changing populations*, in The Kinetics of Cellular Proliferation, Grune and Straton, 1959, pp. 382–407.
64. G. GAUSE, *The struggle for existence*, Williams & Wilkins, 1934.
65. M. A. GILCHRIST AND A. SASAKI, *Modeling host–parasite coevolution: a nested approach based on mechanistic models*, J. Theoret. Biol., 218 (2002), pp. 289–308.
66. A. B. GUMEL, *Causes of backward bifurcations in some epidemiological models*, J. Math. Anal. Appl., 395 (2012), pp. 355–365.
67. K. P. HADELER AND P. VAN DEN DRIESSCHE, *Backward bifurcation in epidemic control*, Math. Biosci., 146 (1997), pp. 15–35.
68. J. W. HAEFNER, *Modeling biological systems: principles and applications*, Springer, New York, second ed., 2005. With 1 CD-ROM (Windows and UNIX).
69. S. M. HAENSCH S, BIANUCCI R AND ET AL., *Distinct clones of* Yersinia pestis *caused the black death*, PLoS Pathog., 6(10) (2010), e1001134.
70. B. HAMILTON, J. MARTIN, AND S. J. VENTURA, *Births: Preliminary data for 2010*, National Vital Statistics Reports, 60 (2011).
71. A. HARDY AND E. MAGNELLO, *Statistical methods in epidemiology: Karl Pearson, Ronald Ross, Major Greenwood and Austin Bradford Hill, 1900–1945*, Soz.- Präventivmed., 47 (2002), pp. 80–89.
72. M. P. HASSELL, *Density-dependence is single species populations*, The Journal of Animal Ecology, 44 (1975), pp. 283–295.
73. H. HETHCOTE, M. ZHIEN, AND L. SHENGBING, *Effects of quarantine in six endemic models for infectious diseases*, Math. Biosci., 180 (2002), pp. 141–160. John A. Jacquez memorial volume.
74. H. W. HETHCOTE, *The mathematics of infectious diseases*, SIAM Review, 42, pp. 599 – 653.
75. ———, *A thousand and one epidemic models*, in In Frontiers in Mathematical Biology, S. Levin, ed., vol. 100 of Lecture Notes in Biomathematics, Springer, Berlin, 1994, pp. 504–515.
76. R. D. HOLT AND J. PICKERING, *Infectious diseases and species coexistence: a model of Lotka–Volterra form*, Am. Nat., 126 (1985), pp. 196–211.
77. C. M. HURVICH AND C.-L. TSAI, *Regression and time series model selection in small samples*, Biometrika, 76 (1989), pp. 297–307.
78. M. IANNELLI, *Mathematical theory of age-structured population dynamics*, Giardini, Pisa, 1995.
79. V. ISHAM, *Stochastic models of epidemics with special reference to AIDS*, The Annals of Applied Probability, 3 (1993), pp. 1–27.
80. S. IWAMI, Y. TAKEUCHI, AND X. LIU, *Avian–human influenza epidemic model*, Math. Biosci., 207 (2007), pp. 1–25.
81. ———, *Avian flu pandemic: can we prevent it?* J. Theoret. Biol., 257 (2009), pp. 181–190.
82. J. D. MURRAY, E. A. STANLEY, AND D. L. BROWN, *On the spatial spread of rabies among foxes*, Proc. Royal Soc. London, B, 229 (1986), pp. 111–150.
83. W. JONES, ed., *Hippocrates collected works I*, Cambridge Harvard University Press, Cambridge, 1868.
84. W. KERMACK AND A. MCKENDRICK, *A contribution to mathematical theory of epidemics*, Proc. Roy. Soc. Lond. A, 115 (1927), pp. 700–721.
85. W. KERMACK AND A. MCKENDRICK, *Contributions to the mathematical theory of epidemics–I*, Bulletin of Mathematical Biology, 53 (1991), pp. 33–55.
86. ———, *Contributions to the mathematical theory of epidemics–II. The problem of endemicity*, Bulletin of Mathematical Biology, 53 (1991).
87. ———, *Contributions to the mathematical theory of epidemics–III. Further studies of the problem of endemicity*, Bulletin of Mathematical Biology, 53 (1991), pp. 89–118.

88. A. KOROBEINIKOV, *Lyapunov functions and global properties for SEIR and SEIS epidemic models*, Math Med Biol., 21 (2004), pp. 75–83.
89. A. KOROBEINIKOV AND P. K. MAINI, *A Lyapunov function and global properties for SIR and SEIR epidemiological models with nonlinear incidence*, Math. Biosci. Eng., 1 (2004), pp. 57–60.
90. M. KOT, *Elements of mathematical ecology*, Cambridge University Press, Cambridge, 2001.
91. N. N. KRASOVSKIĬ, *Nekotorye zadachi teorii ustoichivosti dvizheniya*, Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1959.
92. A. LANGE AND N. M. FERGUSON, *Antigenic diversity, transmission mechanisms, and the evolution of pathogens*, PLoS Comput. Biol., 5 (2009), e1000536.
93. J. P. LASALLE, *Some extensions of Liapunov's second method*, IRE Trans., CT-7 (1960), pp. 520–527.
94. S. LENHART AND J. T. WORKMAN, *Optimal control applied to biological models*, Chapman & Hall/CRC Mathematical and Computational Biology Series, Chapman & Hall/CRC, Boca Raton, FL, 2007.
95. R. LEVINS, *Some demographic and genetic consequences of environmental heterogeneity for biological control*, Bulletin of the Entomological Society of America, (1969), p. 237–240.
96. M. Y. LI AND J. S. MULDOWNEY, *Global stability for the SEIR model in epidemiology*, Math. Biosci., 125 (1995), pp. 155–164.
97. X.-Z. LI, X.-C. DUAN, M. GHOSH, AND X.-Y. RUAN, *Pathogen coexistence induced by saturating contact rates*, Nonlinear Anal. Real World Appl., 10 (2009), pp. 3298–3311.
98. A. M. LIAPUNOV, *Stability of motion*, With a contribution by V. A. Pliss and an introduction by V. P. Basov. Translated from the Russian by Flavian Abramovici and Michael Shimshoni. Mathematics in Science and Engineering, Vol. 30, Academic Press, New York, 1966.
99. T. LIETMAN, T. PORCO, AND S. BLOWER, *Leprosy and tuberculosis: the epidemiological consequences of cross-immunity*, Am. J. Public Health, 87 (1997), pp. 1923–1927.
100. M. LIPSITCH AND M. B. MURRAY, *Multiple equilibria: tuberculosis transmission require unrealistic assumptions*, Theor. Popul. Biol., 63 (2003), pp. 169–170.
101. Q.-X. LIU, Z. JIN, AND M.-X. LIU, *Spatial organization and evolution period of the epidemic model using cellular automata*, Phys. Rev. E (3), 74 (2006), 031110.
102. A. L. LLOYD AND V. A. A. JANSEN, *Spatiotemporal dynamics of epidemics: synchrony in metapopulation models*, Math. Biosci., 188 (2004), pp. 1–16. Topics in biomathematics and related computational problems.
103. J. MA AND Z. MA, *Epidemic threshold conditions for seasonally forced SEIR models*, Math. Biosci. Eng., 3 (2006), pp. 161–172.
104. T. MALTHUS, *An Essay on the Principle of Population*, Printed for J. Johnson, in St. Paul's Church Yard, London, 1798.
105. G. I. MARCHUK, *Mathematical modelling of immune response in infectious diseases*, vol. 395 of Mathematics and Its Applications, Kluwer Academic Publishers Group, Dordrecht, 1997. Translated from the Russian manuscript by Guennadi Kontarev and Igor Sidorov.
106. M. MARTCHEVA, *India's approach to eliminating plasmodium falciparum malaria: A modeling perspective*, J. Biol. Systems, 18 (2010), pp. 867–891.
107. M. MARTCHEVA, *An evolutionary model of influenza A with drift and shift*, J. Biol. Dynamics, 6 (2012), pp. 299–332.
108. ———, *Avian flu: modeling and implications for control*, J. Biol. Systems, 22 (2014), pp. 151–175.
109. M. MARTCHEVA, B. M. BOLKER, AND R. D. HOLT, *Vaccine-induced pathogen strain replacement: what are the mechanisms?*, J. R. Soc. Interface, 5 (2008), pp. 3–13.
110. M. MARTCHEVA AND X.-Z. LI, *Linking immunological and epidemiological dynamics of HIV: The case of super-infection*, J. Biol. Dyn., 7 (2013), pp. 161–182.
111. M. MARTCHEVA AND O. PROSPER, *Unstable dynamics of vector-borne diseases: Modeling through differential-delay equations*, in Dynamic Models of Infectious Diseases: Vol 1. Vector Borne Diseases, S. H. R. Vadrevu and R. Durvasula, eds., Springer, 2012, pp. 43–75.

112. M. MARTCHEVA AND H. R. THIEME, *Progression age enhanced backward bifurcation in an epidemic model with super-infection*, J. Math. Biol., 46 (2003), pp. 385–424.
113. R. MAY AND M. NOWAK, *Coinfection and the evolution of parasite virulence*, Proc. R. Soc. London, B, 261 (1995), pp. 209–215.
114. R. M. MAY AND R. M. ANDERSON, *Population biology of infectious diseases: Part II*, Nature, 280 (1979), pp. 455–461.
115. A. G. MCKENDRICK, *Applications of mathematics to medical problems*, Proc. Edinburgh Math. Soc., 44 (1926), p. 98–130.
116. J. MEDLOCK AND M. KOT, *Spreading disease: integro-differential equations old and new*, Math. Biosci., 184 (2003), pp. 201–222.
117. R. MILLER NEILAN AND S. LENHART, *An introduction to optimal control with an application in disease modeling*, in Modeling paradigms and analysis of disease transmission models, vol. 75 of DIMACS Ser. Discrete Math. Theoret. Comput. Sci., Amer. Math. Soc., Providence, RI, 2010, pp. 67–81.
118. F. A. MILNER AND G. RABBIOLO, *Rapidly converging numerical algorithms for models of population dynamics*, J. Math. Biol., 30 (1992), pp. 733–753.
119. H. MOTULSKY AND A. CHRISTOPOULOS, *Fitting models to biological data using linear and nonlinear regression. A practical guide to curve fitting*, GraphPad Software Inc., San Diego, 2003.
120. A. MUBAYI, C. KRIBS ZALETA, M. MARTCHEVA, AND C. CASTILLO-CHÁVEZ, *A cost-based comparison of quarantine strategies for new emerging diseases*, Math. Biosci. Eng., 7 (2010), pp. 687–717.
121. S. MURPHY, J. XU, AND K. D. KOCHANEK, *Deaths: Preliminary data for 2010*, National Vital Statistics Reports, 60 (2012).
122. U. NATIONS, *The Millemnnium Development Goals Report 2010*, United Nations, New York, 2010.
123. M. E. J. NEWMAN, *Spread of epidemic disease on networks*, Phys. Rev. E (3), 66 (2002), 016128.
124. ———, *Networks*, Oxford University Press, Oxford, 2010. An introduction.
125. G. A. NGWA, *On the population dynamics of the malaria vector*, Bull. Math. Biol., 68 (2006), pp. 2161–2189.
126. M. NOWAK AND R. MAY, *Superinfection and the evolution of parasite virulence*, Proc. R. Soc. London, B, 255 (1994), pp. 81–89.
127. M. A. NOWAK AND R. M. MAY, *Virus dynamics*, Oxford University Press, Oxford, 2000. Mathematical principles of immunology and virology.
128. M. NUÑO, Z. FENG, M. MARTCHEVA, AND C. CASTILLO-CHAVEZ, *Dynamics of two-strain influenza with isolation and partial cross-immunity*, SIAM J. Appl. Math., 65 (2005), pp. 964–982 (electronic).
129. F. NYABADZA, *A mathematical model for combating HIV/AIDS in Southern Africa: Will multiple strategies work?*, J. Biol. Sys., 14 (2006), pp. 357–372.
130. L. F. OLSEN, G. L. TRUTY, AND W. M. SCHAFFER, *Oscillations and chaos in epidemics: a nonlinear dynamic study of six childhood diseases in Copenhagen, Denmark*, Theoret. Population Biol., 33 (1988), pp. 344–370.
131. P. E. OLSON, C. S. HAMES, A. S. BENENSON, AND E. N. GENOVESE, *The Thucydides syndrome: ebola deja vu? (or ebola reemergent?)*, Emerging Infectious Diseases, 2.
132. W. H. ORGANIZATION, *The top 10 causes of death*, Fact Sheet, 310 (July, 2013).
133. T. L. PAEZ, *Introduction to model validation*, in Proceedings of the IMAC-XXVII, Society for Experimental Mechanics, US, 2009, pp. 1–11.
134. M. J. PAPAGRIGORAKIS, C. YAPIJAKIS, P. N. SYNODINOS, AND E. BAZIOTOPOULOU-VALAVANI, *DNA examination of ancient dental pulp incriminates typhoid fever as a probable cause of the plague of Athens*, International Journal of Infectious Diseases, 10 (2006), pp. 206–214.
135. A. PUGLIESE AND A. GANDOLFI, *A simple model of pathogen-immune dynamics including specific and non-specific immunity*, Math. Biosci., 214 (2008), pp. 73–80.

136. L. RASS AND J. RADCLIFFE, *Spatial deterministic epidemics*, vol. 102 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2003.
137. C. RHODES AND R. ANDERSON, *Epidemic thresholds and vaccination in a lattice model of disease spread*, Theoretical Population Biology, (1997), pp. 101–118.
138. W. E. RICKER, *Stock and recruitment*, Journal of the Fisheries Research Board of Canada, 11 (1954).
139. H. J. ROSE, *The use of amantadine and influenza vaccine in a type a influenza epidemic in a boarding school*, Journal of Royal College of General Practitioners, 30 (1980), pp. 619–621.
140. S. RUAN, D. XIAO, AND J. C. BEIER, *On the delayed Ross–Macdonald model for malaria transmission*, Bull. Math. Biol., 70 (2008), pp. 1098–1114.
141. I. SAZONOV, M. KELBERT, AND M. B. GRAVENOR, *The speed of epidemic waves in a one-dimensional lattice of SIR models*, Math. Model. Nat. Phenom., 3 (2008), pp. 28–47.
142. L. M. SCHOULS, A. VAN DER ENDE, I. VAN DE POL, C. SCHOT, L. SPANJAARD, P. VAUTERIN, D. WILDERBEEK, AND S. WITTEVEEN, *Increase in genetic diversity of Haemophilus influenzae serotype B (HIB) strains after introduction of HIB vaccination in The Netherlands*, J. Clin. Microbiol., 43 (2005), pp. 2741–2749.
143. A. SEIERSTAD AND K. SYDSÆTER, *Optimal control theory with economic applications*, vol. 24 of Advanced Textbooks in Economics, North-Holland Publishing Co., Amsterdam, 1987.
144. F. SHARPE AND A.J.LOTKA, *A problem in age distributions*, Phil. Mag., 21 (1911), pp. 435–438.
145. S. C. SHIBOSKI AND N. P. JEWELL, *Statistical analysis of the time dependence of HIV infectivity based on partner study data*, J. Am. Stat. Assoc., 87 (1992), pp. 360–372.
146. Z. SHUAI AND P. VAN DEN DRIESSCHE, *Global stability of infectious disease models using Lyapunov functions*, SIAM J. Appl. Math., 73 (2013), pp. 1513–1532.
147. R. SLIMI, S. EL YACOUBI, E. DUMONTEIL, AND S. GOURBIÈRE, *A cellular automata model for Chagas disease*, Appl. Math. Model., 33 (2009), pp. 1072–1085.
148. S. STROGATZ, *Nonlinear Dynamics and Chaos with Applications to Physics, Biology, Chemistry, and Engineering*, Perseus Books, 2001.
149. L. H. TAYLOR, S. M. LATHAM, AND M. E. WOOLHOUSE, *Risk factors for human disease emergence*, Philosophical Transactions of the Royal Society B: Biological Sciences, 356 (2001), pp. 983–989.
150. H. R. THIEME, *Asymptotically autonomous differential equations in the plane*, Rocky Mountain J. Math., 24 (1994), pp. 351–380. 20th Midwest ODE Meeting (Iowa City, IA, 1991).
151. ———, *Mathematics in population biology*, Princeton Series in Theoretical and Computational Biology, Princeton University Press, Princeton, NJ, 2003.
152. H. R. THIEME, *Pathogen competition and coexistence and the evolution of virulence*, in Mathematics for Life Sciences and Medicine (Y. Takeuchi and Y. Iwasa and K.Sato, eds.), Springer, Berlin, Heidelberg, 2007, pp. 123–153.
153. H. R. THIEME AND C. CASTILLO-CHAVEZ, *How may infection-age-dependent infectivity affect the dynamics of HIV/AIDS?*, SIAM J. Appl. Math., 53 (1993), pp. 1447–1479.
154. H. R. THIEME AND J. YANG, *An endemic model with variable re-infection rate and applications to influenza*, Math. Biosci., 180 (2002), pp. 207–235. John A. Jacquez memorial volume.
155. P. N. V. TU, *Dynamical systems*, Springer-Verlag, Berlin, 1992. An introduction with applications in economics and biology.
156. N. TUNCER AND M. MARTCHEVA, *Analytical and numerical approaches to coexistence of strains in a two-strain SIS model with diffusion*, J. Biol. Dyn., 6 (2012), pp. 406–439.
157. ———, *Modeling seasonality in avian influenza H5N1*, J. Biol. Systems, 21 (2013), pp. 1340004, 30.
158. A. TURING, Phil. Trans. R. Soc. London, 237 (1952), pp. 37–72.
159. P. VAN DEN DRIESSCHE AND J. WATMOUGH, *Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission*, Math. Biosci., 180 (2002), pp. 29–48. John A. Jacquez memorial volume.

160. P. VAN DEN DRIESSCHE AND M. L. ZEEMAN, *Disease induced oscillations between two competing species*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 601–619.
161. E. VENTURINO, *How diseases affect symbiotic communities*, Math. Biosci., 206 (2007), pp. 11–30.
162. E. VYNNYCKY AND R. WHITE, *An introduction to infectious disease modeling,* Oxford University Press, Oxford, 2010
163. W. M. POST, D. L. DE ANGELIS, AND C. C. TRAVIS, *Endemic disease in environments with spatially heterogeneous host populations*, Mathematical Biosciences, 63 (1983), pp. 289–302.
164. W. WANG, Y. CAI, M. WU, K. WANG, AND Z. LI, *Complex dynamics of a reaction–diffusion epidemic model*, Nonlinear Anal. Real World Appl., 13 (2012), pp. 2240–2258.
165. H. WEDEMEYER, E. MIZUKOSHI, A. R. DAVIS, J. R. BENNINK, AND B. REHERMANN, *Cross-reactivity between hepatitis c virus and influenza a virus determinant-specific cytotoxic T cells*, J. Virol., 75 (2001), p. 11392–11400.
166. WHO, *Cumulative number of confirmed human cases of avian influenza A(H5N1) reported to WHO*.
167. WIKI, *List of countries by population*.
168. D. WODARZ, *Killer cell dynamics*, vol. 32 of Interdisciplinary Applied Mathematics, Springer, New York, 2007. Mathematical and computational approaches to immunology.
169. N. YOUSFI, K. HATTAF, AND M. RACHIK, *Analysis of a HCV model with CTL and antibody responses*, Appl. Math. Sci. (Ruse), 3 (2009), pp. 2835–2846.
170. L. ZHOU AND M. FAN, *Dynamics of an SIR epidemic model with limited medical resources revisited*, Nonlinear Anal. Real World Appl., 13 (2012), pp. 312–324.

Index

M. Martcheva, *An Introduction to Mathematical Epidemiology*,
Texts in Applied Mathematics 61, DOI 10.1007/978-1-4899-7612-3