

Texts in Applied Mathematics

Fred Brauer
Carlos Castillo-Chavez

Mathematical Models in Population Biology and Epidemiology

Second Edition



Springer

Texts in Applied Mathematics 40

Editors

S.S. Antman

Department of Mathematics

and

Institute for Physical Science and Technology

University of Maryland

College Park, MD 20742-4015

USA

ssa@math.umd.edu

P. Holmes

Department of Mechanical and Aerospace Engineering

Princeton University

215 Fine Hall

Princeton, NJ 08544

pholmes@math.princeton.edu

L. Sirovich

Laboratory of Applied Mathematics

Department of Biomathematical Sciences

Mount Sinai School of Medicine

New York, NY 10029-6574

lsirovich@rockefeller.edu

K. Sreenivasan

Department of Physics

New York University

70 Washington Square South

New York City, NY 10012

katepalli.sreenivasan@nyu.edu

Advisors

G. Iooss P. Holmes D. Barkley

M. Dellnitz P. Newton

For further volumes:

<http://www.springer.com/series/1214>

Fred Brauer • Carlos Castillo-Chavez

Mathematical Models in Population Biology and Epidemiology

Second Edition



Springer

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

Carlos Castillo-Chavez
Mathematical and Computational
Modeling Sciences Center (MCMSC)
Arizona State University
PO Box 871904
Tempe, Arizona 85287
USA
ccchavez@asu.edu

ISSN 0939-2475
ISBN 978-1-4614-1685-2 e-ISBN 978-1-4614-1686-9
DOI 10.1007/978-1-4614-1686-9
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011941355

Mathematics Subject Classification (2010): 92

© Springer Science+Business Media, LLC 2012

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

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

Preface to the Second Edition

Our textbook *Mathematical Models in Population Biology and Epidemiology* has reached its first decade and in the process it has captured and maintained the interest of a sufficient number of members of the computational, mathematical, modeling, and theoretical biology communities that the writing of a revised, updated and extended edition has gained the support of Springer. The field was already immense a decade ago when we took on the writing of this book, and our choices of what to include in the book were somewhat arbitrary, namely those that satisfied our interests, philosophies, and egos. Today, the research in the topics closer to our interest has grown so much that the writing of a book that heavily intersects with the field of population biology is beyond the confines of a single volume. So following the adage “why fix something that is not broken (even if old),” we have decided to maintain the core of the first edition; correct some of the errors, typos, and confusing paragraphs (while unknowingly introducing new ones); include a new chapter on the spatiotemporal dynamics of populations and expand the sections that focus on disease dynamics and control. We hope that this new edition is not only fatter but also better.

The emergence and/or reemergence of infectious diseases such as SARS, tuberculosis, and influenza are used to justify our substantial expansions of the epidemiology chapters (9 and 10). Further, this volume gives additional emphasis to the study of models that capture the dynamics of single epidemic outbreaks, influenza models, and parameter estimation.

Furthermore, a new chapter (8) is devoted to the study of the dynamics of spatially structured populations. Specifically, we have introduced the basic diffusion, reaction–diffusion, and metapopulation modeling frameworks as a *primer* for readers interested in a central topic in mathematical biology. A substantial number of additional exercises have been added, and the number of projects focused on biology has been doubled in this edition.

As in all new efforts, we have here played the primary role of “collectors” of the insights; questions and thoughts of many individuals, particularly current and former students and alumni of MTBI (Mathematical and Theoretical Biology Institute, <http://mtbi.asu.edu/>). MTBIers have helped shape this book for the past 15 years,

in a variety of ways. For example, many problems and projects have been motivated or have been adapted from the 144 technical reports generated over those 15 years, all collected in an accessible web site (<http://mtbi.asu.edu/research/archive>). We would like to thank particularly the contributions of the following MTBIs, including graduate students and faculty, Juan Aparicio, Leon Arriola, H. T. Banks, Faina Berezovskaya, Naala Brewer, Erika Camacho, Reynaldo Castro, Gerardo Chowell, Ariel Cintron-Arias, Maytee Cruz-Aponte, Mustafa Erdem, Arlene Evangelista, Zhilan Feng, Jose Flores, Luis Gordillo, Christopher Kribs-Zaleta, Raquel Lopez, Dori Luli, Emmanuel Morales, Romarie Morales, Ben Morin, Anuj Mubayi, Miriam Nuno, Anarina Murillo, David Murillo, Dustin Padilla, Kehinde Salau, Fabio Sanchez, Baojun Song, Karyn Sutton, Karen Rios-Soto, Ilyssa Summers, Steve Tennenbaum, Griselle Torres-Garcia, Jose Vega, Xiaohong Wang, Steve Wirkus, and Abdul-Aziz Yakubu. In addition, several individuals and colleagues found typos or mistakes in the first edition. We want to thank Malay Banerjee, Eric Cytrynbaum, Kathleen Dearing, Jonathan Dushoff, David Gerberry, Luis Gordillo, Jeff Moehlis, Steve Krone, Simon Levin, Marcin Mersan, Joseph Mugisha, Mason Porter, Dan Rubin, Hal Smith, V. P. Stokes, and Pauline van den Driessche for their keen eyesight and the kindness used to inform us of these mishaps.

We thank Kamal Barley, who provided invaluable help in redrawing most of the figures and constructing the index. Furthermore, we want to acknowledge particularly Ben Morin, David Murillo, and Sunmi Lee for providing solutions to a large number of problems. Our colleague Sergei Suslov and his students Jose Vega and Raquel Lopez provided a substantial number of problems and projects that are now incorporated in Chapter 8. In addition, David Kramer found many improvements as a copyeditor, and Donna Chernyk, our editor at Springer, supplied a great deal of help.

Some of the work of many of the people who had a part in the production of this book has received support from the Department of Defence, MITACS (Mathematics of Information Technology and Complex Systems), the National Science Foundation, NSERC (Natural Sciences and Engineering Research Council), the Sloan Foundation, and the Offices of the Provost and President of Arizona State University.

Finally, we want to thank our families for their love and patience despite the fact that we shamelessly used weekends and family times in this endeavor.

Vancouver, B.C., Canada
Tempe, AZ, U.S.A.
September 1, 2011.

Fred Brauer
Carlos Castillo-Chavez

Preface to the First Edition

This book is intended to inspire students in the biological sciences to incorporate mathematics in their approach to science. We hope to show that mathematics has genuine uses in biology by describing some models in population biology and the mathematics that is useful in analyzing them, as well as some case studies representing actual, if somewhat idealized, situations. A secondary goal is to expose students of mathematics to the process of modeling in the natural and social sciences.

A realistic background in mathematics for studying this book is a year of calculus, some background in elementary differential equations, and a little matrix theory. The mathematical treatment is based less on techniques for obtaining explicit solutions in “closed form”, to which students in elementary mathematics courses may be accustomed, than on approximate and qualitative methods. The emphasis is on describing the mathematical results to be used and showing how to apply them, rather than on detailed proofs of all results. References to where proofs may be found are given. Our hope is that students in the biological sciences will cover enough mathematics in their first two years of university to make this book accessible in the third or fourth year. Some review notes on the mathematics needed may be found at <http://www.cdm.yorku.ca/rev2.pdf>.

For many problems, the use of a computer algebra system can give many insights into the behavior of a model, especially for generating graphical representations of solutions. Some of the exercises and projects in the book either require the use of a computer algebra system or are simplified considerably by one. At this writing, Maple, Matlab, and Mathematica are widely used systems; students should become proficient in using at least one of them. In addition, the more specialized dynamical systems program XPP or its Windows version WinPP is very useful for studying dynamical systems and is especially valuable for differential–difference equations and equations with time lags. This program, created by Bard Ermentrout, may be downloaded from

<http://ftp.math.pitt.edu/pub/bardware/winpp.zip>

The more elaborate version XPP may be run under Windows on an X-Windows server and may be downloaded from

<http://ftp.math.pitt.edu/pub/bardware/xpp4w32.zip>

Matlab is also suitable for differential–difference equations and equations with time lags.

There are some topics that in earlier times would have been appendices in this book. These include some mathematical ideas such as Taylor approximation and the elements of linear algebra, and also some programs for solving problems with Maple, and WinPP. These topics, which we consider as the components of a virtual appendix, and some detailed solutions to selected exercises will be found in the form of PDF files at a future web site.

Answers to selected exercises are still given in the book as an appendix. Genuine understanding of the material in the book requires working of exercises; this is not a spectator sport. Answers in the back of the book are to be used to *check* your work, not to lead you to the solution.

In addition to exercises, there are several more extended descriptions of models, which call for readers to fill in some gaps. These are designated as Projects, and they may be given as group assignments.

The book concentrates on population biology. One of the practical sides of population biology is resource management. Another aspect is the study of structured population models. Mathematical epidemiology is an example, with populations structured by disease status. The core of the book, which should be included in any beginning modeling course, is Chapters 1, 2, 4, and the first five sections of Chapter 5. These chapters cover elementary continuous and discrete models for single-species populations and interacting populations. They include examples and exercises that may be too simplistic for more experienced students, who may progress through this material a little more rapidly than beginners. Chapters 9 and 10, on mathematical epidemiology, are also on a relatively elementary level and may be studied by students with relatively little background. Chapter 3, on continuous models with delays, Chapter 6, on harvesting and its implications in resource management, Chapter 7, on population models with age structure, and Chapter 8, on population models with spatial structure, as well as the later sections of Chapters 5 and 9, are more demanding mathematically. This material should probably be reserved for students with more mathematical background and some experience in biology.

The bibliography includes not only the books and papers to which reference is made in the text but also related works which pursue further some of the topics in the book. The book is meant to be an introduction to the principles and practice of mathematical modeling in the biological sciences, one which will start students on a path; it is certainly not the last word on the subject.

Vancouver, B.C., Canada
Ithaca, N.Y., U.S.A.
December 30, 2000.

Fred Brauer
Carlos Castillo-Chavez

Acknowledgments (First Edition)

The book is based primarily on notes growing out of lectures given by Fred Brauer in a modeling course at the University of Wisconsin, Madison, beginning in 1981. It reflects the input and criticism, not always followed, of faculty, colleagues, friends and students from biology and mathematics, including those who took Brauer's course. Carlos Castillo-Chavez used the notes in his Mathematical Ecology course, a course for students in the natural and mathematical sciences that he has taught or co-taught for the last 10 years at Cornell University. Castillo-Chavez also used preliminary versions of this manuscript over the last four summers as part of a core intensive seven-week research undergraduate experience (REU) in mathematical and theoretical biology at Cornell's Mathematical and Theoretical Biology Institute (MTBI). MTBI faculty and students contributed to and influenced the content and philosophy of this book. Castillo-Chavez has used portions of this manuscript in the undergraduate differential equations course that he has taught over the last two years at Cornell University's Department of Theoretical and Applied Mechanics. Several of the projects included in this book resulted from undergraduate research carried out at MTBI's Cornell-SACNAS REU program. This research was co-directed by Castillo-Chavez and Abdul-Aziz Yakubu.

The contributions of Stephen P. Blythe, Carlos Castillo-Garsow, Carlos Hernández Suarez, Maia Martcheva, Ricardo Saenz, Baojun Song, Steve Tennenbaum, and Jorge Velasco-Hernández are too many to list. In particular, Section 4.9 and Section 8.4 are modified versions of some of the lectures that Maia Martcheva gave at MTBI and the index was created by Carlos Castillo-Garsow.

Héctor Miguel Cejudo Camacho devoted a considerable amount of time and energy to the drawing or redrawing of most of the figures in this book and he helped put the manuscript into Springer format. In the course of a sequence of revisions Ricardo Oliva facilitated the transfer of files between authors and his expertise with both text and figures helped us turn the various pieces of text and figures into what we hope is a coherent book.

We would also like to acknowledge some individuals who have taught us directly or indirectly, or who have shaped in many ways the material in this book. They are: Zvia Agur, Roy Anderson, Viggo Andreasen, Juan Aparicio, Sally Blower, Stavros

Busenberg, Angel Capurro, Colin Clark, Charles Conley, Kenneth Cooke, Jim Cushing, Odo Diekmann, Klaus Dietz, Leon Glass, Zhilan Feng, Karl Hadeler, Tom Hallam, Hans Heesterbeek, Wenzhang Huang, Herbert Hethcote, John Jacquez, Simon Levin, Jia Li, Donald Ludwig, Leah Keshet, Denise Kirschner, Christopher Kribs, Tom Kurtz, Michael Mackey, Robert May, Charles McCulloch, Roger Nisbet, Joel Robbin, David Sánchez, Lisa Sattenspiel, Lee Segel, Carl Simon, Hal Smith, Avrum Soudack, Steve Strogatz, Hörst Thieme, Pauline van den Driessche, Julio Villarreal, Paul Waltman, Steve Wirkus, Gail Wolkowicz, Abdul-Aziz Yakubu, and others.

We want to thank several anonymous reviewers for their input as well as those members of the editorial board of the Springer-Verlag series of Texts in Applied Mathematics who were involved in the oversight of this book. We thank the editorial staff of Springer-Verlag for their patience and understanding in the preparation of this book. In particular, we would like to recognize the efforts of Frank McGuckin, Springer-Verlag's production editor, who kept this project alive despite our failures in meeting self-imposed deadlines. We also thank Kathy McKenzie for her editing services. Springer-Verlag Senior Editor Achi Dosanjh encouraged and supported this project from the beginning. We are grateful for her patience and great sense of humor.

Finally, we would like to remind the reader that despite everybody's support in the production of this book, at the end, the responsibility for the final product is entirely ours.

Contents

Prologue xvii

Part I Simple Single Species Models

1 Continuous Population Models 3

1.1 Exponential Growth 3

1.2 The Logistic Population Model 8

1.3 The Logistic Equation in Epidemiology 13

1.4 Qualitative Analysis 17

1.5 Harvesting in Population Models 26

1.5.1 Constant-Yield Harvesting 26

1.5.2 Constant-Effort Harvesting 28

1.6 Eutrophication of a Lake: A Case Study 31

1.7 Appendix: Parameters in Biological Systems 38

1.8 Project: The Spruce Budworm 42

1.9 Project: Estimating the Population of the United States..... 45

2 Discrete Population Models 49

2.1 Introduction: Linear Models..... 49

2.2 Graphical Solution of Difference Equations 53

2.3 Equilibrium Analysis 56

2.4 Period-Doubling and Chaotic Behavior 61

2.5 Discrete-Time Metered Models..... 67

2.6 A Two-Age Group Model and Delayed Recruitment 70

2.7 Systems of Two Difference Equations 76

2.8 Oscillation in Flour Beetle Populations: A Case Study 80

2.9 Project: A Discrete *SIS* Epidemic Model 86

2.10 Project: A Discrete-Time Two-Sex Pair-Formation Model 88

3 Continuous Single-Species Population

Models with Delays	91
3.1 Introduction	91
3.2 Models with Delay in Per Capita Growth Rates	93
3.3 Delayed-Recruitment Models	98
3.4 Models with Distributed Delay	104
3.5 Harvesting in Delayed Recruitment Models	108
3.5.1 Constant-Effort Harvesting	108
3.5.2 Constant-Yield Harvesting	109
3.6 Nicholson's Blowflies: A Case Study	112
3.7 Project: A Model for Blood Cell Populations	116
3.8 Project: Some Epidemic Models	119
3.9 Project: A Neuron Interaction Model	119

Part II Models for Interacting Species

4 Introduction and Mathematical Preliminaries	123
4.1 The Lotka–Volterra Equations	123
4.2 The Chemostat	126
4.3 Equilibria and Linearization	128
4.4 Qualitative Behavior of Solutions of Linear Systems	135
4.5 Periodic Solutions and Limit Cycles	148
4.6 Appendix: Canonical Forms of 2×2 Matrices	156
4.7 Project: A Model for Giving Up Smoking	158
4.8 Project: A Model for Retraining of Workers by their Peers	159
4.9 Project: A Continuous Two-Sex Population Model	160
5 Continuous Models for Two Interacting Populations	165
5.1 Species in Competition	165
5.2 Predator–Prey Systems	173
5.3 Laboratory Populations: Two Case Studies	185
5.4 Kolmogorov Models	190
5.5 Mutualism	191
5.6 The Spruce Budworm: A Case Study	199
5.7 The Community Matrix	206
5.8 The Nature of Interactions Between Species	209
5.9 Invading Species and Coexistence	212
5.10 Example: A Predator and Two Competing Prey	214
5.11 Example: Two Predators Competing for Prey	217
5.12 Project: A Simple Neuron Model	218
5.13 Project: A Plant–Herbivore Model	221

6	Harvesting in Two-species Models	223
6.1	Harvesting of Species in Competition	223
6.2	Harvesting of Predator–Prey Systems	229
6.3	Intermittent Harvesting of Predator–Prey Systems	237
6.4	Some Economic Aspects of Harvesting	242
6.5	Optimization of Harvesting Returns	247
6.6	Justification of the Optimization Result	251
6.7	A Nonlinear Optimization Problem	254
6.8	Economic Interpretation of the Maximum Principle	260
6.9	Project: A Harvesting Model	263
6.10	Project: Harvesting of Two Species	264

Part III Structured Population Models

7	Models for Populations with Age Structure	267
7.1	Linear Discrete Models	267
7.2	Linear Continuous Models	273
7.3	The Method of Characteristics	275
7.4	Nonlinear Continuous Models	281
7.5	Models with Discrete Age Groups	288
7.6	Project: Ordinary Differential Equations with Age Structure	290
7.7	Project: Nonlinear Age Structured Population Growth	290
7.8	Project: A Size Structured Population Model	291
8	Models for Populations with Spatial Structure	293
8.1	Introduction	293
8.2	Some Simple Examples of Metapopulation Models	294
8.3	A General Metapopulation Model	297
8.4	A Metapopulation Model with Residence and Travel	299
8.5	The Diffusion Equation	301
8.6	Solution by Separation of Variables	304
8.7	Solutions in Unbounded Regions	314
8.8	Linear Reaction–Diffusion Equations	321
8.9	Nonlinear Reaction–Diffusion Equations	323
8.9.1	Two-Species Interactions	326
8.10	Diffusion in Two Dimensions	330
8.11	Project: Cats and Birds in Space	332
8.12	Project: The Cable Equation	333
8.13	Project: Some Equations of Diffusion Type	335

Part IV Disease Transmission Models

9	Epidemic Models	345
9.1	Introduction to Epidemic Models	345
9.2	The Simple Kermack–McKendrick Epidemic Model	350
9.3	A Branching-Process Disease-Outbreak Model	361
9.3.1	Transmissibility	367
9.4	Network and Compartmental Epidemic Models	369
9.5	More Complicated Epidemic Models	373
9.5.1	Exposed Periods	373
9.5.2	Treatment Models	375
9.5.3	An Influenza Model	376
9.5.4	A Quarantine-Isolation Model	377
9.6	An SIR Model with a General Infectious Period Distribution	382
9.7	The Age of Infection Epidemic Model	384
9.8	Models with Disease Deaths	388
9.9	A Vaccination Model	391
9.10	The Next Generation Matrix	393
9.10.1	A Global Asymptotic Stability Result	403
9.11	Directions for Generalization	404
9.12	Some Warnings	404
9.13	Project: Discrete Epidemic Models	405
9.14	Project: Fitting Data for an Influenza Model	407
9.15	Project: Social Interactions	407
10	Models for Endemic Diseases	411
10.1	A Model for Diseases with No Immunity	411
10.2	The SIR Model with Births and Deaths	414
10.3	Some Applications	420
10.3.1	Herd Immunity	420
10.3.2	Age at Infection	421
10.3.3	The Interepidemic Period	422
10.3.4	“Epidemic” Approach to Endemic Equilibrium	424
10.3.5	The SIS Model with Births and Deaths	425
10.4	Temporary Immunity	427
10.5	Diseases as Population Control	431
10.6	Parameter Estimation: Ordinary Least Squares	434
10.6.1	Connecting Models to Data	434
10.6.2	Ordinary Least Squares (OLS) Estimation	436
10.7	Possible Extensions	441
10.8	Project: Pulse Vaccination	443
10.9	Project: A Model with Competing Disease Strains	445

10.10 Project: An Epidemic Model in Two Patches 447

10.11 Project: Population Growth and Epidemics 448

10.12 Project: Estimating Parameters for Leishmaniasis 453

10.13 Project: Invasive Pneumococcal Disease Surveillance Data 457

Epilogue 465

References 483

Index 501

Prologue: On Population Dynamics

As the world population exceeds the six billion mark, the question, “How many people can the earth support and under what conditions?” becomes at least as pressing as it was when Malthus (1798) posed it at the end of the eighteenth century in *An Essay on the Principle of Population*.

The ability to support growing populations within existing economic systems and environments has been one of the main concerns of societies throughout history. *How Many People Can the Earth Support?* is a recent book by J.E. Cohen (1995), in which he tackles from historical and scientific perspectives possible responses to this question. Historical “solutions” to the question of overpopulation have had as their basis two underlying assumptions: first, that under constant positive per capita rates of population growth a population increases exponentially, that is, population “explosion” is observed; second, that resource limitations necessarily limit or control the magnitude of such an explosion. The usefulness and validity of both assumptions are naturally limited since the environment, often called the “carrying capacity,” does not remain fixed. Per capita rates of population growth are not fixed, but are functions of changing environments. A limiting factor in the development of a useful (in both practical and theoretical terms) theory of population dynamics lies in the inability of theoreticians to provide models and frameworks with environmental plasticity. The environmental landscape in which we live is dynamic and often experiences dramatic shifts due to technological innovations (such as birth control, disease, famine, carbon emissions, and war) that periodically alter the bounds of what we think is possible. Cohen observes that population patterns and hence “growth” rates depend on our scale of observation in both time and space. By some scales, they are definitely not constant. For example, Cohen notes that in the fourteenth century repeated waves of Black Death, a form of bubonic plague, together with wars, heavy taxes, insurrections, and poor and sometimes malicious governments, killed an estimated one third of the population living in India and Iceland, and that the population of Meso-America fell by perhaps 80 or 90 percent during the sixteenth century [pp. 38–41]. However, in spite of such sharp short-term decreases, world population size has actually grown steadily since prehistoric times, although not at a constant rate.

Local human populations have exhibited wide fluctuations throughout time and their growth may still be responding to environmental changes in modern times. Local variability in population growth rates is high during wars, epidemics and famine. It may be affected dramatically by advances in housing, agricultural practices, health care, and so on. In retrospect, it is not surprising to see dramatic changes in the carrying capacity of the earth over time, because these changes are driven by strong environmental shifts and events that include: the effects of the global agricultural revolution observed from the seventeenth through the nineteenth centuries; the public-health transformation experienced over the last five decades via the widespread use of antibiotics and the implementation of large-scale vaccination policies; and the fertility revolution of the last four decades due to the global availability of birth-control measures (sometimes having a dramatic impact on per capita birth rates, as in China). Further, improvements in the economic state and declines in mortality from diseases in developing countries have often led to substantial declines in the birth rate. Since it is no longer necessary for a family to have as many children to ensure the survival of enough children to care for their parents in old age, such improvements in the quality of life may lead to decreases in the rate of population growth. Hence, predicting how many individuals the earth can support becomes a rather complex problem with no simple answers, particularly when different definitions of quality of life (e.g., Bangladesh as compared to Germany) are considered.

Questions and challenges raised by complex demographic processes may be addressed practically and conceptually through the use of mathematical models. Models may be particularly valuable when interactions with demographers, sociologists, economists, and health experts are at the heart of the model-building process. Simple models cannot by their own nature incorporate simultaneously many of the factors described above. However, they often provide useful insights, as will be shown in the following chapters, to help our understanding of complex processes. The usefulness of simple models to predict is limited, and their use often may lead to misleading results in the hands of “black box” users. Simple population models such as the Malthus (exponential) and the Verhulst (logistic) models represent a natural starting point in the study of demographic processes. Their main role here is to help in our *understanding* of the dynamics of basic *idealized* demographic phenomena in the social and natural sciences.

If the population of individuals at time t is denoted by $x(t)$, then Malthus’s law (1798) arises from the solution of the initial value problem

$$\frac{dx}{dt} = rx, \quad x(0) = x_0,$$

where $r = b - \mu$ denotes the constant per capita growth rate of the population, that is, the average per person number of offspring b less the per person average number of deaths μ per unit of time, and $x_0 > 0$ denotes the initial population size. Since Δx denote the change in population from t to $t + \Delta$, the dynamics are approximated over a short time period by

$$\Delta x(t) \approx (\text{births in } (t, t + \Delta)) - (\text{deaths in } (t, t + \Delta)),$$

or under our simplistic modeling assumptions by

$$\Delta x(t) \approx b x(t)\Delta - \mu x(t)\Delta = (b - \mu)x(t)\Delta = rx(t)\Delta.$$

In one of the most influential papers in history, a variant of this model was introduced by Malthus in 1798. The assumption of a constant per capita growth rate leads to the solution $x(t) = x_0 e^{rt}$, which predicts population explosion if $r > 0$, extinction if $r < 0$, or no change if $r = 0$. This model may be useful in situations in which the environment is not being taxed, the time scale of observation is small enough to make it acceptable to assume that r remains nearly constant, resources appear to be unlimited, and x_0 is small. This is a reasonable model in estimating the rate of growth of a parasite when first introduced into the bloodstream of an individual (such as the malaria parasite), in the study of the rate of growth in the number of new cases of infection at the *beginning* of an epidemic, in the estimation of the rate of growth of a pest that has just invaded a field, in estimating the rate of decay of the effect of a drug (antibiotic) in the bloodstream of an individual, or in estimating the rates of extinction of endangered species. The model may be inadequate when the number of generations gets large enough for other factors, such as density dependence, to come into play. The assumption of a constant $r > 0$ implies that a generation not only replaces itself over its life span but also contributes to the growth of its population generation after generation, while the assumption of a constant $r < 0$ implies that generations do not contribute in a significant manner to the future of a population, that is, generations are not capable of replacing themselves. An alternative way of thinking about this demographic process is via the *basic reproductive number* or *ratio* \mathcal{R}_0 . This dimensionless quantity is used to represent the average number of offspring produced by a “typical” member of the population during its reproductive life when resources are unlimited which typically occurs when x_0 is small. Here $\mathcal{R}_0 = b/\mu$, and if $\mathcal{R}_0 > 1$, the population will grow, while if $\mathcal{R}_0 < 1$, the population will eventually become extinct. The case $r = 0$ or $\mathcal{R}_0 = 1$ represents stasis, that is, a situation in which each individual on average replaces itself before it dies, and so the population size on average will not change. The case $r = 0$ represents a transition from $r < 0$ to $r > 0$ (or from $\mathcal{R}_0 < 1$ to $\mathcal{R}_0 > 1$), that is, from population decay to population explosion and vice versa. It is common to see that as a parameter—here the per capita growth rate—crosses a “tipping” or “threshold” value, the population dynamics change drastically from a situation in which we have population extinction to that in which we observe population explosion (and vice versa).

The acknowledgment of the existence of finite resources defined by the carrying capacity of an ecosystem demands the introduction of models that cannot support exponential growth indefinitely. The simplest version is obtained when it is assumed that the per capita growth rate G depends on the size of the population. In mathematical terms, we have the model

$$\frac{dx}{dt} = xG(x), \quad x(0) = x_0.$$

The most common example is provided by the logistic equation (introduced by Verhulst (1838, 1845), in which $G(x)$ is assumed to be a linear function, often parametrized as

$$G(x) = r \left(1 - \frac{x}{K} \right).$$

The ideas embodied in the logistic model (exponential growth when x is small compared to K ; little change when x is near K) led biologists to the formulation of a theory that characterizes environments in terms of those that “favor” r selection, the exponential component of growth, versus those that “favor” K selection, where K is a measure of its carrying capacity. Environments were classified as roughly belonging to two types: those that favored growth versus those that did not (because they were near their carrying capacity). The concept of r - K selection has been applied not only to populations of individuals but also to populations of species (environments in the r -phase can support additional species while environments in the K -phase cannot; see for example [May (1974)]).

In order to give demographic meaning to the definition of $G(x)$, it is often convenient to redefine it as the sum of two functions,

$$G(x) = -\mu(x) + \beta(x),$$

where $\mu(x)$ and $\beta(x)$ denote the per capita mortality and fertility rates, respectively. A constant per capita mortality rate μ may be appropriate in situations in which birth and other processes respond rapidly to changes in the population dynamics while it may not be appropriate when population death rates are affected by demographic factors such as population density.

Thus limited growth may be obtained when the per capita birth rate decreases as the population increases under the assumption of a constant per capita death rate. The simplest mathematical form for a decreasing per capita birth rate is the linear function

$$\beta(x) = a - bx,$$

which leads to a per capita growth rate

$$G(x) = \beta(x) - \mu = (a - bx) - \mu = (a - \mu) - bx,$$

or

$$G(x) = (a - \mu) \left(1 - \frac{b}{a - \mu} x \right),$$

which after renaming parameters

$$r = a - \mu, \quad K = \frac{a - \mu}{b},$$

takes on the familiar logistic form

$$G(x) = rx \left(1 - \frac{x}{K} \right).$$

Malthus's model and the logistic model describe the dynamics of populations with overlapping generations, the kind that are best described using differential equations. However, the dynamics of some populations may not be appropriately described with differential equations. For example, salmon have an annual spawning season, and births take place at essentially the same time each year. Systems with non-overlapping generations, like the salmon, are better described by discrete (difference equation) models, the topic of Chapter 2.

Impact on biological growth rates is not always instantaneous. In fact, it is often experienced after some delay (an egg stage duration, for example), and for these situations, the use of a time lag is often appropriate. The use of lags (or more generally, distributed delays) leads to the study of differential–difference equation models, the topic of Chapter 3.

The material in the first section will help the reader gain understanding of the dynamics of single-species models. The material throughout the book should help the modeler and scientist understand the value of working with simple analyzable “realistic” models rather than with detailed “unsolvable” models.

Part I, therefore, focuses on the study of single-species models including those commonly used to predict the growth of human and animal populations. Through a dissection of the underlying assumptions behind each model, it is possible to determine its usefulness as well as its limitations. Hence, single-population models are the building blocks for additional detail such as population structure, the subject of Part III.

One dimensional models assume implicitly that population growth is affected by intraspecies competition (competition such as for resources among members of the same species) in its various forms, including contest and scramble competition. Models with overlapping generations (here modeled via continuous-time one-dimensional ordinary differential equations) exhibit simple dynamical behavior, while models with discrete, nonoverlapping, generations, may exhibit complex dynamics. Specific time predictions are nearly impossible even when plenty of data is available, as is the case in the lake eutrophication example developed in Chapter 1. Nevertheless, one-dimensional models help clarify the world of possibilities once the emphasis has been shifted from a desire to fit specific (quantitative) dynamics to that of simply studying their qualitative behavior as *parameters are varied*. It should be clear that because most models of Part I are parametrized via the use of biologically derived constants that come from a multi-dimensional parameter space, we are in fact dealing with a world of possibilities with “simple” models. The spruce budworm project of Chapter 1 provides a powerful example of this view and, consequently, of the value of one-dimensional continuous-time systems.

Rich and complex dynamics are obtained in Part I, where the dynamics of discrete-time models with nonoverlapping generations are discussed. Simple non-linear models, such as the discrete logistic equation, generate complex dynamics, including chaotic behavior. Hence, the view (dramatically set in a beautiful paper by May, 1976) that *complex dynamics are not necessarily the result of complex rules of interaction* makes a compelling and powerful argument for the search for “simple” explanations for the phenomena observed in biological systems.

Part I focusses on the description and analysis of simple models, including differential, difference, and delay equations, which have played a useful role in theoretical biology (see [May, (1974)]). Models of Part I do not incorporate age structure, nor do they take into account gender-related factors, such as mating, which are often central to the study of the life history of a population. In other words, single-species models for homogeneously mixing populations can be used only to address limited (albeit often important) life history questions. Part III looks at the role of population structure and spatial heterogeneity on population dynamics in some depth.

Part II of this book looks at the dynamics that result from the interactions of populations that would *typically* behave like those of Part I when in isolation. Mechanisms that drive multispecies interactions include competition, mutualism, and predator–prey interactions. In Part II we often establish conditions for species coexistence (both species survive) or for competitive exclusion (one species survives but the other becomes extinct). The beginning of Chapter 4 is historical in nature, focussing on the study of the Lotka–Volterra equations which provide the model prototype of predator–prey interactions, and the chemostat, a laboratory biological system used to cultivate bacteria and the subject of intense mathematical study [Smith & Waltman (1995)]. These classical models motivate the introduction of two-dimensional systems, where we reintroduce the ideas of stability and instability (including oscillations) in two-dimensional continuous time systems. Chapter 5 focuses on species competition, predator–prey systems, and mutualism, with a brief foray into the much more complicated situation in which more than two species interact. Chapter 6 focuses on intervention and it intersects heavily with the field known as *bioeconomics* (see [Clark (1990)] and references therein).

Chapters 2 and 3 begin to look at the role of population structure in rather simple settings. Specifically, in Chapter 2 we look at populations with two age classes, while in Chapter 3 we consider age in various implicit forms. Individuals are allowed to remain in each stage a variable amount of time; in other words, their history plays a role. In this fashion, we are able to move away from models with no history (exponentially distributed waiting times) to models in which the time spent in a particular state is variable. In Part III, we focus on population structure. We introduce structure in two separate ways. In Chapter 7, basic generalizations that incorporate age structure in the Malthus model in discrete and continuous time are introduced. Chapter 7 focuses on the prototypes of age structured models: the (discrete) Leslie and the (continuous) McKendrick–Von Foerster models. Chapter 8 incorporates spatial structure, also in two separate ways. Metapopulations, or populations of populations, are viewed in terms of populations that have colonized different nonisolated “patches” or municipalities connected by travel routes (networks with nodes denoting local populations). They may be continuous models in each patch and discrete in separation into patches. Local motion of individuals in a population is modeled by diffusion processes, leading to partial differential equations.

Part IV is devoted to models for the spread of diseases, structuring populations by their disease status. Chapters 9 and 10 classify individuals according to their epidemiological status while incorporating additional heterogeneity through the introduction of variables that measure the variability in an individual’s ability to fight

infections. We introduce general waiting times in the context of epidemiologically structured models

An additional objective of Parts III and IV is to provide an introduction to some of the modeling and mathematical challenges faced in the field of structured populations [Webb (1985); Metz & Diekmann (1986); Diekmann & Heesterbeek (2000)]. References are provided for those interested in applications involving realistic biological models that have generated interesting mathematical questions and challenges. Discussing the analysis of these generalizations is beyond the scope of this book. However, we hope that this book and the cited references provide a springboard to these topics.

Part I

Simple Single Species Models

Chapter 1

Continuous Population Models

1.1 Exponential Growth

In this chapter we look at a population in which all individuals develop independently of one another while living in an unrestricted environment where no form of competition is possible. If the initial population size is small then a stochastic model is more appropriate, since the likelihood that the population becomes extinct due to chance must be considered. Deterministic models often provide useful ways of gaining sufficient understanding about the dynamics of populations whenever they are large enough. Furthermore, perturbations to large populations at equilibrium often generate over short time scales independent individual responses, which may be appropriately modeled by deterministic models. For example, the introduction of a single infected individual into a large disease-free population leads to the generation of secondary cases of infection, propagating a disease. The environment is free of interference competition, at least at the beginning of the outbreak, when a large population of susceptibles provides a virtually unlimited supply of hosts. The spread of disease in a large population of susceptibles may be thought of as an invasion process generated by independent contacts between a huge pool of susceptibles and a few infectious individuals.

The population density of a single species at time t will be denoted by $x(t)$, where it is assumed that x is everywhere differentiable, that is, x is a *smooth* function of t . Although unrealistic since $x(t)$ is an integer-valued function and thus not continuous, for populations with a large number of members, the assumptions of continuity and differentiability provide reasonable approximations. In many biological experiments the population biomass, which one might expect to be more nearly described by a smooth function than the population size, is often taken as the definition of $x(t)$.

The rate of change of population density can be computed if the birth, death, and migration rates are known. A *closed* population has, by definition, no migration either into or out of the population. In this case, the population size changes only through births and deaths and the rate of change of population size is simply the birth rate minus the death rate. The formulation of a specific model requires explicit

assumptions on the birth and death rates. Ideally, these assumptions are made with the goal of addressing *specific* biological questions such as under what conditions will interference competition (competition for hosts) and pathogen virulence lead to host-pathogen long-term coexistence.

For microorganisms, which reproduce by splitting, it is reasonable to assume that the rate of birth of new organisms is proportional to the number of organisms present. In mathematical terms, this assumption may be expressed by saying that if the population size at time t is x , then over a short time interval of duration h from time t to time $(t+h)$, the number of births is approximately bhx for some constant b , the *per capita birth rate*. Similarly, we may assume that the number of deaths over the same time interval is approximately μhx for some constant μ , the *per capita death rate*. Hence, the net change in population size from time t to time $(t+h)$, which is $x(t+h) - x(t)$, may be approximated by $[(bh - \mu h)x(t)]$. The duration h of the time interval must be short to ensure that the population size does not change very much and thus that the numbers of births and deaths are approximately proportional to $x(t)$. We obtain the approximate equality

$$x(t+h) - x(t) \approx (b - \mu)x(t)h. \quad (1.1)$$

(The symbol \approx is used to denote approximate equality in a sense that must be specified.) Division by h gives

$$\frac{x(t+h) - x(t)}{h} \approx (b - \mu)x, \quad (1.2)$$

and passage to the limit as $h \rightarrow 0$ gives

$$\frac{dx}{dt} = (b - \mu)x \quad (1.3)$$

under the assumption that the function $x(t)$ is differentiable. The approximate equality in (1.1) means that the difference between the two sides of (1.1) is so small that the result of dividing this difference by h gives a quantity that approaches zero as $h \rightarrow 0$.

If the net growth rate is naturally defined as

$$r \equiv b - \mu,$$

then another way of looking at this model is to observe that if the population size at time t is $x(t)$, then in the next *tiny* time interval of length h the *net* increase in population size due to a single organism will be rh . Since all individuals are independent (no competition in an unrestricted environment), the *net* increase in population due to all $x(t)$ organisms will be $rhx(t)$, and thus we arrive again at the differential equation

$$\frac{dx}{dt} = rx. \quad (1.4)$$

This differential equation has the infinite family of solutions given by the one-parameter family of functions $x(t) = ke^{rt}$; hence, this one parameter family gives a solution of (1.4) for every choice of the constant k . The most convenient way to impose a condition that will describe the population dynamics of a specific population is by specifying the initial population size at time $t = 0$ as

$$x(0) = x_0. \quad (1.5)$$

This choice selects *the* solution, $x(t) = x_0e^{rt}$. Condition (1.5) is called an *initial condition*, and the problem consisting of the differential equation (1.4) together with the initial condition (1.5) is called an *initial value problem*.

As pointed out above, the above initial value problem has the unique solution

$$x(t) = x_0e^{rt},$$

where $r > 0$ (or equivalently $b > \mu$) implies that the population size will grow unboundedly as $t \rightarrow \infty$, while $r < 0$ (or $b < \mu$) implies that the population size will approach zero as $t \rightarrow \infty$. An alternative interpretation can be reached using some of the ideas discussed in detail in the appendix to this chapter. In the absence of births ($b = 0$), the population is depleted by deaths at the rate μ , and consequently, the average life-span of a member of this population is $1/\mu$ (see Section 1.7). If $b > 0$, then the average number of offspring over the lifetime of an average individual under Malthus's model would be b/μ . If this ratio (usually referred to as the basic reproductive number or ratio \mathcal{R}_0) is greater than one, then births exceed deaths and the average number of offspring per person over her lifetime is greater than one, that is, the population explodes; if this ratio is less than one, then deaths exceed births, the average number of offspring per person is less than one, and the population dies out.

The prediction that population size will grow exponentially under these conditions was first stated by Malthus (1798). Malthus predicted disaster, because food supplies could not possibly be increased to keep pace with population growth at a constant positive per capita growth rate. Populations that grow exponentially at first are commonly observed in nature. However, their growth rates usually tend to decrease as population size increases. In fact, exponential growth or decay may be considered *typical* temporary local behavior. In other words, populations dynamics can usually be approximated by this simple model only for short periods of time; that is, the dynamics of a population may be handled well locally with linear models. The assumption that the rate of growth of a population is proportional to its size (linear assumption) is usually unrealistic on longer time scales. The next section considers nonlinear assumptions on the rate of population growth rates, which lead to quite different qualitative predictions.

Exercises

We put an asterisk on exercises that may not be straightforward. This notation is used throughout the text.

In Exercises 1 through 10, assume that the rate of change of population size is proportional to population size.

1. Suppose the growth rate of a population is 0.7944 per member per day. Let the population have 2 members on day zero. Find the population size at the end of 5 days.
2. Suppose a population has 100 members at $t = 0$ and 150 members at the end of 100 days. Find the population at the end of 150 days.
3. Suppose the growth rate of a given population is 0.21 per member per day. If the population size on a particular day is 100, find the population size 7 days later.
- 4*. Suppose a population has 39 members at $t = 8$, and 60 members at $t = 12$. What was the population size at $t = 0$?
5. Suppose a population has 24 members at $t = 5$, and 15 members at $t = 15$. What was the population size at $t = 0$?
6. The population of the earth was about 5×10^9 in 1986. Use an exponential growth model with the rate of population increase of 2 percent per year observed in 1986 to predict the population of the earth in the year 2000.
7. Bacteria are inoculated in a petri dish at a density of 10/ml. The bacterial density doubles in 20 hours. Assume that this situation is described by the differential equation

$$\frac{d}{dt}x = Cx,$$

where x is the bacterial density and C is a constant.

- a. Integrate this equation giving x as a function of time.
- b. Find the value of C .
- c. How long does it take for the density to increase to 8 times its original value? To 10 times?
8. Suppose that a population has a constant growth rate r per member per unit time and that the population size at time $t = t_0$ is x_0 . Show that the population size at time t is $x_0 e^{r(t-t_0)}$.
9. Suppose that a population has a growth rate of r per member in unit time, with $r > 0$. Show that the time required for the population to double its initial size (called the *doubling time*) is $(\log 2)/r$. [Note: We will always use “log” to denote the natural logarithm.]
10. Suppose that a population has a growth rate of r per member per unit time with $r < 0$. Show that the time required for the population to decrease to half its initial size (called the *half-life*) is $-(\log 2)/r$.
- 11*. Suppose that a population has a growth rate $r(t)$, which depends on time, so that population size $x(t)$ is governed by the initial value problem

$$\frac{dx}{dt} = r(t)x, \quad x(0) = x_0.$$

Show that the population size at time t is given by

$$x(t) = x_0 e^{\int_0^t r(s) ds}.$$

- 12*. The method of least squares finds the “best” straight line $y = ax + b$ through a set of data points $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ by choosing a and b to minimize

$$\sum_{i=1}^n (y_i - (ax_i + b))^2,$$

the sum of the squares of the vertical distances between the data and the line. The solution is

$$a = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad b = \bar{y} - a\bar{x},$$

where

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$

Table 1.1 gives the census data of the United States from 1790 to 1990. Assume that these data fit an exponential growth model $x(t) = x_0 e^{rt}$, or $\log x(t) = \log x_0 + rt$. Use the method of least squares on these data to estimate r . *Hint:* Here $\log x(t)$ corresponds to y , and t corresponds to x .

Year	Population
1790	3,900,000
1800	5,300,000
1810	7,200,000
1820	9,600,000
1830	12,900,000
1840	17,100,000
1850	23,100,000
1860	31,400,000
1870	38,600,000
1880	50,200,000
1890	62,900,000

Year	Population
1900	76,000,000
1910	92,000,000
1920	105,700,000
1930	122,800,000
1940	131,700,000
1950	150,700,000
1960	179,000,000
1970	205,000,000
1980	226,500,000
1990	248,700,000

Table 1.1 US census data from 1790 to 1990.

1.2 The Logistic Population Model

As before, $x(t)$ denotes the size of a population at time t , and dx/dt or $x'(t)$ the rate of change of population size. We shall continue to assume that the population growth rate depends only on the population's size. Such an assumption appears to be reasonable for simple organisms such as micro-organisms. For more complicated organisms like plants, animals, or human beings this is obviously an oversimplification since it ignores intra-species competition for resources as well as other significant factors, including age structure (the mortality rate may depend on age rather than on population density, while the birth rate may depend on the adult population size rather than on total population size). Furthermore, the possibility that birth or death rates may be influenced by the size of populations that interact with the population under study must also be considered (competition, predation, mutualism). We shall consider the effects of some of these factors in later chapters.

Here we study models in which the growth rate depends only on population size, because in spite of their shortcomings, these models do predict the qualitative behavior of many real populations. The *per capita growth rate*, or rate of growth per member, is given by $x'(t)/x(t)$, which we are assuming is a function of $x(t)$. In the previous section it was assumed that the total growth rate was proportional to population size (a linear model), or equivalently, we took a constant per capita growth rate. In this section total growth rates that decrease as population size increases are considered.

The simplest population model in which the per capita growth rate is a decreasing function of population size is $\lambda - ax$. This assumption leads to the *logistic* differential equation

$$x' = x(\lambda - ax),$$

first introduced by Verhulst (1838) and later studied further by R. Pearl and L. J. Reed (1920). This equation is commonly written in the form

$$x' = rx \left(1 - \frac{x}{K}\right), \quad (1.6)$$

with parameters $r = \lambda$, $K = \lambda/a$. The parameters r and K , assumed positive, may then be given biological significance. It is observed that $x' \approx rx$ when x is small, and that $x' = 0$ when x is near K . In other words, when x is small the population experiences exponential growth, while when x is near K the population hardly changes.

Separation of variables allows us to rewrite equation (1.6) as

$$\int \frac{dx}{x(K-x)} = \frac{r}{K} \int dt,$$

while using partial fractions, giving

$$\frac{1}{x(K-x)} = \frac{1}{K} \left(\frac{1}{x} + \frac{1}{K-x} \right),$$

allows us to integrate it:

$$\frac{r}{K}t + c = \int \frac{dx}{x(K-x)} = \frac{1}{K} \left(\int \frac{dx}{x} + \int \frac{dx}{K-x} \right) = \frac{1}{K} (\log x - \log(K-x)),$$

where c is the constant of integration.

If the population size at time $t = 0$ is x_0 , substitution of the initial condition $x(0) = x_0$ gives

$$c = \frac{1}{K} (\log x_0 - \log(K-x_0)).$$

We now have

$$\begin{aligned} \frac{1}{K} (\log x - \log(K-x)) &= \frac{r}{K}t + \frac{1}{K} (\log x_0 - \log(K-x_0)), \\ \log \left(\frac{x}{K-x} \right) &= rt + \log \frac{x_0}{K-x_0}, \\ \log \frac{x(K-x_0)}{x_0(K-x)} &= rt \\ \frac{x(K-x_0)}{x_0(K-x)} &= e^{rt}. \end{aligned}$$

Further algebraic simplification gives

$$\begin{aligned} x(K-x_0) &= x_0(K-x)e^{rt} = Kx_0e^{rt} - xx_0e^{rt} \\ x(K-x_0+x_0e^{rt}) &= Kx_0e^{rt} \end{aligned}$$

and finally

$$x(t) = \frac{Kx_0e^{rt}}{K-x_0+x_0e^{rt}} = \frac{Kx_0}{x_0 + (K-x_0)e^{-rt}}. \quad (1.7)$$

The above solution is valid only if $0 < x_0 < K$, so that the logarithms obtained in the integration are defined. To obtain the solution without this restriction, our integration should have given logarithms of absolute values. Nevertheless, the formula (1.7) for the solution of the logistic equation is valid for all x_0 , as could be verified by a more careful analysis.

The expression (1.7) for the solution of the logistic initial value problem shows that the population size $x(t)$ approaches the limit K as $t \rightarrow \infty$ if $x_0 > 0$. The value K is called the *carrying capacity* of the population, because it represents the population size that available resources can continue to support. The value r is called the *intrinsic growth rate*, because it represents the per capita growth rate achieved if the population size were small enough to ensure negligible resource limitations. The logistic model predicts rapid initial growth for $0 < x_0 < K$, then a decrease in growth rate as time passes so that the size of the population approaches a limit (Figure 1.1). This behavior is in agreement with the observed behavior of many populations, and for this reason, the logistic model is often used as a means of describing population size.

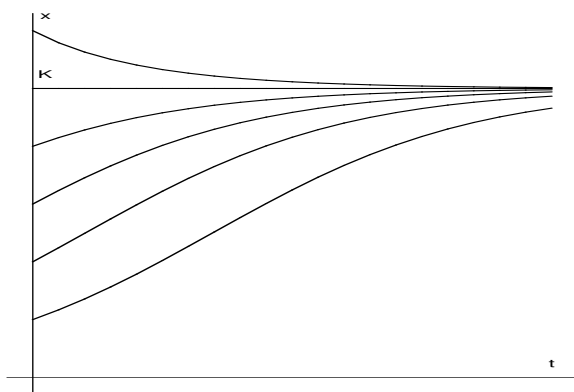


Fig. 1.1 Solution of the logistic equation.

Example 1. Census data for the population of the United States in millions fit the function

$$x(t) = \frac{265}{1 + 69e^{-0.03t}}$$

reasonably well, with the year 1790 taken as $t = 0$. We may compare this with (1.7) in the form

$$x(t) = \frac{K}{1 + \left(\frac{K-x_0}{x_0}\right)e^{-rt}}$$

to see that this expression is a solution of the logistic model with $K = 265$, $r = 0.03$. If we use this logistic model to describe the population of the United States we would predict a carrying capacity of 265,000,000, and a 1990 census total of 226,300,000. (In fact, the population size found in the 1990 census was approximately 250,000,000.) The data from [Table 1.1](#) plotted together with this solution are shown in [Figure 1.2](#).

It is possible to give a derivation of the logistic model based on a specific assumption about the resources on which a population depends. Let C denote the concentration of nutrients and assume that the per capita growth rate r is proportional to C , that is, $r = aC$ for some constant a . Assume also that we begin with a fixed concentration $C(0)$ of nutrients and that the consumption of a unit of nutrient produces b units of population size. Then population size is governed by

$$x' = aCx \tag{1.8}$$

with

$$\frac{1}{b}x' = -C'. \tag{1.9}$$

Integration of (1.9) with respect to t gives

$$C = -\frac{1}{b}x + K,$$

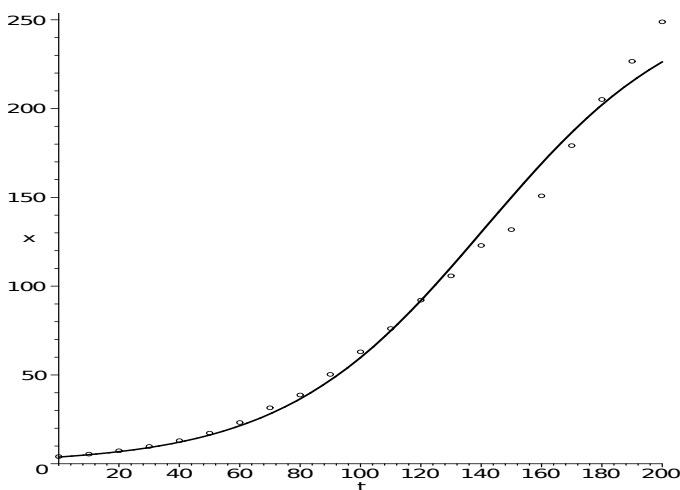


Fig. 1.2 US population size.

and substitution of $x = x_0$, $C = C(0)$ for $t = 0$ enables us to calculate the constant of integration K as

$$K = C(0) + \frac{x_0}{b}.$$

Now (1.8) becomes

$$x' = ax \left(K - \frac{1}{b}x \right) = aKx \left(1 - \frac{x}{bK} \right),$$

and this is the logistic differential equation with intrinsic growth rate aK and carrying capacity bK . The reader should observe that the intrinsic growth rate depends on the initial concentration of nutrients, but the carrying capacity depends only on the rate of conversion of nutrients into population (this derivation was taken from Edelstein-Keshet (1988)).

Exercises

1. Suppose a population satisfies a logistic model with $r = 0.4$, $K = 100$, $x(0) = 5$. Find the population size for $t = 10$.
2. Suppose a population satisfies a logistic model with carrying capacity 100 and that the population size is 10 when $t = 0$ and 20 when $t = 1$. Find the intrinsic growth rate.
3. The Pacific halibut fishery is modeled by the logistic equation with carrying capacity 80.5×10^6 , measured in kilograms, and intrinsic growth rate 0.71 per year. If the initial biomass is one-fourth the carrying capacity, find the biomass

one year later and the time required for the biomass to grow to half the carrying capacity.

4. Use a logistic model with an assumed carrying capacity of 100×10^9 , an observed population of 5×10^9 in 1986, and an observed rate of growth of 2 percent per year when population size is 5×10^9 to predict the population of the earth in the year 2000.
- 5*. Show that for a population that satisfies the logistic model, the maximum rate of growth of population size is $rK/4$, attained when population size is $K/2$.
- 6*. Show that for every choice of the constant c , the function

$$x = \frac{K}{1 + ce^{-rt}}$$

is a solution of the logistic differential equation.

- 7*. Suppose a population satisfies a differential equation having the form of the logistic equation but with an intrinsic growth rate that depends on t :

$$x' = r(t)x\left(1 - \frac{x}{K}\right), \quad x(0) = x_0.$$

Show that the solution is

$$x(t) = \frac{Kx_0}{x_0 + (K - x_0)e^{-\int_0^t r(s)ds}}.$$

[Hint: Since there is an existence and uniqueness theorem that says that the initial value problem has exactly one solution, verification that the given function satisfies the differential equation and initial condition suffices to show that it is the solution. It is not necessary to “solve” the initial value problem and derive this solution.]

8. The subfield of biology that deals with the “comparative growth rates of different biological quantities” is called *allometry*. Let $x(t)$ denote the growth rate of the biological quantity 1 and $y(t)$ the growth rate of the biological quantity 2. Assume that both per capita growth rates are proportional:

$$\frac{1}{y} \frac{dy}{dt} = \alpha \frac{1}{x} \frac{dx}{dt}. \quad (1.10)$$

Show that the following allometric law follows, that is,

$$y = Kx^\alpha,$$

where K is a constant.

- 9*. Metabolic rate (MR) is defined as the rate of heat production per unit time. The food energy required while resting and fasting (the basal metabolic rate) over a window of time T is given by

$$MR \times T.$$

It is suggested that

$$MR = \rho S,$$

where S is the surface area of the body and, according to the allometric law,

$$MR = kW^\alpha,$$

where W is the body weight.

- a. Show that for a spherical cow,

$$MR = k_s W^{2/3}.$$

- b. What would the relationship be for a cubical cow, i.e., is α the same as for a spherical cow and what is the relation between k_s and k_c ?

10. We have a starving animal and assume that its loss of weight is proportional to its metabolic rate, say

$$\frac{dW}{dt} = -\mu MR.$$

- a. Assume that $MR = \frac{1}{3}W^{\frac{3}{4}}$ and set $\mu = 1$. Find the body weight of a 25 kg animal t days after it stops eating.
- b. If death from starvation occurs after an animal has burned off 50% (law of Chossat) of its body substance, how long does it take for an animal to die given the scenario above?

1.3 The Logistic Equation in Epidemiology

The logistic equation is most often associated with the study of the dynamics of a population under density-dependent birth or death rates. However, it also arises naturally in the study of epidemiological systems (as was first shown by Hethcote (1976)). A more thorough description of mathematical models in epidemiology may be found in Chapter 9, but some epidemiological examples will be introduced in earlier chapters. In this section a model for the transmission dynamics of gonorrhea in a homosexually active population is derived. $S(t)$ denotes all sexually active noninfected individuals, and $I(t)$ denotes all sexually active infected (assumed infective) individuals with $N(t) = S(t) + I(t)$ the total population size. It is assumed that new sexually active individuals arrive at the rate $\mu N(t)$ and that none of them are infected; that individuals also leave the sexually active class at the rate $\mu N(t)$ and that individuals leave the infective class by recovery, with return to the susceptible class, at the rate $\gamma I(t)$.

With these assumptions, we arrive at the following model:

$$\begin{aligned}\frac{dS}{dt} &= \mu N(t) - B(S, I) - \mu S + \gamma I, \\ \frac{dI}{dt} &= B(S, I) - (\mu + \gamma)I,\end{aligned}$$

where $B(S, I)$ denotes the incidence rate, that is, the number of *new* cases of infection per unit time. Obviously, $B(S, 0) = B(0, I) = B(0, 0) = 0$, that is, if there are no susceptibles or infected or both, then there should be no new cases of infection. Hence, it is plausible to assume

$$B(S, I) \propto SI.$$

If we let c (assumed to be constant) denote the average number of sexual partners per individual per unit time and ϕ (assumed constant) the average number of contacts per partner, then $\phi c S(t)$ denotes the total number of contacts per unit time of all susceptible individuals at time t . If all individuals mix at random (homogeneous mixing), then the number of sexual contacts per unit time between susceptibles will be

$$\phi c S(t) \frac{S(t)}{N(t)},$$

where $N(t) = S(t) + I(t)$, while the number of sexual contacts per unit time between susceptibles and infectives will then be

$$\phi c S(t) \frac{I(t)}{N(t)},$$

and consequently,

$$B(S, I) \propto \phi c S(t) \frac{I(t)}{N(t)}.$$

In addition, we assume that only a fraction q ($0 \leq q \leq 1$) of these contacts develop into new cases of infection. Defining the transmission rate as $\beta = q\phi c$, we arrive at the following expression for the incidence rate in a randomly mixing population:

$$B(S, I) = \beta S(t) \frac{I(t)}{N(t)}.$$

The model becomes

$$\begin{aligned}\frac{dS}{dt} &= \mu N(t) - \beta S(t) \frac{I(t)}{N(t)} - \mu S + \gamma I, \\ \frac{dI}{dt} &= \beta S(t) \frac{I(t)}{N(t)} - (\mu + \gamma)I,\end{aligned}\tag{1.11}$$

with $N = S + I$, and initial conditions $S(0) = S_0 > 0$, $I(0) = I_0 > 0$.

Since

$$\frac{d}{dt}(S + I) = \frac{d}{dt}N(t) = 0,$$

the birth rate equals the death rate and the total population size is constant. The substitution

$$S(t) = N - I(t),$$

where N is a constant, reduces the solution of the system (1.11) to the solution of the single differential equation

$$\frac{dI}{dt} = \beta(N - I(t))\frac{I}{N} - (\mu + \gamma)I(t), \quad (1.12)$$

or

$$\frac{dI}{dt} = \beta I \left(1 - \frac{I}{N}\right) - (\mu + \gamma)I(t),$$

or

$$\frac{dI}{dt} = (\beta - (\mu + \gamma))I \left(1 - \frac{I}{N(1 - \frac{\mu + \gamma}{\beta})}\right)$$

or

$$\frac{dI}{dt} = \beta \left(1 - \frac{1}{R_0}\right) I \left(1 - \frac{I}{N(1 - \frac{1}{R_0})}\right),$$

where $\mathcal{R}_0 = \beta/(\mu + \gamma)$.

The definitions $r = \beta(1 - 1/R_0)$ and $K = N(1 - 1/\mathcal{R}_0)$ reveal the logistic form

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

Hence if $\mathcal{R}_0 > 1$ then $r > 0$, $K > 0$, and $I \rightarrow K$ (which is less than N). If $\mathcal{R}_0 < 1$, then $r < 0$, $K < 0$, and $I \rightarrow 0$ (see Exercise 1 below). Of course, $K < 0$ makes no biological sense. Therefore we conclude that a positive (endemic) equilibrium $I_\infty > 0$ exists and it is approached by solutions if and only if $\mathcal{R}_0 > 1$; otherwise the only biological equilibrium is $I_\infty = 0$, which is approached by solutions if $\mathcal{R}_0 \leq 1$ but not if $\mathcal{R}_0 > 1$.

Exercises

1. Show that if $r < 0$ and $K < 0$, every solution of the logistic equation with $x(0) \geq 0$ approaches zero as $t \rightarrow \infty$.
- 2*. Consider the system

$$\begin{aligned} \frac{dS}{dt} &= \Lambda - \beta S(t) \frac{I(t)}{N(t)} - \mu S + \gamma I, \\ \frac{dI}{dt} &= \beta S(t) \frac{I(t)}{N(t)} - (\mu + \gamma)I, \end{aligned} \quad (1.13)$$

where Λ denotes the total recruitment rate, assumed constant.

- a. Look at $dN/dt = d(S+I)/dt$ and solve the resulting differential equation for N , obtaining $N(t) = K - (K - N(0))e^{-\mu t}$. This shows that the system is equivalent to the solution of the single *nonautonomous* differential equation

$$\frac{dI}{dt} = \beta(N(t) - I)\frac{I}{N(t)} - (\mu + \gamma)I, \quad (1.14)$$

where

$$N(t) = K - (K - N(0))e^{-\mu t}, \quad (1.15)$$

with $K = \Lambda/\mu$.

- b. Show that $N(t) \rightarrow K$ as $t \rightarrow \infty$.
 c. Choose $K = 1000$, $1/\mu = 10$ years, and two initial population sizes, $N(0) = 1,200$ and $N(0) = 700$. Using a differential equation solver, find $I(10)$, $I(20)$, and $I(50)$ using values of the parameters that give $R_0 > 1$.
 d. If we look at the right side of equation (1.14) and let $t \rightarrow \infty$ and replace $S(t)$ by $K - I$, then we arrive formally at the following “asymptotic” differential equation:

$$\frac{dI}{dt} = \beta(K - I)\frac{I}{K} - (\mu + \gamma)I(t), \quad (1.16)$$

where

$$K \equiv \frac{\Lambda}{\mu} = \lim_{t \rightarrow \infty} N(t).$$

Here, without justification, $S(t)$ has been replaced by $K - I$, and hence equation (1.16) and Equation (1.12) are not the same. However, recent work [Castillo-Chavez and Thieme (1995)] has shown that these equations have the same qualitative dynamics. Compare the values found in (c) with those found using the limiting equation (1.16) numerically.

- 3*. Consider the following SIS epidemic model with variable population size (birth rate is different from death rate):

$$\begin{aligned} \frac{dS}{dt} &= bN - \beta S \frac{I}{N} + \gamma I - \mu S \\ \frac{dI}{dt} &= \beta S \frac{I}{N} - (\gamma + \mu)I. \end{aligned}$$

- a. Check that

$$\frac{dN}{dt} = rN,$$

where $r = (b - \mu)$, and that consequently, $N(t) = N(0)e^{rt}$.

- b. Rescale the above system by introducing the new variables

$$x(t) = \frac{S(t)}{N(t)} \quad y(t) = \frac{I(t)}{N(t)}.$$

Note that

$$\frac{dx}{dt} = \frac{1}{N} \frac{dS}{dt} - \frac{S}{N^2} \frac{dN}{dt}$$

and that $x(t) + y(t) = 1$ for all t .

c. Verify that

$$\frac{dy}{dt} = (\beta - (\gamma + b))y \left(1 - \frac{y}{1 - \frac{(\gamma + b)}{\beta}} \right).$$

d. Let $R_0 = \beta/(\gamma + b)$. Can you interpret R_0 ? What is the qualitative behavior of the above equation?

e. What is the meaning in terms of the original variables S and I as $y \rightarrow 0$ and $y \rightarrow y^* \in (0, 1)$.

1.4 Qualitative Analysis

When using the logistic model in practice, one normally assumes that a population is indeed described by a logistic model and then attempts to choose the parameters r and K and the initial population size x_0 to give the best fit with experimental data. It is important to remember that the values of r , K , and x_0 are therefore subject to error. However, errors in r and x_0 do not affect our prediction of the ultimate population size K . The property that a small change in the initial size x_0 of a solution has only a small effect on the behavior of the solution as $t \rightarrow \infty$ is called *stability* of the solution. We will require stability of any solution to which we ascribe biological significance; if a small disturbance can cause a large change in the solution it is unreasonable to consider the solution meaningful.

It is also important to remember that the logistic model is an assumed form, not a consequence of a fundamental law. We will want to consider larger classes of models and examine properties that are valid for these larger classes rather than those properties that depend on the specifics of the logistic model. A property that holds for a large class of models is said to be *robust*, to indicate that it is more likely, in some sense, to have biological significance.

The information that we derived about the behavior of solutions of the logistic model was obtained from the explicit solution by separation of variables. If we wish to search for robust properties we must learn to deduce properties of solutions from the differential equation directly, without depending on analytic expressions for solutions.

The derivative of a solution $x(t)$ at a point $(t, x(t))$ is the right side of the logistic equation

$$\frac{dx}{dt} = rx \left(1 - \frac{x}{K} \right).$$

This is positive if $0 < x < K$, zero if $x = 0$ or $x = K$, and negative if $x < 0$ or $x > K$. Thus, a solution $x(t)$ is an increasing function of t when $0 < x(t) < K$ and

a decreasing function of t when $x(t) > K$. (We ignore the case $x(t) < 0$ because it has no biological significance.) The constants $x \equiv 0$ and $x \equiv K$ are solutions. Differentiation of the logistic equation with respect to t gives

$$\frac{d^2x}{dt^2} = \frac{d}{dx} \left(rx \left(1 - \frac{x}{K} \right) \right) \frac{dx}{dt} = r \left(1 - \frac{2x}{K} \right) \frac{dx}{dt} = r^2 x \left(1 - \frac{2x}{K} \right) \left(1 - \frac{x}{K} \right).$$

From this we deduce that $(d^2x)/(dt^2)$ changes sign as x crosses the horizontal line $x = K/2$ and thus that a solution that crosses this line has an inflection point at the crossing. The solution curves must be as shown in [Figure 1.3](#).

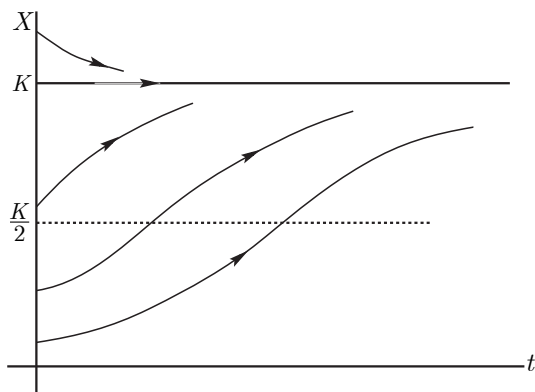


Fig. 1.3 Solution curves of the logistic equation.

If $x(t) < K$ for some t , then the graph of $x(t)$ cannot cross the line $x = K$ (to see this, we must use the *uniqueness of solutions*; if the graph did cross the line $x = K$ there would be two solutions passing through the point of crossing, and this is impossible) and is increasing for all t . Thus, $x(t)$ tends to a limit as $t \rightarrow \infty$, but the only possible limits are values of x for which the right side of the differential equation is zero, namely $x = 0$ or $x = K$. Since solutions near $x = 0$ tend away from $x = 0$, they cannot approach zero, but must tend to K as $t \rightarrow \infty$. A solution $x(t)$ that is above K decreases for all t and by a similar argument must tend to K as $t \rightarrow \infty$. Thus, we see that every nonnegative solution except the constant solution $x \equiv 0$ tends to K as $t \rightarrow \infty$, and we have obtained this information *without explicitly solving the differential equation*. The method we have used may be adapted to more general first-order differential equations.

We will consider *autonomous* first order differential equations, that is, equations of the form

$$x' = f(x), \tag{1.17}$$

in which the right side does not contain the independent variable t . Sometimes we will write the equation in the form

$$x' = xr(x), \quad (1.18)$$

with $r(x)$ representing the per capita growth rate. We define an *equilibrium* of the differential equation (1.17) to be a value x_∞ such that $f(x_\infty) = 0$. An equilibrium corresponds to a constant solution $x(t) \equiv x_\infty$ of the differential equation.

If $x(t)$ is a solution of a differential equation $x' = f(x)$ that tends to a limit as $t \rightarrow \infty$ then it is not difficult to show that its limiting value must be an equilibrium. In fact, for a first-order differential equation, every solution must either tend to an equilibrium as $t \rightarrow \infty$ or be unbounded. However, not every equilibrium is a limit of nonconstant solutions. For example, the only solution of the logistic equation that tends to zero as $t \rightarrow \infty$ is the identically zero solution.

In order to describe the behavior of solutions near an equilibrium, we introduce the process of *linearization*. If x_∞ is an equilibrium of the differential equation $x' = f(x)$, so that $f(x_\infty) = 0$, we make the change of variable $u(t) = x(t) - x_\infty$, representing deviation of the solution from the equilibrium value. Substitution gives

$$u'(t) = f(x_\infty + u(t)),$$

and application of Taylor's theorem gives

$$u'(t) = f(x_\infty) + f'(x_\infty)u(t) + \frac{f''(c)}{2!}(u(t))^2$$

for some c between x_∞ and $x_\infty + u(t)$. We use $f(x_\infty) = 0$ and write $h(u) = (f''(c)/2!)u^2$. Then we may rewrite the differential equation $x' = f(x)$ in the equivalent form

$$u' = f'(x_\infty)u + h(u).$$

The function $h(u)$ is small for small $|u|$ in the sense that $h(u)/u \rightarrow 0$ as $u \rightarrow 0$; more precisely, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|h(u)| < \varepsilon|u|$ whenever $|u| < \delta$. The *linearization* of the differential equation at the equilibrium x_∞ is defined to be the linear homogeneous differential equation

$$v' = f'(x_\infty)v, \quad (1.19)$$

obtained by neglecting the higher-order term $h(u)$ in $u' = f'(x_\infty)u + h(u)$. The importance of the linearization lies in the fact that the behavior of its solutions is easy to analyze, and this behavior also describes the behavior of solutions of the original equation (1.17) near the equilibrium.

Theorem 1.1. *If all solutions of the linearization (1.19) at an equilibrium x_∞ tend to zero as $t \rightarrow \infty$, then all solutions of (1.17) with $x(0)$ sufficiently close to x_∞ tend to the equilibrium x_∞ as $t \rightarrow \infty$.*

We have stated the theorem in a form that generalizes readily to results for systems of differential equations, as well as to results to be given later, when equations with delay are discussed. In the specific situation covered in Theorem 1.1 the condition that all solutions of the linearization tend to zero is $f'(x_\infty) < 0$. For an equilib-

rium x_∞ with $f'(x_\infty) < 0$ we must have $f(x) > 0$ for $x < x_\infty$ and $f(x) < 0$ for $x > x_\infty$ if x is sufficiently close to x_∞ . Thus, the direction field is as shown in Figure 1.4.

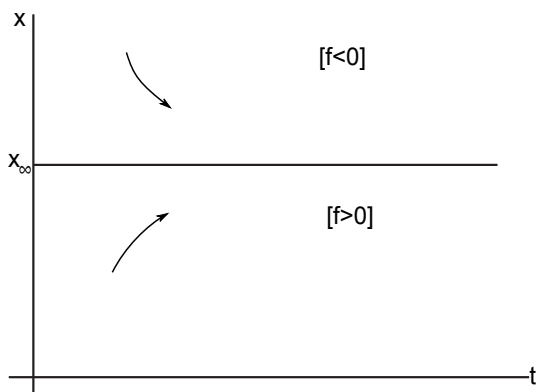


Fig. 1.4 Direction fields of the autonomous equation.

A solution with $x(0) > x_\infty$ is monotone decreasing but bounded below by x_∞ , and therefore tends to a limit as $t \rightarrow \infty$. Since the only possible limits of solutions are equilibria, such a solution must tend to x_∞ (if there are no other equilibria between $x(0)$ and x_∞). By a similar argument a solution with $x(0) < x_\infty$ increases monotonically to x_∞ if there is no other equilibrium between $x(0)$ and x_∞ . Thus, all solutions with $x(0)$ sufficiently close to x_∞ approach x_∞ as $t \rightarrow \infty$. Indeed, for first-order differential equations we can be more precise: If x_∞ is an equilibrium with $f'(x_\infty) < 0$, then every solution whose initial value $x(0)$ is between x_∞ and the next equilibrium, in either direction, must tend to x_∞ as $t \rightarrow \infty$.

An equilibrium x_∞ is said to be *stable* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|x(0) - x_\infty| < \delta$ implies $|x(t) - x_\infty| < \varepsilon$ for all $t > 0$. It is implicit in this definition that the existence of the solution $x(t)$ is required for $0 \leq t < \infty$. An equilibrium x_∞ is said to be *asymptotically stable* if it is stable and if in addition, $|x(0) - x_\infty| < \delta$ implies

$$\lim_{t \rightarrow \infty} x(t) = x_\infty.$$

Thus, stability means roughly that a small change in initial value produces only a small effect on the solution, and this condition is a natural requirement for an equilibrium to be biologically meaningful. It is possible for systems to have equilibria for which all solutions starting near the equilibrium tend toward the equilibrium but only after traveling away from the equilibrium. Such an equilibrium would not be stable, but our definition of asymptotic stability requires stability in order to exclude this possibility. In biological applications we will ordinarily require asymptotic stability rather than stability, both because asymptotic stability can be determined from the linearization, while stability cannot, and because an asymptotically stable equilibrium is not disturbed greatly by a perturbation of the differential equation.

In terms of asymptotic stability we may restate Theorem 1.1 and a corresponding instability result proved in the same way as follows:

Corollary 1.1. *An equilibrium x_∞ of (1.17) with $f'(x_\infty) < 0$ is asymptotically stable, while an equilibrium x_∞ with $f'(x_\infty) > 0$ is unstable.*

Actually, this result can be proved by an examination of the direction field without making use of the linearization. We introduce the linearization approach here because it is essential for the generalizations of Theorem 1.1 in later chapters.

We have already mentioned the logistic model, with $r(x) = r(1 - x/K)$, as an example. Other examples that have been used in population models include

$$r(x) = r \log \frac{K}{x}, \quad [\text{Gompertz (1825)}]$$

$$r(x) = \frac{r(K-x)}{K+ax}, \quad [\text{F. Smith (1963)}]$$

$$r(x) = r \left(1 - \left(\frac{x}{K} \right)^\theta \right), \quad [\text{Ayala, Gilpin, and Ehrenfeld (1973)}]$$

$$r(x) = re^{1-x/K} - d. \quad [\text{Nisbet and Gurney (1982)}]$$

In using a model of this type to study a population problem one would assume a particular form for $r(x)$, conduct experiments, and fit the resulting data to this form to estimate the parameters of the model, and then compare other observations with the predictions of the model to judge its validity.

Every autonomous differential equation of the form $x' = f(x)$ or $x' = xr(x)$ has separable variables and thus can in principle be solved by integration. The reader should observe that for each of the above examples, the necessary integration is sufficiently complicated to make a qualitative approach attractive.

If in the model

$$x' = xr(x) \quad (1.18)$$

the function $r(x)$ is nonnegative and decreasing for $0 \leq x \leq K$, then it is said to be a *compensation model*. If the per capita growth rate $r(x)$ is increasing for small x , the model is said to be a *depensation model*. If the per capita growth rate is actually negative for small x , then the model is said to be a *critical depensation model*. While compensation models are the ones most commonly examined, both depensation and critical depensation models arise in fishery studies.

A compensation model is characterized by the conditions

$$r(x) \geq 0, \quad r'(x) \leq 0 \quad \text{for } 0 \leq x \leq K.$$

For a depensation model we assume

$$\begin{aligned} r(x) &\geq 0, \quad r''(x) \leq 0 \quad \text{for } 0 \leq x \leq K, \\ r'(x) &> 0 \quad \text{for } 0 < x < K^*, \\ r'(x) &< 0 \quad \text{for } K^* < x < K. \end{aligned}$$

Thus, $r(x)$ achieves a maximum at K^* , and since

$$\lim_{x \rightarrow 0} r(x) = \lim_{x \rightarrow 0} f(x)/x = f'(0),$$

it follows that $f'(0) < r(K^*)$. Thus, the line joining the origin to the point $(K^*, f(K^*))$ on the *growth curve* $y = f(x)$, which has slope $r(K^*)$, lies above the tangent to the growth curve at the origin. Further, since

$$f'(x) = xr'(x) + r(x), \quad f''(x) = xr''(x) + 2r'(x),$$

we have $f''(0) = 2r'(0) \geq 0$ and

$$f''(K^*) = K^*r''(K^*) + 2r'(K^*) = K^*r''(K^*) < 0.$$

This shows that the growth curve has an inflection point to the left of K^* .

For a critical depensation model we assume

$$\begin{aligned} f(x) &< 0 \quad \text{for } 0 < x < K_0, \\ f(x) &\geq 0 \quad \text{for } K_0 \leq x \leq K. \end{aligned}$$

Under this assumption it is not difficult to show that the equation $x' = f(x)$ has three equilibria—an unstable equilibrium at K_0 and asymptotically stable equilibria at 0 and K . Then if the initial population size is below K_0 , the population will die out. As we will see in the next section, in the case of critical depensation hunting may drive a population to extinction by bringing the population size below the critical level K_0 , and this trend to extinction will not be reversed if hunting ceases. The extinction through hunting of the passenger pigeon in the nineteenth and early twentieth centuries from an original population of 7 billion may have been an example of critical depensation. This property is sometimes called the *Allee effect* [Allee (1931)].

In the three cases of compensation, depensation, and critical depensation the growth curve has the different forms shown in [Figure 1.5](#).

In all of these models we have been assuming tacitly that the function $f(x)$ on the right side of the differential equation is exact. Any realistic study would recognize that the model can be at best an approximation, and that instead of $x' = f(x)$ we should really be studying a differential equation of the form

$$\frac{dy}{dt} = f(y) + h(y),$$

in which the term $h(y)$ represents the error made in the assumption of the specific form $f(y)$. It is in the nature of $h(y)$ that it cannot be known explicitly. Thus, instead of looking for explicit formulas for solutions of $y' = f(y) + h(y)$, we must be satisfied with qualitative information about the solutions for a given class of functions $h(y)$. It is possible to establish the following result, which justifies our interest in asymptotically stable equilibria of the model $x' = f(x)$.

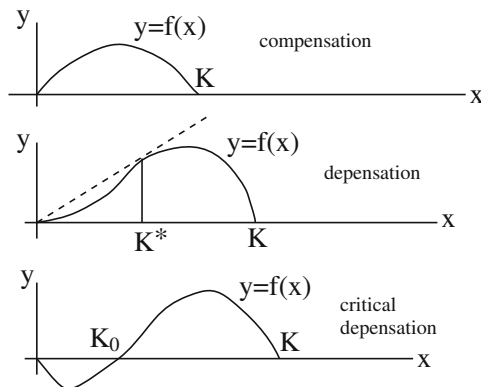


Fig. 1.5 Growth curves for the cases of compensation, depensation and critical depensation.

Theorem 1.2. Let x_∞ be an asymptotically stable equilibrium of $x' = f(x)$ with $f'(x_\infty) < 0$. Then

- (i) If $h(y)/(y - x_\infty) \rightarrow 0$ as $y \rightarrow x_\infty$, and if $|y(0) - x_\infty|$ is sufficiently small, the solution $y(t)$ of $y' = f(y) + h(y)$ tends to x_∞ as $t \rightarrow \infty$, i.e., x_∞ is an asymptotically stable equilibrium of $y' = f(y) + h(y)$.
- (ii) If $|h(y)| \leq A$ for all y and A is sufficiently small, and if $|y(0) - x_\infty|$ is sufficiently small, then $|y(t) - x_\infty| \leq (KA)/|f'(x_\infty)|$ for some constant K , i.e., solutions of $y' = f(y) + h(y)$ are close to x_∞ for all large t .

The essential content of Theorem 1.2 is that a perturbation $h(y)$ that tends to zero more rapidly than $(y - x_\infty)$ as $y \rightarrow x_\infty$ has no effect on the existence and asymptotic stability of an equilibrium, while a bounded perturbation has at worst a bounded effect on solutions. Thus conclusions drawn from analysis of the model $x' = f(x)$ are valid in a sense for a large class of more refined models.

Although we will not cite Theorem 1.2 or its analogues for difference equations, differential–difference equations, or systems of differential equations explicitly, these are the results that justify our focus on asymptotically stable equilibria of biological models. The proofs of the theoretical results cited in this section may be found in books on the qualitative theory of differential equations, [Brauer and Nohel (1989); Hurewicz (1958); Sánchez (1979); Waltman (1986)].

The description of interacting species will require systems of differential equations. The use of vector–matrix notation and the methods of linear algebra will make it possible to develop the theory of equilibria and asymptotic stability in a form analogous to what we have described here. The central result is the analogue of Theorem 1.1 in its form stated originally—that if all solutions of the linearization at an equilibrium tend to zero, then the equilibrium is asymptotically stable.

Exercises

In each of Exercises 1 through 5, find all equilibria and determine which are asymptotically stable.

1. $x' = rx(1 - x/K)$.
2. $x' = rx \log \frac{K}{x}$.
3. $x' = \frac{rx(K-x)}{K+ax}$.
4. $x' = rx \left(1 - \left(\frac{x}{K}\right)^\theta\right) \quad (0 < \theta < 1)$.
5. $x' = x(re^{1-x/K} - d)$.
6. a. A population is governed by the differential equation

$$x' = x(e^{3-x} - 1).$$

Find all equilibria and determine their stability.

- b. A fraction p ($0 < p < 1$) of the population in part (a) is removed in unit time, so that the population size is governed by the differential equation

$$x' = x(e^{3-x} - 1) - px.$$

For what values of p is there an asymptotically stable positive equilibrium?

7. For which initial values $y(0)$ does the solution $y(t)$ of the differential equation

$$y' = y(2y - e^{-y})$$

approach zero as $t \rightarrow \infty$?

- 8*. Let the concentration of a substance in a cell be C and the concentration outside the cell be Γ . Assume that the substance enters the cell by diffusion at a rate βC . This would give a differential equation

$$C' = \beta C(\Gamma - C).$$

Now include an absorption term to give a differential equation

$$C' = \beta C(\Gamma - C) - \frac{C}{1+C},$$

where β is a constant.

- a. Find the equilibria of this model and determine their stability.
 - b. Compare the behavior of this model as $t \rightarrow \infty$ with the behavior of the model without absorption.
9. For the following equations compute the equilibrium points and analyze their local stability using linearization.
 - a. $p'(t) = \beta p(t)(1 - p(t)) - ep(t)$, where β and e are positive constants. Consider the cases $\beta/e > 1$ and $\beta/e < 1$. Sketch a graph of the solution.

- b. $p'(t) = ap(t)e^{-kp(t)/p_0} - mp(t)$, where a , k , m , and p_0 are positive constants.

10. Discuss the model

$$x' = rx \left(1 - \frac{x}{K}\right) \left(\frac{x}{K_0} - 1\right),$$

where $0 < K_0 < K$. Find all limits of solutions with $x(0) > 0$ as $t \rightarrow \infty$ and find the set of initial values corresponding to each limit.

- 11*. a. By sketching direction fields determine whether the equilibrium $x = 0$ of the differential equation $x' = x^3$ is asymptotically stable or unstable.
 b. By sketching direction fields determine whether the equilibrium $x = 0$ of the differential equation $x' = -x^3$ is asymptotically stable or unstable.
 c. Is an equilibrium x_∞ of a differential equation $x' = f(x)$ with $f(x_\infty) = 0$, $f'(x_\infty) = 0$, $f''(x_\infty) = 0$, $f'''(x_\infty) < 0$ asymptotically stable or unstable?
- 12*. Let $g(x)$ be a function such that $g(K) = 0$ and $g(x) > 0$ for $0 < x < K$ and suppose $0 < x_0 < K$.

- a. Show that the function $x(t)$ defined implicitly by the relation

$$\int_{x_0}^{x(t)} \frac{du}{ug(u)} = t \quad (1.20)$$

is a solution of the initial value problem $x' = xg(x)$, $x(0) = x_0$. [Remark: This solution is obtained by separation of variables.]

- b. Show that the integral on the left side of (1.20) is negative if $x(t) < x_0$ and positive if $x(t) > x_0$, and deduce that the solution of the initial value problem must satisfy $x(t) > x_0$ for $t > 0$.
 c. Show that as $t \rightarrow \infty$, the integral on the left side of (1.20) becomes unbounded. Deduce that $g(x(t)) \rightarrow 0$, hence $x(t) \rightarrow K$.
- 13*. Populations of the North American spruce budworm have been modeled assuming that in the absence of predation the population would follow the logistic equation but that the population is subject to predation by birds [Ludwig, Jones, and Holling (1978)]. The resulting model is

$$x' = xp(x) - xq(x)$$

with

$$p(x) = r \left(1 - \frac{x}{K}\right), \quad q(x) = \frac{cx}{x^2 + A^2}.$$

- a. Show that for a differential equation $x' = xp(x) - xq(x)$, the equilibrium $x = 0$ is asymptotically stable if $p(0) < q(0)$ and that an equilibrium $x_\infty > 0$ is asymptotically stable if $p'(x_\infty) < q'(x_\infty)$.
 b. By sketching the graphs of $p(x)$ and $q(x)$ for different values of r show that the system has either one positive equilibrium that is asymptotically stable, or three positive equilibria of which two are asymptotically stable and one is unstable. [You may ignore the value of r for which equilibria

coalesce and there is a transition between one and three equilibria. It will be convenient to use a computer algebra system for making these sketches.]

14*. For the differential equation

$$x' = f(x)$$

with $f(0) = 0$, $f(x) < 0$ ($0 < x < K_0$), $f(K_0) = 0$, $f(x) > 0$ ($K_0 < x < K$), $f(K) = 0$, $f(x) < 0$ ($x > K$) (critical depensation), show that the equilibria at 0 and K are asymptotically stable and that the equilibrium at K_0 is unstable.

15. Imagine a small herd of cows in a field of modest size. The following example shows how the initial condition (the state of the field) might affect the final outcome. R.M. May (1974) developed a theoretical model to describe the dynamics of the amount of vegetation V :

$$\frac{dV}{dt} = G(V) - Hc(V), \quad (1.21)$$

where $G(V) = rV(1 - V/K)$ describes the growth of vegetation, r and K are positive constants; $c(V) = \beta V^2/(V_0^2 + V^2)$ is the consumption of vegetation per cow, β and V_0 are constants; H is the number of cows in the herd.

Choose $r = 1/3$, $K = 25$, $\beta = 0.1$, and $V_0 = 3$.

- (1) Graph the functions $G(V)$ and $Hc(V)$ for different herd sizes: $H = 10, 20, 30$. What conclusions can you draw by examining these graphs?
- (2) For the same parameter values as in part (a), graph dV/dt versus V . Use these graphs to determine all possible steady states of (1.21) and their stability for each value of H .

1.5 Harvesting in Population Models

We wish to study the effect on a population model of the removal of members of the population at a specified rate. If a population modeled by a differential equation

$$x' = f(x)$$

is subject to a harvest at a rate of $h(t)$ members per unit time for some given function $h(t)$ then the harvested population is modeled by the differential equation

$$x' = f(x) - h(t).$$

1.5.1 Constant-Yield Harvesting

If the function $h(t)$ is a constant H , so that members are removed at the constant rate of H per unit time the model is

$$x' = f(x) - H.$$

This type of harvesting is called *constant-rate* or *constant yield* harvesting. It arises when a quota is specified (for example, through permits, as in deer hunting seasons in many states, or by agreement as sometimes occurs in whaling).

If the population is governed by a logistic equation, the model with harvesting is

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{K}\right) - H, \quad (1.22)$$

and equilibria of (1.22) may be found by solving the quadratic equation $rx(1 - x/K) - H = 0$, or $x^2 - Kx + KH/r = 0$. There are two equilibria,

$$x_L = \frac{K - \sqrt{K^2 - \frac{4HK}{r}}}{2}, \quad \text{and} \quad x_U = \frac{K + \sqrt{K^2 - \frac{4HK}{r}}}{2},$$

provided $K^2 - 4HK/r \geq 0$, or $H \leq rK/4$. If $H > rK/4$, both roots are complex, $x'(t) < 0$ for all x , and every solution crashes, hitting zero in finite time. If a solution reaches zero in finite time, we consider the system to have collapsed. If $0 \leq H < rK/4$, there are two equilibria: x_L , which increases from 0 to $K/2$ as H increases from 0 to $rK/4$, and x_U , which decreases from K to $K/2$ as H increases. The stability of an equilibrium x_∞ of $x' = F(x) - H$ requires $F'(x_\infty) < 0$, which for the logistic model means $x_\infty > K/2$. Thus, x_L is always unstable and x_U is always asymptotically stable. When H increases to the critical value $H_c = rK/4$, there is a discontinuity in the behavior of the system—the two equilibria coalesce and annihilate each other. For $H < H_c$ the population size tends to an equilibrium size that approaches $K/2$ as $H \rightarrow H_c$ (provided the initial population size is at least x_L), but for $H > H_c$ the population size reaches zero in finite time for all initial population sizes (Figure 1.6). Such a discontinuity is called a (mathematical) *catastrophe*; the biological implications are catastrophic to the species being modeled [Brauer and Sánchez (1975)].

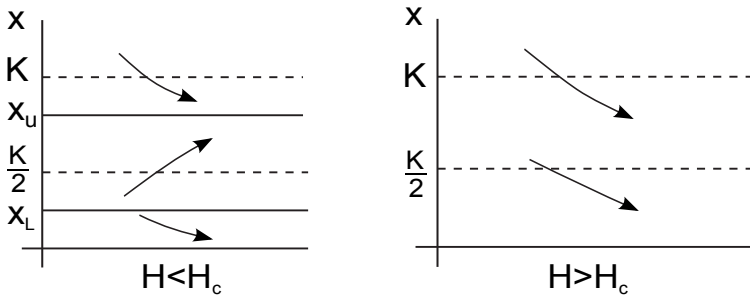


Fig. 1.6 Behavior of solutions subject to constant-yield harvesting.

For a general model $x' = f(x) - H$ equilibria are found by solving $f(x) - H = 0$, that is, by finding values x_∞ of x for which the growth curve $y = f(x)$ and the harvest

curve $y = H$ (a horizontal line) intersect. An equilibrium x_∞ is asymptotically stable if $(f(x) - H)'_{x=x_\infty} = f'(x_\infty) < 0$, that is, if at such an intersection the growth curve crosses the harvest curve from above to below as x increases (Figure 1.7).

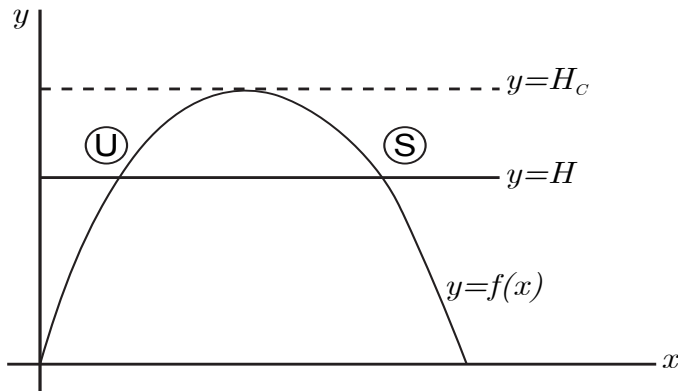


Fig. 1.7 Intersections of the growth curve with the line of constant-yield harvesting.

From Figure 1.7 it is clear that if $H > \max f(x)$, there is no equilibrium, and the critical harvest rate H_c at which two equilibria coalesce and disappear is $\max f(x)$.

1.5.2 Constant-Effort Harvesting

If the function $h(t)$ is a linear function of population size $h(t) = Ex(t)$, the model is

$$x' = f(x) - Ex.$$

This type of harvesting is called *proportional* or *constant-effort harvesting*. It arises in the modeling of fisheries, where it is often assumed that x , the number of fish caught per unit time, is proportional to E , the effort expended in fishing. This fishing effort may be measured, for example, by the number of boats fishing at a given time. The assumption that the catch is proportional to effort may be questioned on the grounds that more effort per fish caught may be necessary if the fish population is very small, but it appears to be a reasonable hypothesis for many actual fisheries.

If the population is governed by a logistic model, the harvested model is

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{K}\right) - Ex,$$

and there are two equilibria, one at $x = 0$ and one obtained by solving $r(1 - x/K) - E = 0$, which we denote by $x_\infty(E) = K(r - E)/r$, provided $0 \leq E \leq r$. It is easy to verify that the equilibrium at $x = 0$ is unstable and the equilibrium at

$x_\infty(E)$ is asymptotically stable for $0 \leq E \leq r$. As the effort increases from zero to r , the equilibrium decreases from K to zero. For a given effort E the yield is $Ex_\infty(E) = KE - KE^2/r$. This yield attains a maximum value of $rK/4$ for $E = r/2$, with $x_\infty(E) = K/2$; increasing the effort beyond $r/2$ is counterproductive in that it decreases the yield.

For a general model $x' = f(x) - Ex$ the equilibria are found by solving $f(x) - Ex = 0$, that is by finding values $x_\infty(E)$ of x for which the growth curve $y = f(x)$ and the harvest curve $y = Ex$ intersect. An equilibrium is asymptotically stable if

$$(f(x) - Ex)'_{x=x_\infty} = f'(x_\infty) - E < 0,$$

that is, if at such an intersection the growth curve crosses the harvest curve from above to below as x increases. If $f(0) = 0$, then $x = 0$ is an equilibrium that is unstable unless $x = 0$ is the only equilibrium.

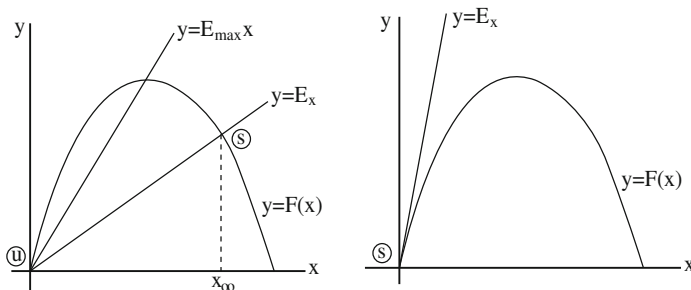


Fig. 1.8 Intersections of the growth curve and the line of constant-effort harvesting.

For a given effort E the yield is $Y = Ex_\infty(E) = F(x_\infty)$, and the maximum yield is $\max F(x)$, obtained with $E = E_{\max}$ chosen so that the line $y = Ex$ passes through the maximum of $F(x)$.

The *yield-effort curve* is the graph of yield against effort. In the case of compensation the yield increases as effort increases to a maximum, called the *maximum sustainable yield (MSY)*, and then decreases continuously to zero, reaching zero when $E = f'(0)$. However, in the case of depensation there is a critical effort $E^* = r(K^*)$ for which the yield drops to zero discontinuously (Figure 1.9). The same happens with critical depensation, but in critical depensation there is the additional property that, if the effort is large enough the population size may fall below K_0 and then be drawn to the asymptotically stable equilibrium at zero.

Exercises

1. A population of sandhill cranes (*Grus canadensis*) has been modeled by a logistic equation with carrying capacity of 194,600 members and intrinsic growth

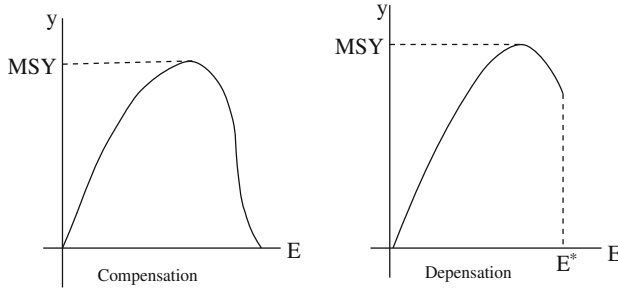


Fig. 1.9 Yield effort curve.

rate 0.0987 year^{-1} . Find the critical harvest rate for which constant-yield harvesting will drive the population to extinction, and find the equilibrium population size under constant-yield harvesting of 3000 birds per year.

- Suppose a population governed by a Gompertz model

$$x' = rx \log \frac{K}{x}$$

[Gompertz (1825)] is subjected to constant-yield harvesting. Find the critical harvest rate in terms of the parameters r and K .

- If the sandhill crane population of Exercise 1 is modeled by a Gompertz equation with the same carrying capacity and intrinsic growth rate, what would be the critical harvest rate?
- Suppose a population governed by a logistic equation with carrying capacity K , intrinsic growth rate r , and initial size K is subjected to constant-effort harvesting. By solving the initial value problem

$$x' = rx \left(1 - \frac{x}{K}\right) - Ex, \quad x(0) = K,$$

analytically, determine the population size $x(t)$ and verify that if $E \leq r$ then

$$\lim_{t \rightarrow \infty} x(t) = K(1 - E/r),$$

while if $E > r$,

$$\lim_{t \rightarrow \infty} x(t) = 0.$$

- Find the maximum sustainable yield for a population governed by a Gompertz model and subjected to harvesting (either constant-yield or constant-effort).
- Find the maximum sustainable yield for a population governed by the model

$$x' = \frac{rx(K-x)}{K+ax}$$

[F.E. Smith (1963)] and subjected to harvesting.

7. Find the maximum sustainable yield for a population governed by the model

$$x' = rxe^{(1-x/K)} - dx$$

[Nisbet and Gurney (1982)] and subjected to harvesting.

- 8*. Show that the maximum sustainable yield under constant-effort harvesting is equal to the maximum yield under constant yield-harvesting.

1.6 Eutrophication of a Lake: A Case Study

A lake is a very complicated ecosystem. A full model for a lake would need to take into account its dimensions, depth, and temperature, the concentrations of a variety of organic and inorganic materials in the lake, the kinds and quantities of vegetation in the lake, and the variety of fish and other animal life in the lake. A lake may be oligotrophic, a state characterized by low inputs of nutrients and levels of plant production and relatively clear water. On the other hand, a lake may be eutrophic, a state characterized by high nutrient input and plant production, murky water, and toxicity. Avoidance or reversal of eutrophication has benefits, both aesthetic and commercial.

Ultimately, eutrophication is caused by excessive inputs of nutrients as a by-product of agriculture, forestry, or urban development. The primary cause of eutrophication is usually excessive inputs of phosphorus, mainly due to runoff from agricultural and urban lands. Phosphorus accumulates in sediment and recycles from sediment to water, and the phosphorus input from this recycling may be a larger contribution than the phosphorus flow into the lake from outside. We shall describe a very simple model which displays many of the behaviors observed in real lakes. This model [Carpenter, Ludwig, and Brock (1999)] focuses on the phosphorus concentration in the lake and is given by the single differential equation

$$\frac{dp}{dt} = L - sp + r \frac{p^q}{m^q + p^q}. \quad (1.23)$$

Here p is the amount of phosphorus in the water. The rate of input of phosphorus from the watershed is L . The rate of loss of phosphorus from sedimentation, outflow, and absorption by consumers or plants is assumed to be proportional to the amount of phosphorus present and is given by sp . Study of limnological mechanisms suggest that the recycling rate is a sigmoid

$$\frac{rp^q}{m^q + p^q}.$$

Here the exponent q ($q \geq 2$) describes the steepness of this function at its inflection point. The value of q may range from 20 for a shallow warm lake to 2 for a deep

cold lake. The parameter r is the maximum recycling rate of phosphorus and m is the concentration of phosphorus at which recycling is half its maximum rate.

An equilibrium of the equation (1.23) is an intersection of the curve

$$Q = L + r \frac{p^q}{m^q + p^q}, \quad (1.24)$$

representing phosphorus input and the line

$$Q = sp, \quad (1.25)$$

representing phosphorus outflow. Because the line (1.25) starts at zero when $p = 0$ and is unbounded while the curve (1.24) has a nonnegative value L when $p = 0$ and is bounded as $p \rightarrow \infty$, there is at least one equilibrium. There may be as many as three equilibria. A “typical” situation is that of a small value of L and an oligotrophic equilibrium; with a large value of L there is a eutrophic equilibrium; and with an intermediate value of L there are three equilibria: an oligotrophic equilibrium, a eutrophic equilibrium, and an intermediate unstable equilibrium that separates the domains of attraction of the other two equilibria.

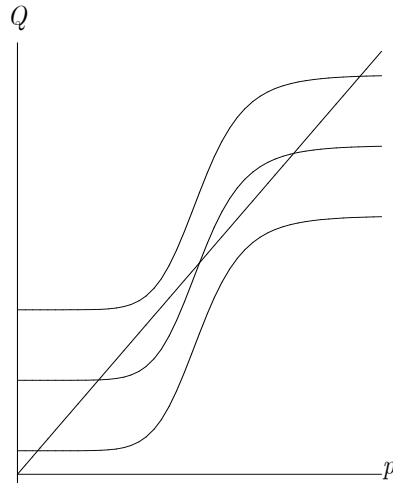
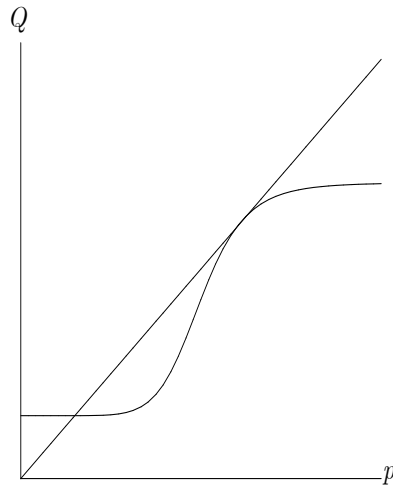
The challenge of water quality management is to control a lake that is at a eutrophic equilibrium and move it to an oligotrophic equilibrium. Sometimes this may be accomplished by reducing the external phosphorus input. However, in some lakes reducing external input alone cannot reverse eutrophication due to the amount of recycling. In order to improve the water quality of such a lake additional methods of intervention to decrease recycling or increase sedimentation would be needed. Whether such additional methods are feasible depends on the properties of the lake.

We shall consider only what can be accomplished by reducing phosphorus input. In terms of the model (1.23), we locate equilibria as intersections of the curve (1.24) and the line (1.25). Decreasing the external phosphorus input corresponds to moving the curve (1.24) down. In [Figure 1.10](#) there is a eutrophic equilibrium for large L and an oligotrophic equilibrium for small L , while for intermediate values of L both equilibria are present.

In order to manage the lake to the oligotrophic equilibrium when both oligotrophic and eutrophic equilibria are present it is necessary to bring the phosphorus concentration below the unstable equilibrium. This may require additional intervention methods. In [Figure 1.11](#) there is an input level for which the eutrophic and unstable equilibria coalesce and for which line and curve are tangent.

When L is reduced below this critical level only the oligotrophic equilibrium remains. The equilibrium jumps to this level and the eutrophication of the lake is reversed. This equilibrium jump is known as *hysteresis*, and a lake displaying this property is said to be *hysteretic*.

The minimum phosphorus input to a lake is determined by factors such as soil chemistry, and this minimum input may not be low enough to move a hysteretic lake out of its eutrophic equilibrium. Even if a hysteretic lake could be moved to the oligotrophic equilibrium, changes in conditions may increase the minimum attainable phosphorus input and lead to eutrophication.

**Fig. 1.10** Two eutrophic equilibria.**Fig. 1.11** Coalescence of eutrophic equilibria.

A hysteretic lake that is at its oligotrophic equilibrium, if disturbed by a large input of phosphorus (even a single input and not a change in the normal input), may be moved to a state that is in the domain of attraction of the eutrophic equilibrium. Thus, it may switch rather rapidly from an oligotrophic state to a eutrophic state, and this switch may be quite difficult to reverse. Sudden shifts in the state of an ecosystem can lead to severe consequences. One much-studied example is the collapse of a fishery due to a shift that does not appear to influence the fishery.

If the slope of the line (1.25) is greater than the maximum slope of the curve (1.24) the system (1.23) has only an oligotrophic equilibrium (Figure 1.12). In this

case the lake resists eutrophication, and a large input of phosphorus can be absorbed without significant harm. Such a lake is said to be *reversible*.

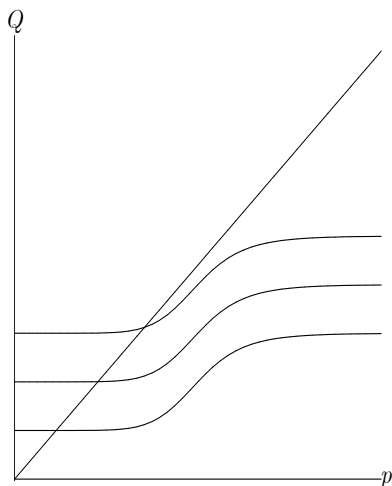


Fig. 1.12

On the other hand, it is possible that even for the minimum possible phosphorus input the line (1.25) is below the curve (1.24) for small p , and the system (1.23) has only a eutrophic equilibrium (Figure 1.13). Such a lake is said to be *irreversible*; it is not possible to bring it to an oligotrophic equilibrium by reducing the input of phosphorus.

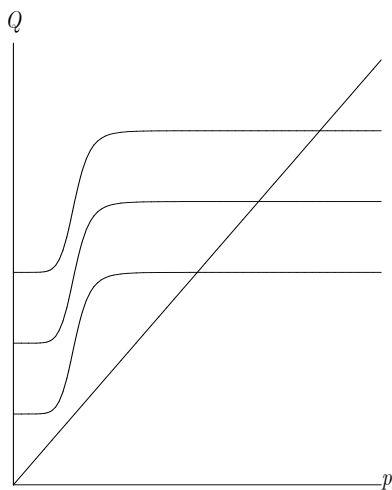


Fig. 1.13 No oligotrophic equilibrium.

Roughly speaking, a lake is reversible if the slope s of the line (1.25), representing the rate of loss of phosphorus, is sufficiently large. If s is sufficiently small, a lake is irreversible, and intermediate values of s correspond to hysteretic lakes. The classification of a lake will depend both on s and on the minimum achievable input L .

One of the most studied lakes in the world is Lake Mendota, which abuts the campus of the University of Wisconsin–Madison. Detailed measurements of phosphorus input and phosphorus mass have been recorded there for more than 20 years, and estimates have been made of the parameters s , r , m , and q in equation (1.23) [Carpenter, Brock, and Ludwig(1999)], namely

$$\begin{aligned} s &= 0.817/\text{year}, \\ r &= 731,000 \text{ kg/year}, \\ m &= 116,000 \text{ kg}, \\ q &= 7.88 \end{aligned} \tag{1.26}$$

Unfortunately, there is a great deal of uncertainty in these estimates, with a 100-fold range for m and a 10-fold range for r . This means that an estimate of the state of the lake based on these parameters and the model (1.23) ranges from reversible to irreversible. Thus, even with this much experimental data there is much uncertainty about the lake's potential response to management.

Let us imagine that the parameter values given by (1.26) are correct and then use the model (1.23) to analyze the state of the lake. It is convenient to rescale the model by setting

$$p = mx$$

to give

$$m \frac{dx}{dt} = L - smx + r \frac{m^q x^q}{m^q + m^q x^q} = L - smx + r \frac{x^q}{1 + x^q}.$$

The equilibrium condition becomes

$$L - smx + r \frac{x^q}{1 + x^q} = 0.$$

We set

$$L = ra, \quad s = \frac{rb}{m},$$

to make the equilibrium condition

$$a + \frac{x^q}{1 + x^q} = bx. \tag{1.27}$$

Thus, we attempt to determine the state of the lake by finding the intersections of the curve $y = a + x^q/(1 + x^q)$ and the line $y = bx$.

The model to which we have reduced (1.23) is

$$x' = \frac{r}{m} \left(a - bx + \frac{x^q}{1+x^q} \right). \quad (1.28)$$

The rate r/m affects the dynamics of the system, particularly the rate of approach to equilibrium but not the location or stability of equilibria. Equilibria depend on the three dimensionless parameters a , b , q . We may think of b and q as properties of the lake being studied (although some management methods may be able to change b) and a as a control parameter. We should note that uncertainty in the measurement of r and m translates into even greater uncertainty in the value of their ratio and thus into the value of b .

If we use the parameter values for Lake Mendota given by (1.26), we obtain $b = 0.130$. The minimum feasible phosphorus input rate is 3800 kg/year, corresponding to a value $a = 3800/731,000 = 0.0052$. With these values of a and b , and $q = 7.88$, we may plot the curve $y = a + x^q/(1+x^q)$ and the line $y = bx$ using a computer algebra system such as Maple or Mathematica to see that there are three equilibria. It is helpful to zoom in on the portion of the graphs for small x to identify the oligotrophic and unstable equilibria, as well as the overall structure (Figures 1.14, 1.15).

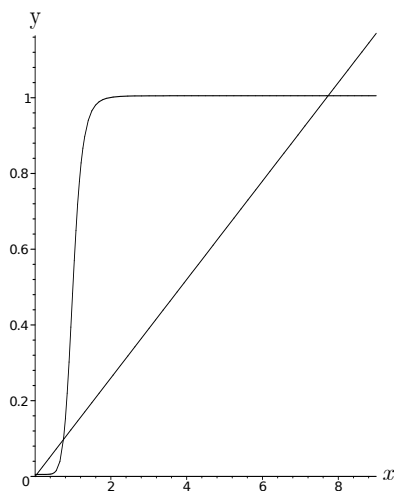


Fig. 1.14 Lake Mendota.

We may estimate these equilibrium values as 0.04, 0.75, and 7.73. These correspond to phosphorus values of 4640 kg (oligotrophic equilibrium), 87,000 kg (unstable equilibrium), and 897,000 kg (eutrophic equilibrium), respectively. The lake is currently at an equilibrium of 57,000 kg, which corresponds to $x = 0.49$. This is less than the unstable equilibrium of the system with minimum phosphorus input, and thus it is in the domain of attraction of the oligotrophic equilibrium. This indicates that the lake is hysteretic, but could be moved to an oligotrophic equilibrium.

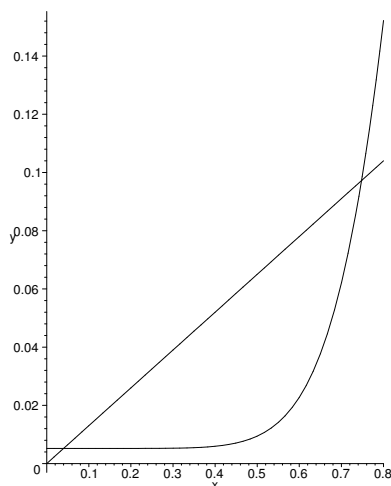


Fig. 1.15 Lake Mendota closeup.

However, in the real world of substantial uncertainty about parameter values (not to mention model simplifications) it is not possible to draw an unequivocal conclusion.

Activities such as agriculture, forestry, and urban development that produce excessive nutrient inputs into a lake and lead to eutrophication have direct economic benefits. There is a tradeoff between these benefits and the costs from damage to the lake. The various interested groups, including farmers, foresters, developers, and environmental activists, are likely to have different estimates of the overall profits and losses of an activity. For each estimate one could formulate an economic optimization model, but there are also political questions involved in the decision of which model to use. As with any political question there are substantial opportunities for misunderstanding or misinterpretation of models and data.

Exercises

1. Suppose the values of s , r , and m obtained for Lake Mendota are correct, but a huge error has been made in the estimate of q , so that actually $q = 2$. What is the state of the lake and can it be brought to an oligotrophic equilibrium?
2. Suppose the values of s and q obtained for Lake Mendota are correct but the actual values of m and r are $m = 100,000$ kg and $r = 800,000$ kg/year. What is the state of the lake and can it be saved?
3. Suppose the parameter values obtained for Lake Mendota are correct. A developer makes a one-time dump of 40,000 kg of phosphorus. What is the effect on the lake?

4. Suppose the parameter values obtained for Lake Mendota are correct. A new development raises the minimum phosphorus input to 8000 kg per year. What is the effect on the lake?

1.7 Appendix: Parameters in Biological Systems

A challenge in mathematical biology is to develop models that incorporate parameters with biological meaning. Models are used to address specific biological questions whose “answers” must be provided in biological terms. The exponential distribution plays an important role. In this section its usefulness is illustrated in a variety of settings.

Consider a death process. Let the constant μ denote the per capita death rate of a population. Then its dynamics are modeled by the equation

$$\frac{dN(t)}{dt} = -\mu N(t), \quad 0 \leq \mu < \infty, \quad N(0) = N_0,$$

where $N(t)$ denotes the population size at time t . Hence,

$$\frac{N(t)}{N_0} = e^{-\mu t}, \quad \text{for } t \geq 0,$$

that is, $e^{-\mu t}$ denotes the proportion of individuals who were alive at time $t = 0$ and who are still alive at time $t = t$ or, in probabilistic language, $e^{-\mu t}$ denotes the probability of a person being alive at time $t \geq 0$ given that he was alive at time $t = 0$.

Consequently,

$$F(t) = \begin{cases} 1 - e^{-\mu t} & \text{for } t \geq 0, \\ 0 & \text{for } t < 0, \end{cases}$$

gives the proportion of individuals who have died in $[0, t]$, or in probabilistic lingo, $F(t)$ denotes the probability of dying in the time interval $[0, t]$. $F(t)$ is a *probability distribution*, that is, $F(t)$ satisfies the properties

- (i) $F(t) \geq 0$,
- (ii) $\lim_{t \rightarrow -\infty} F(t) = 0$,
- (iii) $\lim_{t \rightarrow \infty} F(t) = 1$.

In fact $F(t)$ is the exponential cumulative probability distribution or exponential cumulative distribution function.

Probability distributions are often associated with random variables, that is, variables that take a value or a set of values with some probability. If we let X denote the time to death of an individual, then it is reasonable to assume that such a time is a random event and that X takes on the values $[0, \infty)$ with some probability. Therefore X is a continuous random variable.

Modeling the time to death X with an exponential probability distribution is equivalent to the following probability statement:

$$Prob[X \leq t] \equiv F(t) = \begin{cases} 1 - e^{-\mu t} & \text{for } t \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

From the above expression we recover the probability density associated with $F(t)$ after we observe that

$$Prob[t < X \leq t + \Delta] \approx F(t + \Delta) - F(t)$$

or, approximately, when Δ is small, that

$$Prob[t < X \leq t + \Delta] \approx \Delta \left(\lim_{\Delta \rightarrow 0} \left(\frac{F(t + \Delta) - F(t)}{\Delta} \right) \right) = \Delta f(t),$$

where $f(t) = dF/dt$ is the probability density function of X . It satisfies the properties

- (i) $f(t) \geq 0$,
- (ii) $\int_{-\infty}^{\infty} f(t) dt = 1$,
- (iii) $Prob[t < X \leq t + \Delta] = \int_t^{t+\Delta} f(e) de \approx f(t)\Delta$.

Therefore, when $F(t)$ denotes the exponential distribution,

$$f(t) = F'(t) = \begin{cases} \mu e^{-\mu t} & \text{for } t \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$Prob[\text{dying in } (t, t + \Delta)] \approx \mu e^{-\mu t} \Delta.$$

Furthermore, the average time before death, or the life expectancy, is given by the mean or expected value of X , that is

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

where

$$f(t) = \begin{cases} \mu e^{-\mu t} & \text{for } t \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Using $E[X] = \int_0^{\infty} t e^{-\mu t} dt$, after we integrate by parts we find that

$$E[X] = \frac{1}{\mu}.$$

We can do a little more, and Bayes's theorem is useful in this respect. If A and B are two probabilistic events then

$$\text{Prob}[A|B] \cdots \text{Prob}[B] \equiv \text{Prob}[A \cap B],$$

or in words, “The probability of the event A given that B has occurred times the probability of the event B equals the probability of the simultaneous occurrence of A and B ”.

If we let $B = \{X > t\}$, that is, the event that the time to death is greater than t ; and $A = \{X \leq t + \Delta\}$, that is, the event time to death is less than or equal to $t + \Delta$, then the probability that one dies before $t + \Delta$ given that one was alive at time t is computed as follows:

$$\text{Prob}[X \leq t + \Delta | X > t] \cdot \text{Prob}[X > t] = \text{Prob}[t < X \leq t + \Delta],$$

or

$$\text{Prob}[X \leq t + \Delta | X > t] = \frac{\text{Prob}[t < X \leq t + \Delta]}{\text{Prob}[X > t]},$$

or

$$\text{Prob}[X \leq t + \Delta | X > t] \approx \frac{f(t)\Delta}{1 - F(t)}.$$

In the case of the exponential distribution,

$$\text{Prob}[X \leq t + \Delta | X > t] \approx \frac{\mu e^{-\mu t} \Delta}{e^{-\mu t}} = \mu \Delta.$$

We note that $\text{Prob}[X \leq t + \Delta | X > t]$ is independent of time and approximately proportional to the length Δ of the time interval. This is why the constant of proportionality μ is referred to as the “probability of dying per unit time” ($0 \leq \mu < \infty$).

Example 1. In our study of the logistic equation with harvesting we found that

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K} \right) - \alpha N.$$

Hence $1/\alpha$ is the average time before being harvested, while when N is small compared to K , we have that

$$\frac{dN}{dt} \approx rN$$

with

$$N(t) = N_0 e^{rt}, \quad 0 \leq t < t_0,$$

with t_0 and N_0 sufficiently small. Hence, r denotes the “intrinsic” growth rate, that is, the per capita growth rate in the absence of (interference) competition.

If we define

$$\mathcal{R}_0 = \frac{r}{\alpha} = (\text{intrinsic growth rate}) \times (\text{average time before death}),$$

then when $\alpha = r$ or $\mathcal{R}_0 = 1$, we see that

$$\begin{aligned} \frac{dN}{dt} &= -\frac{r}{K}N^2, \\ \frac{dN}{N^2} &= -\frac{r}{K}dt, \\ \int_{N_0}^{N(t)} N^{-2} dN &= -\frac{r}{K}t, \\ \frac{1}{N_0} - \frac{1}{N(t)} &= -\frac{r}{K}t, \\ N(t) &= \frac{1}{\frac{1}{N_0} + \frac{r}{K}t} \rightarrow 0 \text{ as } t \rightarrow \infty, \end{aligned}$$

while when $\mathcal{R}_0 \neq 1$, we have that

$$\frac{dN}{dt} = r \left(1 - \frac{1}{\mathcal{R}_0}\right) N \left(1 - \frac{N}{K(1 - \frac{1}{\mathcal{R}_0})}\right)$$

and

$$N(t) \rightarrow K^* = K \left(1 - \frac{1}{\mathcal{R}_0}\right) > 0$$

as $t \rightarrow \infty$ when $\mathcal{R}_0 > 1$, while $N(t) \rightarrow 0$, the only biological equilibrium, as $t \rightarrow \infty$ when $\mathcal{R}_0 \leq 1$.

Consequently, \mathcal{R}_0 denotes the number of descendants of the initial small (compared to K) population N_0 of founders. If $\mathcal{R}_0 > 1$, then $N(t)$ grows and establishes itself at the level $K^* = K^*(\mathcal{R}_0)$, while if $\mathcal{R}_0 \leq 1$, then $N(t) \rightarrow 0$, and the population becomes extinct.

In the epidemiological (*SIS*) model for the transmission dynamics of gonorrhea in a homosexually active population (Section 1.3) we concluded that the dynamics were governed by the single equation

$$\frac{dI}{dt} = \beta I \left(1 - \frac{I}{K}\right) - \alpha I,$$

where $\beta = q\phi c$ denoted the transmission coefficient and α the treatment or recovery rate.

We observe that

$$\mathcal{R}_0 = \frac{\beta}{\alpha} = (\text{effective contacts}) \times (\text{effective infective period}).$$

Hence,

$$\mathcal{R}_0 > 1 \Rightarrow I(t) \rightarrow K \left(1 - \frac{1}{\mathcal{R}_0}\right) > 0 \text{ as } t \rightarrow \infty,$$

while

$$\mathcal{R}_0 \leq 1 \Rightarrow I(t) \rightarrow 0 \text{ as } t \rightarrow \infty.$$

Hence, if the number of secondary infections (descendents) generated by the small (when compared to K) initial population of infectives is greater than one, then there is an epidemic: $I(t)$ grows initially, and the disease establishes itself; in fact, $I(t) \rightarrow K^* = K(1 - 1/\mathcal{R}_0)$ as $t \rightarrow \infty$, while if $\mathcal{R}_0 \leq 1$, then $I(t) \rightarrow 0$ as $t \rightarrow \infty$, and the disease dies out.

Exercises

1. Show that

$$f(t) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq t \leq b, a < b, \\ 0 & \text{otherwise,} \end{cases}$$

is a probability density. Compute its mean value and its associated cumulative distribution.

2. Show that

$$f(t) = \begin{cases} 0.005e^{-0.005t} & \text{for } t \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

is a probability density. Compute its mean value as well as $\text{Prob}[X < 5]$, $\text{Prob}[X \leq 5]$, $\text{Prob}[2 < X < 7]$, $\text{Prob}[X > 8]$, and compute its cumulative distribution. What is $\text{Prob}[X \leq 7.001|X > 7]$? What is $\text{Prob}[X \leq 8|X > 7]$?

3. Suppose that a disease has two stages, a low infective stage and a high infective stage. Suppose that we begin with $I_1(0) = I_0 > 0$ individuals at time $t = 0$ all in stage 1 and that their disease progression from low to high to no disease is modeled by

$$\begin{aligned} \frac{dI_1}{dt} &= -\alpha_1 I_1, \\ \frac{dI_2}{dt} &= \alpha_1 I_1 - \alpha_2 I_2, \end{aligned}$$

where $1/\alpha_1$ denotes the average time in stage one and $1/\alpha_2$ denotes the average time in stage two (before total recovery). What is the average infective period? What is the probability density associated with the times as infective?

1.8 Project: The Spruce Budworm

The spruce budworm is an insect that lives in the spruce and fir forests of eastern Canada and the northeastern United States. Normally, the spruce budworm exists in low numbers in these forests, kept in check by birds. However, every 40 years or

so there is an outbreak of these pests and their numbers can defoliate and damage most of the fir trees in a forest in about 4 years. The trees (if they are not killed) can replace their foliage in about 7 to 10 years and their life span is about 100 to 150 years. The budworms, on the other hand, live for only a few months and can increase their numbers five-fold in the course of a summer. Also, since the birds do not feed exclusively on budworms, their numbers are for the most part independent of the budworm population [Yodzis (1989); Strogatz (1994)]. D. Ludwig, D.D. Jones, and C.S. Holling (1978) proposed the following model of the forest budworm system:

$$\frac{dN}{dt} = RN \left(1 - \frac{N}{K} \right) - BP \left(\frac{N^2}{A^2 + N^2} \right). \quad (1.29)$$

The first term on the right-hand side is just logistic growth for the spruce budworm population, N is the density of the budworms. R is the maximum growth rate of the spruce budworm population, K is the carrying capacity of the forest (here we can assume that it is directly proportional to the leaf biomass).

The second term is a functional response of the predator birds times the number of predators P , B is the maximum predation rate of an individual bird (on average), and A is the budworm population when the predation rate is at half the maximum.

Part 1

1. The forest biomass is directly proportional to K , which in this model is constant. Give a reasonable rationale for this, i.e., what assumptions do you think Ludwig et al. used to justify this.
2. The predator (bird) population is also taken as a constant. Explain.

Part 2

1. In order to more easily analyze the model we first nondimensionalize it. We can write without loss of generality the following dynamically equivalent equation:

$$\frac{dx}{dt} = rx \left(1 - \frac{x}{k} \right) - \left(\frac{x^2}{1 + x^2} \right). \quad (1.30)$$

Notice that we now have to deal with only two parameters. Also, all the parameters and variables are dimensionless, which means that we have to know only relative values of things rather than absolute values, which may depend on the scale of measurement and other factors. What are the appropriate transformations of the variables N and t such that the model has only the two parameters r and k , and what are r and k in terms of the original parameters (R, K, A, B , and P)?

2. Interpret in general terms (in the sense of “something relative to something else”) the rescaled variables x and t and parameters r and k . [Hint: What has “disappeared” from the functional response?]

Part 3

1. What is the equilibrium that always exists?
2. Find the other equilibria graphically by plotting on the same graph

$$f(x) = \frac{x}{1+x^2} \quad \text{and} \quad g(x) = r \left(1 - \frac{x}{k}\right).$$

Show one example of each qualitatively different situation. Try $0 < x < 25$.

3. Show the stability of the equilibria found in Part (2) by plotting $\dot{x}/x^{3/4}$ versus $x/x^{3/4}$ and drawing (by hand is OK) arrows along the x -axis showing movement toward or away from the equilibrium points. (Dividing both the x - and y -axis by $x^{3/4}$ will rescale so that the picture is clearer).
4. What is the minimum value of r and the maximum value of k for which it is *impossible* to have three nontrivial equilibria? You can find them analytically or graphically (to two decimal places).

Part 4

1. Below is a plot (Figure 1.16) of the nontrivial equilibria x^* versus k . Draw a few vectors for Δx on both sides of the curve. This curve is called a *bifurcation curve*. A bifurcation is a point in parameter space where equilibria appear, disappear, or change stability, and the bifurcation curve indicates the parameter values for which a change may occur.

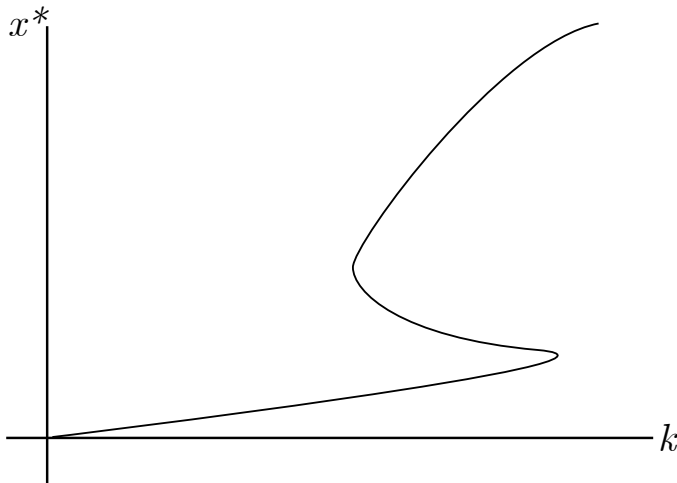


Fig. 1.16 Spruce budworm model bifurcation curve.

2. Indicate where the curve represents the set of stable and unstable equilibria. (You should be able to do this using Part (1) without any calculations.)
3. Assume that the population tends to stay at equilibrium but that k (essentially the amount of leaf biomass) will drift up for low stable x^* and will drift down

for high stable x^* . Show what happens to the population over time on the graph and explain.

Part 5

1. Compute and plot the bifurcation curves for this equation. This can be done most easily in parametric form with $r = r(x)$ and $k = k(x)$. You can see this easily by looking at your graph in Part 2. There are two conditions for the bifurcations,

- a. x is at an equilibrium, i.e., $dx/d\tau = 0 \Rightarrow f(x) = g(x)$, and
- b. $g(x)$ is tangent to $f(x)$, or $(d/dx)g(x) = (d/dx)f(x)$.

You should get two equations, one for $r(x)$ and the other for $k(x)$, which will define two separate curves in the (r, k) -plane (remember to restrict the ranges of r and k to positive values). Plot the curves in the (r, k) -plane and label the different regions with number of equilibria and whether they are “outbreak”, “refuge”, or “bistable.”

2. The point where the two curves intersect and disappear is called the *cusp*. Calculate its coordinates.

1.9 Project: Estimating the Population of the United States.

Table 1.1 in Exercise 3, Section 1.1, gives census data for the United States from 1790 to 1990. In this project, we will explore the question of fitting these data to a logistic model.

Method 1: We may rewrite the solution(1.7) of the logistic model as

$$\frac{K-x}{x} = \frac{K-x_0}{x_0} e^{-rt},$$

or taking natural logarithms,

$$\log \frac{K-x}{x} = \log \frac{K-x_0}{x_0} - rt. \quad (1.31)$$

Question 1

Derive the relation (1.31) from (1.7)

Thus, if we plot $\log \frac{K-x}{x}$ against t , we should obtain a straight line. However, there is a problem: We do not know the value of K . We may try to avoid this problem by estimating K by eye from the graph of the data points (t, x) . If we obtain similar curves when we fit the data to a logistic curve for several different values of K , so that our results are not very sensitive to changes in K , then we may have some confidence in these results.

Question 2

For each of the values $K = 200, K = 250, K = 300$, use the data of [Table 1.1](#) to plot $\log \frac{K-x}{x}$ as a function of t .

Question 3

For each of the values $K = 200, K = 250, K = 300$, use the method of least squares (Exercise 3, Section 1.1) to estimate the slope $-r$ and the intercept $\log \frac{K-x_0}{x_0}$.

Question 4

For each of the values $K = 200, K = 250, K = 300$, use the results obtained in Question 3 to give a function describing the population size, and use this function to predict the result of the year 2000 census.

Method II: From the logistic model, we have

$$\frac{x'}{x} = r \left(1 - \frac{x}{K} \right),$$

Thus if we plot $\frac{x'(t)}{x(t)}$ as a function of x , we should obtain a straight line with x -intercept K and slope $\frac{-r}{K}$. The problem here is that our data describe x , not x' . However, we can use the data to estimate x' . If x_i and x_{i+1} are two consecutive measurements taken with a time interval h , we may approximate x'_i by $\frac{x_{i+1} - x_i}{h}$. Actually, the approximation $\frac{x_{i+1} - x_{i-1}}{2h}$ is an approximation to x'_i which is significantly more accurate.

Question 5

Use the data of [Table 1.1](#) to estimate x'_i and then plot $\frac{x'_i}{x_i}$ against x_i .

Question 6

Use the method of least squares to estimate r and K , and use your result to predict the result of the year 2000 census.

In Question 6 you should find that the data points lie quite close to your straight line up to 1940 but not for 1950 and later.

Question 7

Using only the data from 1950 on, estimate x'_i and plot $\frac{x'_i}{x_i}$ against x_i . Then use the method of least squares to estimate r and K , and use your result to predict the result of the year 2000 census.

Question 8

What reasons might there be for the apparent jump in carrying capacity between 1940 and 1950?

Over the past 150 years, immigration to the United States has varied considerably from year to year but has averaged more than 150,000 per year. The logistic model does not include any immigration.

Question 9

Suggest a modification of the logistic model that would include immigration. What would you expect to be the effect of this modification on the estimate of carrying capacity that would be obtained using the same data?

Question 10

Taking into account all that you have done in this project, what is your best guess for the result of the year 2000 census?

Chapter 2

Discrete Population Models

2.1 Introduction: Linear Models

In this chapter we shall consider populations with a fixed interval between generations or possibly a fixed interval between measurements. Thus, we shall describe population size by a *sequence* $\{x_n\}$, with x_0 denoting the initial population size, x_1 the population size at the next generation (at time t_1), x_2 the population size at the second generation (at time t_2), and so on. The underlying assumption will always be that population size at each stage is determined by the population sizes in past generations, but that intermediate population sizes between generations are not needed. Usually the time interval between generations is taken to be a constant.

For example, suppose the population changes only through births and deaths, so that $x_{n+1} - x_n$ is the number of births minus the number of deaths over the time interval from t_n to t_{n+1} . Suppose further that the birth and death rates are constants b and d , respectively (that is, if the population size is x then there are bx births and dx deaths in that generation). Then

$$x_{n+1} - x_n = (b - d)x_n,$$

or

$$x_{n+1} = x_n + (b - d)x_n = (1 + b - d)x_n.$$

We let $r = 1 + b - d$ and obtain the *linear homogeneous difference equation*

$$x_{n+1} = rx_n.$$

This together with the prescribed initial population size x_0 determines the population size in each generation. By a *solution* of the difference equation $x_{n+1} = rx_n$ with initial value x_0 we mean a sequence $\{x_n\}$ such that $x_{n+1} = rx_n$ for $n = 0, 1, 2, \dots$, with x_0 as prescribed.

It is easy to solve the difference equation $x_{n+1} = rx_n$ algebraically. We begin by observing that $x_1 = rx_0$, $x_2 = rx_1 = r^2x_0$, $x_3 = rx_2 = r^3x_0$, and then we guess (and prove by induction) that the unique solution is $x_n = r^n x_0$ ($n = 0, 1, 2, \dots$). It follows

that if $|r| < 1$, then $x_n \rightarrow 0$ as $n \rightarrow \infty$, while if $|r| > 1$, then x_n grows unboundedly as $n \rightarrow \infty$. More precisely, if $0 \leq r < 1$, x_n decreases monotonically to zero; if $-1 < r < 0$, x_n oscillates, alternating between positive and negative values, but tends to zero; if $r > 1$, x_n increases to $+\infty$; if $r < -1$, x_n oscillates unboundedly. Negative values of x_n for this difference equation have no biological meaning, but we soon will consider difference equations in which the unknown is a deviation from equilibrium (which may be either positive or negative) rather than a population size. For this reason we have used the difference equation $x_{n+1} = rx_n$ as our first example, even though a more plausible model for a real population might be

$$x_{n+1} = \begin{cases} rx_n & \text{for } x_n > 0, \\ 0 & \text{for } x_n \leq 0, \end{cases}$$

which says that the population becomes extinct once it becomes zero in any generation. This will occur if and only if $r \leq 0$. The model $x_{n+1} = rx_n$ also arises under the assumption that all members of each generation die, but there is a constant birth rate b to form the next generation. In this case $d = 1$, so that $r = b$. We may form a different model by allowing migration and assuming a constant migration rate β per generation, with positive β denoting immigration and negative β denoting emigration. This leads to the linear inhomogeneous difference equation

$$x_{n+1} = rx_n + \beta,$$

which may also be solved iteratively,

$$\begin{aligned} x_1 &= rx_0 + \beta, \\ x_2 &= rx_1 + \beta = r(rx_0 + \beta) + \beta = r^2x_0 + r\beta + \beta, \\ x_3 &= rx_2 + \beta = r(r^2x_0 + r\beta + \beta) + \beta = r^3x_0 + r^2\beta + r\beta + \beta \\ &\vdots \end{aligned}$$

Again we may guess, and then prove by induction, that

$$\begin{aligned} x_n &= r^n x_0 + \beta(r^{n-1} + r^{n-2} + \dots + r + 1) \\ &= r^n x_0 + \frac{\beta(1 - r^n)}{1 - r} = \left(x_0 - \frac{\beta}{1 - r}\right)r^n + \frac{\beta}{1 - r}. \end{aligned}$$

If $r > 1$, then x_n grows unboundedly for $\beta > -(r-1)x_0$, but x_n reaches zero if $\beta < -(r-1)x_0$; thus sufficiently large emigration will wipe out a population that would otherwise grow unboundedly. If $0 < r < 1$, then x_n tends to the limit $\beta/(1-r) > 0$ for $\beta > 0$, while x_n reaches zero for $\beta < 0$. Thus, immigration may produce survival of a population that would otherwise become extinct.

The assumption of a constant growth rate independent of population size is unlikely to be reasonable for real populations except possibly while the population size is small enough not to be subject to the effects of overcrowding. Various nonlinear

difference equation models have been proposed as more realistic. For example, the difference equations

$$x_{n+1} = \frac{rx_n}{x_n + A} \quad [\text{Verhulst (1845)}]$$

and

$$x_{n+1} = \frac{rx_n^2}{x_n^2 + A}$$

have been suggested as descriptions of populations that die out completely in each generation and have birth rates that saturate for large population sizes. The difference equations

$$x_{n+1} = x_n + rx_n \left(1 - \frac{x_n}{K}\right) \quad \text{and} \quad x_{n+1} = rx_n \left(1 - \frac{x_n}{K}\right),$$

both called the *logistic* difference equation, and essentially equivalent, describe populations with growth rates that decrease to zero as the population grows large. Neither should be taken too seriously for large population sizes since x_{n+1} becomes negative if x_n is too large. Another form, which could with some justification also be called the logistic equation, is

$$x_{n+1} = x_n e^{r(1-x_n/K)}.$$

Here the growth rate decreases to zero as $x_n \rightarrow \infty$, but x_{n+1} cannot become negative. Other difference equations, which have in fact been used as models to try to fit field data, are

$$x_{n+1} = rx_n(1 + \alpha x_n)^{-\beta} \quad [\text{Hassell (1975)}]$$

and

$$x_{n+1} = \begin{cases} r\epsilon^\beta x_n^{1-\beta} & \text{for } x_n > \epsilon \\ rx_n & \text{for } x_n < \epsilon. \end{cases}$$

It should be recognized that none of these models is derived from actual population growth laws. Rather, they are attempts to give quantitative expression to rough qualitative ideas about the biological laws governing the population. For this reason, we should be skeptical of the biological significance of any deduction from a specific model that holds only for that model. Our goal should be to formulate principles that are *robust*, that is, valid for a large class of models (ideally for all models that embody some set of qualitative hypotheses). In Section 2.5 we will describe some difference equation models that have been used to model fish populations and that are based on biological assumptions. Such models give some insight into the types of qualitative hypotheses that may be realistic.

Exercises

1. Find the solution of the difference equation $x_{n+1} = \frac{1}{2}x_n$, $x_0 = 2$.
2. Find the solution of the difference equation $x_{n+1} = \frac{1}{2}x_n + 1$, $x_0 = 2$.
3. Find by calculating recursively the solution of the second-order difference equation $x_{n+2} = \frac{1}{2}x_n$, $x_0 = 1$, $x_1 = -1$.
4. Consider the second order difference equation

$$x_{n+2} - 3x_{n+1} + 2x_n = 0.$$

- a. Show that the general solution to the equation is of the form

$$x_n = A_1 + 2^n A_2,$$

where A_1 and A_2 are any constants.

- b. Suppose that x_0 and x_1 are given. Then A_1 and A_2 must satisfy the system of equations

$$A_1 + A_2 = x_0,$$

$$A_1 + 2A_2 = x_1.$$

- c. From the general solution, solve for the specific solution with initial conditions $x_0 = 10$ and $x_1 = 20$.
5. Find by calculating recursively the solution of the second-order difference equation $x_{n+2} = rx_n$, $x_0 = 1$, $x_1 = -1$.
 6. Find the general form of the solution of the difference equation

$$x_{n+1} = c - x_n$$

with c arbitrary for an arbitrary initial value $x_0 = a$.

7. Consider the model

$$x_{n+1} = rx_n \left(1 - \frac{x_n}{K}\right), \quad r > 0.$$

- a. Show that $x_{n+1} < 0$ if and only if $x_n > K$.
 - b. Show that $x_{n+1} > K$ is possible with $0 < x_n < K$ only for $r > 4$.
 - c. What conditions on x_0 are necessary and sufficient to guarantee $x_n > 0$ for $n = 1, 2, 3, \dots$?
8. Find the general form of the solution of the difference equation

$$x_{n+1} = 1 - x_n$$

for an arbitrary initial value $x_0 = a$.

- 9*. The solution of the difference equation $x_{n+2} = x_n + x_{n+1}$, $x_0 = 0$, $x_1 = 1$ is called the *Fibonacci sequence* (originally formulated by Leonardo Fibonacci

(1202) to describe the number of pairs of rabbits under the hypothesis that each pair of rabbits reproduces only at age one month and age two months and produces exactly one pair of offspring on each of these two occasions, with all rabbits living exactly two months).

- a. Calculate the first eight terms of the Fibonacci sequence.
- b. Suppose it can be shown that the ratio of successive terms x_{n+1}/x_n of the Fibonacci sequence tends to a limit τ as $n \rightarrow \infty$. Show that $\tau = \frac{1}{\tau} + 1$.
- c. Deduce that $\tau = \frac{1+\sqrt{5}}{2}$.

10*. For a general, not necessarily linear, first-order difference equation

$$x_{n+1} = f(x_n),$$

show that if a solution $\{x_n\}$ approaches a limit x_∞ as $n \rightarrow \infty$, then the limit x_∞ must satisfy the equation

$$x_\infty = f(x_\infty).$$

2.2 Graphical Solution of Difference Equations

There is a way of solving difference equations graphically, called the *cobwebbing method*, which we illustrate for the simple linear homogeneous example $x_{n+1} = rx_n$. We begin by drawing the *reproduction curve* $y = rx$ in the (x, y) -plane. Then we mark x_0 , go vertically to the reproduction curve, and from there horizontally to the line $y = x$ at the point (x_1, x_1) . Then we go vertically to the reproduction curve and from there horizontally to the line $y = x$ at the point (x_2, x_2) , and so on. There are four separate cases: $r > 1$, $0 < r < 1$, $-1 < r < 0$, and $r < -1$, corresponding to different relative positions of the reproduction curve $y = rx$ and the line $y = x$. In each case, the graphical solution illustrates the behavior already obtained analytically (Figure 2.1).

The cobwebbing method can be applied to any difference equation of the form $x_{n+1} = f(x_n)$ using the reproduction curve $y = f(x)$ and the line $y = x$; it gives information about the behavior of solutions. This is particularly useful for difference equations whose analytic solution is complicated. We give two more illustrative examples.

Example 1. (Verhulst equation) For the equation

$$x_{n+1} = \frac{rx_n}{x_n + A},$$

the reproduction curve is $y = rx/(x + A)$. Its slope is given by $dy/dx = rA/(x + A)^2$, which has the value r/A at $x = 0$. This means that we must distinguish the cases $r < A$, for which the line $y = x$ lies below the reproduction curve, and $r > A$, for which the line $y = x$ intersects the reproduction curve (Figure 2.2).

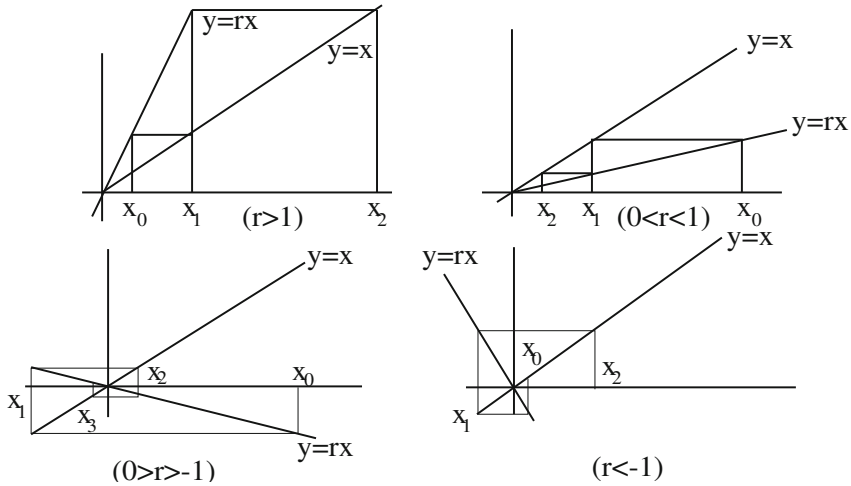


Fig. 2.1 Cobwebbing for a linear difference equation.

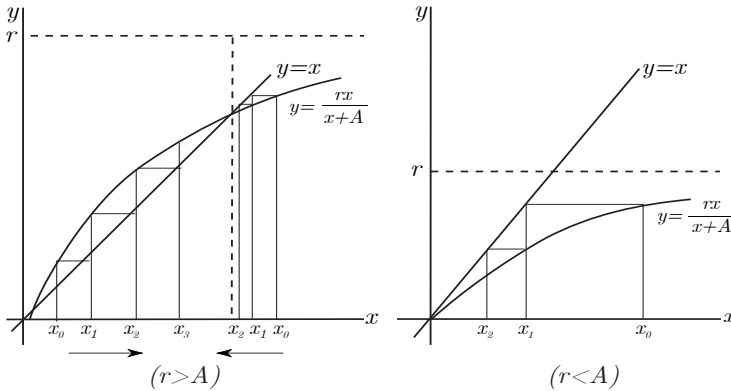


Fig. 2.2 Cobwebbing for the Verhulst equation

If $r > A$, every solution, regardless of the initial value x_0 , tends to the limit $x_\infty = r - A$ where the line $y = x$ and the reproduction curve $y = rx/(x + A)$ intersect. If $r < A$, every solution tends to the limit zero.

Example 2. For the equation

$$x_{n+1} = \frac{rx_n^2}{x_n^2 + A},$$

the reproduction curve is $y = rx^2/(x^2 + A)$, which intersects the line $y = x$ at $x = 0$ and at $x = (r \pm \sqrt{r^2 - 4A})/2$. Thus for $r > 2\sqrt{A}$ there are three real intersections, and for $r < 2\sqrt{A}$ the only real intersection is at $x = 0$ (Figure 2.3).

If $r > 2\sqrt{A}$, all solutions with $x_0 < (r - \sqrt{r^2 - 4A})/2$ tend to zero, and solutions with $x_0 > (r - \sqrt{r^2 - 4A})/2$ tend to the limit $x_\infty = (r + \sqrt{r^2 - 4A})/2$. If $r < 2\sqrt{A}$,

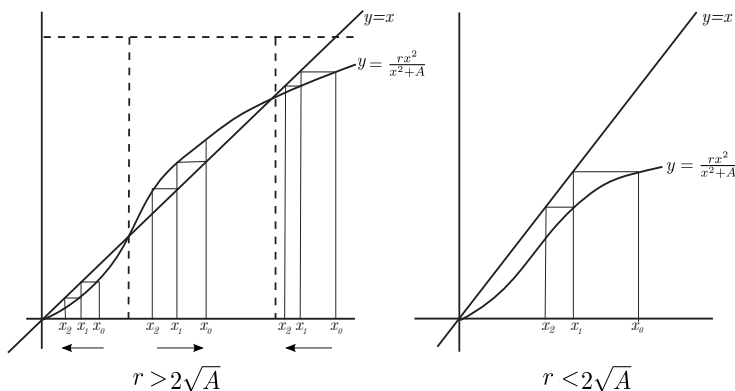


Fig. 2.3 Cobwebbing for model with Allee effect.

all solutions tend to the limit zero. This model attempts to describe populations that collapse if their initial size is too small but survive if their initial size is large enough. This is analogous to the critical depensation model, or Allee effect, described for continuous population models in Section 1.4.

Exercises

1. Use the cobwebbing method to sketch the first few terms of the solution of

$$x_{n+1} = x_n + x_n(1 - x_n), \quad x_0 = \frac{1}{2}.$$

2. Use the cobwebbing method to sketch the first few terms of the solution of

$$x_{n+1} = x_n + 2.2x_n(1 - x_n), \quad x_0 = \frac{1}{2}.$$

3. Use the cobwebbing method to sketch the first few terms of the solution of

$$x_{n+1} = x_n + 3x_n(1 - x_n), \quad x_0 = \frac{1}{2}.$$

4. Consider the difference equation

$$x_{n+1} = \begin{cases} x_n^{1/2} & \text{for } x_n > \frac{1}{4}, \\ 2x_n & \text{for } x_n < \frac{1}{4}. \end{cases}$$

Sketch the solutions for several different choices of x_0 between zero and one.

5. [Kaplan & Glass (1995)] Assume that the density of flies in a swamp is described by the equation

$$x_{n+1} = Rx_n - \frac{R}{2000}x_n^2.$$

Consider three values of R , where one value of R comes from each of the following ranges:

- (1) $1 \leq R < 3$,
- (2) $3 \leq R < 3.449$,
- (3) $3.57 \leq R < 4$.

For each value of R graph x_{n+1} as a function of x_n . Using the cobwebbing method, follow x_n for several generations. Describe the qualitative behavior found for $R = 2$.

2.3 Equilibrium Analysis

In the examples of the preceding section we observed a tendency for solutions to approach a limit that is the x -coordinate of an intersection of the reproduction curve and the line $y = x$. Such a value of x is a constant solution of the difference equation. This motivates the following definition of *equilibrium* of a difference equation:

$$x_{n+1} = f(x_n). \quad (2.1)$$

Definition 2.1. An *equilibrium* of a difference equation (2.1) is a value x_∞ such that $x_\infty = f(x_\infty)$, so that $x_n = x_\infty$ ($n = 0, 1, 2, \dots$) is a constant solution of the difference equation.

In order to describe the behavior of solutions near an equilibrium, we introduce the process of *linearization* just as we did in Section 1.4 for first-order differential equations. If x_∞ is an equilibrium of the difference equation $x_{n+1} = f(x_n)$, so that $x_\infty = f(x_\infty)$, we make the change of variable $u_n = x_n - x_\infty$ ($n = 0, 1, 2, \dots$). Thus u_n represents deviation from the equilibrium value. Substitution gives

$$x_\infty + u_{n+1} = f(x_\infty + u_n),$$

and application of Taylor's theorem gives

$$x_\infty + u_{n+1} = f(x_\infty + u_n) = f(x_\infty) + f'(x_\infty)u_n + \frac{f''(c_n)}{2!}u_n^2$$

for some c_n between x_∞ and $x_\infty + u_n$. We write $h(u_n) = f''(c_n)u_n^2/2!$ and use the relation $x_\infty = f(x_\infty)$ to form the difference equation equivalent to the original difference equation (2.1),

$$u_{n+1} = f'(x_\infty)u_n + h(u_n). \quad (2.2)$$

The function $h(u)$ is small for u , small in the sense that $|h(u)/u| \rightarrow 0$ as $|u| \rightarrow 0$; more precisely, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|h(u)| < \varepsilon|u|$ whenever $|u| < \delta$. The *linearization* of the difference equation $x_{n+1} = f(x_n)$ at the equilibrium

x_∞ is defined to be the linear homogeneous difference equation

$$v_{n+1} = f'(x_\infty)v_n, \quad (2.3)$$

obtained by neglecting the higher-order term $h(u_n)$ in (2.2). The importance of the linearization lies in the fact that the behavior of its solutions describes the behavior of solutions of the original equation (2.1) near the equilibrium. The behavior of solutions of the linearization has been described completely in Section 2.1. The following result explains the significance of the linearization at an equilibrium.

Theorem 2.1. *If all solutions of the linearization (2.3) at an equilibrium x_∞ tend to zero as $n \rightarrow \infty$, then all solutions of (2.1) with x_0 sufficiently close to x_∞ tend to the equilibrium x_∞ as $n \rightarrow \infty$.*

Proof. For convenience we write $\rho = |f'(x_\infty)|$. The assumption that all solutions of the linearization tend to zero is equivalent to $\rho < 1$. Now choose $\varepsilon > 0$ so that $\rho + \varepsilon < 1$. The difference equation $x_{n+1} = f(x_n)$ is equivalent to $u_{n+1} = f'(x_\infty)u_n + h(u_n)$. Then

$$|u_{n+1}| \leq |f'(x_\infty)||u_n| + |h(u_n)| < \rho|u_n| + \varepsilon|u_n|$$

provided $|u_n| < \delta$, where δ is determined by the condition that $|h(u)| < \varepsilon|u|$ for $|u| < \delta$. Thus, $|u_{n+1}| \leq (\rho + \varepsilon)|u_n|$ provided $|u_n| < \delta$. If $|u_0| < \delta$, it is easy to show by induction that $|u_{n+1}| < \delta$ for $n = 0, 1, 2, \dots$. This establishes $|u_{n+1}| \leq (\rho + \varepsilon)|u_n|$ for $n = 0, 1, 2, \dots$. Now it is easy to show, again by induction, that

$$|u_n| \leq (\rho + \varepsilon)^n |u_0|, \quad n = 0, 1, 2, \dots$$

Since $\rho + \varepsilon < 1$, it follows that $u_n \rightarrow 0$, and thus that $x_n \rightarrow x_\infty$ as $n \rightarrow \infty$.

In Section 2.1 we observed that if $|f'(x_\infty)| < 1$, then the solutions of $v_{n+1} = f'(x_\infty)v_n$ all tend to zero, and further that this approach is monotone if $0 < f'(x_\infty) < 1$ and oscillatory if $-1 < f'(x_\infty) < 0$. It is possible to refine Theorem 2.1 to show that the approach to an equilibrium x_∞ of $x_{n+1} = f(x_n)$ is monotone if $0 < f'(x_\infty) < 1$ and oscillatory if $-1 < f'(x_\infty) < 0$. That this is true is suggested by the cobwebbing method (Figure 2.4).

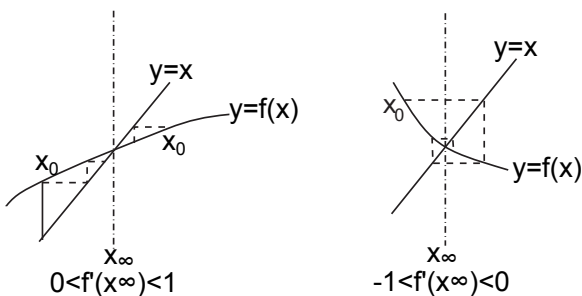


Fig. 2.4 Stability of equilibrium.

The content of Theorem 2.1 is that an equilibrium x_∞ with $|f'(x_\infty)| < 1$ has the property that every solution with x_0 close enough to x_∞ remains close to x_∞ and tends to x_∞ as $n \rightarrow \infty$. This property is called *asymptotic stability* of the equilibrium x_∞ . The condition $f'(x_\infty) < 1$ means that the curve $y = f(x)$ crosses the line $y = x$ from above to below as x increases, while the condition $f'(x_\infty) > -1$ means that the curve $y = f(x)$ cannot be too steep at the crossing. If $|f'(x_\infty)| > 1$, it is not difficult to show that except for the constant solution $x_n = x_\infty$ ($n = 0, 1, 2, \dots$), solutions cannot remain close to x_∞ . This property is called *instability* of the equilibrium x_∞ . An unstable equilibrium has no biological significance, since any deviation, however small, is enough to force solutions away.

We emphasize that Theorem 2.1 applies to solutions whose initial value x_0 is close enough to the equilibrium x_∞ . This is because the nonlinear term $h(u_n)$ in the difference equation $u_{n+1} = f'(x_\infty)u_n + h(u_n)$ is small enough to have an almost negligible effect on the solution only near the equilibrium x_∞ . Theorem 2.1 gives no explicit method of computing how close to x_∞ is close enough for the solution with a given initial value to tend to x_∞ . Often this can be seen in practice using the cobwebbing method of constructing solutions graphically, as we have shown in Section 2.2. Proofs of the theorems in this section may be found in such books as [Elaydi (1996)] and [Sandefur(1990)].

Example 1. For the *logistic* difference equation

$$x_{n+1} = x_n + rx_n \left(1 - \frac{x_n}{K}\right)$$

with $f(x) = (1+r)x - rx^2/K$ and $f'(x) = (1+r) - 2rx/K$, it is easy to find equilibria by solving the quadratic equation $x = x + rx(1 - x/K)$ and obtaining the roots $x = 0$ and $x = K$. Since $f'(0) = 1 + r$, the equilibrium $x = 0$ is asymptotically stable if $-1 < 1 + r < 1$, or $-2 < r < 0$. Since $r > 0$ in applications, this means that the equilibrium $x = 0$ is unstable. Since $f'(K) = 1 - r$, the equilibrium $x = K$ is asymptotically stable if $0 < r < 2$. It is not difficult to show that for $0 < r < 2$, every solution tends to the equilibrium K . If $r > 2$, the equilibrium $x = K$ is unstable and there is no asymptotically stable equilibrium to which solutions can tend. In the following section, we shall explore the behavior of solutions if $r > 2$ in more detail.

The logistic difference equation is sometimes presented in the form

$$x_{n+1} = rx_n \left(1 - \frac{x_n}{K}\right).$$

The study of the equation in this form is quite similar to the previous discussion; there is an equilibrium at $x = 0$ that is asymptotically stable if $r < 1$, in which case every solution tends to zero, and an equilibrium at $x = K(1 - 1/r)$ that is asymptotically stable if $1 < r < 3$, in which case every solution tends to $K(1 - 1/r)$, and if $r > 3$ there is no asymptotically stable equilibrium.

Example 2. For the Verhulst equation

$$x_{n+1} = \frac{rx_n}{x_n + A},$$

we have $f(x) = rx/(x+A)$; $f'(x) = rA/(x+A)^2$. The solution of $x = rx/(x+A)$ gives two roots, $x = 0$ and $x = r - A$. Thus, if $r < A$ the only equilibrium corresponding to a nonnegative population size is $x = 0$. Since $f'(0) = r/A < 1$, this equilibrium is asymptotically stable and every solution tends to zero. If $r > A$ there are two equilibria, $x = 0$, and $x = x_\infty = r - A$. Since $f'(0) = r/A > 1$, the equilibrium at $x = 0$ is unstable. Since $f'(x_\infty) = A/r < 1$, the equilibrium x_∞ is asymptotically stable. We have seen in Section 2.2 (graphically) that in fact every solution approaches x_∞ , that is, that the equilibrium x_∞ is globally asymptotically stable.

Exercises

In Exercises 1 through 9 find each equilibrium of the given difference equation and determine whether it is asymptotically stable or unstable.

1. $x_{n+1} = \frac{rx_n^2}{x_n^2 + A}$ (r and A are nonnegative).
2. $x_{n+1} = x_n e^{r(1-x_n/K)}$.
3. $x_{n+1} = rx_n(1 + \alpha x_n)^{-\beta}$.
4. $x_{n+1} = \begin{cases} x_n^{1/2} & \text{for } x_n > \frac{1}{4}, \\ 2x_n & \text{for } x_n < \frac{1}{4}. \end{cases}$
5. $x_{n+1} = \frac{2x_n}{1+x_n}$.
6. $x_{n+1} = x_n \log x_n$.
7. $x_{n+1} = \left(\frac{1}{\alpha x_n^b}\right)(\lambda x_n)$.
8. $x_{n+1} = x_n \exp(r(1 - \frac{x_n}{K}))$.
9. $x_{n+1} = \frac{\lambda x_n}{(1 + \alpha x_n)^b}$.
10. a. A population is governed by the difference equation

$$x_{n+1} = x_n e^{3-x_n}.$$

Show that all equilibria are unstable.

- b. The population of part (a) is to be stabilized by removing a fraction p ($0 < p < 1$) of the population in each time period after all births and deaths have taken place, to give the model

$$x_{n+1} = (1-p)x_n e^{3-(1-p)x_n}.$$

For what values of p does the population have an asymptotically stable positive equilibrium?

11. a. In the Fibonacci equation (see Exercise 9, Section 2.1) $x_{n+2} = x_n + x_{n+1}$, make the change of variable $u_n = x_{n+1}/x_n$ and obtain the transformed difference equation $u_{n+1} = 1 + 1/u_n$.
 b. Find all equilibria of the transformed difference equation of part (a) and determine which are asymptotically stable.
12. a. Find the nonnegative equilibria of a population governed by

$$x_{n+1} = \frac{3x_n^2}{x_n^2 + 2}$$

and check for stability.

- b. Suppose a fraction a is removed from the population in each generation, so that the model becomes

$$x_{n+1} = \frac{3x_n^2}{x_n^2 + 2} - ax_n.$$

For what values of a is there a stable equilibrium only at $x = 0$?

13. [Kaplan & Glass (1995)] The following equation plays a role in the analysis of nonlinear models of gene and neural networks:

$$x_{n+1} = \frac{\alpha x_n}{1 + \beta x_n},$$

where α and β are positive numbers and $x_n \geq 0$.

- a. Algebraically determine the fixed points. For each fixed point give the range of α and β for which it exists, indicate whether the fixed point is stable or unstable, and state whether the dynamics in the neighborhood of the fixed point are monotonic or oscillatory. For parts (b) and (c) assume $\alpha = \beta = 1$.
 - b. Sketch the graph of x_{n+1} as a function of x_n . Graphically iterate the equation starting from the initial condition $x_0 = 10$. What happens as the number of iterates approaches ∞ ?
 - c. Algebraically determine x_{n+2} as a function of x_n , and x_{n+3} as a function of x_n . Based on these computations what is the algebraic expression for x_{n+k} as a function of x_n ? What is the behavior of x_{n+k} as $k \rightarrow \infty$? This should agree with what you found in part (b).
14. Consider the following pair of difference equations:

$$\begin{aligned} x_{n+1} &= f(n, x_n), \\ y_{n+1} &= g(n, y_n), \end{aligned}$$

where f and g are nonnegative functions defined on $[0, \infty)$. Assume that $f(n, x_n) \leq g(n, x_n)$ for each nonnegative integer n and each nonnegative x_n and $g(n, y(n))$ is nondecreasing with respect to the second argument y_n . Prove that if $\{x_n\}_{n \geq 0}$ is a solution of the first equation and $\{y_n\}_{n \geq 0}$ is a solution of the

second equation with $x_0 \leq y_0$, then $x_n \leq y_n$ for all $n = 0, 1, 2, 3, \dots$.

15. Consider the single-species discrete-time population model

$$x_{n+1} = x_n e^{\beta \frac{1-x_n}{1+x_n}}$$

where $x_n \geq 0$ is the nonnegative population density in generation n , and the positive constant β is greater than 4. Let $\{x_1, x_2\}$ be a 2-cycle of this model, where $x_1 > 0$, $x_2 > 0$ and $x_1 \neq x_2$. Decide whether the 2-cycle $\{x_1, x_2\}$ is asymptotically stable. Explain.

2.4 Period-Doubling and Chaotic Behavior

For the logistic difference equation

$$x_{n+1} = x_n + rx_n \left(1 - \frac{x_n}{K}\right),$$

we have seen that the equilibrium $x_\infty = K$ is asymptotically stable for $0 < r < 2$. How do solutions behave if $r > 2$? We may think of r as a parameter that may be varied, and as r passes through the value 2, there must be a fundamental change in the behavior of solutions. While there is an equilibrium of K for all r , every solution tends to this equilibrium if $0 < r < 2$, but no solution other than the constant solution $x_n = K$ ($n = 0, 1, 2, \dots$) tends to this equilibrium if $r > 2$. What happens when r increases past 2 is that a solution of period 2 appears. By this we mean that there are two values, x^+ and x^- , with $f(x^+) = x^-$, $f(x^-) = x^+$ such that the alternating sequence x^+, x^-, x^+, \dots is a solution of the difference equation.

To establish the existence of this periodic solution, we take

$$f(x) = x + rx \left(1 - \frac{x}{K}\right) = (1+r)x - \frac{r}{K}x^2$$

and define

$$\begin{aligned} f_2(x) &= f(f(x)) = (1+r)f(x) - \frac{r}{K}(f(x))^2 \\ &= (1+r)^2x - \frac{r(1+r)}{K}x^2 - \frac{r}{K}\left((1+r)x - \frac{r}{K}x^2\right)^2 \\ &= (1+r)^2x - \frac{r(1+r)(2+r)}{K}x^2 + \frac{2r^2}{K^2}(1+r)x^3 - \frac{r^3}{K^3}x^4. \end{aligned}$$

We now look for equilibria of the second-order difference equation

$$x_{n+2} = f_2(x_n).$$

Such equilibria give solutions of period 2 for the original difference equation $x_{n+1} = f(x_n)$. These equilibria are solutions of the fourth-degree polynomial equation

$$x = (1+r)^2x - \frac{r(r+1)(r+2)}{K}x^2 + \frac{2r^2(1+r)}{K^2}x^3 - \frac{r^3}{K^3}x^4,$$

giving

$$x \left(r^3 \left(\frac{x}{K} \right)^3 - 2r^2(1+r) \left(\frac{x}{K} \right)^2 + r(r+1)(r+2) \left(\frac{x}{K} \right) - r(r+2) \right) = 0$$

or

$$x \left(\left(\frac{x}{K} \right) - 1 \right) \left(r^2 \left(\frac{x}{K} \right)^2 - r(r+2) \left(\frac{x}{K} \right) + (r+2) \right) = 0.$$

There are four roots, namely $x = 0$, $x = K$, and the roots x_+ , x_- of the quadratic equation $r^2(x/K)^2 - r(r+2)(x/K) + (r+2) = 0$. Thus

$$x_+ = \frac{(r+2) + \sqrt{r^2-4}}{2r}K, \quad x_- = \frac{(r+2) - \sqrt{r^2-4}}{2r}K,$$

and these roots are real if $r \geq 2$. We also have

$$\begin{aligned} f(x_+) &= (1+r)x_+ - \frac{r}{K}x_+^2 \\ &= (r+1)\frac{r+2}{2r}K + (r+1)\frac{\sqrt{r^2-4}}{2r}K \\ &\quad - \frac{r}{K}\frac{K^2}{4r^2} \left((r+2)^2 + r^2 - 4 + 2(r+2)\sqrt{r^2-4} \right) \\ \frac{2r}{K}f(x_+) &= (r+1)(r+2) + (r+1)\sqrt{r^2-4} \\ &\quad - \frac{1}{2} \left(r^2 + 4r + 4 + r^2 - 4 + 2(r+2)\sqrt{r^2-4} \right) \\ &= (r+2) - \sqrt{r^2-4} = \frac{2r}{K}x_-. \end{aligned}$$

Thus $f(x_+) = x_-$, and since $f_2(x_+) = f(f(x_+)) = x_+$, we have $f(x_-) = x_+$. We have now shown that if $r > 2$, there is a periodic solution of period 2 of $x_{n+1} = f(x_n)$ given by $x_n = x_+$ (if n is odd), $x_n = x_-$ (if n is even).

In order to test the stability of this periodic solution, we must compute $f_2'(x_+)$, which may be done by starting with

$$\begin{aligned} f_2(x) - x &= -rx \left(\left(\frac{x}{K} \right) - 1 \right) \left(r^2 \left(\frac{x}{K} \right)^2 - r(r+2) \left(\frac{x}{K} \right) + (r+2) \right) \\ &= r \left(x - \frac{x^2}{K} \right) \left(r^2 \left(\frac{x}{K} \right)^2 - r(r+2) \left(\frac{x}{K} \right) + (r+2) \right). \end{aligned}$$

Differentiation (using the product rule) gives

$$\begin{aligned} f_2'(x) - 1 &= r \left(1 - \frac{2x}{K} \right) \left(r^2 \left(\frac{x}{K} \right)^2 - r(r+2) \left(\frac{x}{K} \right) + (r+2) \right) \\ &\quad + r \left(x - \frac{x^2}{K} \right) \left(2r \frac{x}{K^2} - \frac{r(r+2)}{K} \right). \end{aligned}$$

Since $r^2(x_+/K)^2 - r(r+2)(x_+/K)r(r+2) = 0$, we have

$$\begin{aligned} f_2'(x_+) - 1 &= r \left(x_+ - \frac{x_+^2}{K} \right) \left(2r \frac{x_+}{K^2} - \frac{r(r+2)}{K} \right) \\ &= \frac{r(r+2) + r\sqrt{r^2-4}}{2} \left(1 - \frac{(r+2) + \sqrt{r^2-4}}{2r} \right) \sqrt{r^2-4} \\ &= \frac{1}{4} \left((r+2) + \sqrt{r^2-4} \right) \left((r-2) - \sqrt{r^2-4} \right) \sqrt{r^2-4} \\ &= 4 - r^2. \end{aligned}$$

We now have $f_2'(x_+) = 5 - r^2$. If we accept the theorem that a constant solution $x_n = \bar{x}$ ($n = 1, 2, \dots$) of the second-order difference equation $x_{n+2} = f_2(x_n)$ is asymptotically stable if $|f_2'(\bar{x})| < 1$, a theorem analogous to the one established in Section 2.3 for first-order difference equations (which will be described further in Exercises 2 and 3 below), then we see that this periodic solution is asymptotically stable if $-1 < 5 - r^2 < 1$, or $2 < r < \sqrt{6} = 2.449$. Thus, if $2 < r < 2.449$, there is a solution of period 2 to which every solution of $x_{n+1} = f(x_n)$ tends.

For $r > \sqrt{6}$, the solution of period 2 is unstable, but it can be shown that a solution of period 4 appears and that this solution is asymptotically stable if $\sqrt{6} < r < 2.544$. When it becomes unstable, a solution of period 8 appears, which is asymptotically stable for $2.544 < r < 2.564$. This *period-doubling* phenomenon continues until $r = 2.570$, when periodic solutions whose periods are not powers of 2 begin to appear, but these solutions are unstable. In addition, for many values of $r > 2.570$ solutions are *aperiodic*, that is, they never settle down to either an equilibrium or a periodic orbit [Strogatz (1994)]. It is possible to show analytically that a solution of period 3 appears when $r = \sqrt{8} = 2.828$ [Saha and Strogatz (1995)]. For $r > \sqrt{8}$ there is a periodic solution of period k for every integer k , but different initial values give different solutions. There are also solutions whose behavior is apparently random; such solutions are called *chaotic* (see Figure 2.5, a bifurcation diagram generated by a program in (the virtual) Appendix C). The existence of a solution of period 3 implies chaotic behavior [Li and Yorke (1975)].

These facts, whose proofs require a close examination of the properties of continuous functions and fixed points of iterates of continuous functions, are not restricted to the logistic difference equation. It is a remarkably robust fact that for every difference equation $x_{n+1} = rf(x_n)$ with $f(x)$ a function increasing to a unique maximum and then decreasing, the period-doubling phenomenon and the onset of chaos occur. In fact, if r_n is the value of r for which the asymptotically stable solution of period

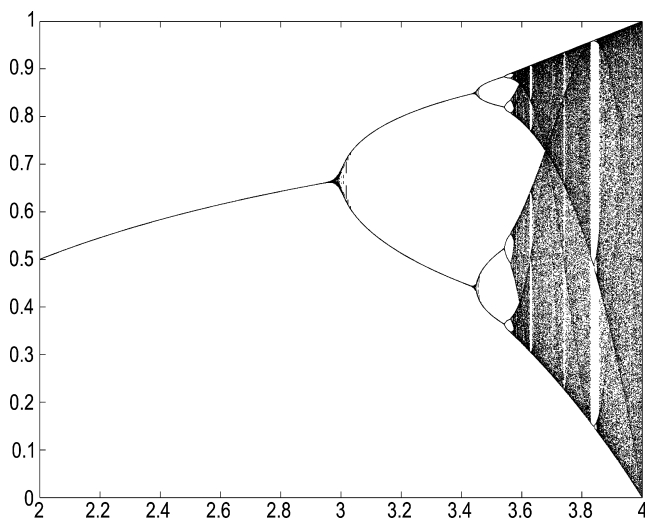


Fig. 2.5 Bifurcation diagram

2^n appears, then

$$\lim_{n \rightarrow \infty} \frac{r_{n+1} - r_n}{r_{n+2} - r_{n+1}} = 4.6692016\dots,$$

the *Feigenbaum constant*. For the logistic equation, $r_1 = 2.000$, $r_2 = 2.449$, $r_3 = 2.544$, and $(r_2 - r_1)/(r_3 - r_2) = 4.73$; usually the limiting value is approached very rapidly. This says that the period-doubling values of r occur closer and closer together [Feigenbaum (1980)].

From a biological point of view, these results are also remarkable. One interpretation is that even very simple models can produce apparently unpredictable behavior and this suggests the possibility that the governing laws may be relatively simple and therefore discoverable [May (1976)]. There do appear to be experimental observations supporting the possibility of chaotic behavior [Gurney, Blythe, and Nisbet (1980)]. On the other hand, the fact that such simple models lead to unpredictable results suggests that experimental results and observations may not be repeatable. This suggests that one should focus on the range of values of r in which the behavior is predictable and in the chaotic ranges look for properties of solutions, such as upper and lower bounds, that are verifiable.

For models of the form $x_{n+1} = rf(x_n)$ with $f(x)$ a bounded monotone increasing functions, such as the Verhulst equation

$$x_{n+1} = \frac{rx_n}{x_n + A},$$

it is easy to verify that since $rf(x)$ is bounded while x is not, there is a largest equilibrium x_∞ at which $y = rf(x)$ crosses the line $y = x$ from above to below. This implies $0 < rf'(x_\infty) < 1$, and shows that the equilibrium x_∞ is asymptotically stable; in fact solutions approach x_∞ monotonically. Thus, there is no possibility of period-

doubling or chaotic behavior, or even of stable oscillations. This fact is also robust in that it is valid for all bounded increasing recruitment functions $f(x)$. The biological significance of the difference between recruitment functions that are monotone increasing and recruitment functions that rise to a maximum and then fall involves the nature of the intraspecies competition for resources. Recruitment functions with a maximum correspond to *scramble* competition, in which resources are divided among all members and excessive population sizes reduce the survival rate, while monotone recruitment functions correspond to *contest* competition, in which some members obtain enough resources for survival, while others do not and die as a result. We now have a legitimate example of a biological assumption leading to qualitative predictions of behavior that might be experimentally observable.

Exercises

1. For what value of r does a solution of period 2 appear for the difference equation

$$x_{n+1} = rx_n e^{1-x_n}?$$

2. Let $\{x_+, x_-\}$ be a solution of period 2 of the difference equation

$$x_{n+1} = f(x_n).$$

Show that both x_+ and x_- are equilibria of the second-order difference equation

$$x_{n+2} = f(f(x_n)).$$

3. Define a new index $k = n/2$ for n even and the iterated function

$$f_2(x) = f(f(x)).$$

Show that x_+ and x_- from Exercise 2 are equilibria of the first order difference equation

$$x_{k+1} = f_2(x_k).$$

[*Remark:* Exercise 3 together with Theorem 2.1 of Section 2.3 shows that an equilibrium x^* of the second order equation $x_{n+2} = f_2(x_n)$ is asymptotically stable if $|f_2'(x^*)| < 1$. Exercise 8 below gives another stability criterion for the asymptotic stability of a solution of period 2 of the difference equation $x_{n+1} = f(x_n)$.]

4. [Kaplan & Glass(1995)] Consider an ecological system described by the finite difference equation

$$x_{n+1} = Cx_n^2(2 - x_n), \quad \text{for } 0 \leq x_n \leq 2,$$

where x_n is the population density in year n and C is a positive constant that we assume is equal to $25/16$.

- a. Sketch the graph of the right hand side of this equation. Indicate the maxima, minima, and inflection points.
 - b. Determine the fixed points of this system.
 - c. Determine the stability at each fixed point and describe the dynamics in a neighborhood of the fixed points.
 - d. In a brief sentence or two describe the expected dynamics starting from initial values of $x_0 = 1/3$ and also $x_0 = 1$ in the limit as $n \rightarrow \infty$. In particular, comment on the possibility that the population may go to extinction or to chaotic dynamics in the limit $n \rightarrow \infty$.
5. [Kaplan and Glass(1995)] The following finite difference equation has been considered as a mathematical model for a periodically stimulated biological oscillator [Bélair and Glass (1983)].

$$\phi_{n+1} = \begin{cases} 6\phi_n - 12\phi_n^2 & \text{for } 0 \leq \phi_n < 0.5, \\ 12\phi_n^2 - 18\phi_n + 7 & \text{for } 0.5 \leq \phi_n \leq 1. \end{cases}$$

- a. Sketch ϕ_{n+1} as a function of ϕ_n for $0 \leq \phi_n \leq 1$. Be sure to show all maxima and minima and compute the values of ϕ_{n+1} at these extreme points.
 - b. Compute all fixed points. What are the qualitative dynamics in a neighborhood of each fixed point?
 - c. If you have done part (a) right, you should be able to find a cycle of period 2. What is this cycle? Show it on your sketch.
6. For the logistic difference equation $x_{n+1} = x_n + rx_n(1 - x_n/K)$ with $r > 2$, show that

$$0 < x_- < x_\infty < x_+ < K.$$

- 7*. a. Let $\{x_+, x_-\}$ be a solution of period 2 of the difference equation $x_{n+1} = f(x_n)$. Use the chain rule of calculus to show that if $f_2(x) = f(f(x))$, then

$$f'_2(x_+) = \left. \frac{d}{dx} f(f(x)) \right|_{x=x_+} = f'(x_-)f'(x_+).$$

- b. Deduce from part (a) that the solution of period 2 is asymptotically stable if

$$|f'(x_-)| \cdot |f'(x_+)| < 1.$$

8. [Kaplan and Glass (1995)] The finite difference equation

$$x_{n+1} = 0.5 + \alpha \sin(2\pi x_n), \quad 0 \leq x < 1,$$

where $0 \leq \alpha < 0.5$, has been used as a mathematical model for periodic stimulation of biological oscillators .

- a. There is one steady state. Determine this steady state and its stability as a function of α .
 - b. Sketch x_{n+1} as a function of x_n for $\alpha = 0.25$. Be sure to indicate all maxima, minima, and inflection points.
 - c. For $\alpha = 0.25$ there is a stable orbit of period 2. What is it?
- 9*. For what value of r does a solution of period 2 appear for the difference equation

$$x_{n+1} = rx_n e^{1-x_n}?$$

Hint: Let $f(x)$ be the right-hand side function, i.e., $f(x) = rxe^{1-x}$. Find the condition for r under which $f(f(x))$ has fixed points. If you find that the condition is $r > e$, you will find the solution.

10. The population of a species is described by the finite difference equation

$$x_{n+1} = ax_n \exp(-x_n) \text{ for } x_n \geq 0,$$

where a is a positive constant.

- a. Determine the fixed points.
 - b. Evaluate the stability of the fixed points.
 - c. For what value of a is there a period-doubling bifurcation (using the conclusion of the previous exercise).
 - d. For what values of a will the population become extinct starting from any initial condition?
11. The objective of this problem is to get you to read and think about some of the work on difference-equation models in population biology. Read [May (1976)]. Write a summary that deals with critical ideas, methods, and presentation in that article. The questions you might wish to answer are these:
- a. What is the main focus of this article? Is a particular question being addressed?
 - b. Do the mathematical models help illuminate the topics? If so, in what ways?
 - c. Are there alternative methods or approaches that might have been suitable for answering the questions the author addressed?

2.5 Discrete-Time Metered Models

In many populations there is a recruitment cycle in which the population size at each stage is a function of the population size at the previous stage, but the form of this function is determined by a continuous birth and death process. In this case the population size is given by a difference equation

$$x_{n+1} = f(x_n)$$

describing what may be called the long-term dynamics of the model. The function $f(x)$ is constructed from assumptions on births and deaths occurring continuously in the intervals between stages and incorporates the short-term dynamics of the model. Such models are called *metered models*. As difference equations they may be analyzed by the methods of this chapter. What is new in this section is the use of models of the type considered in Chapter 1 to establish specific forms for the reproductive curve $y = f(x)$. In many fish populations there is an annual birth process, with the number of births depending on the adult population size at the time, followed by a continuous death rate until the next birth cycle, at which time the survivors make up the adult population. Such populations lend themselves naturally to metered models.

To describe the general form of a metered model, we let x_n be the size of the adult population at the n th stage. Suppose this parent stock gives rise to B_n young and the survivors of this class at time T (often one year for fish populations) are the x_{n+1} adults at the next stage. More generally, we may assume that there are R_n surviving recruits of whom H_n are harvested with the remainder $R_n - H_n$ forming the adult population x_{n+1} at the next stage. This parent stock x_{n+1} is often called the *escapement* by fishery biologists. This description assumes that none of the adults' x_n parents survive to the next stage, but it is not difficult to relax this restriction. It also assumes that harvesting occurs just before the reproductive stage.

We shall assume constant fertility, that is, that the number of births B_n is proportional to the number of adults x_n , that is,

$$B_n = \alpha x_n.$$

We also assume that between birth times there is a per capita death rate that is a function of the number of survivors from the B_n newborn members. This means that if there are $z(t)$ survivors at time t then there is a function $\phi(z)$ (the per capita death rate) such that

$$\frac{dz}{dt} = -z\phi(z).$$

Then the recruitment R_n is the value for $t = T$ of the solution of the initial value problem

$$\frac{dz}{dt} = -z\phi(z), \quad z(0) = B_n = \alpha x_n. \quad (2.4)$$

Formally, we can solve by separation of variables, obtaining

$$\int_{\alpha x_n}^{R_n} \frac{dz}{z\phi(z)} = -T.$$

The function f in the metered model $x_{n+1} = f(x_n)$ is given implicitly by the relation

$$\int_{\alpha x_n}^{f(x_n)} \frac{dz}{z\phi(z)} = -T.$$

Under harvesting, the model is

$$x_{n+1} = R_n - H_n = f(x_n) - H_n,$$

with the same function f .

Example 1. (The Beverton and Holt stock recruitment model) In some bottom-feeding fish populations, including the North Atlantic plaice and haddock studied by Beverton and Holt (1957), recruitment appears to be essentially unaffected by fishing, and this is true over a wide range of fishing effort. These species have very high fertility rates and very low survivorship to adulthood. The Beverton and Holt model assumes a linear per capita mortality rate, so that the differential equation describing survivorship has the form

$$\frac{dz}{dt} = -z(\mu_1 + \mu_2 z),$$

with μ_1 and μ_2 positive constants. Explicit solution of the initial value problem (2.4) leads to a recruitment function of the form

$$R_n = \frac{ax_n}{1 + bx_n},$$

where a and b are positive constants related to μ_1 and μ_2 . In fact, the same form is valid if μ_1 and μ_2 are arbitrary nonnegative functions of t . This leads to the Beverton and Holt metered model

$$x_{n+1} = \frac{ax_n}{1 + bx_n}.$$

The reader should observe that this is equivalent to the Verhulst model

$$x_{n+1} = \frac{rx_n}{x_n + A}$$

described earlier with $a = r/A$, $b = 1/A$. As we have seen for the Verhulst equation, there is an asymptotically stable positive equilibrium only if $r > A$, or equivalently if $a > 1$.

Example 2. (The Ricker stock recruitment model) It was observed by Ricker (1954, 1958) that some species of fish, including salmon, habitually cannibalize their eggs and young. The Ricker model assumes a per capita death rate proportional to the initial size of the young population. Then the survivorship differential equation has the form

$$\frac{dz}{dt} = -zB_n = -\alpha x_n z, \quad z(0) = \alpha x_n.$$

This has the solution

$$z = \alpha x_n e^{-\alpha x_n t}$$

and therefore

$$R_n = \alpha x_n e^{-\alpha T x_n},$$

which we write

$$R_n = \alpha x_n e^{-\beta x_n}$$

by letting $\beta = \alpha T$. This leads to the Ricker metered model

$$x_{n+1} = \alpha x_n e^{-\beta x_n}.$$

Exercises

1. a. Show that the Ricker model $x_{n+1} = \alpha x_n e^{-\beta x_n}$ has an equilibrium $x = 0$ and a positive equilibrium $x_\infty = \log \alpha / \beta$ if $\alpha > 1$.
 b. Determine the range of values of the parameter α for which each of these equilibria is asymptotically stable.
2. In the Beverton and Holt model

$$x_{n+1} = \frac{ax_n}{1 + bx_n},$$

determine the constants a and b in terms of α and T if the survivorship differential equation is

$$\frac{dz}{dt} = -dz^2.$$

(Or $\mu_1 = 0, \mu_2 = d$.)

3. Analyze the behavior of the continuous analogue of the metered Ricker model,

$$\frac{dx}{dt} = \alpha x e^{-\beta x} - x,$$

and compare with the behavior of the metered model.

4. Analyze the behavior of the continuous analogue of the metered Beverton and Holt model

$$\frac{dx}{dt} = \frac{ax}{1 + bx} - x.$$

2.6 A Two-Age Group Model and Delayed Recruitment

Suppose we are interested in studying a population that in the n th generation contains x_n immature members and y_n adult members, with a birth rate depending on the number of adult members and a transition rate from immature to adult members depending on the number of immature members. If the birth rate is α and the rate of transition is β , we are led to a system of two difference equations,

$$\begin{aligned} x_{n+1} &= \alpha y_n \\ y_{n+1} &= \beta x_n, \end{aligned}$$

assuming no survival of adult members to the next generation. Graphical methods of solving this system are cumbersome, but the method of analytic solution is easy. Iteration gives $x_1 = \alpha y_0$, $y_1 = \beta x_0$; $x_2 = \alpha y_1 = \alpha\beta x_0$, $y_2 = \beta x_1 = \alpha\beta y_0$, $x_3 = \alpha y_2 = \alpha^2\beta y_0$, $y_3 = \beta x_2 = \alpha\beta^2 x_0$. The pattern becomes apparent if we introduce vector–matrix notation. Define the two-dimensional column vector

$$\mathbf{z}_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$$

and the *reproduction matrix*

$$A = \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}.$$

Then the system can be written

$$\mathbf{z}_{n+1} = A\mathbf{z}_n,$$

and now iterative solution gives

$$\mathbf{z}_n = A^n \mathbf{z}_0,$$

where A^n is the n th power of the matrix A . More generally, we could assume a nonlinear birth function $B(y)$ and a nonlinear mortality function $D(y)$, that is, a nonlinear system

$$\begin{aligned} x_{n+1} &= B(y_n) \\ y_{n+1} &= \alpha x_n - D(y_n). \end{aligned}$$

An equilibrium of this system is a solution (x_∞, y_∞) of the system $x_\infty = B(y_\infty)$, $y_\infty = \alpha x_\infty - D(y_\infty)$. We may linearize about the equilibrium and examine the asymptotic stability of the equilibrium by studying the linearized system

$$\begin{aligned} u_{n+1} &= B'(y_\infty)v_n, \\ v_{n+1} &= \alpha u_n - D'(y_\infty)v_n, \end{aligned}$$

with coefficient matrix

$$A = \begin{pmatrix} 0 & B'(y_\infty) \\ \alpha & -D'(y_\infty) \end{pmatrix}.$$

Such a study requires the machinery of linear algebra, which we shall not undertake here. Models with a larger number of age groups are also natural and their study leads to systems of difference equations each with dimension equal to the number of age groups. Again, the use of linear algebra is essential. In order to study a two-age group model such as

$$x_{n+1} = B(y_n), \quad y_{n+1} = \alpha x_n - D(y_n),$$

without being forced to use linear algebra, we may eliminate by substituting $B(y_{n-1})$ for x_n in the second equation. We then obtain a single *second-order* difference equation

$$y_{n+1} = \alpha B(y_{n-1}) - D(y_n),$$

using the relation $x_{n+1} = B(y_n)$ to determine x_n once this second-order equation has been solved. An equilibrium of this second order equation is a value y_∞ such that

$$y_\infty = \alpha B(y_\infty) - D(y_\infty).$$

The linearization at the equilibrium is the second-order linear homogeneous difference equation

$$u_{n+1} = \alpha B'(y_\infty)u_{n-1} - D'(y_\infty)u_n.$$

In order to study the stability of an equilibrium of a difference equation of order higher than one, we first state the following linearization theorem without proof.

Theorem 2.2. *If x_∞ is an equilibrium of the difference equation*

$$x_{n+k} = f(x_{n+k-1}, x_{n+k-2}, \dots, x_{n+1}, x_n)$$

of order k , so that

$$x_\infty = f(x_\infty, x_\infty, \dots, x_\infty),$$

the equilibrium is asymptotically stable if all solutions of the linearization at the equilibrium

$$u_{n+k} = \sum_{j=1}^k a_j u_{n+k-j}$$

(with $a_j = f_j(x_\infty, x_\infty, \dots, x_\infty)$ and f_j denoting the partial derivative with respect to the j th variable) tend to zero.

In order to determine whether all solutions of a linear difference equation tend to zero, we look for solutions of the form $x_n = \lambda^n x_0$ and obtain a *characteristic equation*—a polynomial equation of degree k —for λ . For the difference equation $u_{n+k} = \sum_{j=1}^k a_j u_{n+k-j}$, this characteristic equation is $\lambda^{n+k} = \sum_{j=1}^k a_j \lambda^{n+k-j}$, or

$$\lambda^k - \sum_{j=1}^k a_j \lambda^{k-j} = 0.$$

If the roots of this characteristic equation, say $\lambda_1, \lambda_2, \dots, \lambda_k$, are distinct, then every solution of the difference equation $u_{n+k} = \sum_{j=1}^k a_j u_{n+k-j}$ is a linear combination of $\lambda_1^n, \lambda_2^n, \dots, \lambda_k^n$. If the characteristic equation has multiple roots, then there are also terms $\lambda_j^n \log \lambda_j$, but in any case the condition that all solutions of a linear homogeneous difference equation tend to zero is that all roots λ_j of the characteristic equation satisfy $|\lambda_j| < 1$.

Combination of this information about the solutions of linear difference equations with Theorem 2.2 gives the following extension of Theorem 2.1 of Section 2.3.

Theorem 2.3. *Let x_∞ be an equilibrium of the following difference equation of order k :*

$$x_{n+k} = f(x_{n+k-1}, x_{n+k-2}, \dots, x_{n+1}, x_n).$$

If all roots of the characteristic equation

$$\lambda^k - \sum_{j=1}^k f_j(x_\infty, x_\infty, \dots, x_\infty) \lambda^{k-j} = 0$$

of the linearization at this equilibrium satisfy $|\lambda| < 1$, then the equilibrium x_∞ is asymptotically stable.

For a first-order difference equation $x_{n+1} = f(x_n)$ the characteristic equation is $\lambda - f'(x_\infty) = 0$, and thus the condition for asymptotic stability is $|f'(x_\infty)| < 1$, as given in Theorem 2.1 of Section 2.3. For the equilibrium x_+ of the second order difference equation

$$x_{n+2} = f_2(x_n)$$

considered in Section 2.4 the characteristic equation is $\lambda^2 - f'_2(x_+) = 0$, with roots $\lambda = \pm \sqrt{|f'_2(x_+)|}$ or $\lambda = \pm i \sqrt{|f'_2(x_+)|}$, depending on whether $f'_2(x_+) > 0$ or $f'_2(x_+) < 0$. In either case the condition for asymptotic stability is $|f'_2(x_+)| < 1$, a fact used without proof in Section 2.4.

The results developed in Theorem 2.3 would enable us to study the delayed recruitment model $y_{n+1} = \alpha B(y_{n-1}) - D(y_n)$ formulated at the beginning of this section. However, we shall instead consider the model

$$x_{n+1} = ax_n + F(x_{n-\tau}),$$

which is often used to study whale populations. Here x_n represents the adult breeding population, a ($0 \leq a \leq 1$) the survival coefficient, and $F(x_{n-\tau})$ the recruitment to the adult stage with a delay of τ years. Equilibrium population size is obtained by solving

$$x_\infty = ax_\infty + F(x_\infty),$$

or $F(x_\infty) = (1-a)x_\infty = Mx_\infty$, where $M = 1-a$ is the annual mortality rate. More generally, we could consider a model of the form $x_{n+1} = G(x_n) + F(x_{n-\tau})$ with equilibrium population size determined from $x_\infty = G(x_\infty) + F(x_\infty)$. To study the stability of equilibrium, we linearize about the equilibrium by setting $x_n = u_n + x_\infty$ and neglecting higher order terms, obtaining

$$u_{n+1} = au_n + F'(x_\infty)u_{n-\tau}.$$

We let $b = F'(x_\infty)$ to write this in the form

$$u_{n+1} = au_n + bu_{n-\tau}.$$

The characteristic equation is

$$\lambda^{\tau+1} - a\lambda^\tau - b = 0,$$

and asymptotic stability of equilibrium requires $|\lambda| < 1$ for all roots of this equation.

If $\tau = 0$, the characteristic equation is $\lambda - a - b = 0$ and the stability condition is $|a + b| < 1$, or $-1 - a < b < 1 - a$.

If $\tau = 1$, the characteristic equation is $\lambda^2 - a\lambda - b = 0$, which has roots

$$\lambda = a \pm \frac{\sqrt{a^2 + 4b}}{2}.$$

If $a^2 + 4b \geq 0$, these roots are real, and the condition $|\lambda| < 1$ is equivalent to $a + \sqrt{a^2 + 4b} < 2$ and $a - \sqrt{a^2 + 4b} > -2$. These conditions give $\sqrt{a^2 + 4b} < 2 - a$ and $\sqrt{a^2 + 4b} < 2 + a$. Since $2 - a \leq 2 + a$, we have the single condition $\sqrt{a^2 + 4b} < 2 - a$, or $b < 1 - a$. If $a^2 + 4b < 0$, the roots are complex and $|\lambda|^2 = a^2/4 + (-a^2 - 4b)/4 = -b$. Since $b < 0$, we must have $-1 < b < 0$. Combining the cases $a^2 + 4b < 0$ and $a^2 + 4b \geq 0$, we see that for $\tau = 1$, the equilibrium x_∞ is asymptotically stable if $-1 < b < 1 - a$.

For values of $\tau > 1$, the stability condition is more difficult to analyze, but it is possible to establish the following result [Levin and May (1976)].

Theorem 2.4. *There is a function $z_\tau(a) \leq -1 + a$ with $z_\tau(a) \nearrow -1 + a$ as $\tau \rightarrow \infty$ such that the equilibrium x_∞ is asymptotically stable if*

$$z_\tau(a) < b < 1 - a.$$

We have shown that $z_0(a) = -1 - a$, $z_1(a) = -1$. Since $z_\tau(a) < -1 + a$ for all τ , the equilibrium is certainly asymptotically stable if $-1 + a < b < 1 - a$, or $|b| < 1 - a$.

The population of the Antarctic fin whale has been studied using this model with $F(x) = rx(1 - x/K)$, $r = 0.12$, $a = 0.96$, $k = 600,000$, $\beta = 5$. The equilibrium population size is given by $rx_\infty(1 - x_\infty/K) = (1 - a)x_\infty$, or $x_\infty = K(1 - (1 - a)/r)$. If we use $M = 1 - a$, we have $x_\infty = K(1 - M/r)$. Since $F'(x) = 4 - 2rx/K$, $F'(x_\infty) = 2M - r$. The equilibrium is certainly asymptotically stable if $|2M - r| < 1 - a = M$, or $M < r < 3M$. With $K = 600,000$, $M = 0.04$, $r = 0.12$, this condition is not satisfied, since $r = 3M$. However, since $z_\tau(a)$ is actually less than $-1 + a$, the stability condition is satisfied.

Discrete single-species models do not involve merely first-order difference equations. As we have seen, age-class models lead to systems of difference equations and delayed-recruitment models lead to higher-order difference equations. For a unified treatment, we would have to show how to write a difference equation of order k as a system of k first-order difference equations, and then use vector-matrix notation and methods of linear algebra to develop the theory of equilibria and asymptotic stability.

Exercises

1. Convert the system of difference equations

$$x_{n+1} = 2y_n, \quad y_{n+1} = 3x_n$$

to a second-order difference equation and find the first three terms of the solution with $x_0 = y_0 = 1$.

2. Solve the second order difference equation

$$x_{n+2} - x_n = 0$$

with $x_0 = 1, x_1 = -1$.

3. Solve the second-order difference equation

$$x_{n+2} + x_n = 0$$

with $x_0 = 1, x_1 = -1$.

The Jury criterion states that the eigenvalues of a 2×2 matrix M have magnitude less than one if and only if $|tr(M)| < det(M) + 1 < 2$. Use it in problems 4 and 5.

4. Assume that the population (P) of a parasite and that of its host population (H) are modeled by the difference equations

$$\begin{aligned} P_{t+1} &= \alpha H_t (1 - e^{-aP_t}), \\ H_{t+1} &= \alpha H_t e^{-aP_t}, \end{aligned}$$

where α, a are positive.

- a. Calculate the equilibrium population sizes and show that they are positive only if $\alpha > 1$.
- b. Use the Jury criterion to show that if $\alpha > 1$, then the equilibrium is unstable. *Hint:* The following relation holds:

$$\frac{\alpha}{\alpha - 1} \ln \alpha > 1 \text{ for all } \alpha > 1.$$

- 5*. Determine all equilibria and the stability of each equilibrium for the system

$$\begin{aligned} x_{n+1} &= ax_n e^{-by_n}, \\ y_{n+1} &= cx_n (1 - e^{-by_n}). \end{aligned}$$

(This system is known as the Nicholson and Bailey model (1935) for a host-parasite system; x_n denotes the number of hosts and y_n the number of parasites.)

6. One of the common discrete-time models for the growth of a single species is the Pielou logistic equation

$$x_{n+1} = \frac{\alpha x_n}{1 + \beta x_n},$$

where $x_n \geq 0$ is the size of the population at generation n , $\alpha > 1$, and $\beta > 0$. If we assume that there is a delay of time period 1 in the response of growth rate per individual to density change, we obtain the delay difference equation model

$$x_{n+1} = \frac{\alpha x_n}{1 + \beta x_{n-1}}.$$

Determine the stability of all the nonnegative fixed points of this equation.

2.7 Systems of Two Difference Equations

In Section 2.6, we examined a system of two difference equations by reducing it to a single second order difference equation. In Section 2.8 we shall examine a system that cannot be reduced to a single equation of higher order. In this section, we shall outline the main results of the analysis of stability of an equilibrium of a system of two first-order difference equations.

We begin with a system of two difference equations,

$$\begin{aligned} x_{n+1} &= f(x_n, y_n), \\ y_{n+1} &= g(x_n, y_n). \end{aligned} \tag{2.5}$$

An equilibrium of the system (2.5) is a solution (x_∞, y_∞) of the system

$$f(x, y) = x, \quad g(x, y) = y.$$

Generally, $f(x, y) = x$ and $g(x, y) = y$ are represented by curves in the (x, y) -plane, and an equilibrium is an intersection of the two curves. If (x_∞, y_∞) is an equilibrium of (2.5), then the system (2.5) has a constant solution $x_n = x_\infty, y_n = y_\infty$ ($n = 1, 2, \dots$).

The description of the behavior of solutions near an equilibrium parallels the description given in Section 2.3 for a single first-order difference equation. If (x_∞, y_∞) is an equilibrium of the system (2.5), we make the change of variables $u_n = x_n - x_\infty, v_n = y_n - y_\infty$ ($n = 0, 1, 2, \dots$), so that (u_n, v_n) represents deviation from the equilibrium. We then have the system

$$\begin{aligned} u_{n+1} &= f(x_\infty + u_n, y_\infty + v_n) - x_\infty = f(x_\infty + u_n, y_\infty + v_n) - f(x_\infty, y_\infty) \\ v_{n+1} &= g(x_\infty + u_n, y_\infty + v_n) - y_\infty = g(x_\infty + u_n, y_\infty + v_n) - g(x_\infty, y_\infty). \end{aligned} \tag{2.6}$$

If we use Taylor's theorem to approximate the functions $f(x_\infty + u_n, y_\infty + v_n)$ and $g(x_\infty + u_n, y_\infty + v_n)$ by their linear terms and neglect the remainder terms, we obtain a linear system

$$\begin{aligned} u_{n+1} &= f_x(x_\infty, y_\infty)u_n + f_y(x_\infty, y_\infty)v_n, \\ v_{n+1} &= g_x(x_\infty, y_\infty)u_n + g_y(x_\infty, y_\infty)v_n. \end{aligned} \quad (2.7)$$

called *the linearization of the system (2.5) at the equilibrium (x_∞, y_∞)* , which approximates the system (2.5) near the equilibrium. The analogue of Theorem 2.1, Section 2.3, which explains the significance of the linearization at an equilibrium, is valid.

Theorem 2.5. *If all solutions of the linearization (2.7) of the system (2.5) at an equilibrium (x_∞, y_∞) tend to zero as $n \rightarrow \infty$, then all solutions of (2.7) with x_0 and y_0 sufficiently close to x_∞ and y_∞ respectively tend to the equilibrium (x_∞, y_∞) as $n \rightarrow \infty$.*

The proof is more complicated than that given in Section 2.3 for $n = 1$, and we shall omit it.

The next problem is to determine conditions under which all solutions of the linear system (2.7) approach zero. The idea behind the solution of this problem, although there are some technical complications, is to look for solutions of the form $u_n = u_0\lambda^n$, $v_n = v_0\lambda^n$ and then determine conditions under which all values of λ for which this is possible satisfy $|\lambda| < 1$. (Recall that if $|\lambda| < 1$, then $\lambda^n \rightarrow 0$ as $n \rightarrow \infty$.) The basic fact is that all solutions of the linear system (2.7) approach zero if all roots of the characteristic equation

$$\lambda^2 - \text{tr}A(x_\infty, y_\infty)\lambda + \det A(x_\infty, y_\infty) = 0$$

satisfy $|\lambda| < 1$. Here, $\text{tr}A$ and $\det A$ are the trace and determinant of the 2×2 matrix

$$A(x_\infty, y_\infty) = \begin{pmatrix} f_x(x_\infty, y_\infty) & f_y(x_\infty, y_\infty) \\ g_x(x_\infty, y_\infty) & g_y(x_\infty, y_\infty) \end{pmatrix}.$$

This characteristic equation may also be written as a determinant, namely as

$$\det(A(x_\infty, y_\infty) - \lambda I) = 0, \quad (2.8)$$

where

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

the identity matrix. It arises from the condition that

$$A(x_\infty, y_\infty) \begin{pmatrix} u \\ v \end{pmatrix} = \lambda \begin{pmatrix} u \\ v \end{pmatrix}$$

has a nontrivial solution for the vector

$$\begin{pmatrix} u \\ v \end{pmatrix}.$$

In this vector–matrix form, the stability result generalizes to systems of arbitrary order.

Theorem 2.6. *If all roots of the characteristic equation (2.8) at an equilibrium satisfy $|\lambda| < 1$, then all solutions of the system (2.5) with initial values sufficiently close to an equilibrium approach the equilibrium.*

A proof of this result may be found in books that explore the theory of difference equations, such as Elaydi (1996) and Sandefur (1990).

The characteristic equation for a system of k difference equations at an equilibrium is a polynomial equation of degree k . Conditions are known under which all roots of a polynomial equation have absolute value less than 1. These conditions were originally derived to analyze some economic models [Samuelson (1941)]. For a quadratic equation

$$f(\lambda) = \lambda^2 + a_1\lambda + a_2 = 0, \quad (2.9)$$

both roots satisfy $|\lambda| < 1$ if and only if

$$1 + a_1 + a_2 > 0, \quad 1 - a_1 + a_2 > 0, \quad 1 + a_2 > 0.$$

These three conditions can be combined and written as

$$0 < |a_1| < a_2 + 1 < 2, \quad (2.10)$$

which is the *Jury criterion* (Exercise 3, Section 2.6).

To establish the Jury criterion, we begin by noting that $f(\lambda) \rightarrow +\infty$ as $\lambda \rightarrow \infty$ and $\lambda \rightarrow -\infty$. If $f(-1) < 0$, there is a root less than -1 , and if $f(1) < 0$ there is a root greater than 1. Further, the product of the roots of (2.9) is a_2 ; thus we must have $|a_2| < 1$, $f(-1) > 0$, $f(1) > 0$ in order to have all roots of (2.9) satisfy $|\lambda| < 1$. We may rewrite these conditions as $-1 < a_2 < 1$, or $0 < a_2 + 1 < 2$, $f(-1) = 1 - a_1 + a_2 > 0$, $f(1) = 1 + a_1 + a_2 > 0$. The conditions $f(-1) > 0$ and $f(1) > 0$ may be combined to give

$$-(1 + a_2) < a_1 < 1 + a_2,$$

or $|a_1| < 1 + a_2$. Thus, in order to have the roots of (2.9) satisfy $|\lambda| < 1$, the conditions in (2.10) must be satisfied.

To prove that the conditions in (2.10) imply that the roots of (2.9) satisfy $|\lambda| < 1$, we consider first the case that the roots of (2.9) are complex conjugate. In this case, both roots have the same absolute value, and $|a_2| < 1$ implies that this absolute value is less than 1. If the roots of (2.9) are real and $f(-1) > 0$, $f(1) > 0$, then either both roots are less than -1 (contradicted by $|a_2| < 1$), or both roots are greater than 1 (contradicted by $|a_2| < 1$), or both roots are between -1 and 1. Thus the conditions in (2.10) imply that both roots satisfy $|\lambda| < 1$, and the Jury criterion is established.

In the next section, we will examine a system of three first order-difference equations. It can be shown that the conditions under which the roots of a cubic equation

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

satisfy $|\lambda| < 1$ are

$$\begin{aligned} 1 + a_1 + a_2 + a_3 &> 0, & 1 - a_1 + a_2 - a_3 &> 0, \\ 3 + a_1 - a_2 - 3a_3 &> 0, & 1 + a_1a_3 - a_2 - a_3^2 &> 0 \end{aligned} \quad (2.11)$$

[Samuelson (1941)]. We will make use of this result.

Exercises

1. Find all equilibria of the system

$$\begin{aligned} x_{n+1} &= B(y_n), \\ y_{n+1} &= \alpha x_n - D(y_n), \end{aligned}$$

treated as a single second-order difference equation in Section 2.6, and establish conditions for their stability.

2. For the delayed-recruitment model

$$x_{n+1} = ax_n + F(x_{n-r}),$$

with $0 \leq a \leq 1$, considered in Section 2.6, the characteristic equation at an equilibrium x_∞ , that is, a solution of $F(x_\infty) = (1 - a)x_\infty$, is

$$\lambda^{r+1} - a\lambda^r - F'(x_\infty) = 0.$$

Determine the conditions on a and $F'(x_\infty)$ for stability of equilibrium if $r = 2$ and write them in the form

$$z_r(a) < F'(x_\infty) < 1 - a,$$

i.e., determine the function $z_r(a)$.

3. Consider the two-dimensional system

$$\begin{aligned} x_{n+1} &= \frac{\alpha y_n}{1 + (x_n)^2}, \\ y_{n+1} &= \frac{\beta x_n}{1 + (y_n)^2}, \end{aligned}$$

where α and β are positive constants. If $\alpha^2 < 1$ and $\beta^2 < 1$, prove that the origin $(0, 0)$ is globally asymptotically stable. [Global asymptotic stability means that every solution approaches the origin, not just solutions starting close to the origin.]

4. Consider the single-species, age-structured population model

$$\begin{aligned}x_{n+1} &= y_n \exp(r - ax_n - y_n) \\ y_{n+1} &= x_n,\end{aligned}$$

where $x_n \geq 0$, $y_n \geq 0$, and the constants a , r are positive. Show that all the solutions are bounded. Interpret your result.

2.8 Oscillation in Flour Beetle Populations: A Case Study

Some recent experimental studies of flour beetles (*Tribolium castaneum*) have indicated a possibility of behavior in the laboratory that appears to be chaotic [R.F. Costantino, R.A. Desharnais, J.M. Cushing, B. Dennis, (1997), (1995)]. We shall describe and attempt to analyze a model for such behavior, taking note of the properties of the life cycle of the flour beetle.

The life cycle consists of larval and pupal stages, each lasting approximately two weeks, followed by an adult stage. Both larvae and adults are cannibalistic, consuming eggs and thus reducing larval recruitment. In addition, there is adult cannibalism of pupae. We take two weeks as the unit of time and formulate a discrete model describing the larval population L , pupal population P , and adult population A at two-week intervals.

If there were no cannibalism, we could begin with a linear model

$$\begin{aligned}L_{n+1} &= bA_n, \\ P_{n+1} &= (1 - \mu_L)L_n, \\ A_{n+1} &= (1 - \mu_P)P_n + (1 - \mu_A)A_n,\end{aligned}$$

where b is the larval recruitment rate per adult in unit time, and μ_L , μ_P , μ_A are the death rates in the respective stages. However, in practice, $\mu_P = 0$ since there is no pupal mortality except for cannibalism. We assume that cannibalistic acts occur randomly as the organisms move through the container of flour that forms their environment. This suggests a metered model with cannibalism rates proportional to the original size of the group being cannibalized, as in the Ricker fish model (Section 2.5). We are led to a model

$$\begin{aligned}L_{n+1} &= bA_n e^{-c_{EA}A_n} e^{-c_{EL}L_n}, \\ P_{n+1} &= (1 - \mu_L)L_n, \\ A_{n+1} &= P_n e^{-c_{PA}A_n} + (1 - \mu_A)A_n,\end{aligned} \tag{2.12}$$

with “cannibalism coefficients” c_{EA} , c_{EL} , c_{PA} . The fractions $e^{-c_{EA}A_n}$ and $e^{-c_{EL}L_n}$ are the probabilities that an egg is not eaten in the presence of A_n adults and L_n larvae

through the larval stage. The fraction $e^{-c_{PA}A_n}$ is the survival probability of a pupa through the pupal stage in the presence of A_n adults.

Equilibria of our basic model (2.12) are solutions (L, P, A) of the system of equations

$$\begin{aligned} L e^{c_{EL}L} &= b A e^{-c_{EA}A}, \\ P &= (1 - \mu_L)L, \\ \mu_A e^{c_{PA}A} &= P. \end{aligned} \tag{2.13}$$

This system has a solution $(0, 0, 0)$ corresponding to extinction and also has a solution with $L > 0$, $P > 0$, $A > 0$ corresponding to survival for some sets of parameter values. We are unable to find this survival equilibrium analytically, but we may solve numerically for a given choice of parameters.

We may rewrite the equilibrium conditions by eliminating P as

$$\begin{aligned} (1 - \mu_L)L &= \mu_A A e^{c_{PA}A}, \\ L e^{c_{EL}L} &= b A e^{-c_{EA}A}. \end{aligned}$$

Division of the second equation by the first gives

$$e^{c_{EL}L} = \frac{b(1 - \mu_L)}{\mu_A} e^{-(c_{EA} + c_{PA})A}. \tag{2.14}$$

The left side of (2.14) increases with L and is greater than one for all positive L , while the right side of (2.14) decreases with A and is between $b(1 - \mu_L)/\mu_A$ and zero. Thus, if the quantity θ , defined by

$$\theta = \frac{b(1 - \mu_L)}{\mu_A}, \tag{2.15}$$

is less than one, there cannot be a solution of (2.14) and thus there cannot be a survival equilibrium. On the other hand, if θ is greater than one, the equation (2.14) represents a straight line from $(0, \log \theta / c_{EL})$ to $(\log \theta / (c_{EA} + c_{PA}), 0)$ in the (A, L) -plane. An equilibrium is an intersection of this line with the curve $(1 - \mu_L)L = \mu_A e^{c_{PA}A}$, which starts from the origin and grows as A increases. Thus if $\theta > 1$, there is always a survival equilibrium.

Some experiments have been carried out with flour beetle populations and fit to the model (2.12) with the parameter values $b = 7$, $c_{EA} = c_{EL} = 0.01$, $c_{PA} = 0.005$, $\mu_L = 0.2$, $\mu_A = 0.01$ [Costantino, Desharnais, Cushing, and Dennis (1997), (1995)]. Since experimental data are inevitably noisy, it is not possible to determine parameters exactly, but it is possible to obtain a confidence interval for the parameters. We take these values as a baseline and then compare the model with the experiment when some of the parameters are manipulated. For example, we may remove (harvest) some adults at each census and thus set μ_A arbitrarily. In real life outside the laboratory, adult mortality may be changed by spraying with a pesticide. It is also

possible to manipulate the cannibalism coefficient c_{PA} by changing the supply of food; increasing the food supply reduces the rate of cannibalism of pupae by adults.

With the parameter values given above, we find $\theta = 560$, and there is a survival equilibrium $L = 36$, $P = 29$, $A = 398$, as well as the extinction equilibrium $L = 0$, $P = 0$, $A = 0$. In order to determine the stability of these equilibria, we must compute the matrix of partial derivatives at an equilibrium and form the characteristic equation as in Section 2.7.

At an equilibrium (L, P, A) this matrix is

$$\begin{pmatrix} -c_{EL}bAe^{-c_{EA}A}e^{-c_{EL}L} & 0 & be^{-c_{EL}L}e^{-c_{EA}A}(1-c_{EA}A) \\ 1-\mu_L & 0 & 0 \\ 0 & e^{-c_{PA}A} & 1-\mu_A-c_{PA}Pe^{-c_{PA}A} \end{pmatrix}. \quad (2.16)$$

At the extinction equilibrium $(0, 0, 0)$, it reduces to

$$\begin{pmatrix} 0 & 0 & b \\ 1-\mu_L & 0 & 0 \\ 0 & 1 & 1-\mu_A \end{pmatrix}.$$

The characteristic equation at $(0, 0, 0)$ is (after some manipulation of signs)

$$\det \begin{pmatrix} -\lambda & 0 & b \\ 1-\mu_L & -\lambda & 0 \\ 0 & 1 & 1-\mu_A-\lambda \end{pmatrix} = \lambda^2(\lambda - (1-\mu_A)) - b(1-\mu_L) = 0,$$

or $\lambda^3 - (1-\mu_A)\lambda^2 - b(1-\mu_L) = 0$. Thus, it has the form $\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0$ with $a_1 = -(1-\mu_A)$, $a_2 = 0$, $a_3 = -b(1-\mu_L)$. The conditions for asymptotic stability ((2.11), Section 2.7) are

$$\begin{aligned} 1 + a_1 + a_2 + a_3 &> 0, & 1 - a_1 + a_2 - a_3 &> 0, \\ 3 + a_1 - a_2 - 3a_3 &> 0, & 1 + a_1a_3 - a_2 - a_3^2 &> 0, \end{aligned} \quad (2.17)$$

and these become

$$\begin{aligned} \mu_A - b(1-\mu_L) &> 0, \\ 2 - \mu_A + b(1-\mu_L) &> 0, \\ 3 - (1-\mu_A) + 3b(1-\mu_L) &> 0, \\ 1 + b(1-\mu_L)(1-\mu_A) - b^2(1-\mu_L)^2 &> 0. \end{aligned}$$

Because $0 \leq \mu_A \leq 1$, $0 \leq \mu_L \leq 1$, the second and third of these conditions are satisfied automatically. The first condition is satisfied if and only if $\mu_A > b(1-\mu_L)$, which is equivalent to $\theta < 1$.

The last condition,

$$(b(1-\mu_L))^2 - b(1-\mu_L)(1-\mu_A) < 1,$$

is satisfied as well, since

$$(b(1 - \mu_L))^2 - b(1 - \mu_L)(1 - \mu_A) < \mu_A^2 - b(1 - \mu_L)(1 - \mu_A) < \mu_A^2 < 1.$$

Thus, the extinction equilibrium is asymptotically stable if and only if $\theta < 1$, that is, if and only if the extinction equilibrium is the only equilibrium.

At a survival equilibrium we may use the equilibrium conditions (2.13) to simplify the coefficient matrix (2.16) to

$$\begin{pmatrix} -c_{EL}L & 0 & \frac{L}{A} - c_{EA}L \\ 1 - \mu_L & 0 & 0 \\ 0 & e^{c_{PA}A} & 1 - \mu_A - \mu_A c_{PA}A \end{pmatrix}.$$

In this case, the characteristic equation is

$$\lambda(\lambda + c_{EL}L)(\lambda - (1 - \mu_A - \mu_A c_{PA}A)) - \left(\frac{L}{A} - c_{EA}L\right)(1 - \mu_L)e^{-c_{PA}A} = 0,$$

or

$$\begin{aligned} \lambda^3 + (c_{EL}L + \mu_A c_{PA}A - (1 - \mu_A))\lambda^2 \\ - c_{EL}L(1 - \mu_A)\lambda - \left(\frac{L}{A} - c_{EA}L\right)(1 - \mu_L)e^{-c_{PA}A} = 0, \end{aligned}$$

that is, a cubic equation $\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0$ with coefficients

$$\begin{aligned} a_1 &= c_{EL}L + \mu_A c_{PA}A - (1 - \mu_A), \\ a_2 &= -c_{EL}L(1 - \mu_A), \\ a_3 &= -\left(\frac{L}{A} - c_{EA}L\right)(1 - \mu_L)e^{-c_{PA}A}. \end{aligned} \tag{2.18}$$

We are unable to analyze the stability of the survival equilibrium in general, but for a particular choice of parameters $b, c_{EA}, c_{EL}, c_{PA}, \mu_L, \mu_A$ we can calculate the survival equilibrium (L, P, A) numerically and then use the values given by (2.18) to check the stability condition (2.17).

With the baseline parameters $b = 7, c_{EA} = c_{EL} = 0.01, c_{PA} = 0.005, \mu_L = 0.2, \mu_A = 0.01$, the survival equilibrium is $(36, 29, 398)$, and we find from (2.18) that $a_1 = -0.61, a_2 = -0.36, a_3 = 0.43$. The stability conditions (2.17) are satisfied, and thus the survival equilibrium is asymptotically stable. This agrees with experimental observations. However, this does not validate the model, since the parameter values were chosen to fit the experimental data.

To obtain some validation of the basic model, we must manipulate some of the parameter values and see whether experimental observations still agree with model predictions. Thus, we set $\mu_A = 0.96, c_{PA} = 0.5$ by harvesting adults and reducing the flour supply. With these parameter values, the model predicts a survival equilibrium $(12, 10, 3)$, and (2.18) gives $a_1 = 1.52, a_2 = -0.005, a_3 = -0.69$ in the cubic

characteristic equation. Now the stability condition $1 + a_1 a_3 - a_2 - a_3^2 > 0$ is violated, and our model predicts instability of the survival equilibrium. A more detailed study of the model indicates that with $\mu_A = 0.96$, the dynamics are very sensitive to changes in the cannibalism rate c_{PA} . For $c_{PA} = 0.5$ there is a solution of period 3 and a chaotic attractor, while for $c_{PA} = 0.55$ there are two attractors and a solution of period 8. Experimental observations indicate chaotic behavior, but it is not possible to be specific about the nature of the dynamics. Nevertheless, this does indicate some validity for the model and supplies what appears to be genuinely chaotic behavior in the laboratory.

Another way in which it is possible to perturb the model is to introduce periodic forcing by varying the volume of flour. Experiments indicate that cannibalism rates are inversely proportional to flour volume. Thus we may assume

$$c_{EL} = \frac{k_{EL}}{V}, \quad c_{EA} = \frac{k_{EA}}{V}, \quad c_{PA} = \frac{k_{PA}}{V},$$

where V is the volume of flour. We make flour volume oscillate with period 2 and amplitude αV_0 about a mean V_0 , so that $V_n = V_0(1 + \alpha(-1)^n)$. Then the cannibalism coefficients at stage n are

$$c_{EL} = \frac{k_{EL}}{V_0(1 + \alpha(-1)^n)}, \quad c_{EA} = \frac{k_{EA}}{V_0(1 + \alpha(-1)^n)}, \quad c_{PA} = \frac{k_{PA}}{V_0(1 + \alpha(-1)^n)}.$$

If we let c_{EL} , c_{EA} , c_{PA} denote the cannibalism coefficients in the average flour volume, $c_{EL} = k_{EL}/V_0$, $c_{EA} = k_{EA}/V_0$, $c_{PA} = k_{PA}/V_0$, we obtain the periodic model

$$\begin{aligned} L_{n+1} &= bA_n \exp\left(-\frac{c_{EL}L_n + c_{EA}A_n}{1 + \alpha(-1)^n}\right), \\ P_{n+1} &= (1 - \mu_L)L_n, \\ A_{n+1} &= P_n \exp\left(-\frac{c_{PA}A_n}{1 + \alpha(-1)^n}\right) + (1 - \mu_A)A_n. \end{aligned} \tag{2.19}$$

This model has an extinction equilibrium $(0, 0, 0)$, which may be shown to be asymptotically stable if $\theta < 1$. If $\theta > 1$, there is a solution of period 2 that is asymptotically stable if θ is close to 1, but for larger values of θ the dynamics may be chaotic. In addition, population sizes are considerably larger than in the unforced case, and this is borne out by experiment.

There are two important lessons that may be drawn from this model. The first is that trying to control a pest population by removing adults may have unintended consequences such as large fluctuations in the pest population size. The second is that periodic variation in the environment may produce substantial increases in population size.

The analysis of the model (2.12) becomes considerably simpler if we neglect cannibalism of eggs by larvae. Mathematically, this means taking $c_{EL} = 0$. The recruitment of larvae at equilibrium is then changed from $bA_n e^{-c_{EA}A_n} e^{-c_{EL}L_n}$ to $bA_n e^{-c_{EA}A_n}$. In order to make the parameter values correspond, we should replace

b by $be^{-c_{EL}L^*}$, where c_{EL} is the original cannibalism coefficient and L^* is the equilibrium larval population. With our baseline parameters this would mean replacing $b = 7$ by $b^* = 7e^{-0.36} = 4.88$. We will not carry out the analysis of this reduced model

$$\begin{aligned} L_{n+1} &= b^* A_n e^{c_{EA} A_n}, \\ P_{n+1} &= (1 - \mu_L) L_n, \\ A_{n+1} &= P_n e^{-c_{PA} A_n} + (1 - \mu_A) A_n, \end{aligned} \tag{2.20}$$

but will indicate it in a sequence of exercises.

Exercises

1. a. Show that the survival equilibrium of (2.20) has

$$A = \frac{1}{c_{EA} + c_{PA}} \log \theta,$$

and once A has been calculated,

$$L = b^* A e^{-c_{EA} A}, \quad P = b^* (1 - \mu_L) A e^{-c_{EA} A}.$$

- b. Calculate the equilibrium population sizes for the parameter values $b^* = 4.88$, $c_{EA} = 0.01$, $\mu_L = 0.2$, and (i) $\mu_A = 0.81$, $c_{PA} = 0.005$, (ii) $\mu_A = 0.96$, $c_{PA} = 0.5$.
2. Show that the extinction equilibrium of the model (2.20) is asymptotically stable if and only if $\theta < 1$.
3. Show that at a survival equilibrium of (2.20) the characteristic equation is a cubic polynomial with

$$\begin{aligned} a_1 &= \mu_A c_{PA} A - (1 - \mu_A), \\ a_2 &= 0, \\ a_3 &= - \left(\frac{L}{A} - c_{EA} L \right) (1 - \mu_L) e^{-c_{PA} A}. \end{aligned}$$

4. Show that with parameter values $b^* = 4.88$, $c_{EA} = 0.01$, $\mu_L = 0.2$, the survival equilibrium of (2.20) is asymptotically stable if $\mu_A = 0.01$, $c_{PA} = 0.005$, and unstable if $\mu_A = 0.96$, $c_{PA} = 0.5$.
5. Run simulations to compare the behaviors of the models (2.12) and (2.20) with the two sets of parameter values used in this section and a variety of initial values.
6. Show that it is possible to eliminate L and P from the model (2.20) and obtain a single third-order difference equation,

$$A_{n+3} = b^*(1 - \mu_L)A_n e^{-c_{EA}A_n} e^{-c_{PA}A_{n+2}} + (1 - \mu_A)A_{n+2}.$$

2.9 Project: A Discrete SIS Epidemic Model

In this project we outline of analysis of an *SIS* (susceptible infective suceptible) discrete epidemic model in a human population with variable size. The *SIS* model is given by the system

$$\begin{aligned} S_{n+1} &= f(T_n) + S_n\pi(n, n+1)h(I_n) + I_n\pi(n, n+1)[1 - \xi(n, n+1)], \\ I_{n+1} &= S_n\pi(n, n+1)[1 - h(I_n)] + I_n\pi(n, n+1)\xi(n, n+1)\zeta(n, n+1), \end{aligned}$$

with

$$T_n = S_n + I_n = f(T_n) + T_n\pi(n, n+1) + I_n\pi(n, n+1)[\zeta(n, n+1) - 1],$$

where $\pi(n, n+1)$, $\xi(n, n+1)$, $\zeta(n, n+1)$ are assumed to be constants with α, μ, σ positive constants, that is,

$$1 - \pi(n, n+1) = 1 - e^{-\mu}$$

is the probability of death due to natural causes,

$$1 - \xi(n, n+1) = 1 - e^{-\sigma}$$

is the probability of recovering,

$$1 - \zeta(n, n+1) = 1 - e^{-\rho}$$

is the probability of death due to infection,

$$h(I_n) = e^{-\alpha I_n}$$

is the probability of not becoming infected, and $f(T_n)$ is the birth or immigration rate (two cases). In this project we take $\rho = 0$, that is, the disease is not fatal.

The model assumes that the time step is one generation; from generation n to $n+1$, infections occur before deaths; there are no infected offspring, that is, all newborns or recruits enter into the susceptible class; in the case of nonconstant recruitment, if there are no people, then there are no births, that is, $f(0) = 0$; if there are too many people, then there are not enough resources to sustain further reproduction, that is, $\lim_{T_n \rightarrow \infty} f(T_n) = 0$; the probability of not becoming infected when there are no people is one, that is, $h(0) = 1$; the probability of not becoming infected as the number of infected increases is a strictly decreasing function, $h'(I_n) < 0$; as the number of infected people increases, the probability of not becoming infected goes to zero, that is, $\lim_{I_n \rightarrow \infty} h(I_n) = 0$.

Case A. Assume a constant recruitment rate, that is, let

$$f(T_n) = \Lambda > 0,$$

where Λ is a constant (immigration rate).

1. Show that the model becomes:

$$\begin{aligned} S_{n+1} &= \Lambda + S_n e^{-\mu} e^{-\alpha I_n} + I_n e^{-\mu} [1 - e^{-\sigma}] \\ I_{n+1} &= S_n e^{-\mu} [1 - e^{-\alpha I_n}] + I_n e^{-\mu} e^{-\sigma}, \end{aligned} \quad (2.21)$$

with

$$T_{n+1} = \Lambda + T_n e^{-\mu}.$$

2. Show that

$$T_n = e^{-\mu n} \left(T_0 - \frac{\Lambda}{1 - e^{-\mu}} \right) + \frac{\Lambda}{1 - e^{-\mu}}$$

and that

$$\lim_{n \rightarrow \infty} T_n = \frac{\Lambda}{1 - e^{-\mu}} \equiv T_\infty.$$

3. Set $T_0 = T_\infty$ (initial population size). This simply means that the population starts at its asymptotic limit, that is, the population is at a demographic equilibrium. Now substitute S_n by $S_n = T_\infty - I_n$ into (2.21) and show that

$$I_{n+1} = (T_\infty - I_n) e^{-\mu} [1 - e^{-\alpha I_n}] + I_n e^{-(\mu + \sigma)}. \quad (2.22)$$

4. Show that $I^* = 0$ is a fixed point of (2.22) and also show that the basic reproductive number is

$$R_0 = \frac{\alpha T_\infty e^{-\mu}}{1 - e^{-(\mu + \sigma)}}.$$

5. Explain the biological meaning of \mathcal{R}_0 .
6. Show that if $\mathcal{R}_0 < 1$, then $I^* = 0$ is a global attractor of $u(I_n)$.
7. Show that if $\mathcal{R}_0 > 1$, then the disease-free equilibrium is unstable.
8. Show that there exists a unique fixed point $I^* > 0$ of $v(I_n)$ for $\mathcal{R}_0 > 1$.

Case B. Assume a nonconstant recruitment rate of Ricker type, that is, let

$$f(T_n) = \beta T_n e^{-\gamma T_n}.$$

Then

$$\begin{aligned} S_{n+1} &= \beta T_n e^{-\gamma T_n} + S_n e^{-\alpha I_n} e^{-\mu} + e^{-\mu} [1 - e^{-\sigma}] I_n, \\ I_{n+1} &= S_n e^{-\mu} [1 - e^{-\alpha I_n}] + I_n e^{-\mu} e^{-\sigma}, \\ T_{n+1} &= S_{n+1} + I_{n+1} = \beta T_n e^{-\gamma T_n} + T_n e^{-\mu} \end{aligned} \quad (2.23)$$

where β = maximal birth rate/person/generation

Let $\mathcal{R}_d = \beta / (1 - e^{-\mu})$.

1. Show that if $\mathcal{R}_d < 1$, then there is no positive fixed point.
2. Show that if $\mathcal{R}_d > 1$, then there exist two fixed points, $T'_\infty = 0$ and $T_\infty^2 > 0$.
3. Show that if $1 < \mathcal{R}_d < e^{2/(1-e^{-\mu})}$, then T_∞^2 is locally stable.
4. What is the biological interpretation of \mathcal{R}_d ? Assume that (2.23) can be “reduced” to a single “equivalent” limiting equation

$$I_{n+1} = (T_\infty - I_n)e^{-\mu}(1 - e^{-\alpha I_n}) + I_n e^{-\mu} e^{-\sigma} \quad (2.24)$$

when $1 < \mathcal{R}_d < e^{2/(1-e^{-\mu})}$ and where $T_\infty^2 = \frac{1}{\gamma} \log(\mathcal{R}_d)$.

5. Show that the basic reproductive number is

$$\mathcal{R}_0 = \frac{\alpha e^{-\mu} \frac{1}{\gamma} \log R_d}{1 - e^{-(\mu+\sigma)}}.$$

6. Show that T_∞^2 is a global attractor if $\mathcal{R}_0 < 1$. Show that if $\mathcal{R}_0 > 1$, then the endemic equilibrium of (2.24) is a global attractor.
7. Simulate the full system (2.23) in the region $1 < R_d < e^{2/(1-e^{-\mu})}$, where T_∞ is a fixed point, and in the regions $R_d > e^{2/(1-e^{-\mu})}$, where period-doubling bifurcation occurs on the route to chaos. Does the demography drive the disease dynamics?

References: Castillo-Chavez and Yakubu (2000b, 2000c, 2000d), Barrera, Cintron-Arias, Davidenko, Denogean, and Franco (2000).

2.10 Project: A Discrete-Time Two-Sex Pair-Formation Model

1. Consider the following discrete-time two-sex pair-formation model:

$$\begin{aligned} x(t+1) &= (\beta_x \mu_x \mu_y + (1 - \mu_y) \mu_x + (1 - \sigma) \mu_x \mu_y) p(t) \\ &\quad + \mu_x x(t) G(x(t), y(t), p(t)), \\ y(t+1) &= (\beta_y \mu_y \mu_x + (1 - \mu_x) \mu_y + (1 - \sigma) \mu_x \mu_y) p(t) \\ &\quad + \mu_y y(t) H(x(t), y(t), p(t)), \\ p(t+1) &= \sigma \mu_x \mu_y p(t) + \mu_x x(t) (1 - G(x(t), y(t), p(t))), \end{aligned}$$

where the functions $G : [0, \infty) \times [0, \infty) \times [0, \infty) \rightarrow [0, 1]$ and $H : [0, \infty) \times [0, \infty) \times [0, \infty) \rightarrow [0, 1]$ denote the state-dependent probability functions and satisfy the equation

$$\mu_x x(t) (1 - G(x(t), y(t), p(t))) = \mu_y y(t) (1 - H(x(t), y(t), p(t)))$$

and where $\beta_x, \beta_y, \mu_x, \mu_y$, and σ are constants in the interval $[0, 1]$.

(a) Given that

$$G(x(t), y(t), p(t)) = \frac{p(t)}{y(t) + p(t)},$$

where $(x(t), y(t), p(t))$ belong to the set Ω , where

$$\Omega := \left\{ (x(t), y(t), p(t)) \mid 0 \leq \frac{x(t)}{y(t)} \leq \frac{\mu_y}{\mu_x(1 - G(x(t), y(t), p(t)))} \right\}$$

find

$$H(x(t), y(t), p(t)).$$

(b) For the given function in (a) find the marriage function $\phi : [0, \infty) \times [0, \infty) \times [0, \infty) \rightarrow [0, 1]$ that satisfies the equation

$$\begin{aligned} \phi(x(t), y(t), p(t)) &\equiv \mu_x x(t)(1 - G(x(t), y(t), p(t))) \\ &= \mu_y y(t)(1 - H(x(t), y(t), p(t))). \end{aligned}$$

(c) Show that the marriage function in (b) satisfies the following properties for all $(x(t), y(t), p(t)) \in \Omega$ and the constant $k \in [0, \infty)$:

(i)

$$\phi(x(t), y(t), p(t)) \geq 0,$$

(ii)

$$\phi(kx(t), ky(t), kp(t)) = k\phi(x(t), y(t), p(t)),$$

(iii)

$$\phi(x(t), 0, p(t)) = \phi(0, y(t), p(t)) = 0.$$

(d) If $\beta_x = \beta_y = \mu_x = \mu_y = \sigma$, use the marriage function in (b) to solve the following equation for the characteristic equation $\lambda = \lambda^*$:

$$-\sigma\mu_x\mu_y + \lambda = \phi\left(\frac{\beta_x\mu_x\mu_y}{\lambda - \mu_x} - 1, \frac{\beta_y\mu_y\mu_x}{\lambda - \mu_y} - 1, 1\right),$$

where

$$\frac{\beta_x\mu_x\mu_y}{\lambda - \mu_x} - 1 > 0 \quad \text{and} \quad \frac{\beta_y\mu_y\mu_x}{\lambda - \mu_y} - 1 > 0.$$

2. Use the marriage function in (b) with $\varepsilon = 0$ and $\beta_x = \beta_y = \mu_x = \mu_y = \sigma$ to find a positive fixed point $[\xi_0, \eta_0, 1] \in \Omega$ of the following system (if one exists):

$$\begin{aligned} \xi(t+1) &= \frac{\beta_x\mu_x\mu_y + \mu_x + \mu_x\xi(t)}{\sigma\mu_x\mu_y + \phi(\xi(t), \eta(t), 1)} - 1, \\ \eta(t+1) &= \frac{\beta_y\mu_x\mu_y + \mu_y + \mu_y\eta(t)}{\sigma\mu_x\mu_y + \phi(\xi(t), \eta(t), 1)} - 1, \\ \varsigma(t+1) &= 1. \end{aligned}$$

3. Use the Jury test to find values of σ (if any exist) for which the fixed point $[\xi_0, \eta_0, 1]$ is stable.

References: Castillo-Chavez and Yakubu (2000e, 2000f).

Chapter 3

Continuous Single-Species Population Models with Delays

3.1 Introduction

Up to now in our study of continuous population models we have been assuming that $x'(t)$, the growth rate of population size at time t , depends only on $x(t)$, the population size at the same time t . However, there are situations in which the growth rate does not respond instantaneously to changes in population size. One of the first models incorporating a delay was proposed by Volterra (1926) to take into account the delay in response of a population's death rate to changes in population density caused by an accumulation of pollutants in the past. Other causes of response delays that have been mentioned in the biological literature include differences in resource consumption with respect to age structure, migration and diffusion of populations, gestation and maturation periods, delays in behavioral response to environmental changes, and dependence of a population on a food supply that requires time to recover from grazing. In deriving a mathematical model to reflect a particular biological delay mechanism, one must consider carefully how this mechanism affects the growth rate. One approach to modeling delays that has been used is formulation of a discrete model (or difference equation) and consideration of the delay in the time between steps. While this is appropriate for populations with a discrete reproduction cycle, such as many fish populations, it does not accurately model populations with continuous growth and time lags. The metered models studied in Section 2.5 allow for a continuous death process but involve a discrete reproduction stage. To describe populations with a continuous reproduction stage we will need to introduce more general classes of models involving *differential–difference equations* or, even more generally, *integro-differential equations*. We will consider models in which the per capita growth rate $x'(t)/x(t)$ at time t is a function of the population size $x(t - T)$, that is, at T time units previous) and also models with *distributed delay* in which the per capita growth rate is an integral over the past. In addition, we will examine models in which the birth rate at time t is a function of past population size, either at a single time ($t - T$) or an integral over the past history, and the death rate is a function of the population size at time t . Different biological mechanisms may

lead to either of these types of models; the two types are *not* equivalent, although their mathematical analyses follow similar paths.

The assumption of a fixed delay T will lead to a differential–difference equation of the form $x'(t) = f(x(t), x(t-T))$. For such an equation one must prescribe initial data for a time interval of length T rather than merely an initial value $x(0)$. In other words, one specifies a function $x_0(t)$ for $-T \leq t \leq 0$ and then requires the initial condition $x(t) = x_0(t)$ for $-T \leq t \leq 0$. The assumption of a distributed delay will lead to an integrodifferential equation of the form

$$x'(t) = \int_{-\infty}^t f(x(t), x(s))p(t-s)ds = \int_0^{\infty} f(x(t), x(t-u))p(u)du$$

for which one must prescribe initial data on $-\infty \leq t \leq 0$. More precisely, if $p(u) \equiv 0$ for $u \geq U$ one must prescribe initial data on $-U \leq t \leq 0$.

Analytic solutions, even for the simplest differential–difference equations, are in general hopeless. For example, suppose we try to solve the equation $x'(t) = -\alpha x(t-1)$, α constant with $x(t) = \varepsilon \geq 0$ for $-1 \leq t \leq 0$. Then $x'(t) = -\alpha\varepsilon$ for $0 \leq t \leq 1$, which gives $x(t) = \varepsilon - \alpha\varepsilon t$ for $0 \leq t \leq 1$, and $x(1) = \varepsilon(1-\alpha)$. We could use $x(t) = \varepsilon(1-\alpha t)$ for $0 \leq t \leq 1$ to obtain $x'(t) = -\alpha\varepsilon[1-\alpha(t-1)]$ for $1 \leq t \leq 2$ and then integrate to find $x(t)$ for $1 \leq t \leq 2$. This technique is feasible in principle, but the analytic form of the solution tends to become more and more complicated as we proceed. For this reason, we will concentrate on other ways of attempting to describe solutions. In this example, if $\alpha < 1$, then $x(1) \geq 0$. If $\alpha > 1$, then $x(1) < 0$, and we can see that $x'(t)$ will become positive, suggesting that the solution will oscillate. In fact, if $0 \leq \alpha \leq 1/e$, $x(t)$ remains positive for all t and decreases monotonically to zero. If $1/e \leq \alpha \leq \pi/2$, $x(t)$ oscillates but tends to zero as $t \rightarrow \infty$, while if $\alpha > \pi/2$, $x(t)$ oscillates but does not tend to zero, tending instead to a periodic solution. Some additional properties of the solution are established in the exercises below.

Because of the impossibility of obtaining explicit analytic solutions of equations with delays, we will concentrate on qualitative analysis. In Chapters 1 and 2 we obtained some qualitative results for the behavior of solutions of differential equations and difference equations. Even though explicit solution was possible in some cases, the qualitative approach enabled us to obtain information about a larger class of problems. For differential–difference equations the qualitative approach follows a similar line, but there are additional technical problems.

Another way of obtaining information about solutions employed in Chapters 1 and 2 was to use a computer algebra system to solve problems numerically and graph the solutions. We will use this approach for differential–difference equations as well, but there are difficulties. A program in Mathematica for solving differential–difference equations is given in (virtual) Appendix C, but it is somewhat cumbersome. Two programs in Maple given in Appendix C are somewhat less cumbersome but run extremely slowly. For differential–difference equations, Matlab and the dynamical systems program XPP, or its Windows analogue WinPP (whose acquisition is described in the preface to this book), are considerably more efficient. The so-

lution graphs in this chapter were obtained using WinPP. For example, Figure 3.1 shows the graph of the solution of $x'(t) = -x(t-1)$, with $x(t) = 1$ for $-1 \leq t \leq 0$.

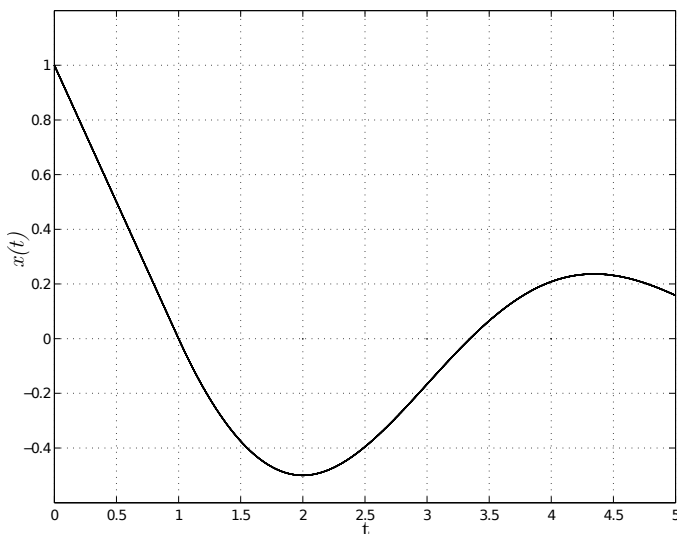


Fig. 3.1 Solution of $x'(t) = -x(t-1)$.

Exercises

In Exercises 1 through 5, consider the differential–difference equation $x'(t) = -\alpha x(t-1)$ with initial data $x(t) \equiv 1$ for $-1 \leq t \leq 0$.

1. Show that $x'(t) \equiv 0$ for $t < 0$ but $x'(t) \equiv -\alpha$ for $0 \leq t \leq 1$, so that $x'(t)$ is discontinuous at $t = 0$.
2. Find the solution for $1 \leq t \leq 2$ and show that $x(2) = \frac{1}{2}(\alpha^2 - 4\alpha + 2)$.
3. Show that $x'(t)$ approaches the limit $-\alpha$ as t approaches 1 through values greater than 1, so that $x'(t)$ is continuous at $t = 1$.
4. Use a computer algebra system to sketch the solutions with $\alpha = 0.3$ and $\alpha = 4$.
- 5*. Show that $x''(t)$ is discontinuous at $t = 1$.

3.2 Models with Delay in Per Capita Growth Rates

If we assume that the per capita growth rate $x'(t)/x(t)$ is a function of $x(t-T)$, as might be appropriate for example in modeling a population whose food supply

requires a time T to recover from grazing, causing the food supply at time t to depend on the population size at time $(t - T)$, we are led to a model of the form

$$x'(t) = x(t)g(x(t - T)),$$

a differential–difference equation. For example, the delay logistic equation is

$$x'(t) = rx(t)\left(1 - \frac{x(t - T)}{K}\right). \quad (3.1)$$

This equation was introduced by Hutchinson (1948) to describe the dynamics of animal populations. More generally, we could assume a delay distributed over time. If the probability that the delay is between u and $u + \Delta u$ is approximately $p(u)\Delta u$, where $p(u)$ is a nonnegative function with $\int_0^\infty p(u)du = 1$ then we are led to the integrodifferential equation

$$x'(t) = x(t) \int_0^\infty g(x(t - u))p(u)du,$$

which is transformed by the change of variable $t - u = s$ to the equivalent form

$$x'(t) = x(t) \int_{-\infty}^t g(x(s))p(t - s)ds.$$

The average time delay will then be $\int_0^\infty up(u)du$. One form of continuous delay frequently used in population models is

$$p(u) = \frac{u}{T^2}e^{-u/T},$$

for which it is not difficult to verify that $\int_0^\infty p(u)du = 1$, $\int_0^\infty up(u)du = 2T$, and $p(u)$ has a maximum for $u = T$ (see Exercises 3, 4 below).

Because the analytic solution of differential–difference equations of the form $x'(t) = x(t)g(x(t - T))$ is not possible in general, we attempt instead to describe the behavior of solutions in terms of equilibria and asymptotic stability just as we did for differential equations in Chapter 1 and for difference equations in Chapter 2.

Definition 3.1. An *equilibrium* of the differential–difference equation

$$x'(t) = x(t)g(x(t - T)) \quad (3.2)$$

is a value x_∞ such that $x_\infty g(x_\infty) = 0$, so that $x(t) \equiv x_\infty$ is a constant solution of the differential–difference equation.

Observe that the form of the differential–difference equation (3.2) implies that $x = 0$ is always an equilibrium. The delay logistic equation has two equilibria, $x = 0$ and $x = K$, just as does the (undelayed) logistic differential equation $x'(t) = rx(t)(1 - x(t)/K)$, which is the case $T = 0$ of the delay logistic equation (3.1). For the differential equation $x' = xg(x)$, which is the case $T = 0$ of the differential–

difference equation $x'(t) = x(t)g(x(t-T))$, an equilibrium x_∞ is asymptotically stable if and only if

$$(xg(x))' \Big|_{x=x_\infty} = x_\infty g'(x_\infty) + g(x_\infty) < 0,$$

so that the equilibrium $x = 0$ is asymptotically stable if $g(0) < 0$ and an equilibrium $x_\infty > 0$ is asymptotically stable if $g'(x_\infty) < 0$ (because $g(x_\infty) = 0$). The asymptotic stability of an equilibrium x_∞ of the differential-difference equation $x'(t) = x(t)g(x(t-T))$ requires $(xg(x))' \Big|_{x=x_\infty} < 0$ and an additional condition. In order to describe the additional condition required, we proceed much as we did in Chapter 1 for differential equations. We begin by linearizing about the equilibrium x_∞ with the aid of Taylor's theorem. That is, we let $u(t) = x(t) - x_\infty$ and obtain the equivalent differential-difference equation

$$\begin{aligned} u'(t) &= (x_\infty + u(t))g(x_\infty + u(t-T)) \\ &= (x_\infty + u(t)) \left(g(x_\infty) + g'(x_\infty)u(t-T) + \frac{g''(c)}{2!}u(t-T)^2 \right) \\ &= x_\infty g(x_\infty) + g(x_\infty)u(t) + x_\infty g'(x_\infty)u(t-T) + h(u(t), u(t-T)) \\ &= g(x_\infty)u(t) + x_\infty g'(x_\infty)u(t-T) + h(u(t), u(t-T)), \end{aligned}$$

where c is between x_∞ and $x_\infty + u(t-T)$ and

$$h(u(t), u(t-T)) = g'(x_\infty)u(t)u(t-T) + x_\infty \frac{g''(c)}{2!}u(t-T)^2$$

is “small” when $u(s)$ is small for $t-T \leq s \leq t$. The *linearization* of the differential-difference equation $x'(t) = x(t)g(x(t-T))$ is defined to be the linear differential-difference equation

$$v'(t) = g(x_\infty)v(t) + x_\infty g'(x_\infty)v(t-T), \quad (3.3)$$

obtained by neglecting higher-order terms collected as $h(u(t), u(t-T))$. Just as with differential equations and difference equations, the importance of the linearization lies in the fact that the behavior of its solutions describes the behavior of solutions of the original equation $x'(t) = x(t)g(x(t-T))$ near the equilibrium. The original equation is equivalent to

$$u'(t) = g(x_\infty)u(t) + x_\infty g'(x_\infty)u(t-T) + h(u(t), u(t-T)),$$

and h is small in the sense that for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|h(y, z)| < \varepsilon(|y| + |z|)$ whenever $|y| < \delta, |z| < \delta$. The following result, which we state without proof, is valid. A proof may be found in Bellman and Cooke (1963).

Theorem 3.1. *If all solutions of the linearization*

$$v'(t) = g(x_\infty)v(t) + x_\infty g'(x_\infty)v(t-T)$$

at an equilibrium x_∞ tend to zero as $t \rightarrow \infty$, then every solution $x(t)$ of $x'(t) = x(t)g(x(t-T))$ with $|x(t) - x_\infty|$ sufficiently small for $-T \leq t \leq 0$ tends to the equilibrium x_∞ as $t \rightarrow \infty$.

For the equilibrium $x = 0$, the linearization is $v'(t) = g(0)v(t)$. Since $g(0) > 0$ for most models of this type, the equilibrium $x = 0$ is unstable. For an equilibrium $x_\infty > 0$, $g(x_\infty) = 0$ and the linearization is $v'(t) = x_\infty g'(x_\infty)v(t-T)$. The delay logistic equation (3.1), for example, has linearization $v'(t) = -rv(t-T)$ at the equilibrium $x_\infty = K$, since $g(x) = r(1-x/K)$, $g'(x) = -r/K$, and thus $x_\infty g'(x_\infty) = -r$. In order to determine whether all solutions of a linear differential-difference equation

$$v'(t) = bv(t-T) \quad (3.4)$$

tend to zero as $t \rightarrow \infty$, we look for the condition on the parameter λ (possibly complex) that $v(t) = ce^{\lambda t}$ be a solution ($c \neq 0$). Substitution of $v(t) = ce^{\lambda t}$ into (3.4) gives

$$c\lambda e^{\lambda t} = bce^{\lambda(t-T)} = bce^{\lambda t}e^{-\lambda T}.$$

Since $e^{\lambda t} \neq 0$, we may divide both sides by $ce^{\lambda t}$, obtaining

$$\lambda = be^{-\lambda T}. \quad (3.5)$$

This is a transcendental equation for λ having infinitely many roots. A basic result, which we shall state without proof, is that if all roots of the characteristic equation (3.5) have negative real part, then all solutions of the differential-difference equation (3.4) tend to zero as $t \rightarrow \infty$. This result is analogous to the corresponding result for differential equations (Section 1.1); a proof may be found in Bellman and Cooke (1963). However, it is much more difficult to analyze the transcendental equation (3.5) in the delay case. The undelayed case is the special case $T = 0$ of (3.4) for which the characteristic equation becomes $\lambda = b$ and the condition that all roots have negative real part is $b < 0$. For the delay case, with $T > 0$, it is possible to show that the condition that all roots of the characteristic equation (3.5) have negative real part is

$$0 < -bT < \frac{\pi}{2}. \quad (3.6)$$

This was first established by Hayes (1950). The condition (3.6) requires $b < 0$, and in addition, that the time lag T not be too large. If the condition (3.6) is violated, the characteristic equation (3.5) has complex roots with real part zero if $-bT = \pi/2$ and positive real part if $-bT > \pi/2$. The appearance of roots $\lambda = iy$ with real part zero when $-bT = \pi/2$ is reflected in the differential-difference equation (3.4) by the appearance of periodic solutions of the form $v(t) = ce^{iyt}$, or $v(t) = c \cos yt$ in real form. Combining this analysis with Theorem 3.1, we see that an equilibrium $x_\infty > 0$ of the differential-difference equation $x'(t) = x(t)g(x(t-T))$ is asymptotically stable if $0 < -x_\infty g'(x_\infty)T < \pi/2$. For the delay logistic equation this stability condition is $0 < rT < \pi/2$. Thus, in addition to the stability condition $(xg(x))'|_{x=x_\infty} < 0$ for stability of equilibrium for ordinary differential equations, we have an additional requirement that the delay T be sufficiently small.

A graphic display of a solution is often helpful for obtaining insights into the behavior of a solution, even if the graph does not prove results. For example, [Figure 3.2](#) displays the solution of the delay logistic equation (3.1) with $r = 1$, $K = 2$, $T = 1$, and $x(t) = 1$ for $-1 \leq t \leq 0$.

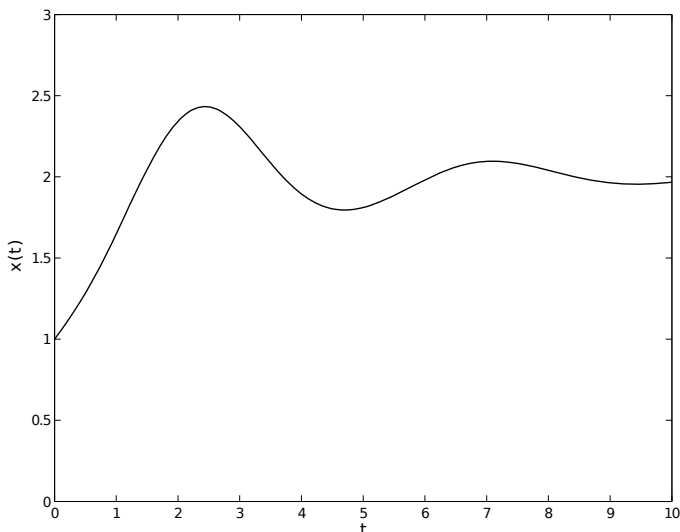


Fig. 3.2 Solution of the delay logistic equation.

Exercises

1. Show that the equilibrium $x = K$ of the delay logistic equation is asymptotically stable if $0 \leq rT < \frac{\pi}{2}$.
2. Show that the equilibrium $x = K$ of the differential–difference equation

$$x'(t) = rx(t) \log \left(\frac{K}{x(t-T)} \right)$$

is asymptotically stable if $0 \leq rT < \frac{\pi}{2}$.

3. Show that the function

$$p(u) = \frac{ue^{-u/T}}{T^2}$$

has a maximum for $u = T$.

4. With the aid of integration by parts show that

$$\int_0^\infty p(u)du = 1 \text{ and } \int_0^\infty up(u)du = 2T$$

for the function $p(u)$ of Exercise 3.

5. Use a computer algebra system to display the solutions of

$$x'(t) = rx(t)\left(1 - x(t-1)\right)$$

with $r = 1$, $r = 2$, $r = 3$.

- 6*. Find all equilibria of the differential–difference equation

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

- 7*. Find the characteristic equation at each equilibrium of the differential–difference equation of Exercise 5.

3.3 Delayed-Recruitment Models

If R is the rate at which new members are recruited into a population and D is the mortality rate, then we have

$$x'(t) = R - D.$$

We will consider models in which $x(t)$ represents the number of adult members and there is a fixed age T at which members mature and are recruited into the population. If preadult mortality depends only on age, there is a constant probability of survival to adulthood. If the birth rate at a given time depends only on the adult population size at that time, then the recruitment rate at time t is a function of the population size at time $(t - T)$. If the death rate at a given time also depends only on the adult population size at that time, then the death rate at time t is a function of the population size at time t . Thus, under the above hypotheses, we have a population model of the form

$$x'(t) = R(x(t-T)) - D(x(t)). \quad (3.7)$$

We will assume that the function $D(x)$ is a monotone function, with $D(0) = 0$ and $\lim_{x \rightarrow \infty} D(x) = \infty$. With regard to the function $R(x)$, we will consider two different types of recruitment function corresponding to different types of competition for resources among either adults or preadults. Under *scramble competition* the available resources are partitioned equally among all individuals. The birth rate is zero when the population size is zero and will also tend to zero for large population sizes, with some maximum birth rate occurring at an intermediate population size. Under *contest competition* a certain number of individuals can be maintained at the expense of the others. The birth rate then increases with population size but is bounded as the population size becomes large. Thus, we make one of the following sets of hypotheses on the recruitment function $R(x)$: (1) $R(0) = 0$, (2) R increases to a unique maxi-

imum value ρ for $x = \xi$ and then decreases to zero as $x \rightarrow \infty$ (scramble competition), (3) $R(0) = 0$, $R'(x) > 0$ for $0 < x < \infty$, and $\lim_{x \rightarrow \infty} R(x) = \rho$ (contest competition). We may also admit the possibility of an Allee effect by allowing $R(x) < D(x)$ on some interval $0 < x < \alpha$, with $R(\alpha) = D(\alpha)$, $R'(\alpha) > D'(\alpha)$. We will carry out an equilibrium analysis, following the kind of approach we have used in several settings previously. However, we shall first state some results relating the possible range of population size to the initial data; proofs may be found in Brauer (1986).

Theorem 3.2. *Define M by the condition*

$$D(M) = \rho = \max_{0 \leq x \leq \infty} R(x).$$

Then, if the initial data satisfy $0 \leq x(t) \leq M$ for $-T \leq t \leq 0$, the solution $x(t)$ of the differential–difference equation (3.7) satisfies $0 \leq x(t) \leq M$.

For studying the possibility of extinction of a population, it is useful to have a criterion for which the population size has a positive lower bound.

Theorem 3.3. *In the case of scramble competition, define m by the condition*

$$D(m) = R(M).$$

If $M > \xi$, and if the initial data satisfy $0 < m \leq x(t) \leq M$ for $-T \leq t \leq 0$, then the solution $x(t)$ of (3.7) satisfies $m \leq x(t) \leq M$ for $t \geq 0$.

This lower-bound result does not cover the possibility of an Allee effect, contest competition, or scramble competition with $M \leq \xi$. For differential–difference equation models, unlike differential equation models, it is possible for the population size to drop to zero without signifying extinction because there may be juvenile members present who will mature into adult members. Extinction is guaranteed, however, if the population size remains zero for a time interval of length T . The start of the qualitative analysis of the delayed-recruitment model (3.7) follows our standard pattern.

Definition 3.2. An *equilibrium* of the differential–difference equation

$$x'(t) = R(x(t-T)) - D(x(t))$$

is a value x_∞ such that $R(x_\infty) = D(x_\infty)$, an intersection of the recruitment curve $y = R(x)$ and the mortality curve $y = D(x)$.

The linearization at an equilibrium, obtained in the usual way by letting $u = x - x_\infty$, substituting, expanding by Taylor's theorem, and neglecting higher-order terms, is

$$u'(t) = -D'(x_\infty)u(t) + R'(x_\infty)u(t-T). \quad (3.8)$$

The following result, analogous to Theorem 3.1 of the preceding section, whose proof may be found in Bellman and Cooke (1963), is valid.

Theorem 3.4. *If all solutions of the linearization (3.8) at an equilibrium x_∞ tend to zero as $t \rightarrow \infty$, then every solution $x(t)$ of the differential–difference equation (3.7) with $|x(t) - x_\infty|$ sufficiently small for $-T \leq t \leq 0$ tends to the equilibrium x_∞ as $t \rightarrow \infty$.*

In order to describe the behavior of solutions of the linearization, we must study a problem more general than that considered in the preceding section, namely the behavior of solutions of the linear differential–difference equation

$$u'(t) = au(t) + bu(t - T). \quad (3.9)$$

(In the preceding section, we considered the special case $a = 0$.) We look for solutions of the form $u(t) = ce^{\lambda t}$ and obtain the characteristic equation $\lambda = a + be^{-\lambda T}$. In order that all solutions of $u'(t) = au(t) + bu(t - T)$ tend to zero as $t \rightarrow \infty$, all solutions of the characteristic equation must have negative real part. It is possible to prove the following result [Hayes (1950)].

Theorem 3.5. *Let $p = aT$, $q = bT$. Then all solutions of (3.9) tend to zero as $t \rightarrow \infty$ if*

$$p < -q < p \sec z = \frac{z}{\sin z},$$

where z is the solution in $(0, \pi)$ of $z = p \tan z$ if $p \neq 0$, and $z = \pi/2$, so that $z/\sin z = \pi/2$ if $p = 0$.

In fact, if $p < -q < e^{p-1}$, the convergence of solutions to zero is ultimately monotone, while if $p < e^{p-1} < -q < p \sec z$, the convergence of solutions is oscillatory. In particular, if $p < -q < -p$, all solutions tend to zero as $t \rightarrow \infty$ for every value of T . The stability region for (3.9) in the (p, q) -parameter plane is shown in Figure 3.3.

For the delayed-recruitment model $x'(t) = R(x(t - T)) - D(x(t))$, the linearization is

$$u'(t) = -D'(x_\infty)u(t) + R'(x_\infty)u(t - T),$$

and we have $p = -D'(x_\infty)T < 0$, $q = R'(x_\infty)T$. Combining Theorems 3.4 and 3.5, we obtain the following conclusions (Figure 3.3):

1. At an equilibrium x_∞ with $R'(x_\infty) > 0$, or $q > 0$, we have exponential asymptotic stability if $p < -q$, or $D'(x_\infty) > R'(x_\infty)$ and exponential instability if $p > -q$, or $D'(x_\infty) < R'(x_\infty)$. (This covers contest competition; scramble competition with $M < \xi$, so that $x_\infty < \xi$, the lowest equilibrium when there is an Allee effect; and the case $D(x) > R(x)$ for all $x > 0$, so that $x = 0$ is the only equilibrium. In particular, there can be no oscillation in these situations.)
2. At an equilibrium x_∞ with $R'(x_\infty) < 0$, or $q < 0$, asymptotic stability can be either exponential or oscillatory. There is numerical evidence to indicate that in case of instability there may be periodic solutions, period-doubling, and eventually chaotic behavior, as is the case for difference equations.
3. An equilibrium x_∞ with $|R'(x_\infty)| < D'(x_\infty)$ is asymptotically stable for all T , $0 \leq T < \infty$. This property is known as *absolute stability*.

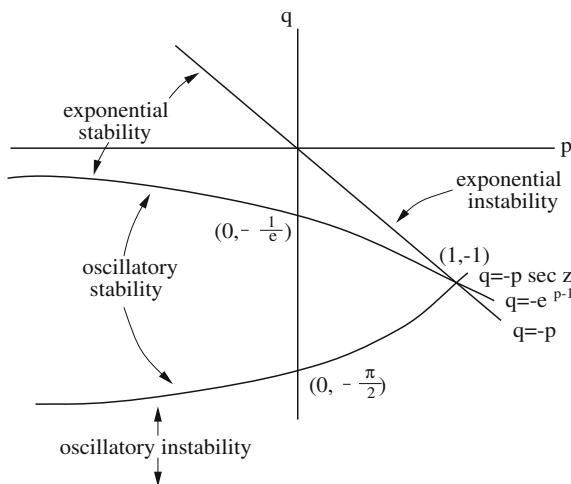


Fig. 3.3 Stability regions for $u'(t) = pu(t) + qu(t-T)$.

In fact, it is possible to establish the stronger result of Cooke and van den Driessche (1986):

Theorem 3.6. *For an equilibrium x_∞ of the differential–difference equation (3.7), one of the following three possibilities must occur:*

- i. *The equilibrium x_∞ is asymptotically stable for all values of T , $0 \leq T < \infty$ (absolute stability).*
- ii. *The equilibrium x_∞ is asymptotically stable for T less than some value T^* and unstable for T greater than T^* .*
- iii. *The equilibrium x_∞ is unstable for every $T \geq 0$.*

Theorem 3.6 implies that increasing the time lag cannot stabilize an equilibrium that is unstable when the time lag is small, but can destabilize an equilibrium that is stable when the time lag is small.

In contest competition, since $D(x) \rightarrow \infty$ while $R(x)$ is bounded as $x \rightarrow \infty$, there must be a largest equilibrium that is asymptotically stable, indeed exponentially asymptotically stable. The smallest asymptotically stable equilibrium (if there should be more than one) serves effectively as a positive lower bound for solutions. In scramble competition with $M < \xi$, there is an exponentially asymptotically stable equilibrium $x_\infty < \xi$, which serves as an effective positive lower bound. When there is an Allee effect with smallest equilibrium α , solutions with initial data less than α tend monotonically to zero, but solutions with initial data greater than α are forced upward, and α serves as a lower bound. If $D(x) > R(x)$ for all $x > 0$, there is no positive lower bound for solutions; indeed, all solutions tend to the exponentially asymptotically stable equilibrium $x = 0$.

Example 1. Consider the recruitment function

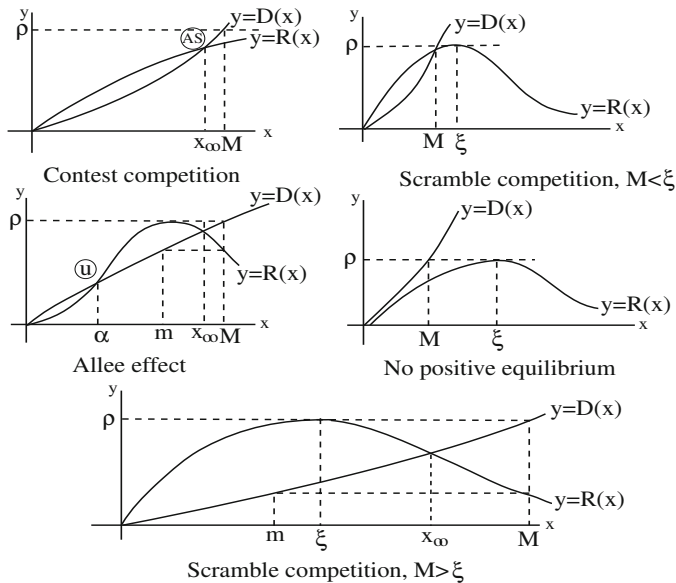


Fig. 3.4 Equilibrium possibilities.

$$R(x) = rxe^{-x/A}$$

and the mortality function $D(x) = dx$. Then $R(x)$ has a maximum of rA/e attained when $x = A$. Thus, the model

$$x'(t) = rx(t-T)e^{-x(t-T)/A} - dx(t) \quad (3.10)$$

represents scramble competition. There is always an equilibrium at $x = 0$ and there is a positive equilibrium x_∞ given by $re^{-x_\infty/A} = d$, or $x_\infty = A \log r/d$, if and only if $r > d$. Then $R'(x) = re^{-x/A}(1 - x/A)$ and $D'(x) = d$, so that $R'(0) = r$, $D'(0) = d$, and $R'(x_\infty) = d(1 - x_\infty/A) = d(1 - \log r/d)$, $D'(x_\infty) = d$. Theorem 3.5 shows that the equilibrium $x = 0$ is asymptotically stable if and only if $r < d$, that is, if and only if there is no positive equilibrium. If $r > d$, so that the equilibrium $x = 0$ is unstable, we may analyze the stability of the equilibrium $x_\infty > 0$ by applying Theorem 3.5 with $p = -D'(x_\infty)T = -dT$, $q = R'(x_\infty)T = dT(1 - \log r/d)$. We can deduce that the equilibrium x_∞ is absolutely stable if $-1 < 1 - x_\infty/A < 1$, which reduces to $\log(r/d) < 2$ or $r < de^2$. More generally, the equilibrium x_∞ is asymptotically stable if

$$-dT < -dT \left(1 - \log \frac{r}{d}\right) < -dT \sec z, \quad (3.11)$$

where $z = -dT \tan z$. The condition (3.11) reduces to

$$1 > 1 - \log \frac{r}{d} > \sec z,$$

or

$$0 < \log \frac{r}{d} < 1 - \sec z.$$

This implies that a large ratio r/d tends to make for instability, at least for large values of T . Figure 3.5 shows the solution of the differential–difference equation $x'(t) = 10x(t-1)e^{-x(t-1)} - x(t)$ with $x(t) = 1$ for $-1 \leq t \leq 0$.

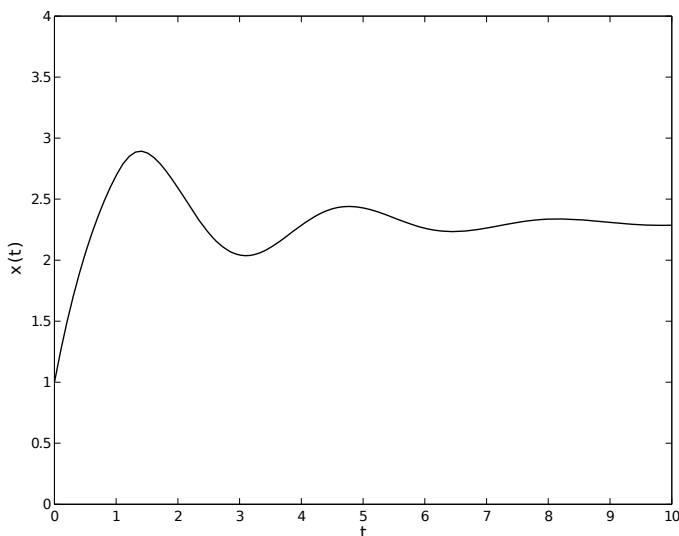


Fig. 3.5 Solution of $x'(t) = 10x(t-1)e^{-x(t-1)} - x(t)$.

Exercises

1. Find all equilibria of

$$x'(t) = \frac{rx(t-T)}{x(t-T)+A} - dx(t)$$

and determine the values of T for which they are asymptotically stable. Use a computer algebra system to graph some solutions for $A = 1$, $r = 2$, $d = 1$.

2. Find all equilibria of

$$x'(t) = \frac{r[x(t-T)]^2}{(x(t-T))^2 + A^2} - dx(t)$$

and determine the values of T for which they are asymptotically stable. Use a computer algebra system to graph some solutions for $A = 1$, $r = 2$, $d = 1$.

- 3*. (i) A population is governed by the delayed-recruitment model

$$x'(t) = x(t - T)e^{3-x(t-T)} - x(t).$$

Find all equilibria and determine their stability.

- (ii) A fraction p ($0 < p < 1$) of the population in part (a) is removed per unit time, so that the population is modeled by

$$x'(t) = x(t - T)e^{3-x(t-T)} - x(t) - px(t).$$

Find the relation between the removal fraction p and the largest delay T for which there is an asymptotically stable positive equilibrium.

- (iii) Use a computer algebra system to graph solutions with $T = 1$ for $p = 0.2$, $p = 0.5$, $p = 0.8$.
4. Use a computer algebra system to graph solutions of the differential-difference equation $x'(t) = rx(t - 1)e^{-x(t-1)} - x(t)$ with $r = 2$, $r = 10$.

3.4 Models with Distributed Delay

In Section 3.2 we suggested that the model $x'(t) = x(t)g(x(t - T))$ with the per capita growth rate depending on population size with delay T might be generalized to the form

$$x'(t) = x(t) \int_0^\infty g(x(t - s))p(s)ds,$$

describing a *distributed delay*. Here $p(s)\Delta s$ represents the probability of a delay between s and $s + \Delta s$, so that $\int_0^\infty p(s)ds = 1$. The *average delay* is then, by definition, $\int_0^\infty sp(s)ds$. An equilibrium of the integro-differential equation

$$x'(t) = x(t) \int_0^\infty g[x(t - s)]p(s)ds$$

is a value x_∞ such that

$$x_\infty \int_0^\infty g(x_\infty)p(s)ds = x_\infty g(x_\infty) = 0,$$

so that $x = 0$ is an equilibrium and equilibria $x_\infty > 0$ are given by $g(x_\infty) = 0$. To linearize about an equilibrium x_∞ , we let $u(t) = x(t) - x_\infty$, so that

$$u'(t) = (x_\infty + u(t)) \int_0^\infty g(x_\infty + u(t - s))p(s)ds,$$

and expand using Taylor's theorem, obtaining

$$\begin{aligned}
u'(t) &= (x_\infty + u(t)) \int_0^\infty (g(x_\infty) + g'(x_\infty)u(t-s) + \cdots) p(s) ds \\
&= (x_\infty + u(t)) \left(g(x_\infty) + g'(x_\infty) \int_0^\infty u(t-s) p(s) ds + \cdots \right) \\
&= x_\infty g(x_\infty) + g(x_\infty)u(t) + x_\infty g'(x_\infty) \int_0^\infty u(t-s) p(s) ds + \cdots,
\end{aligned}$$

where all quadratic and higher-order terms in u have been suppressed. The linearization at the equilibrium is thus

$$v'(t) = g(x_\infty)v(t) + x_\infty g'(x_\infty) \int_0^\infty v(t-s) p(s) ds.$$

As with other types of equations we have studied, such as differential equations, difference equations, and differential–difference equations, the behavior of solutions near an equilibrium is described by the behavior of solutions of the linearization at the equilibrium for integrodifferential equations. Therefore, we are led to study the linear integrodifferential equation

$$v'(t) = av(t) + b \int_0^\infty v(t-s) p(s) ds$$

with $p(s) \geq 0$ for $0 \leq s < \infty$ and $\int_0^\infty p(s) ds = 1$. In the special case we are considering, $a = g(x_\infty)$ and $b = x_\infty g'(x_\infty)$, so that either $a = 0$ (if $x_\infty > 0$) or $b = 0$ (if $x_\infty = 0$), but we shall consider the general case in which both a and b may be different from zero because it will arise later. It is possible to prove that solutions tend to zero as $t \rightarrow \infty$ if $b > 0$ and $a + b < 0$; or if $b < 0$, $a + b < 0$, and the average delay $\int_0^\infty s p(s) ds$ is small enough; but that solutions cannot tend to zero if $a + b \geq 0$. The necessary (but not sufficient) stability condition $a + b < 0$ is in fact the necessary and sufficient condition for asymptotic stability in the undelayed case, corresponding to an ordinary differential equation $x' = xg(x)$ with linearization

$$v' = (xg(x))'|_{x=x_\infty} v = (g(x_\infty) + x_\infty g'(x_\infty))v = (a + b)v.$$

To study the behavior of solutions of

$$v'(t) = av(t) + b \int_0^\infty v(t-s) p(s) ds$$

for a specific kernel $p(s)$, we look for solutions $v(t) = ce^{\lambda t}$ and construct a characteristic equation, obtaining

$$\begin{aligned}
\lambda ce^{\lambda t} &= ace^{\lambda t} + b \int_0^\infty ce^{\lambda(t-s)} p(s) ds, \\
\lambda &= a + b \int_0^\infty e^{-\lambda s} p(s) ds = a + bL\{p(\lambda)\},
\end{aligned}$$

where $L\{p(\lambda)\}$ denotes the Laplace transform of the function p evaluated at λ . We will consider two specific choices of p , both normalized so that $\int_0^\infty p(s)ds = 1$ and $\int_0^\infty sp(s)ds = T$, making use of the formulas

$$\int_0^\infty e^{-\alpha s} ds = \frac{1}{\alpha}, \quad \int_0^\infty se^{-\alpha s} ds = \frac{1}{\alpha^2}, \quad \int_0^\infty s^2 e^{-\alpha s} ds = \frac{2}{\alpha^3}.$$

Our first example is $p_1(s) = 4/T^2 se^{-2s/T}$, with $p_1(t)$ vanishing at zero, rising to a maximum at $s = T/2$, and then falling off exponentially. We have

$$\begin{aligned} L\{p_1(\lambda)\} &= \int_0^\infty e^{-\lambda s} p_1(s) ds = \frac{4}{T^2} \int_0^\infty se^{-(\lambda + \frac{2}{T})s} ds \\ &= \frac{4}{T^2} \frac{1}{(\lambda + \frac{2}{T})^2} = \frac{4}{T^2 \lambda^2 + 4T\lambda + 4}. \end{aligned}$$

The characteristic equation is

$$a + \frac{4b}{\lambda^2 T^2 + 4T\lambda + 4} = \lambda,$$

or

$$\lambda^3 + \left(\frac{4T - aT^2}{T^2}\right)\lambda^2 + \left(\frac{4 - 4aT}{T^2}\right)\lambda - \frac{4a + 4b}{T^2} = 0.$$

Our problem now is that we need conditions on the coefficients of a polynomial equation that ensure that all roots have negative real part. This is provided by the *Routh–Hurwitz conditions*, which for a cubic equation $\lambda^3 + \alpha\lambda^2 + \beta\lambda + \gamma = 0$ say that all roots have negative real part if and only if $\alpha > 0$, $\gamma > 0$, $\alpha\beta > \gamma$. Here, $\alpha = \frac{4-aT}{T}$, $\beta = \frac{4-4aT}{T^2}$, $\gamma = -\frac{4(a+b)}{T^2}$, and the stability conditions are $a + b < 0$, $aT < 4$, $-bt < (2 - aT)^2$. Now take $a = g(x_\infty)$, $b = x_\infty g'(x_\infty)$. If $x_\infty = 0$, so that $b = 0$, these conditions reduce to $g(x_\infty) < 0$, which is satisfied in population models only if there is an Allee effect. If $x_\infty > 0$ so that $a = 0$ the stability conditions reduce to $0 < -x_\infty g'(x_\infty)T < 4$, quite similar to the single-delay case, where the stability condition was $0 < -x_\infty g'(x_\infty)T < \pi/2$.

Our second example is $p_2(s) = 1/Te^{-s/T}$ with p_2 , decreasing exponentially from $1/T$ to zero, rather than rising to a maximum like $p_1(s)$. We have

$$L\{p_2(\lambda)\} = \int_0^\infty e^{-\lambda s} p_2(s) ds = \frac{1}{T} \int_0^\infty e^{-(\frac{1}{T} + \lambda)s} ds = \frac{1}{\lambda T + 1},$$

and the characteristic equation is $a + \frac{b}{\lambda T + 1} = \lambda$, or $\lambda^2 + \frac{1-aT}{T^2}\lambda - \frac{a+b}{T^2} = 0$. The stability condition that both roots of this quadratic equation have negative real part is $1 - aT > 0$, $-(a+b) > 0$. Now take $a = g(x_\infty)$, $b = x_\infty g'(x_\infty)$; if $x_\infty = 0$, these conditions reduce to $g(x_\infty) < 0$, which is not satisfied. If $x_\infty > 0$, the stability condition is just $g'(x_\infty) < 0$, since $a = 0$, exactly as if there were no delay; there is no requirement that the average delay not be too large. The point is that with distributed delay

each delay kernel must be examined in its own right. *It is not true that increasing the average delay always destroys stability.*

Delayed recruitment populations with variable maturation ages can also be described by integrodifferential equations, now of the form

$$x'(t) = \int_0^\infty \frac{1}{\pi} B(x(t-s)) p(s) ds - D(x(t)).$$

Here $B(x(t-s))$ is the number born at time $(t-s)$, reaching age s at time t ; $p(s)\Delta s$ is the fraction of those members who reach maturity, doing so at age s to $s+\Delta s$; π is the ratio of the number of members who reach maturity to the number born, so that $p(s)\Delta s/\pi$ is the fraction of members born who reach maturity at age s to $s+\Delta s$; and $D(x(t))$ is the number dying at time t . An equilibrium of this integrodifferential equation is a solution x_∞ of the equation

$$D(x_\infty) = \frac{1}{\pi} B(x_\infty) \int_0^\infty p(s) ds = \frac{1}{\pi} B(x_\infty).$$

The by-now familiar process yields the form of the linearization at an equilibrium as

$$v'(t) = -D'(x_\infty)v(t) + \frac{1}{\pi} B'(x_\infty) \int_0^\infty v(t-s)p(s) ds.$$

This is of the form

$$v'(t) = av(t) + b \int_0^\infty v(t-s)p(s) ds,$$

with characteristic equation $\lambda = a + b \int_0^\infty e^{-\lambda s} p(s) ds$ already described. As before, it can be shown that the equilibrium is asymptotically stable if $(1/\pi)B'(x_\infty) < D'(x_\infty)$ and $B'(x_\infty) > 0$ or $B'(x_\infty) < 0$ and $\int_0^\infty sp(s) ds$ is small enough, but further information depends on the form of $p(s)$.

Exercises

1. Find all equilibria of

$$x'(t) = \frac{1}{\pi} \int_0^\infty \frac{rx(t-s)}{x(t-s)+A} p(s) ds - dx(t)$$

and determine the values of T for which they are asymptotically stable if

(i)

$$p(s) = \frac{4se^{-2s/T}}{T^2},$$

(ii)

$$p(s) = \frac{e^{-s/T}}{T}.$$

2. Find all equilibria of

$$x'(t) = \frac{1}{\pi} \int_0^\infty \frac{r(x(t-s))^2}{(x(t-s))^2 + A^2} p(s) ds - dx(t)$$

and determine the values of T for which they are asymptotically stable for each of the choices of $p(s)$ in Exercise 1.

3. Find all equilibria of

$$x'(t) = \frac{1}{\pi} \int_0^\infty rx(t-s)e^{-x(t-s)}p(s)ds - dx(t),$$

and determine the value of T for which they are asymptotically stable for each of the choices of $p(s)$ in Exercise 1.

3.5 Harvesting in Delayed Recruitment Models

In Section 1.4 we investigated the effect of harvesting on a continuous population model. Here we wish to study whether a continuous model with a time lag, such as a delayed-recruitment model of the type introduced in Section 3.3, responds to harvesting in the same way as a model without a time lag. Thus we could consider a population model

$$x'(t) = R(x(t-T)) - D(x(t)) - Ex(t) \quad (3.12)$$

to model constant-effort harvesting, or a model

$$x'(t) = R(x(t-T)) - D(x(t)) - H \quad (3.13)$$

to model constant-yield harvesting.

3.5.1 Constant-Effort Harvesting

The analysis of the constant-effort harvesting model (3.12) is exactly the same as the analysis of the unharvested model with the mortality function $D(x)$ replaced by $D(x) + Ex$. Rather than carrying out the analysis in the general case, we examine only the special case $R(x) = rxe^{-x/A}$, $D(x) = dx$ considered as an example in Section 3.3. By following the analysis of this example we can see that the equilibrium $x = 0$ is asymptotically stable if and only if $r < d + E$, or $E > r - d$. If $0 \leq E < r - d$, there is a positive equilibrium x_∞ that is absolutely stable if $r < (d + E)e^2$, or $E > r/e^2 - d$. In general, increasing the effort tends to stabilize the equilibrium, but too high an

effort will wipe out the population. If $r > de^2$, so that the stability of the positive equilibrium depends on the delay T , then increasing the harvest effort tends to stabilize the equilibrium, but increasing the delay tends to destabilize it. The maximum sustainable yield is found exactly as for the undelayed model in Section 1.5, but the introduction of a time lag into the model raises the possibility of instability of the equilibrium corresponding to the maximum sustainable yield. For the special case $R(x) = rxe^{-x/A}$, $D(x) = dx$, the maximum sustainable yield is achieved when the mortality–harvest curve $y = (d + E)x$ passes through the maximum point $(A, rA/e)$ of the recruitment curve $y = rxe^{-x/A}$. This requires $\frac{rA}{e} = (d + E)A$, or $E = r/e - d$. Since $r/e - d > r/e^2 - d$, the equilibrium corresponding to maximum sustainable yield is actually absolutely stable. The behavior of more general compensation models under constant-effort harvesting is similar to the above special case. We shall not explore the more complicated situation for depensation models.

3.5.2 Constant-Yield Harvesting

The model (3.13) must be modified slightly in order to rule out the possibility that $x'(t) < 0$ when $x(t) = 0$, since negative population sizes have no biological significance. To be realistic (and also to avoid important errors in carrying out numerical approximations), we should specify the model by

$$x'(t) = \begin{cases} R(x(t-T)) - D(x(t)) - H & \text{for } x(t) \geq 0, \\ \max\{R(x(t-T)) - D(x(t)) - H, 0\} & \text{for } x(t) = 0. \end{cases} \quad (3.14)$$

Because of the time lag, it is possible for $x(t)$ to reach zero without the collapse of the system; even if the adult population is wiped out, there may still be immature members who will reach the adult stage. Of course, if $x(t) = 0$ on a time interval of length T , there is no possibility of regeneration of the adult population, and we shall consider the population to have become extinct. Under the hypotheses made in Section 3.3 on the recruitment function $R(x)$ and the mortality function $D(x)$, it is possible to show that the bounds on solutions obtained in Section 3.3 are also valid for the constant-yield harvesting model (3.14). An equilibrium of the model (3.14) is defined to be a value x_∞ for which

$$R(x_\infty) = D(x_\infty) + H,$$

which is the abscissa of an intersection of the recruitment curve $y = R(x)$ and the mortality–harvesting curve $y = D(x) + H$. We may sometimes use the notation $x_\infty(H)$ to emphasize the dependence on H . The effect of increasing the harvest rate H is to move the mortality–harvesting curve upward. For sufficiently large H , the mortality–harvesting curve is above the recruitment curve, and there are no equilibria. It is possible to prove that if the model (3.14) has no equilibrium $x_\infty > 0$, then every solution tends to zero as $t \rightarrow \infty$ if $H = 0$ and reaches zero in finite time and

remains at zero if $H > 0$. There is a critical harvest rate H_c such that there is an equilibrium $x_\infty(H) > 0$ if $H < H_c$ but there is no equilibrium $x_\infty(H)$ if $H > H_c$ (Figure 3.6). When $H = H_c$ the recruitment curve $y = R(x)$ and the mortality–harvesting curve $y = D(x) + H$ are tangent at $x_\infty(H_c)$.

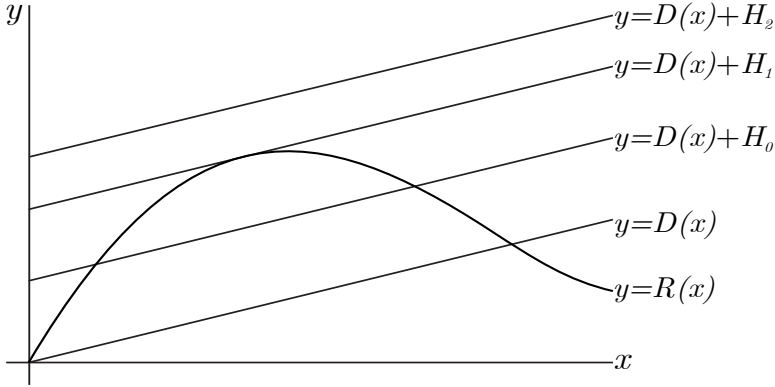


Fig. 3.6 Equilibria under harvesting.

The linearization of (3.14) at an equilibrium x_∞ is

$$u'(t) = -D'(x_\infty)u(t) + R'(x_\infty)u(t - T),$$

which is exactly the same as in the case of no harvesting studied in Section 3.3. The same analysis of the characteristic equation as that in Section 3.3 leads to the following conclusions:

- (i) An equilibrium x_∞ with $|R'(x_\infty)| < D'(x_\infty)$ is absolutely stable.
- (ii) An equilibrium x_∞ with $-R'(x_\infty) > D'(x_\infty) > 0$ is asymptotically stable for small T and unstable for large T .
- (iii) An equilibrium x_∞ with $R'(x_\infty) > D'(x_\infty) > 0$ is unstable for all T .

These conclusions are valid provided $R(x)$ and $D(x)$ satisfy the hypotheses of Section 3.3, namely

$$D(0) = 0, \quad D'(x) \geq 0, \quad \lim_{x \rightarrow \infty} D(x) = +\infty,$$

and either

$$R(0) = 0, \quad R'(x) \geq 0 \text{ for } x \leq \xi, \quad R'(x) \leq 0 \text{ for } x \geq \xi, \quad \lim_{x \rightarrow \infty} R(x) = 0$$

or

$$R(0) = 0, \quad R'(x) \geq 0 \text{ for } 0 \leq x < \infty, \quad \lim_{x \rightarrow \infty} R(x) = r < \infty.$$

Under these hypotheses, $D(x) > R(x)$ for large x , and this implies that if there are any positive equilibria, there must be a largest equilibrium x_∞ for which $D'(x_\infty) >$

$R'(x_\infty)$, and this equilibrium must be asymptotically stable at least for sufficiently small T . Because

$$R'(x_\infty(H_c)) = D'(x_\infty(H_c)) > 0$$

the equilibrium $x_\infty(H)$ must be absolutely stable when H is close to H_c . Thus, the effect of harvesting is to stabilize a population system, even one that may behave chaotically without harvesting. However, overharvesting can lead to catastrophe and the wiping out of the population. Figure 3.7 shows the solution of the differential-difference equation $x'(t) = 10x(t-1)e^{-x(t-1)} - x(t) - 2$ with $x(t) = 1$ for $-1 \leq t \leq 0$. The effect of harvesting becomes clear when Figure 3.7 is compared to Figure 3.5.

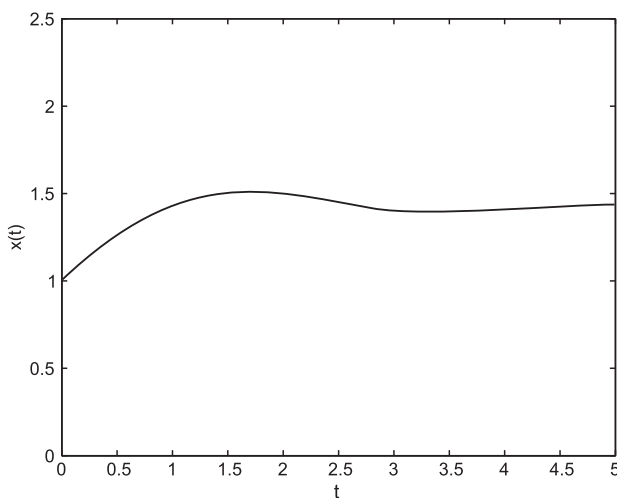


Fig. 3.7 Solution of $x'(t) = 10x(t-1)e^{-x(t-1)} - x(t) - 2$.

Exercises

1. A population under harvesting is governed by the delayed-recruitment model

$$x'(t) = x(t-T)e^{3-x(t-T)} - x(t) - H.$$

Find the relation between the harvesting rate H and the largest delay T for which there is an asymptotically stable positive equilibrium. Use a computer algebra system to graph the solutions with $T = 1$ and with $H = 1$, $H = 5$, $H = 10$.

2. Find the critical harvest rate H_c for the model

$$x'(t) = rx(t-T)e^{-x(t-T)/A} - dx(t) - H.$$

3*. For the general model

$$x'(t) = R(x(t-T)) - D(x(t)) - Ex(t),$$

show that the equilibrium $x = 0$ is unstable if $0 \leq E < R'(0) - D'(0)$, and absolutely stable if $E > R'(0) - D'(0)$.

3.6 Nicholson's Blowflies: A Case Study

In 1954, A.J. Nicholson conducted some experiments on the Australian sheep blowfly (*Lucillia cuprina*) in which he observed oscillations of large amplitude in the population size. In these experiments the population was controlled by a single factor, either the rate at which food was supplied to the adult population (in which case the adult population showed oscillations with almost discrete generations and there was a pattern of breeding activity with periods of essentially zero reproduction alternating with relatively long bursts of continuous reproductive activity) or the rate at which food was supplied to the larvae (in which case each burst of reproductive activity showed two clearly discrete generations, and this reproductive activity occurred in alternate population cycles). It is reasonable to ascribe the pattern of oscillations to some kind of delay mechanism appearing in the per capita growth rate as described in Section 3.2 or in the recruitment rate as described in Section 3.3. Which mechanism is more plausible as an explanation for the oscillations? Suppose we attempt to use the delay logistic equation

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

to describe this population, as proposed by May (1974). It is convenient to rescale the model in terms of dimensionless variables. We first let $x = Ky$ to transform the model to $y'(t) = ry(t)(1 - y(t-T))$. Then we make the change of independent variable $t = Ts$ and let $y(t) = z(s)$, so that $y(t-T) = z(s-1)$ and $y'(t) = (dy/ds)(ds/dt) = (1/T)z'(s)$ to obtain the dimensionless form

$$z'(s) = (rT)z(s)(1 - z(s-1))$$

containing a single free parameter rT . For this differential-difference equation there is an equilibrium of $z = 1$, and the linearization at this equilibrium is

$$u'(s) = -rTu(s-1).$$

We know that the equilibrium $z = 1$ is asymptotically stable if $rT < \pi/2$ and that there is a stable periodic solution if $rT > \pi/2$. Numerical simulation for $rT > \pi/2$ indicates relationships among the value of the parameter rT , the ratio of maximum to minimum population size, and the period of the periodic solution. From the observed ratio of maximum to minimum population size, May estimated $rT \approx 2.1$,

which implies a period of 4.54 in dimensionless units, or $4.54T$ in days. Then he deduced from the observed period that the time delay was approximately 9 days, compared to the experimentally observed larva-to-adult maturation time of 11 days. From these data, he concluded that the agreement between observation and the rather crude model used was sufficiently good to indicate that delay in the per capita growth rate might explain the observed behavior.

In 1980, May's calculations were reexamined by W.S.C. Gurney, S. P. Blythe, and R.M. Nisbet (1980). They observed a maximum to minimum ratio of population sizes in Nicholson's adult-food-limited experiment (the only one for which May made calculations) of 36 ± 17 , which would imply $1.880 < rT < 2.035$ (the maximum to minimum population size ratio is difficult to measure accurately, but rT is not very sensitive to changes in this ratio), and that the period of oscillations in the solution is between $4.26T$ and $4.47T$. The fact that the observed period was 38.1 ± 1.5 days implies that the delay T must be between $36/4.47 = 8.2$ days and $39.6/4.26 = 9.3$ days. This is really not very close to the observed delay, for Nicholson observed an egg to reproductive maturity time of between 14.3 and 15.6 days, and since our model describes adult population size this is the appropriate delay. The data also give an independent verification of this delay, since there are clear bursts of reproduction preceding each adult population peak by 14.8 ± 0.4 days. Since egg-to-adult survival is high (approximately 90 percent) and constant in the adult-food-limited experiments, this lag should be the same as the delay T . The discrepancy between a predicted delay of 8.2 to 9.3 days and the observed delay of 14.3 to 15.6 days indicates that the phenomenon observed cannot be explained satisfactorily by use of the delay logistic equation as a model.

A more general model of the form

$$x'(t) = x(t)g(x(t-T))$$

would be no more satisfactory than the delay logistic equation, because linearization about a positive equilibrium would produce the same linearized equation with the same parameters. Thus we conclude that delay in per capita growth rates does not explain the observed behavior, and we are led to search for a more satisfactory model. In fitting data to experiments, more accurate measurements often indicate discrepancies not disclosed by rough measurements. In searching for a model, we will continue to assume that the oscillations are the result of some delay mechanism. While our model will be somewhat more complicated than the delay logistic equation, we would like to have only a small number of controlling parameters. A model with only one controlling parameter in dimensionless variables will always have a linearization about equilibrium either of the form $z'(s) = az(s)$, whose solutions are exponential functions, or of the form $z'(s) = bz(s-1)$, which has been tried and found wanting in our examination of the delay logistic equation. Thus we must expect at least two controlling parameters in a satisfactory model. Because the time required for eggs to mature to adults appears to be related to the oscillatory behavior, it is natural to consider a delayed recruitment model of the form

$$x'(t) = R(x(t-T)) - D(x(t)),$$

in which T is the maturation time, $R(x)$ is the number of eggs produced per unit time when the adult population size is x multiplied by the probability of survival to adulthood, and $D(x)$ is the number of adults dying per unit time when the population size is x . Data about egg production, survival to adulthood, and death rates suggest use of the functions

$$R(x) = rxe^{-x/A}, \quad D(x) = dx,$$

to give the model

$$x'(t) = rx(t-T)e^{-x(t-T)/A} - dx(t), \quad (3.15)$$

both for adult-food-limited and for larva-food-limited studies, even though the egg production rates and survival probabilities have different forms in the two cases. The changes of variable $x = Ay$, $t = Ts$, with the definition $y(t) = z(s)$ transforms the model to the dimensionless form

$$z'(s) = (rT)z(s-1)e^{-z(s-1)} - (dT)z(s), \quad (3.16)$$

having two controlling parameters rT and dT . We will assume $r > d$ to allow the possibility of a positive equilibrium z_∞ obtained by solving $rTe^{-z} = dT$, so that $z_\infty = \log(r/d) > 0$. The linearization about the equilibrium z_∞ is

$$u'(s) = dT \left(1 - \log \frac{r}{d} \right) z(s-1) - dTz(s),$$

and we can describe conditions for the stability of the equilibrium z_∞ in terms of the values of rT and dT . In fact, we can plot in the (dT, rT) -plane the regions of exponential asymptotic stability, oscillatory asymptotic stability, and instability with a stable periodic solution, much as we plotted these regions in the (aT, bT) plane for the equation $u'(t) = au(t) + bu(t-T)$ in Section 3.3. With the aid of numerical simulations we can go further: In the region of instability we can plot the contours along which the ratio of maximum to minimum population size is constant and the contours along which the ratio of period of the solution to delay is constant. The confidence limits for these two observable quantities give four contours that define a confidence region in the (dT, rT) -plane and thus estimate rT and dT . The estimates used earlier (36 ± 17 and 4.6 ± 0.1 , respectively) lead us to the estimates $rT = 150 \pm 70$, $dT = 2.9 \pm 0.5$.

We may also estimate r and d independently. During the portion of each cycle when the population size is approaching its minimum, the rate of adult recruitment is close to zero. Thus, on these intervals the model is essentially $z'(s) = -dz(s)$, and $z(s)$ behaves like ce^{-ds} . If we plot $\log z$ against s on these intervals, then d is the negative of the slope of the “best straight line” in these plots; this procedure leads to the estimate $d = 0.27 \pm 0.025$, $dT = 4.0 \pm 0.5$, which does not quite agree with the model prediction but at least does not disagree violently. No doubt it would be possible to fudge the data to give agreement. We may estimate r from the fact that the maximum rate of egg production multiplied by the constant survival proba-

bility divided by the population size at maximum production is r/e (the maximum value of the function $rx e^{-x}$ attained for $x = 1$). From observation of production rates we estimate $7.4 < r < 14$, $100 < rT < 160$, and this is quite compatible with the model estimate $80 < rT < 220$. We conclude that the delayed-recruitment model is a plausible description and may use the model to make further predictions about the behavior of the population. Not only do we obtain population oscillations very similar to those observed experimentally, but we also deduce a pattern of breeding activity with bursts of continuous activity alternating with periods of essentially zero reproduction. If we fit the same model to the observed data for the larva-food-limited experiments, we obtain $rT = 170 \pm 50$, $dT = 5.5 \pm 1.6$ for one run and $rT = 380 \pm 250$, $dT = 4.7 \pm 0.7$ for a second. Both give separate bursts of reproduction with two clearly discrete generations, different from the prediction of the same model for the adult food-limited experiment but in accordance with the observed differences between the two experiments. This reinforces our acceptance of the model. There is another possible explanation in both models, namely that there is in fact an asymptotically stable equilibrium that is approached (possibly very slowly) in an oscillatory manner. However, for the delay logistic model, if this were the case, the period of oscillation could not be less than $4T$ for any value of the controlling parameter rT , and this is quite incompatible with observation. For the delayed-recruitment model the asymptotic stability hypothesis leads to the estimates $rT = 23.5 \pm 4.5$, $dT = 30 \pm 0.7$ for the controlling parameters. The estimate for rT is radically different from the independent estimate, and thus we may reject this explanation for either model. Figure 3.8 shows the solution of the delayed recruitment model with the parameter values $r = 10$, $d = 0.2$, $T = 15$ and with $x(t) = 1$ for $-1 \leq t \leq 0$, which are close to our estimates for the adult food-limited experiment.

When we transformed the delayed-recruitment model (3.15) to the dimensionless form (3.16) we were following a procedure that is often useful in constructing and analyzing models. In (3.15) the population size x is measured in number of members or in biomass, and the time t is measured in units of time. Let us suppose that x is measured in grams and t is measured in days. Then $x'(t)$ is measured in gram day^{-1} . Each term of the equation (3.15) must be measured in the same units. Since both x and A are population sizes measured in grams, the quotient x/A is *dimensionless*. This means that a change in units of mass will not alter the value of x/A . Then $e^{-x(t-T)/A}$ is also dimensionless. Since $x'(t)$ is measured in gram day^{-1} , each term on the right side of (3.15) must also have these dimensions. Since z is measured in grams, the coefficients r and d must have units of days^{-1} . The change of dependent variable $x = Ay$ gives a new variable y , which is dimensionless; y measures population size in units of size A . Similarly, the change of independent variable $t = Ts$ gives a dimensionless variable s measuring time in units of size T . The model (3.16) is equivalent to (3.15) but is in terms of dimensionless independent and dependent variables. It involves two parameters rT and dT , which are dimensionless. The original model (3.15) contained four parameters, r, d, T, A . In general, for every reduction from a variable to a dimensionless variable we may expect to remove one parameter. The dimensionless form (3.16) is more convenient than the form (3.15) for fitting data because its parameters are no longer coupled together.

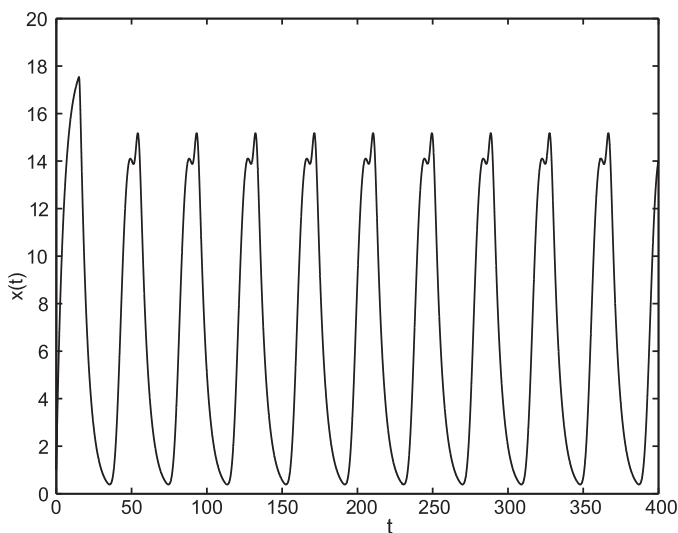


Fig. 3.8 Behavior of the delay-recruitment model.

3.7 Project: A Model for Blood Cell Populations

Many physiological systems normally display predictable patterns, either remaining almost constant or having regular oscillations. There are diseases, called dynamical diseases, that are marked by changes in these patterns such as changes in the nature of the oscillations. One area in which dynamical diseases have been observed experimentally is the blood system. Some forms of anemia and leukemia have been identified as dynamical diseases. A physiological change may cause a change in parameters of a model, which can alter the qualitative nature of the solutions.

Most types of blood cells are formed from primitive stem cells in the bone marrow, a development process that takes about six days. The cell-production process is not well understood, but it appears that the production rate should be small for small density levels and should increase to a maximum and then decrease to zero as cell density increases. One form, which has been used successfully in fitting experimental data [Mackey and Glass (1977); Glass and Mackey (1979), (1988)] is that if the cell density is y , the production rate has the form

$$p(y) = \frac{b\theta^n y}{\theta^n + y^n},$$

with positive constants b , θ , $n > 1$.

In addition, there is a steady elimination process whose rate depends on the type of cell. For granulocytes (a type of white blood cell) the cell destruction rate is about 10 percent per day. We assume that the elimination rate has the form cy . If there were no time lag in the production of cells, we would model the process by

the differential equation

$$y' = p(y) - cy = \frac{b\theta^n y}{\theta^n + y^n} - cy. \quad (3.17)$$

Question 1

Show that the change of variable $y = u\theta$ transforms (3.17) to the differential equation

$$u' = \frac{bu}{1 + u^n} - cu. \quad (3.18)$$

This means that θ can be eliminated from the model by choosing θ as the unit of measurement for blood-cell density.

Question 2

Find all equilibria of (3.18) and determine which are asymptotically stable and which are unstable. (You will need to distinguish between the cases $b > c$ and $b < c$.)

Question 3

Use a computer algebra system to graph some solutions of (3.18) with the parameter values

$$c = 0.1 \text{ day}^{-1}, \quad b = 0.2 \text{ day}^{-1}, \quad n = 10$$

(which are appropriate for granulocytes in healthy humans).

You should recognize that the model (3.18) cannot produce irregular behavior for any choice of the parameters b and c , and thus cannot explain dynamical disease. In fact, it cannot even produce regular oscillations. By overlooking the time lag in the production process, we have lost an essential feature of the system. This suggests using a differential–difference equation model

$$u'(t) = \frac{bu(t - \tau)}{1 + [u(t - \tau)]^n} - cu(t), \quad (3.19)$$

with a delay τ corresponding to the cell-development process.

Question 4

Find all equilibria of (3.19), and show that the equilibrium $u = 0$ is asymptotically stable if $b < c$ and unstable if $b > c$.

While it would be possible to obtain a stability criterion for the positive equilibrium of (3.19) if $b > c$, this is quite complicated technically. Instead, we will use a computer algebra system to learn about the behavior of solutions; we may use the resources mentioned in Section 3.1.

Question 5

Use the parameter values

$$c = 0.1 \text{ day}^{-1}, \quad b = 0.2 \text{ day}^{-1}, \quad n = 10, \quad \tau = 6 \text{ days}$$

and sketch some solutions of (3.19).

In chronic myelogeneous leukemia, it is thought that the production time of blood cells increases, and that this may lead to irregular behavior.

Question 6

Use the parameter values

$$c = 0.1 \text{ day}^{-1}, \quad b = 0.2 \text{ day}^{-1}, \quad n = 10, \quad \tau = 20 \text{ days}$$

and sketch some solutions of (3.19).

In order to avoid the complications that arise in the analysis of differential–difference equations, one might try to use a difference equation model to incorporate a time lag. If the production and elimination process occurred at discrete intervals, we might try to use a model

$$u_{k+1} = u_k + \frac{\beta u_k}{1 + u_k^n} - \xi u_k = (1 - \xi)u_k + \frac{\beta u_k}{1 + u_k^n}. \quad (3.20)$$

Now the production delay τ corresponds to one time unit in the difference equation (3.20).

Question 7

Show that to make (3.20) correspond to the differential–difference equation (3.19), we should take $\beta = \tau b$. Show also that we should take $\xi = 1 - (1 - c)^\tau$. (Elimination of a fraction c per day gives a remainder of $(1 - c)^\tau$ after τ days.)

Question 8

Find all equilibria for (3.20) and obtain conditions in terms of β , ξ , and n for their asymptotic stability.

Question 9

Determine whether the positive equilibrium of (3.20) is asymptotically stable for the parameter values corresponding to

$$c = 0.1 \text{ day}^{-1}, \quad b = 0.2 \text{ day}^{-1}, \quad n = 10, \quad \tau = 6 \text{ day}$$

and also for the parameter values corresponding to

$$c = 0.1 \text{ day}^{-1}, \quad b = 0.2 \text{ day}^{-1}, \quad n = 10, \quad \tau = 20 \text{ day}.$$

Use a computer algebra system to sketch some solutions for each of these sets of parameter values, and compare with the sketches obtained in Questions 5 and 6.

3.8 Project: Some Epidemic Models

We formulate several epidemic models for which the infective period has a fixed length T and for which new infections occur at a rate βSI (mass action), where S is the number of susceptible individuals and I is the number of infective individuals. There are no disease deaths, and the total population size is a constant N . Then I satisfies the equation

$$I'(t) = \beta S(t)I(t) - \beta S(T - T)I(t - T);$$

a more convenient form for this equation is

$$I(t) = \int_{t-T}^t \beta S(x)I(x)dx. \quad (3.21)$$

We consider the following possibilities;

1. Infective individuals recover with no immunity against reinfection. These individuals return to the susceptible class. Then $S = N - I$, and the model consists of the single equation (3.21) with S replaced by $(N - I)$.
2. Infective individuals all die of the disease, and the model is given by the equation (3.21) together with

$$S' = -\beta SI. \quad (3.22)$$

3. Infective individuals all recover and go into a removed class R . The model is given by the equations (3.21), (3.22), together with a third equation for R , namely

$$R'(t) = \beta S(t - T)I(t - T). \quad (3.23)$$

For each of these three different scenarios, determine the asymptotic behavior of the model and compare the behaviors of the three models. Be careful; there may be some cases in which this depends on the values of the parameters β, T .

3.9 Project: A Neuron Interaction Model

There is a coupling of particular pairs of neurons in recurrent synaptic feedback in the vertebrate nervous system. Neuron 1 sends an action potential down its axon, which branches so that one neighbor connects with a near neighbor, neuron 2. Some of the synapses of neuron 2 impinge on neuron 1 and may either excite or inhibit it. We give a model of this [Plant (1981)] based on the Fitzhugh–Nagumo equations.

These describe the time dependence of the action potential $v(t)$. There is also an inactivation variable $w(t)$ necessary to bring the membrane potential back to a resting value after the action potential is stimulated. The equations contain a feedback term depending on the deviation of v from v_0 at an earlier time and have the form

$$v'(t) = v - \frac{v^3}{3} + w + m[v(t-T) - v_0], \quad (3.24)$$

$$w'(t) = r(v + a - bw), \quad (3.25)$$

where v_0 satisfies

$$\frac{v_0^3}{3} - v_0 \left(\frac{1}{b} + 1 \right) - \frac{a}{b} = 0.$$

In the original model without delay, the relaxation time of $v(t)$ is of order $1/br$, and we assume that T exceeds the typical relaxation time $1/br$. The equations linearized about (v_0, w_0) are

$$x'(t) = (1 - v_0^2)(x(t) - y(t) + mx(t-T)), \quad (3.26)$$

$$y'(t) = r[x(t) - by(t)]. \quad (3.27)$$

Our goal is to demonstrate that there can be oscillatory solutions for both excitatory ($m > 0$) and inhibitory ($m < 0$) feedback.

1. Obtain the equations satisfied by v_0, w_0 when $a = 0.8, b = 0.7, r = 0.08$.
2. Obtain the characteristic equation for $T = 0$.
3. Show that the equation

$$\frac{v_0^3}{3} - v_0 \left(\frac{1}{b} + 1 \right) - \frac{a}{b} = 0$$

with the given values of a, b, r has one negative root and estimate its value.

4. Obtain the characteristic equation for $T > 0$ and show that if $T > 1/rb$, there is a positive real root if m exceeds a certain positive value m_0 . [This shows that any crossing with $T > 1/rb$ cannot mean a stability change unless $m < m_0$.]

Part II

Models for Interacting Species

Chapter 4

Introduction and Mathematical Preliminaries

4.1 The Lotka–Volterra Equations

In the 1920's Vito Volterra was asked whether it would be possible to explain the fluctuations that had been observed in the fish population of the Adriatic sea—fluctuations that were of great concern to fishermen in times of low fish populations. Volterra (1926) constructed the model that has become known as the Lotka–Volterra model (because A.J. Lotka (1925) constructed a similar model in a different context about the same time), based on the assumptions that fish and sharks were in a predator–prey relationship.

Here is a description of the model suggested by Volterra. Let $x(t)$ be the number of fish and $y(t)$ the number of sharks at time t . We assume that the plankton, which is the food supply for the fish, is unlimited, and thus that the per capita growth rate of the fish population in the absence of sharks would be constant. Thus, if there were no sharks the fish population would satisfy a differential equation of the form $dx/dt = \lambda x$. The sharks, on the other hand, depend on fish as their food supply, and we assume that if there were no fish, the sharks would have a constant per capita death rate; thus, in the absence of fish, the shark population would satisfy a differential equation of the form $dy/dt = -\mu y$. We assume that the presence of fish increases the shark growth rate, changing the per capita shark growth rate from $-\mu$ to $-\mu + cx$. The presence of sharks reduces the fish population, changing the per capita fish growth rate from λ to $\lambda - by$. This gives the *Lotka–Volterra equations*

$$\frac{dx}{dt} = x(\lambda - by), \quad \frac{dy}{dt} = y(-\mu + cx). \quad (4.1)$$

We cannot solve this system of equations analytically, but we can obtain some information about the behavior of its solutions. Instead of trying to solve for x and y as functions of t , we eliminate t and look for the relation between x and y . In geometric terms, we study the (x, y) *phase plane*. We look for *orbits*, or *trajectories* of solutions (that is, curves in the phase plane representing the functional relation between x and y with the time t as the parameter). We may eliminate t from the

Lotka–Volterra equations by division:

$$\frac{dy/dt}{dx/dt} = \frac{dy}{dx} = \frac{y(-\mu + cx)}{x(\lambda - by)}.$$

We may solve this differential equation by separation of variables:

$$\begin{aligned} \int \frac{-\mu + cx}{x} dx &= \int \frac{\lambda - by}{y} dy, \\ -\mu \log x + cx &= \lambda \log y - by + h, \end{aligned}$$

where h is a constant of integration, or

$$-\mu \log x - \lambda \log y + cx + by = h.$$

The minimum value of the function

$$V(x, y) = -\mu \log x - \lambda \log y + cx + by$$

is obtained by setting $\partial V / \partial x = 0$, $\partial V / \partial y = 0$. Then $c - \mu/x = 0$, $b - \lambda/y = 0$, or $x = \mu/c$, $y = \lambda/b$. This is an equilibrium of the Lotka–Volterra system (i.e., a constant solution $x \equiv \mu/c = x_\infty$, $y \equiv \lambda/b = y_\infty$), which may also be described by the equation

$$\begin{aligned} V(x_\infty, y_\infty) &= h_0 = -\mu \log x_\infty - \lambda \log y_\infty + cx_\infty + by_\infty \\ &= -\mu \log \frac{\mu}{c} - \lambda \log \frac{\lambda}{b} + \mu + \lambda. \end{aligned}$$

Every orbit of the system is given implicitly by an equation $V(x, y) = h$ for some constant $h \geq h_0$, which is determined by the initial conditions. We make the change of variable $x = x_\infty + u = \mu/c + u$, $y = y_\infty + v = \lambda/b + v$, obtaining

$$V(x, y) = -\mu \log \left(\frac{\mu}{c} + u \right) - \lambda \log \left(\frac{\lambda}{b} + v \right) + c \left(\frac{\mu}{c} + u \right) + b \left(\frac{\lambda}{b} + v \right) = h.$$

We observe that

$$\log \left(\frac{\mu}{c} + u \right) = \log \frac{\mu}{c} + \log \left(1 + \frac{cu}{\mu} \right),$$

and if $h - h_0$ is small, we may use the approximation $\log(1 + x) \approx x - x^2/2$ to approximate this expression by

$$\log \frac{\mu}{c} + \frac{cu}{\mu} - \frac{c^2 u^2}{\mu^2}.$$

Similarly, we may approximate $\log(\lambda/b + v)$ by $\log \lambda/b + bv/\lambda - b^2 v^2/\lambda^2$. Then the orbits $V(x, y) = h$ are approximated by

$$-\mu \log \frac{\mu}{c} - cu + \frac{c^2}{\mu} u^2 - \lambda \log \frac{\lambda}{b} - bv + \frac{b^2}{\lambda} v^2 + \mu + cu + \lambda + bv = h,$$

or

$$\frac{c^2}{\mu} u^2 + \frac{b^2}{\lambda} v^2 = h + \mu \log \frac{\mu}{c} + \lambda \log \frac{\lambda}{b} - \mu - \lambda = h - h_0,$$

which represents an ellipse (if $h > h_0$) with the equilibrium (x_∞, y_∞) as its center. This shows that for $h - h_0$ small and positive, the orbits are closed curves around the equilibrium; since the solutions run around closed orbits, they must be *periodic*. Thus the Lotka–Volterra model predicts the fluctuations that had been observed experimentally. It is possible to show that the period of oscillation is approximately $2\pi/\lambda\mu$, and it is easy to see from the phase portrait that the maximum prey population comes one quarter of a cycle before the maximum predator population (Figure 4.1).

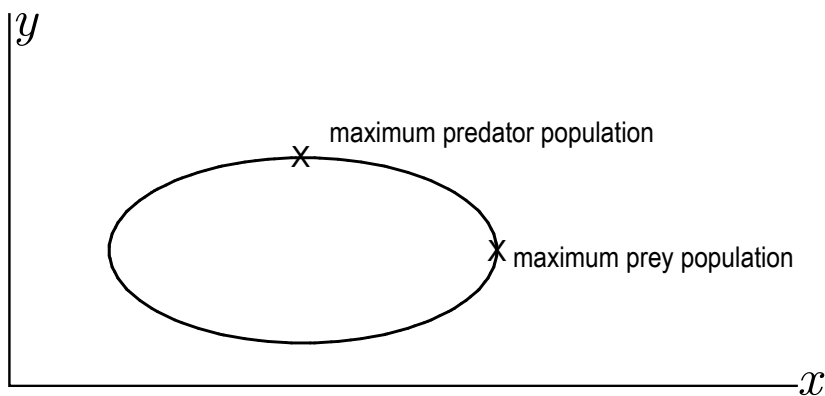


Fig. 4.1

The Lotka–Volterra model represented one of the triumphs of early attempts at mathematical modeling in population biology. However, it turns out that there are serious flaws in the model. Any attempt at refinement by introducing self-limiting terms in the per capita growth rates such as in the logistic equation for single populations will lead to qualitatively different behavior of the solutions, orbits that spiral in toward the equilibrium rather than periodic orbits. The price of refinement of the model is loss of agreement with observation. In order to construct a model that predicts periodic solutions, we will have to assume nonlinear per capita growth rates for the two species.

Example 1. Describe the orbits of the system

$$x' = y, \quad y' = -x$$

Solution. If we consider y as a function of x , we have

$$\int y dy = - \int x dx \quad \frac{y^2}{2} = -\frac{x^2}{2} + c.$$

Thus every orbit is a circle $x^2 + y^2 = 2c$ with center at the origin. Alternatively, we may define the function $V(x, y) = x^2 + y^2$ and calculate

$$\frac{d}{dt}V[x(t), y(t)] = 2x \frac{dx}{dt} + 2y \frac{dy}{dt} = 2xy - 2xy = 0$$

to see that $V(x, y) = x^2 + y^2$ is constant on every orbit.

Exercises

In each of Exercises 1 through 4, describe the orbits of the given system.

1. $x' = xy^2, y' = yx^3$.
2. $x' = e^{-y}, y' = e^x$.
3. $x' = \sin y, y' = x$.
4. $x' = xe^y, y' = xye^{-x}$.
5. Show that the orbits of the Lotka–Volterra system (4.1) are traversed in a counterclockwise direction as t increases. [*Hint:* For a point on the orbit with $\lambda/b, x > \mu/c$ we have $x' = 0$ and $y' > 0$. Thus, at this point x is a maximum and y is increasing.]

4.2 The Chemostat

A chemostat is a piece of laboratory apparatus used to cultivate bacteria. It consists of a reservoir containing a nutrient, a culture vessel in which the bacteria are cultivated, and an output receptacle. Nutrient is pumped from the reservoir to the culture vessel at a constant rate, and the bacteria are collected in the receptacle by pumping the contents of the culture vessel out at the same constant rate. The process is called a *continuous culture of bacteria*, in contrast to a *batch culture*, in which a fixed quantity of nutrient is supplied and bacteria are harvested after a growth period. We wish to describe the behavior of the chemostat by modeling the number of bacteria and the nutrient concentration in the culture vessel, and we shall sketch the classical theory of the simple chemostat due to Novick and Szilard (1950) and Monod (1950). We will obtain a model for two interacting populations that describes a laboratory realization of a very simple lake. More complicated chemostats, in which two or more cultures are introduced, give multispecies models representing more complicated real-world situations. We let y represent the number of bacteria and C the concentration of nutrient in the chemostat, both functions of t . Let V be the vol-

ume of the chemostat and Q the rate of flow into the chemostat from the nutrient reservoir and also the rate of flow out from the chemostat. The fixed concentration of nutrient in the reservoir is a constant $C^{(0)}$. We assume that the average per capita bacterial birth rate is a function $r(C)$ of the nutrient concentration and that the rate of nutrient consumption of an individual bacterium is proportional to $r(C)$, say $\alpha r(C)$. Then the rate of change of population size is the birth rate $r(C)y$ of bacteria minus the outflow rate Qy/V . It is convenient to let $q = Q/V$, so that this outflow rate becomes qy . The rate of change of nutrient volume is the replenishment rate $QC^{(0)}$ minus the outflow rate QC minus the consumption rate $\alpha r(C)y$. This gives the pair of differential equations

$$\begin{aligned}\frac{dy}{dt} &= r(C)y - qy, \\ \frac{d(CV)}{dt} &= Q(C^{(0)} - C) - \alpha r(C)y.\end{aligned}$$

We divide the second equation by the constant V and let $\beta = \alpha/V$ to give the system

$$\begin{aligned}y' &= r(C)y - qy, \\ C' &= q(C^{(0)} - C) - \beta r(C)y.\end{aligned}\tag{4.2}$$

The system (4.2) describes the chemostat, but we must still make some plausible assumptions on the dependence of the bacterial birth rate (or equivalently the nutrient consumption rate) on the nutrient concentration C . It is reasonable to assume that the function $r(C)$ is zero if $C = 0$ and that it saturates (approaches a limit) when C becomes large. The simplest function with these properties is

$$r(C) = \frac{aC}{C+A},$$

where a and A are constants, and this was the choice originally made by Monod. The explicit chemostat model is now

$$\begin{aligned}y' &= \frac{aCy}{C+A} - qy, \\ C' &= q(C^{(0)} - C) - \frac{\beta aCy}{C+A},\end{aligned}\tag{4.3}$$

where a, A, q , and β are constants. Of these constants, q depends on the flow rate and thus can be adjusted. We may inquire whether it is possible to adjust the flow rate or other parameters in the system so that the chemostat will settle down to an equilibrium. If the flow rate is too large, the whole culture will be washed out; a more interesting equilibrium state would involve both bacteria and nutrient in the culture. Because the system (4.3) cannot be solved analytically, we will have to study it by qualitative methods. This will be done in Section 5.2. The qualitative study of systems of differential equations is much more complicated than the study of a single differential equation. In the remainder of this chapter, we shall describe

some useful qualitative results for systems of differential equations, especially for systems of two differential equations for which much more precise results are true than are valid for systems of more than two differential equations. For the most part, we shall state results without proof. Proofs may be found in books on the qualitative theory of differential equations, for example [Brauer and Nohel (1989), Hurewicz (1958), Sánchez (1979), and Waltman (1986)].

4.3 Equilibria and Linearization

One of the main tools in studying continuous models for two interacting populations is linearization at equilibria, just as for models for single species. However, since linearization results can give information only about behavior of solutions near an equilibrium, they will not enable us to examine such questions as the existence of periodic orbits. As might be expected from our discussion of the Lotka–Volterra model, we will want to consider periodic orbits when we study predator–prey systems in Chapter 5, and at that time it will be necessary to make use of some global results not related to linearization. Here we shall describe the main results on linearization for two-dimensional systems. The basic fact is the analogue of the result established for one-dimensional systems (or first-order equations), namely that the behavior of solutions near an equilibrium is determined by the behavior of solutions of the linearization at the equilibrium. Since there is a variety of possible behaviors of solutions of the linearization, we will also have to carry out a classification of equilibria according to the behavior of solutions at the equilibrium.

We will consider populations of two interacting species with population sizes $x(t)$ and $y(t)$ respectively. As in our study of continuous single species models, we will assume that $x(t)$ and $y(t)$ are continuously differentiable functions of t whose derivatives are functions of the two population sizes at the same time. Thus, our models will be systems of two first order differential equations,

$$x' = F(x, y), \quad y' = G(x, y). \quad (4.4)$$

As in our study of single-species models, the assumptions that lead to this form neglect many factors of importance in many real populations, but the model is a useful first step and may model some real populations quite well. An *equilibrium* is a solution (x_∞, y_∞) of the pair of equations $F(x_\infty, y_\infty) = 0$, $G(x_\infty, y_\infty) = 0$. Thus an equilibrium is a constant solution of the system of differential equations. Geometrically, an equilibrium is a point in the phase plane that is the orbit of a constant solution. If (x_∞, y_∞) is an equilibrium, we make the change of variables $u = x - x_\infty$, $v = y - y_\infty$, obtaining the system

$$\begin{aligned} u' &= F(x_\infty + u, y_\infty + v), \\ v' &= G(x_\infty + u, y_\infty + v). \end{aligned}$$

Using Taylor's theorem for functions of two variables, we may write

$$\begin{aligned} F(x_\infty + u, y_\infty + v) &= F(x_\infty, y_\infty) + F_x(x_\infty, y_\infty)u + F_y(x_\infty, y_\infty)v + h_1, \\ G(x_\infty + u, y_\infty + v) &= G(x_\infty, y_\infty) + G_x(x_\infty, y_\infty)u + G_y(x_\infty, y_\infty)v + h_2, \end{aligned}$$

where h_1 and h_2 are functions that are small for small u, v in the sense that

$$\lim_{\substack{u \rightarrow 0 \\ v \rightarrow 0}} \frac{h_1(u, v)}{\sqrt{u^2 + v^2}} = \lim_{\substack{u \rightarrow 0 \\ v \rightarrow 0}} \frac{h_2(u, v)}{\sqrt{u^2 + v^2}} = 0.$$

The *linearization* of the system, obtained by using

$$F(x_\infty, y_\infty) = 0, \quad G(x_\infty, y_\infty) = 0,$$

and neglecting the higher-order terms $h_1(u, v)$ and $h_2(u, v)$, is defined to be the two-dimensional linear system

$$\begin{aligned} u' &= F_x(x_\infty, y_\infty)u + F_y(x_\infty, y_\infty)v, \\ v' &= G_x(x_\infty, y_\infty)u + G_y(x_\infty, y_\infty)v. \end{aligned} \tag{4.5}$$

The coefficient matrix of the system (4.5),

$$\begin{bmatrix} F_x(x_\infty, y_\infty) & F_y(x_\infty, y_\infty) \\ G_x(x_\infty, y_\infty) & G_y(x_\infty, y_\infty) \end{bmatrix},$$

is called the *community matrix* of the system at the equilibrium (x_∞, y_∞) . It describes the effect of the size of each species on the growth rate of itself and the other species at equilibrium. Frequently we shall write a system in the form

$$x' = xf(x, y), \quad y' = yg(x, y), \tag{4.6}$$

so that $f(x, y)$ and $g(x, y)$ are the per capita growth rates of the two species. The community matrix at equilibrium then has the form

$$\begin{bmatrix} x_\infty f_x(x_\infty, y_\infty) + f(x_\infty, y_\infty) & x_\infty f_y(x_\infty, y_\infty) \\ y_\infty g_x(x_\infty, y_\infty) & y_\infty g_y(x_\infty, y_\infty) + g(x_\infty, y_\infty) \end{bmatrix}.$$

There are four distinct kinds of possible equilibria, as follows:

(i) $(0, 0)$, with community matrix

$$\begin{bmatrix} f(0, 0) & 0 \\ 0 & g(0, 0) \end{bmatrix}.$$

(ii) $(K, 0)$ with $K > 0$, $(K, 0) = 0$, having community matrix

$$\begin{bmatrix} Kf_x(K, 0) & Kf_y(K, 0) \\ 0 & g(K, 0) \end{bmatrix}.$$

(iii) $(0, M)$ with $M > 0$, $g(0, M) = 0$, having community matrix

$$\begin{bmatrix} f(0, M) & 0 \\ Mg_x(0, M) & Mg_y(0, M) \end{bmatrix}.$$

(iv) (x_∞, y_∞) with $x_\infty > 0$, $y_\infty > 0$, $f(x_\infty, y_\infty) = 0$, $g(x_\infty, y_\infty) = 0$, having community matrix

$$\begin{bmatrix} x_\infty f_x(x_\infty, y_\infty) & x_\infty f_y(x_\infty, y_\infty) \\ y_\infty g_x(x_\infty, y_\infty) & y_\infty g_y(x_\infty, y_\infty) \end{bmatrix}.$$

We should remark that since from a biological point of view, only nonnegative population sizes are of interest, we consider only equilibria having nonnegative coordinates and we are concerned with only the first quadrant of the phase plane. In the case (iv) of coexistence of the two species, the terms $f_x(x_\infty, y_\infty)$ and $g_y(x_\infty, y_\infty)$ in the community matrix are self-regulating terms that are normally nonpositive.

The terms $f_y(x_\infty, y_\infty)$ and $g_x(x_\infty, y_\infty)$ are *interaction* terms. There are three possible sign combinations, which we shall study separately. If both interaction terms are negative, the two species are said to be in *competition*. If there is one positive and one negative interaction term, the species are said to be in a *predator-prey* relation. Such systems include herbivore-vegetationsystems, in which we may be interested mainly in the herbivore species but include its food supply in the model for greater realism. The simple chemostat modeled in the previous section is an example of this type. A system in which both interaction terms are positive is called *mutualistic*. We will consider each of these types of interaction in Chapter 5.

An equilibrium (x_∞, y_∞) is said to be *stable* if every solution $(x(t), y(t))$ with $(x(0), y(0))$ sufficiently close to the equilibrium remains close to the equilibrium for all $t \geq 0$. An equilibrium (x_∞, y_∞) is said to be *asymptotically stable* if it is stable and if, in addition, solutions with $(x(0), y(0))$ sufficiently close to the equilibrium tend to the equilibrium as $t \rightarrow \infty$. These definitions are the natural analogues of the definitions given previously for first-order equations. The analogue of the linearization theorem is also true, although the proof is more complicated; we shall state it without proof.

Theorem 4.1. *If (x_∞, y_∞) is an equilibrium of the system (4.4) and if all solutions of the linearization at the equilibrium (4.5) tend to zero as $t \rightarrow \infty$, then the equilibrium (x_∞, y_∞) is asymptotically stable.*

It might appear that stability of an equilibrium is a natural requirement for biological meaning, but we shall require asymptotic stability because of the following perturbation result, analogous to a result given for single-species models, which we state without proof.

Theorem 4.2. *Under the hypotheses of Theorem 4.1:*

(i) If

$$\frac{P(x, y)}{((x - x_\infty)^2 + (y - y_\infty)^2)^{1/2}}$$

and

$$\frac{Q(x, y)}{((x - x_\infty)^2 + (y - y_\infty)^2)^{1/2}}$$

tend to zero as $(x, y) \rightarrow (x_\infty, y_\infty)$, then solutions of the perturbed system

$$x' = F(x, y) + P(x, y), \quad y' = G(x, y) + Q(x, y)$$

starting close enough to (x_∞, y_∞) tend to (x_∞, y_∞) as $t \rightarrow \infty$.

(ii) If $|P(x, y)| \leq A$, $|Q(x, y)| \leq A$ for all y and A sufficiently small, then solutions of the perturbed system remain within KA , for some constant K , of solutions of the unperturbed system $x' = F(x, y)$, $y' = G(x, y)$ for $t \geq 0$.

The content of Theorem 4.2 is that an asymptotically stable equilibrium has biological significance, being relatively insensitive to both changes in initial population size and small additional forces.

Example 1. Find the linearization at each equilibrium of the Lotka–Volterra system (4.1)

Solution. The equilibria are the solutions of $x(\lambda - by) = 0$, $y(-\mu + cx) = 0$. Because the partial derivatives of the functions on the right side of the system are, respectively,

$$\begin{aligned} \frac{\partial}{\partial x}[x(\lambda - by)] &= \lambda - by, & \frac{\partial}{\partial y}[x(\lambda - by)] &= -bx, \\ \frac{\partial}{\partial x}[y(-\mu + cx)] &= cy, & \frac{\partial}{\partial y}[y(-\mu + cx)] &= -\mu + cx, \end{aligned}$$

the linearization at an equilibrium (x_∞, y_∞) is

$$u' = (\lambda - by_\infty)u - bx_\infty v, \quad v' = cy_\infty u + (-\mu + cx_\infty)v$$

One equilibrium is $(0, 0)$ with linearization

$$u' = \lambda u, \quad v' = -\mu v.$$

A second equilibrium, obtained by solving $\lambda - by = 0$, $-\mu + cx = 0$, is $(\mu/c, \lambda/b)$. The linearization at this equilibrium is

$$u' = -\frac{b\mu}{c}v, \quad v' = \frac{c\lambda}{b}u.$$

Exercises

In Exercises 1 through 8, find the linearization of each given system at each equilibrium

1. $x' = x - y, y' = x + y - 2$.
2. $x' = y, y' = x + y - 1$.
3. $x' = y + 1, y' = x^2 + y$.
4. $x' = y^2 - 8x, y' = x - 2$.
5. $x' = e^{-y}, y' = e^{-x}$.
6. $x' = \sin y, y' = 2x$.
7. $x' = x(\lambda - ax - by), y' = y(\mu - cx - dy)$.
8. $x' = x(\lambda - ax + by), y' = y(\mu + cx - dy)$.
9. The following two-dimensional nonlinear ordinary differential equation has been proposed as a model for cell differentiation:

$$\begin{aligned}\frac{dx}{dt} &= y - x, \\ \frac{dy}{dt} &= \frac{5x^2}{4 + x^2} - y.\end{aligned}\tag{4.7}$$

- (i) Sketch the curves $y = x$ and $y = 5x^2/(4 + x^2) = 0$ in the positive quadrant of the (x, y) -plane.
 - (ii) Determine the equilibrium points.
 - (iii) Linearize the system of differential equations (4.7) at each equilibrium point.
 - (iv) Determine the local stability of each positive equilibrium point and classify the equilibrium points.
 - (v) Use a computer algebra system to find the numerical solution of system (4.7), with initial points $(1.1, 1.2)$ and $(4.5, 3.9)$, respectively, and plot your solution.
10. [Eldestein-Keshet (1988)] In this problem we shall examine a plant-herbivore model. Let q represent the chemical state of the plant with q low meaning that the plant is toxic and q high meaning that it is good as a food source for the herbivore. Therefore, q is an index of plant quality for the herbivore. Suppose that plant quality is enhanced when herbivory is low or moderate and declines when herbivory is high. Assume that herbivores are small immobile insects with density I . Suppose further that their growth rate depends on the quality of food that they consume. The model equations are

$$\begin{aligned}\frac{dq}{dt} &= k_1 - k_2 q I (I - I_0), \\ \frac{dI}{dt} &= k_3 I \left(1 - \frac{k_4 I}{q} \right).\end{aligned}\tag{4.8}$$

- (i) Explain the equations, and suggest possible meanings for k_1, k_2, k_3, k_4 , and I_0 .
- (ii) Show that the equations can be written in the following dimensionless form:

$$\begin{aligned}\frac{dx}{d\tau} &= 1 - kxy(y-1), \\ \frac{dy}{d\tau} &= \alpha y \left(1 - \frac{y}{x}\right).\end{aligned}\tag{4.9}$$

Determine k and α in terms of the original parameters.

- (iii) Show that there is only one equilibrium.
 - (iv) Determine its stability.
11. Let S, I , and R represent the densities of individuals subject to a disease who are susceptible, infective, and recovered respectively. Suppose that recovered individuals can become susceptible again after some time. The model equations are as follows:

$$\begin{aligned}\frac{dS}{dt} &= -\beta S \frac{I}{N} + \gamma R, \\ \frac{dI}{dt} &= \beta S \frac{I}{N} - \nu I, \\ \frac{dR}{dt} &= \nu I - \gamma R,\end{aligned}$$

where $N = S + I + R$, and β, γ , and ν are the infection, loss of immunity, and recovery rates, respectively.

- (i) Reduce this model to a two-dimensional system of equations.
 - (ii) Find the equilibrium points. Is there a disease-free equilibrium (an equilibrium with $I = 0$)? Is there an endemic equilibrium (an equilibrium with $I > 0$)?
 - (iii) Determine the local stability of each of the equilibrium points found in (b).
12. Let $S(t)$ be the number of susceptible individuals at time t , $I(t)$ the number of infective individuals at time t , $V(t)$ the number of vaccinated or recovered individuals at time t , and $N(t) = S(t) + I(t) + V(t)$. Consider the model

$$\begin{aligned}S'(t) &= \mu N - \beta S \frac{I}{N} - \mu S, \\ I'(t) &= \beta S \frac{I}{N} - (\mu + \gamma)I, \\ V'(t) &= \gamma I - \mu V,\end{aligned}\tag{4.10}$$

where μ is the per capita death rate, γ is the per capita recovery rate, $\beta = pc$ with c the average number of contacts per unit time, and p is the probability of transmission per contact by a susceptible with an infective individual.

- (i) Find the basic reproductive number R_0 .
- (ii) Find the equilibria.
- (iii) Compute the stability of each equilibria.

Note that the birth rate (also known as recruitment rate) μN into the susceptible class is constant. Why?

13. Consider the model (4.10) but change the recruitment rate from the constant μN to the constant Λ . This means that a certain fixed number of individuals join or arrive into the susceptible class per unit time. The model becomes

$$\begin{aligned} S'(t) &= \Lambda - \beta S \frac{I}{N} - \mu S, \\ I'(t) &= \beta S \frac{I}{N} - (\mu + \gamma)I, \\ R'(t) &= \gamma I - \mu R, \end{aligned} \tag{4.11}$$

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

- (i) What are the units of Λ , $\beta SI/N$, μ , γ , β , μS ?
- (ii) Find the equation satisfied by $N(t) = S(t) + I(t) + R(t)$.
- (iii) Solve the equation for $N(t)$ and observe that the population size for this model is not constant.
- (iv) Show that $N(t) \rightarrow \Lambda/\mu$ as $t \rightarrow \infty$.
- (v) Let $K = \Lambda/\mu$. Consider the limiting system

$$\begin{aligned} S'(t) &= \Lambda - \beta S \frac{I}{K} - \mu S, \\ I'(t) &= \beta S \frac{I}{K} - (\mu + \gamma)I, \\ R'(t) &= \gamma I - \mu R. \end{aligned} \tag{4.12}$$

To study the dynamics of this limiting system, it is enough to consider the first two equations since $R(t) = K - S(t) - I(t)$.

- (i). Find \mathcal{R}_0
- (ii). Find all equilibria.
- (iii). Make a stability analysis of the equilibria.

Some recent results in dynamical systems guarantee that the limiting system and the original system have the same qualitative dynamics [Castillo-Chavez and Thieme (1995)].

14. Consider the model with vaccination

$$\begin{aligned}
S'(t) &= \mu N - \beta S \frac{I}{N} - (\mu + \phi)S, \\
I'(t) &= \beta S \frac{I}{N} - (\mu + \gamma)I, \\
V'(t) &= \gamma I + \phi S - \mu V.
\end{aligned} \tag{4.13}$$

Here $V(t)$ denotes the number of vaccinated or recovered members at time t .

- (i) $dN/dt = 0$. What does this imply?
 - (ii) Discuss why it suffices to study the first two equations.
 - (iii) Compute $R_0(\phi)$. What is the value of $R_0(0)$? Compare $R_0(\phi)$ with $R_0(0)$.
 - (iv) Compute the equilibria.
 - (v) Do the stability analysis for the disease-free equilibrium.
15. If vaccination strategies are incorporated for newborns, we consider that not all new births are susceptible. Suppose that the vaccination rate per capita is p ; then a newborn becomes vaccinated with probability p . The modified model is of the new form

$$\begin{aligned}
\frac{dS}{dt} &= (1-p)\mu N - \beta S \frac{I}{N} - \mu S = f_1, \\
\frac{dI}{dt} &= \beta S \frac{I}{N} - (\mu + \gamma)I = f_2, \\
\frac{dV}{dt} &= p\mu N + \gamma I - \mu V = f_3.
\end{aligned} \tag{4.14}$$

- (i) What is the Jacobian matrix of differential equation (4.14) at the disease-free equilibrium points?
- (ii) Find the corresponding eigenvalues of the matrix above.
- (iii) Find the basic reproductive numbers (R_0). Compare R_0 in model (4.10) with that in model (4.14).
- (iv) Study the stability of the disease-free equilibrium points of the model (4.14).

4.4 Qualitative Behavior of Solutions of Linear Systems

In the previous section we reduced the analysis of the stability of an equilibrium (x_∞, y_∞) of a system of differential equations

$$x' = F(x, y), \quad y' = G(x, y)$$

to the determination of the behavior of solutions of the linearization at the equilibrium

$$\begin{aligned}u' &= F_x(x_\infty, y_\infty)u + F_y(x_\infty, y_\infty)v, \\v' &= G_x(x_\infty, y_\infty)u + G_y(x_\infty, y_\infty)v.\end{aligned}\tag{4.15}$$

Next we shall analyze the various possibilities for the behavior of solutions of the two-dimensional linear homogeneous system with constant coefficients

$$\begin{aligned}x' &= ax + by, \\y' &= cx + dy,\end{aligned}\tag{4.16}$$

where a, b, c , and d are constants, in order that we might describe the behavior of solutions of the linearization at an equilibrium. We will then be able to state some refinements of Theorem 4.1 of Section 4.3 that give more specific information about the behavior of solutions near an equilibrium as determined by the community matrix at the equilibrium. We will assume throughout that $ad - bc \neq 0$. This implies that the origin is the only equilibrium of the system (4.16). If this system is the linearization at an equilibrium (x_∞, y_∞) of a nonlinear system (4.15), then this equilibrium is *isolated*, meaning that there is a disc centered at (x_∞, y_∞) containing no other equilibrium of the nonlinear system. It is convenient to use vector-matrix notation: We let \mathbf{x} denote the column vector $\begin{bmatrix} x \\ y \end{bmatrix}$, \mathbf{x}' the column vector $\begin{bmatrix} x' \\ y' \end{bmatrix}$, and A the 2×2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then, using the properties of matrix multiplication, we may rewrite the linear system $x' = ax + by$, $y' = cx + dy$ in the form

$$\mathbf{x}' = A\mathbf{x}.$$

A linear change of variable $\mathbf{x} = P\mathbf{u}$, with P a nonsingular 2×2 matrix, (which represents a rotation of the axes and a change of scale along the axes) transforms the system to $P\mathbf{u}' = AP\mathbf{u}$, or

$$\mathbf{u}' = P^{-1}AP\mathbf{u} = B\mathbf{u}.$$

This is of the same type as the original system $\mathbf{x}' = A\mathbf{x}$, and its coefficient matrix $B = P^{-1}AP$ is *similar* to A . If we can solve for \mathbf{u} , we can then reconstruct $\mathbf{x} = P\mathbf{u}$, and in fact the qualitative properties of solutions for \mathbf{u} are preserved in this reconstruction. Thus, we may describe the various possible phase portraits of $\mathbf{x}' = A\mathbf{x}$ by listing the various possible canonical forms of A under similarity and constructing the phase portrait for each possibility.

Theorem 4.3. *The matrix*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

with $\det A = ad - bc \neq 0$, is similar under a real transformation to one of

- (i) $\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$, $\lambda > \mu > 0$ or $\lambda < \mu < 0$.
- (ii) $\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$, $\lambda > 0$ or $\lambda < 0$.
- (iii) $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$, $\lambda > 0$ or $\lambda < 0$.
- (iv) $\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$, $\lambda > 0 > \mu$
- (v) $\begin{bmatrix} 0 & \beta \\ -\beta & 0 \end{bmatrix}$, $\beta \neq 0$.
- (vi) $\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$, $\alpha > 0, \beta \neq 0$ or $\alpha < 0, \beta \neq 0$.

This theorem will be proved in Section 4.6, the appendix to this chapter.

We now describe the phase portraits for each of these cases in turn.

Case (i). The transformed system is $u' = \lambda u$, $v' = \mu v$, with solution $u = u_0 e^{\lambda t}$, $v = v_0 e^{\mu t}$. If $\lambda < \mu < 0$ then u and v both tend to zero as $t \rightarrow \infty$ and $v/u = v_0 e^{(\mu-\lambda)t}/u_0 \rightarrow +\infty$. Thus, every orbit tends to the origin with infinite slope (except if $v_0 = 0$, in which case the orbit is on the u -axis), and the phase portrait is as shown in [Figure 4.2](#). If $\lambda > \mu > 0$ the portrait is the same, except that the arrows are reversed.

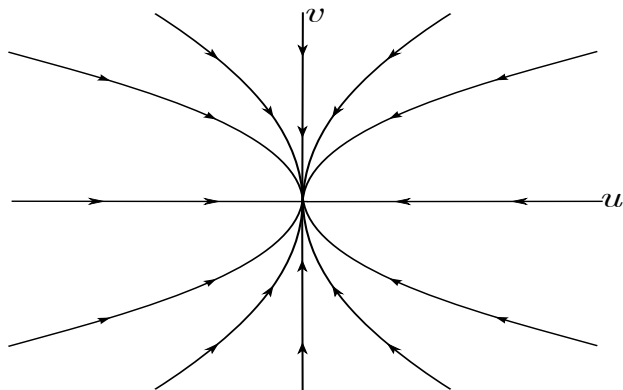


Fig. 4.2 Case i of Theorem 4.3.

Case (ii). The system is $u' = \lambda u$, $v' = \lambda v$, with solution $u = u_0 e^{\lambda t}$, $v = v_0 e^{\lambda t}$. If $\lambda < 0$ both u and v tend to zero as $t \rightarrow \infty$ and $v/u = v_0/u_0$. Thus every orbit is a straight line going to the origin, and all slopes as the orbit approaches the origin are possible. The phase portrait is as shown in [Figure 4.3](#). If $\lambda > 0$, the portrait is the same except that the arrows are reversed.

Case (iii). The system is $u' = \lambda u + v$, $v' = \lambda v$. The solution of the second equation is $v = v_0 e^{\lambda t}$, and substitution into the first equation gives the first-order linear equation

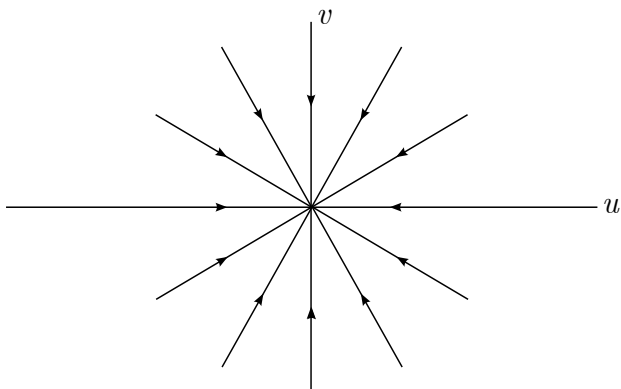


Fig. 4.3 Case ii of Theorem 4.3.

$u' = \lambda u + v_0 e^{\lambda t}$ whose solution is $u = (u_0 + v_0 t) e^{\lambda t}$. If $\lambda < 0$, both u and v tend to zero as $t \rightarrow \infty$, and $v/u = v_0/(u_0 + v_0 t)$ tends to zero unless $v_0 = 0$, in which case the orbit is the u -axis. Because $u = (u_0 + v_0 t) e^{\lambda t}$, we have $du/dt = ((u_0 \lambda + v_0) + \lambda v_0 t) e^{\lambda t}$, and therefore $du/dt = 0$ when $t = -(v_0 + u_0 \lambda)/\lambda v_0$. Thus except for the orbits on the u -axis, every orbit has a maximum or minimum u -value and then turns back toward the origin. The phase portrait is as shown in [Figure 4.4](#). If $\lambda > 0$, the portrait is the same except that the arrows are reversed.

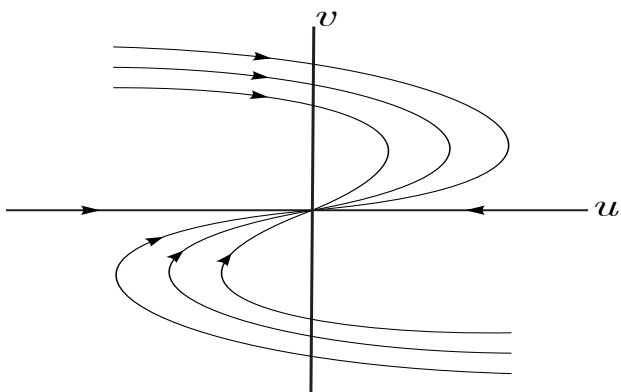


Fig. 4.4 Case iii of Theorem 4.3.

Case (iv). The solution is $u = u_0 e^{\lambda t}$, $v = v_0 e^{\mu t}$ just as in Case (i), but now u is unbounded and $v \rightarrow 0$ as $t \rightarrow \infty$, unless $u_0 = 0$, in which case the orbit is on the v -axis. The phase portrait is as shown in [Figure 4.5](#).

Case (v). The system is $u' = \beta v$, $v' = -\beta u$. Then $u'' = \beta v' = -\beta^2 u$ and $u = A \cos \beta t + B \sin \beta t$ for some A, B . Thus $v = u'/\beta = -A \sin \beta t + B \cos \beta t$ and $u^2 + v^2 =$

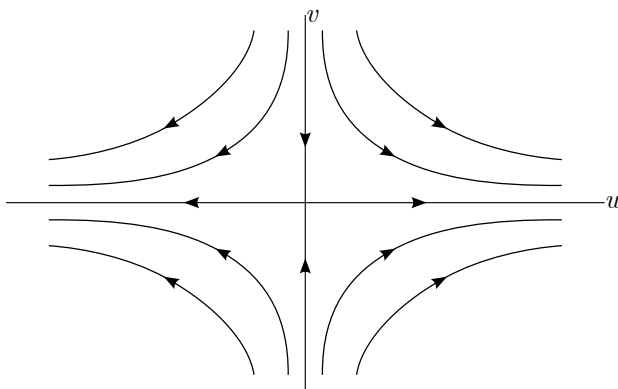


Fig. 4.5 Case iv of Theorem 4.3.

$A^2 + B^2$. Every orbit is a circle, clockwise if $\beta > 0$ and counterclockwise if $\beta < 0$. The phase portrait for $\beta > 0$ is as shown in Figure 4.6.

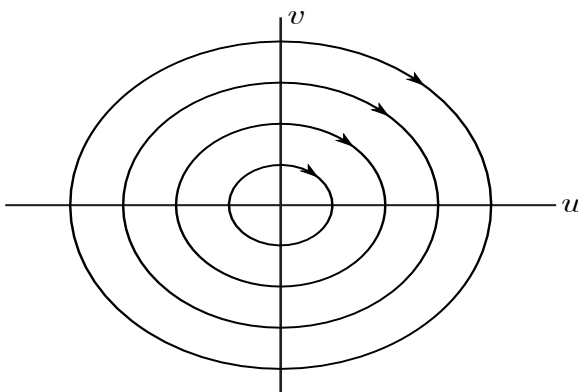


Fig. 4.6 Case v of Theorem 4.3.

Case (vi). The system is $u' = \alpha u + \beta v$, $v' = -\beta u + \alpha v$. We make the change of variables $u = e^{\alpha t} p$, $v = e^{\alpha t} q$, so that $u' = \alpha e^{\alpha t} p + e^{\alpha t} p'$, $v' = \alpha e^{\alpha t} q + e^{\alpha t} q'$ reduces the system to $p' = \beta q$, $q' = -\beta p$, which has been solved in Case (v). Thus, $u = e^{\alpha t} (A \cos \beta t + B \sin \beta t)$, $v = e^{\alpha t} (-A \sin \beta t + B \cos \beta t)$, and $u^2 + v^2 = e^{2\alpha t} (A^2 + B^2)$. If $\alpha < 0$, $u^2 + v^2$ decreases exponentially, and the orbits are spirals inward to the origin, clockwise if $\beta > 0$ and counterclockwise if $\beta < 0$. The phase portrait is as shown in Figure 4.7. If $\alpha > 0$, the portraits are the same except that the arrows are reversed.

These six cases may be classified as being of four distinct types. In Cases (i), (ii), and (iii) all orbits approach the origin as $t \rightarrow +\infty$ (or as $t \rightarrow -\infty$ depending on the

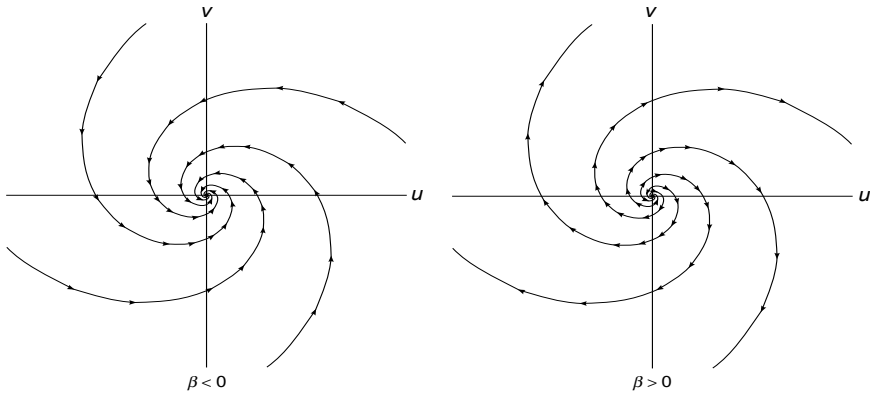


Fig. 4.7 Case vi of Theorem 4.3.

signs of λ and μ) with a limiting direction, and the origin is said to be a *node* of the system. In Case (iv), only two orbits approach the origin as $t \rightarrow +\infty$ or as $t \rightarrow -\infty$, and all other orbits move away from the origin. In this case the origin is said to be a *saddle point*. In Case (vi) every orbit winds around the origin in the sense that its angular argument tends to $+\infty$ or to $-\infty$, and the origin is said to be a *vortex*, *spiral point*, or *focus*. In Case (v), every orbit is periodic; in this case the origin is said to be a *center*. According to Theorem 4.3, asymptotic stability of the origin for the linearization implies asymptotic stability of an equilibrium of a nonlinear system. In addition, instability of the origin for the linearization implies instability of an equilibrium of a nonlinear system. The asymptotic stability or instability of the origin for a linear system is determined by the *eigenvalues* of the matrix A , defined to be the roots of the *characteristic equation*

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - (a + d)\lambda + (ad - bc) = 0. \end{aligned}$$

The sum of the eigenvalues is the *trace* of the matrix A , namely $a + d$, and the product of the eigenvalues is the determinant of the matrix A , namely $ad - bc$. A similarity transformation preserves the trace and determinant and therefore does not change the eigenvalues. Thus, the eigenvalues of A are λ, μ in Cases (i) and (iv); λ (a double eigenvalue) in Cases (ii) and (iii); the complex conjugates $\pm i\beta$ in Case (v); and $\alpha \pm i\beta$ in Case (vi). Examination of the phase portraits in the various cases shows that the origin is asymptotically stable in Case (i) if $\lambda < \mu < 0$; in Cases (ii) and (iii) if $\lambda < 0$; and in Case (vi) if $\alpha < 0$. Similarly the origin is unstable in Case (i) if $\lambda > \mu > 0$; in Cases (ii) and (iii) if $\lambda > 0$; in Case (iv); and in Case (vi) if $\alpha > 0$. The origin is stable but not asymptotically stable in Case (v) (center). A simpler description is that the origin is asymptotically stable if both eigenvalues have negative real part and unstable if at least one eigenvalue has positive real part. If both eigenvalues have real part zero, the origin is stable but not asymptotically

stable. Our assumption that $ad - bc = \det A \neq 0$ rules out the possibility that $\lambda = 0$ is an eigenvalue; thus, eigenvalues with real part zero can occur only if the eigenvalues are pure imaginary, as in Case (v).

Combining this analysis with the linearization result (4.1), we have the following result:

Theorem 4.4. *If (x_∞, y_∞) is an equilibrium of the system (4.15) and if all eigenvalues of the coefficient matrix of the linearization at this equilibrium have negative real part, specifically if*

$$\begin{aligned} \operatorname{tr} A(x_\infty, y_\infty) &= F_x(x_\infty, y_\infty) + G_y(x_\infty, y_\infty) < 0, \\ \det A(x_\infty, y_\infty) &= F_x(x_\infty, y_\infty)G_y(x_\infty, y_\infty) - F_y(x_\infty, y_\infty)G_x(x_\infty, y_\infty) > 0, \end{aligned}$$

then the equilibrium (x_∞, y_∞) is asymptotically stable.

We can be more specific about the nature of the orbits near an equilibrium. In terms of the elements of the matrix A , we may characterize the cases as follows, using the remark that the eigenvalues are complex if and only if

$$\Delta = (a + d)^2 - 4(ad - bc) = (a - d)^2 + 4bc < 0.$$

1. If $\det A = ad - bc < 0$, the origin is a saddle point.
2. If $\det A > 0$ and $\operatorname{tr} A = a + d < 0$, the origin is asymptotically stable, a node if $\Delta \geq 0$ and a spiral point if $\Delta < 0$.
3. If $\det A > 0$ and $\operatorname{tr} A > 0$, the origin is unstable, a node if $\Delta \geq 0$ and a spiral point if $\Delta < 0$.
4. If $\det A > 0$ and $\operatorname{tr} A = 0$, the origin is a center.

It is possible to show that in general, the phase portrait of a nonlinear system at an equilibrium is similar to the phase portrait of the linearization at the equilibrium, *except possibly if the linearization has a center*. This is true under the assumption that the functions $F(x, y)$ and $G(x, y)$ in the system (4.15) are smooth enough that Taylor's theorem is applicable, so that the terms neglected in the linearization process are of higher order. If the linearization at an equilibrium has a *node*, then the equilibrium of the nonlinear system is also a node, defined to mean that every orbit tends to the equilibrium (either as $t \rightarrow \infty$ or as $t \rightarrow -\infty$) with a limiting direction. If the linearization at an equilibrium has a *spiral point*, then the equilibrium of the nonlinear system is also a spiral point, defined to mean that every orbit tends to the equilibrium (either as $t \rightarrow \infty$ or as $t \rightarrow -\infty$) with its angular variable becoming infinite. If the linearization at an equilibrium has a *saddle point*, then the equilibrium of the nonlinear system is also a saddle point. A saddle point is defined by the characterization that there is a curve through the equilibrium such that orbits starting on this curve tend to the equilibrium but orbits starting off this curve cannot stay near the equilibrium. An equivalent formulation is that there are two orbits tending to the equilibrium as $t \rightarrow +\infty$ and two orbits tending away from the equilibrium, or tending to the equilibrium as $t \rightarrow -\infty$. These orbits are called *separatrices*; the two orbits tending to the saddle point are the stable separatrices, while the two orbits tending

away from the saddle point are the unstable separatrices. Other orbits appear like hyperbolas (Figure 4.8).

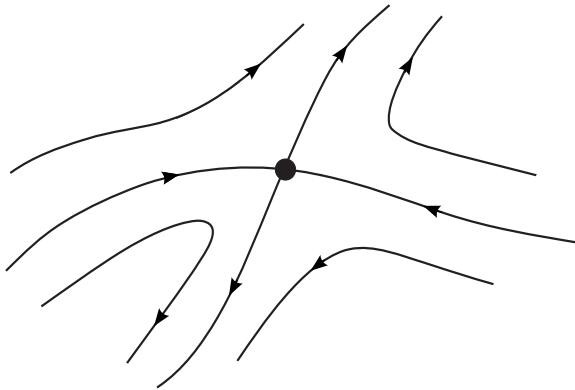


Fig. 4.8 Separatrices at a saddle point.

A *center* is defined to be an equilibrium for which there is an infinite sequence of periodic orbits around the equilibrium with the orbits approaching this equilibrium. If the linearization at an equilibrium has a center then the equilibrium of the nonlinear system *may* be a center, but is not necessarily a center: It could be an asymptotically stable spiral point or an unstable spiral point.

Example 1. If we linearize the Lotka–Volterra system $x' = x(\lambda - by)$, $y' = y(-\mu + cx)$ about the equilibrium (x_∞, y_∞) , with $x_\infty = \mu/c$, $y_\infty = \lambda/b$ we obtain the system

$$\begin{aligned} u' &= -\frac{\mu b}{c}v, \\ v' &= \frac{\lambda c}{b}u, \end{aligned}$$

with coefficient matrix

$$A = \begin{bmatrix} 0 & -\frac{\mu b}{c} \\ \frac{\lambda c}{b} & 0 \end{bmatrix}$$

with trace zero and determinant $\lambda\mu > 0$. Thus the linearization has a center. To solve the linearization we write

$$\begin{aligned} u'' &= -\frac{\mu b}{c}v' = -\lambda\mu u, \\ v'' &= \frac{\lambda c}{b}u' = -\lambda\mu v, \end{aligned}$$

from which we see that both u and v are periodic with frequency $\sqrt{\lambda\mu}$, period $2\pi/\lambda\mu$. We have shown in the preceding section that all solutions of the Lotka–Volterra system near the equilibrium are periodic and thus that the equilibrium is a

center, but study of the linearization is not enough to show this. The addition of a small perturbation to the Lotka–Volterra model could change the center to a spiral point, one that may be either asymptotically stable or unstable. Because the Lotka–Volterra model is so sensitive to perturbations, it is not really suitable as a population model. When we examine predator–prey systems in more detail we will attempt to refine the model.

Example 2. Determine whether each equilibrium of the system

$$x' = y, \quad y' = 2(x^2 - 1)y - x,$$

is asymptotically stable or unstable.

Solution. The equilibria are the solutions of $y = 0, 2(x^2 - 1)y - x = 0$, and thus the only equilibrium is $(0, 0)$. The community matrix at (x_∞, y_∞) is

$$\begin{pmatrix} 0 & 1 \\ 4x_\infty y_\infty - 1 & 2(x_\infty^2 - 1) \end{pmatrix},$$

and thus the community matrix at $(0, 0)$ is

$$\begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix},$$

with trace $-2 < 0$ and the determinant $1 > 0$. Thus the equilibrium $(0, 0)$ is asymptotically stable.

Example 3. Determine the asymptotic stability or instability of each equilibrium of the Lotka–Volterra system (4.1).

Solution. We have seen in Example 1, Section 4.3, that at the equilibrium $(0, 0)$ the community matrix is

$$\begin{pmatrix} \lambda & 0 \\ 0 & -\mu \end{pmatrix}.$$

Since this has negative determinant, $(0, 0)$ is unstable. The community matrix at the equilibrium $(\mu/c, \lambda/b)$ is

$$\begin{pmatrix} 0 & -\frac{b\mu}{c} \\ \frac{c\lambda}{b} & 0 \end{pmatrix}$$

with trace zero. Thus Theorem 4.4 does not settle the question. However, as we have seen in Section 4.1, orbits of the system neither move away from the equilibrium nor approach it. Thus, this equilibrium is neither asymptotically stable nor unstable. [Such an equilibrium is sometimes said to be neutrally stable.]

Exercises

In Exercises 1 through 6, for each equilibrium of the given system determine whether the equilibrium is asymptotically stable or unstable:

1. $x' = x - y, y' = x + y - 2$ (see Exercise 1, Section 4.3).
2. $x' = y, y' = x + y - 1$ (see Exercise 2, Section 4.3).
3. $x' = y + 1, y' = x^2 + y$ (see Exercise 3, Section 4.3).
4. $x' = y^2 - 8x, y' = x - 2$ (see Exercise 4 section 4.3).
5. $x' = e^{-y}, y' = e^{-x}$ (see Exercise 5, Section 4.3).
6. $x' = \sin y, y' = 2x$ (see Exercise 6, Section 4.3).
- 7.* Determine the behavior of orbits of the system

$$\begin{aligned}x' &= x(\lambda - ax - by) \\ y' &= y(\mu - cx - dy).\end{aligned}$$

- 8.* Determine the behavior of orbits of the system

$$\begin{aligned}x' &= x(2 - x) - \frac{xy}{x + 1}, \\ y' &= \frac{xy}{x + 1} - y.\end{aligned}$$

- 9.* Determine the behavior of orbits of the system

$$\begin{aligned}x' &= x(\lambda - ax - by), \\ y' &= y(-\mu + cx).\end{aligned}$$

- 10.* Consider the system

$$\begin{aligned}x' &= y, \\ y' &= -x - y^3.\end{aligned}$$

- (i) Show that $(0, 0)$ is the only equilibrium.
 - (ii) Show that the function $V(x, y) = x^2 + y^2$ decreases along every orbit and tends to zero, showing that $(0, 0)$ is asymptotically stable.
11. [Kaplan & Glass(1995)] A chemotherapeutic agent is being used to treat an intracranial tumor. Let x be the number of molecules of the agent in the blood and y the number of molecules that have crossed the blood–brain barrier. When $t = 0$ we have $x = N$ and $y = 0$. The dynamics are described by the differential equations

$$\begin{aligned}\frac{d}{dt}x &= \alpha(y-x) - \gamma x, \\ \frac{d}{dt}y &= -\alpha(y-x).\end{aligned}$$

- (i) Show that if γ is much larger than α , the roots of the characteristic equation are approximately $-\gamma$ and $-\alpha$.
 - (ii) Use the result from part (i) to replace the x equation by $dx/dt = -\gamma x$ and solve the equation for y as a function of time for $\alpha = 10^{-3} \text{ hr}^{-1}$ and $\gamma = 1 \text{ hr}^{-1}$.
 - (iii) For the values of α and γ in part (ii) compute the time when y is a maximum. What is the approximate value of y at this time?
12. [Kaplan & Glass(1995)] Limpets and seaweed live in a tide pool. The dynamics of this system are given by the differential equations

$$\begin{aligned}\frac{ds}{dt} &= s - s^2 - sl, \\ \frac{dl}{dt} &= sl - \frac{l}{2} - l^2, \quad l \geq 0, s \geq 0,\end{aligned}$$

where the densities of seaweed and limpets are given by s and l , respectively.

- (i) Determine all equilibria of this system.
 - (ii) For each nonzero equilibrium determined in part (a), evaluate the stability and classify it as a node, focus, or saddle point.
 - (iii) Sketch the flows in the phase plane.
 - (iv) What will the dynamics be in the limit as $t \rightarrow \infty$ for initial conditions
 - (i) $s(0) = 0, l(0) = 0$?
 - (ii) $s(0) = 0, l(0) = 15$?
 - (iii) $s(0) = 2, l(0) = 0$?
 - (iv) $s(0) = 2, l(0) = 15$?
13. The following system of differential equation is an SIR model for infectious diseases (with vital dynamics):

$$\begin{aligned}\frac{dS}{dt} &= -\beta SI + \mu - \mu S + \theta R, \\ \frac{dI}{dt} &= \beta SI - \gamma I - \mu I, \\ R &= 1 - S - I,\end{aligned}$$

with $S(0) = S_0, I(0) = I_0, R(0) = R_0$, where $S = S(t)$ is the fraction of susceptible individuals at time t , $I = I(t)$ is the fraction of infective individuals at time t , and $R = R(t)$ is the fraction of recovered individuals, μ is the per capita birth

and death rate, γ and θ are the rates at which individuals leave the infective and recovered classes, respectively. The *basic reproductive number* (the number of secondary cases produced by one infective individual in a population of susceptibles) in this case is given by the formula $\sigma = \beta/(\mu + \gamma)$.

- (i) Choose $\mu = 0.0001$, $\gamma = 0.2$, $\theta = 0.02$, $\beta = 0.15$. Then $\sigma < 1$. Use a computer algebra system to observe that the disease dies out as time goes to infinity.
- (ii) Increase β to $\beta = 0.4$, then $\sigma > 1$.
 - (i) Plot $I(t)$ versus time for initial values $S_0 = 0.2$, $I_0 = 0.8$, $R_0 = 0$.
 - (ii) Plot $I(t)$ versus $S(t)$ for the following initial conditions (on one graph) and explain the biological meaning suggested by the graph.

$$S_0 = 0.2, I_0 = 0.8, R_0 = 0;$$

$$S_0 = 0.4, I_0 = 0.6, R_0 = 0;$$

$$S_0 = 0.6, I_0 = 0.4, R_0 = 0;$$

$$S_0 = 0.8, I_0 = 0.2, R_0 = 0.$$

14. Discuss the qualitative behavior of the following epidemiological models (SIR and SIS without vital dynamics):

(i)

$$\begin{aligned}\frac{dS}{dt} &= -\beta S \frac{I}{N} + \gamma I, \\ \frac{dI}{dt} &= \beta S \frac{I}{N} - \gamma I, \\ N &= S + I.\end{aligned}$$

(ii)

$$\begin{aligned}\frac{dS}{dt} &= -\beta S \frac{I}{N}, \\ \frac{dI}{dt} &= \beta S \frac{I}{N} - \gamma I, \\ \frac{dR}{dt} &= \gamma I, \\ N &= S + I + R.\end{aligned}$$

15. [May (1981)] A modification of the logistic equation describing two species interacting mutualistically results in the model

$$\begin{aligned}\frac{dx}{dt} &= rx \left(1 - \frac{x}{K_1 + ay} \right), \\ \frac{dy}{dt} &= ry \left(1 - \frac{y}{K_2 + bx} \right),\end{aligned}$$

with a, b, K_1, K_2 positive constants. Show that the equilibrium populations are positive if $ab < 1$ and that if this condition is satisfied, then the equilibrium is locally asymptotically stable.

16. Sketch the phase-plane behavior of the following systems of ordinary differential equations and classify the stability of the equilibrium point $(0, 0)$.

(i)

$$\begin{aligned}\frac{dx}{dt} &= 3x - 2y, \\ \frac{dy}{dt} &= 4x + y.\end{aligned}$$

(ii)

$$\begin{aligned}\frac{dx}{dt} &= 5x + 2y, \\ \frac{dy}{dt} &= -13x - 5y.\end{aligned}$$

(iii)

$$\begin{aligned}\frac{dx}{dt} &= -x - y, \\ \frac{dy}{dt} &= \frac{8}{5}x - y.\end{aligned}$$

17. We are going to study an ecological system consisting of two species. The relationship between the two species is one of three kinds: a predator-prey system, a competitive system, or a mutualistic system depending on the nature of the interactions between species. We will use the Lotka-Volterra model to study the system. Let $x(t)$ and $y(t)$ be the population sizes (or densities) at t for the two populations. The key assumption in the Lotka-Volterra model is that the per capita growth rate for each population is a linear function of population sizes. We write the ordinary differential equation model for the system

$$\begin{aligned}\frac{1}{x(t)} \frac{dx}{dt} &= r_x + a_{11}x + a_{12}y, \\ \frac{1}{y(t)} \frac{dy}{dt} &= r_y + a_{21}x + a_{22}y,\end{aligned}\tag{4.17}$$

where r_x, r_y , and the a_{ij} are constants.

Let $r_x = r_y = 1$, $a_{11} = a_{22} = -1$, and $a_{12} = a_{21} = -1/2$ in the model (4.17). The specific model is

$$\frac{1}{x(t)} \frac{dx}{dt} = 1 - x - \frac{y}{2},$$

$$\frac{1}{y(t)} \frac{dy}{dt} = 1 - \frac{x}{2} - y.$$

Identify the system and make a guess about the fate of population sizes.

18. Consider the system of differential equations

$$\frac{dx}{dt} = xy - y,$$

$$\frac{dy}{dt} = xy - x.$$

- (i) Compute the equilibria.
 - (ii) Linearize the equations around each of the equilibria found in part (a). Write the Jacobian matrix and the linear system approximation for each case.
 - (iii) Draw the phase portrait of each linearized system.
19. Consider the system of differential equations

$$\frac{dx}{dt} = xy - y$$

$$\frac{dy}{dt} = xy - x.$$

- (i) Compute the equilibria.
- (ii) Linearize the equations around each of the equilibria found in part (a). Write the Jacobian matrix and the linear system approximation for each case.
- (iii) Draw the phase portrait of each linearized system.

4.5 Periodic Solutions and Limit Cycles

In the preceding section we analyzed the behavior of solutions starting near an equilibrium. When we studied this question for first-order differential equations this information was enough to describe the behavior of all solutions, since every solution was either unbounded or tended to an equilibrium. However, for two-dimensional systems there are other possibilities, and our results are valid only locally. We must consider what can happen to a solution that does not begin near an equilibrium; in particular, we will want to examine the behavior of solutions of systems that have no asymptotically stable equilibrium. Such systems can arise as models for predator–prey systems. We wish to study the behavior of solutions of a two-dimensional system

$$\begin{aligned}x' &= F(x, y), \\y' &= G(x, y),\end{aligned}\tag{4.18}$$

by studying the phase portrait in the phase plane. Let $(x(t), y(t))$ be a solution that is bounded as $t \rightarrow \infty$. The *positive semiorbit* C^+ of this solution is defined to be the set of points $(x(t), y(t))$ for $t \geq 0$ in the (x, y) plane. The *limit set* $L(C^+)$ of the semiorbit is defined to be the set of all points (\bar{x}, \bar{y}) such that there is a sequence of times $t_n \rightarrow \infty$ with $x(t_n) \rightarrow \bar{x}$, $y(t_n) \rightarrow \bar{y}$ as $n \rightarrow \infty$. For example, if the solution $(x(t), y(t))$ tends to an equilibrium (x_∞, y_∞) as $t \rightarrow \infty$, then the limit set consists of the equilibrium (x_∞, y_∞) . If $(x(t), y(t))$ is a periodic solution, so that the semiorbit C^+ is a closed curve, then the limit set $L(C^+)$ consists of all points of the semiorbit C^+ . It is not difficult to show that the limit set of a bounded semiorbit is a closed, bounded, connected set. An *invariant* set for the system (4.18) is a set of points in the plane that contains the positive semiorbit through every point of the set. Thus, for example, an equilibrium is an invariant set, and a closed orbit corresponding to a periodic solution is an invariant set. It is possible to prove, making use of the continuous dependence of solutions of differential equations on initial conditions, that the limit set of a bounded semiorbit is an invariant set.

The results stated above are valid for autonomous systems of differential equations in all dimensions but in two dimensions, more detailed information on the structure of limit sets is available. The reasons for this involve the topological properties of the plane, especially the Jordan curve theorem, which states that a simple closed curve in the plane divides the plane into two disjoint regions, which is not valid in more than two dimensions. The fundamental result on the behavior in the large of solutions of autonomous systems in the plane is the *Poincaré–Bendixson theorem* [Poincaré (1881, 1882, 1885, 1886); Bendixson (1901)], which states that if C^+ is a bounded semiorbit whose limit set $L(C^+)$ contains no equilibrium points, then either C^+ is a periodic orbit and $L(C^+) = C^+$, or $L(C^+)$ is a periodic orbit, called a *limit cycle*, which C^+ approaches spirally (either from inside or outside). We shall not go into the proof, which may be found in many books on differential equations; a particularly readable elementary source is the book by W. Hurewicz (1958).

The properties of limit sets enable us to describe limit sets that do contain equilibrium points. For example, a limit set may consist of a single equilibrium point to which the orbit tends. If a limit set contains a single equilibrium point but contains other points as well, then it must consist of the equilibrium point together with an orbit running from the equilibrium point to itself in order to be connected and invariant. Thus, the equilibrium must be a saddle point and the limit set must be a *separatrix* that runs from the saddle point to itself as t runs from $-\infty$ to $+\infty$ (Figure 4.9).

If a limit set contains more than one equilibrium, then it must also contain orbits joining these equilibria. In essence we can say that a bounded solution tends either to an equilibrium or to a limit cycle, overlooking such “unlikely coincidences” as the possibility of a separatrix running from a saddle point to itself. Thus, if we can show that for a given system all solutions are bounded but that there are no

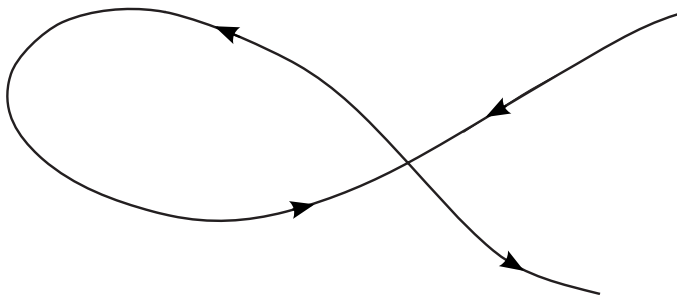


Fig. 4.9 A separatrix.

asymptotically stable equilibrium points, we can deduce that there must be at least one periodic orbit. This situation will arise in our study of predator–prey systems. If there is only one periodic orbit, then it must be *globally asymptotically stable* in the sense that every orbit tends to it. If there is more than one periodic orbit, each must be asymptotically stable from at least one side: orbits may spiral toward it from the inside, from the outside, or both.

In situations in which we know that there is a periodic solution, it may be possible to use the concept of the index of a simple closed curve and the index of an equilibrium to help locate the periodic solution. An autonomous system of differential equations (4.18) defines a *vector field* by associating with each point (x,y) a vector with components $(F(x,y), G(x,y))$. A solution of the system has its orbit tangent to this vector field at every point. For any simple, i.e., not self-intersecting, closed curve S we define the *index* of S with respect to the system to be $1/2\pi$ times the change in the angle of the vector (F,G) as the curve is traversed once in the positive (counterclockwise) direction. Thus, if S is a periodic orbit, it is easy to see that the index of S is $+1$ (Figure 4.10).

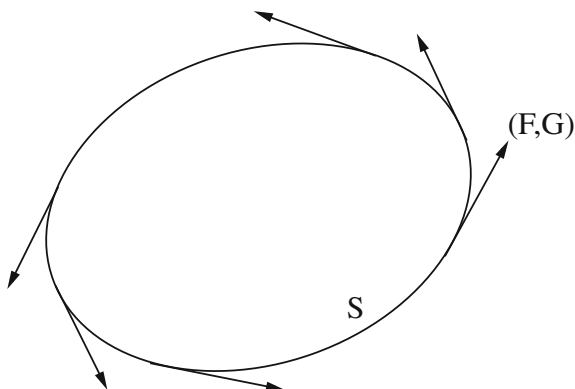


Fig. 4.10 Index of a periodic orbit.

Theorem 4.5. *The index of a simple closed curve S whose interior contains no equilibrium points is zero.*

Proof (Outline of proof). Write S as a sum of small simple closed curves, with cancellation of all interior portions of their boundaries (Figure 4.11). Over a sufficiently

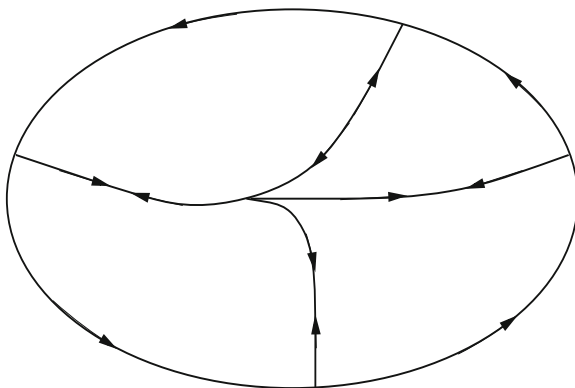


Fig. 4.11 Subdivision of a simple closed curve.

small simple closed curve, the angle of the vector (F, G) cannot change by as much as 2π , by continuity of the vector field. Thus, the index of every sufficiently small simple closed curve is zero. Since the index of S is the sum of the indices of the small curves, the index of S must be zero. The hypothesis that the interior of S contains no equilibrium points is needed to deduce that the vector field is continuous.

An immediate consequence of Theorem 4.5 is that a periodic orbit must contain an equilibrium in its interior. Another consequence is that it is possible to give an unambiguous definition of the *index of an equilibrium point*. Let P be an equilibrium point and let S be a simple closed curve around P , but whose interior does not contain any other equilibrium points. If \tilde{S} is another simple closed curve around P inside S (and thus not containing any other equilibrium points in its interior), the simple closed curve formed by going around S counterclockwise, straight in to \tilde{S} , clockwise, and then back out to S is a simple closed curve whose interior contains no equilibrium and that therefore has index zero (Figure 4.12).

On the other hand, the index of this curve is the index of S minus the index of \tilde{S} ; thus S and \tilde{S} have the same index. Therefore, we define the index of an equilibrium point P to be the index of any simple closed curve whose interior contains P but contains no other equilibrium. This definition is independent of the choice of curve. For any simple closed curve the index of the curve is equal to the sum of the indices of all equilibria in the interior of the curve. The index of an isolated equilibrium is either $+1$ or -1 , with a saddle point having index -1 and all other equilibria having index $+1$. To demonstrate this, we would first show that the vector field $(F(x, y), G(x, y))$ near an equilibrium (which may be assumed to be at the origin)

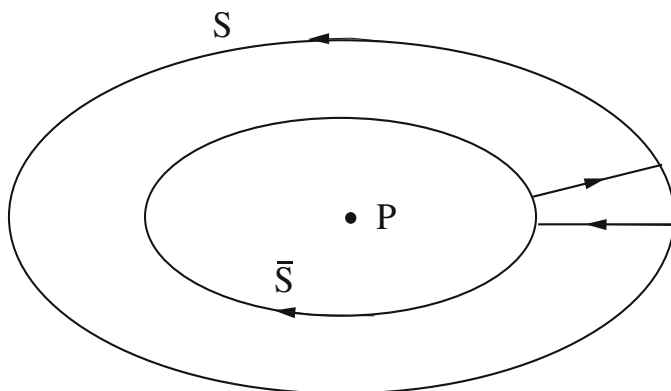


Fig. 4.12

may be replaced by its linear part, $(ax + by, cx + dy)$ because for a sufficiently small curve the vectors $(F(x, y), G(x, y))$ and $(ax + by, cx + dy)$ must have directions similar enough that their indices must be the same. The remainder of the proof consists in an examination of each type of equilibrium (see, for example, Figure 4.13).

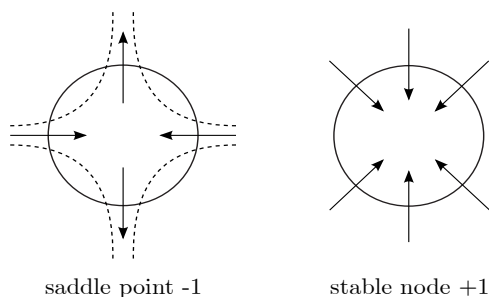


Fig. 4.13

We now deduce that the number of equilibria that are not saddle points minus the number of saddle points in the interior of a periodic orbit is one. In particular, if a system has several equilibria all but one of which are saddle points, then any periodic orbit must go around the equilibrium that is not a saddle point but cannot go around any saddle point.

We conclude this section by giving some criteria implying that there *cannot* be a periodic orbit in a given region. Such results are of interest in situations in which we know that there is an asymptotically stable equilibrium and wish to conclude that all orbits tend to it. We shall make use of Green's theorem in the plane.

Theorem 4.6 (Green). *Let D be a simple connected region in the plane with boundary C , and suppose that $P(x, y)$ and $Q(x, y)$ are continuously differentiable in the closure of D . Then*

$$\int_D \int (\mathcal{Q}_x(x, y) - P_y(x, y)) dA = \int_c (P(x, y) dx + \mathcal{Q}(x, y) dy).$$

Theorem 4.7 (Bendixson (1901)). *Suppose that $F_x(x, y) + G_y(x, y)$ is either strictly positive or strictly negative in a simply connected region D . Then there is no periodic orbit of $x' = F(x, y)$, $y' = G(x, y)$ in D .*

Proof. Suppose that C is a closed orbit in D corresponding to a periodic solution $(x(t), y(t))$ with period T , and let Ω be the interior of C . We apply Green's theorem with $\mathcal{Q}(x, y) = F(x, y)$, $P(x, y) = -G(x, y)$, obtaining

$$\begin{aligned} \int_{\Omega} \int (F_x(x, y) + G_y(x, y)) dA &= \int_c (-G(x, y) dx + F(x, y) dy) \\ &= \int_0^T (-y'(t)x'(t) + x'(t)y'(t)) dt = 0. \end{aligned}$$

But the left side cannot be zero, and we have a contradiction. Thus there can be no periodic orbit in D .

A useful generalization of Bendixson's theorem is obtained by application of Green's theorem with

$$\mathcal{Q}(x, y) = \beta(x, y)F(x, y), \quad P(x, y) = -\beta(x, y)G(x, y)$$

for some function $\beta(x, y)$ that is continuously differentiable in D .

Theorem 4.8 (Dulac (1934)). *Let $\beta(x, y)$ be continuously differentiable and suppose that*

$$\frac{\partial}{\partial x} (\beta(x, y)F(x, y)) + \frac{\partial}{\partial y} (\beta(x, y)G(x, y))$$

is either strictly positive or strictly negative in a simply connected region D . Then there is no periodic orbit of $x' = F(x, y)$, $y' = G(x, y)$ in D .

In the following chapter, we shall be studying models for populations of two interacting species of the form

$$\begin{aligned} x' &= xf(x, y) \\ y' &= yg(x, y). \end{aligned}$$

We may apply Dulac's criterion with $\beta(x, y) = 1/xy$, $F(x, y) = xf(x, y)$, $G(x, y) = yg(x, y)$, so that

$$\frac{\partial}{\partial x} (\beta F) + \frac{\partial}{\partial y} (\beta G) = \frac{\partial}{\partial x} \left(\frac{1}{y} f \right) + \frac{\partial}{\partial y} \left(\frac{1}{x} g \right) = \frac{f_x(x, y)}{y} + \frac{g_y(x, y)}{x}.$$

This gives the following result, which will apply to our models for species in competition and for mutualism, but not to predator-prey models.

Theorem 4.9. Suppose $f_x(x, y) < 0$, $g_y(x, y) < 0$ for $x > 0$, $y > 0$. Then there is no periodic orbit of the system $x' = xf(x, y)$, $y' = yg(x, y)$ in the interior of the first quadrant of the phase plane.

Example 1. Determine the qualitative behavior of solutions of the system

$$\begin{aligned}x' &= x \left(1 - \frac{x}{30} \right) - \frac{xy}{x+10}, \\y' &= y \left(\frac{x}{x+10} - \frac{1}{3} \right).\end{aligned}$$

Solution. Equilibria are solutions of the pair of equations

$$\begin{aligned}x \left(1 - \frac{x}{30} \right) - \frac{xy}{x+10} &= 0 \\y \left(\frac{x}{x+10} - \frac{1}{3} \right) &= 0.\end{aligned}$$

There is an equilibrium, $(0, 0)$. If $y = 0$ the first equation gives $x = 0$ or $x = 30$, and thus there is a second equilibrium, $(30, 0)$. If $x \neq 0$, $y \neq 0$, we must solve $1 - x/30 = y/(x+10)$, $x/(x+10) = 1/3$, which gives a third equilibrium $(5, 12.5)$. The community matrix at $(0, 0)$ is

$$\begin{bmatrix} 1 & 0 \\ 0 & -1/3, \end{bmatrix}$$

and this equilibrium is unstable because the determinant is negative. The community matrix at $(30, 0)$ is

$$\begin{bmatrix} -1 & -3/4 \\ 0 & 5/12 \end{bmatrix}$$

and this equilibrium is also unstable because the determinant is negative. The community matrix at $(5, 12.5)$ is

$$\begin{pmatrix} 1/9 & -1/3 \\ 5/9 & 0 \end{pmatrix}$$

with positive trace, and $(5, 12.5)$ is also unstable. If we add the two equations of the model we obtain $(x+y)' = x(1-x/30) - y/3$. Thus $(x+y)$ is decreasing except in the bounded region $y/3 < x(1-x/30)$. In order that an orbit may be unbounded, we must have $x+y$ unbounded. However, this is impossible, since $x+y$ is decreasing whenever $x+y$ is large. Thus every orbit of the system is bounded. Since all equilibria are unstable, the Poincaré–Bendixson theorem implies that there must be a periodic orbit (around $(5, 12.5)$) to which every orbit tends as $t \rightarrow \infty$.

Example 2. Determine the behavior as $t \rightarrow \infty$ of solutions in the first quadrant of the system

$$\begin{aligned}x' &= x(100 - 4x - 2y), \\y' &= y(60 - x - y).\end{aligned}$$

Solution. There are equilibria at $(0, 0)$, $(25, 0)$, and $(0, 60)$. There is no equilibrium within $x > 0, y > 0$, because such an equilibrium would be found by solving $4x + 2y = 100, x + y = 60$, and subtraction of twice the first equation from the second equation gives $2x = -20$. The equilibria $(0, 0)$ and $(25, 0)$ are unstable because their community matrices are, respectively,

$$\begin{bmatrix} 100 & 0 \\ 0 & 60 \end{bmatrix}$$

and

$$\begin{bmatrix} -100 & -50 \\ 0 & 35 \end{bmatrix}.$$

The equilibrium $(0, 60)$ is asymptotically stable because it has community matrix

$$\begin{pmatrix} -20 & 0 \\ -60 & -60 \end{pmatrix}.$$

In order to show that every orbit approaches $(0, 60)$, we must show that there are no periodic orbits. This follows from Dulac's criterion with $\beta(x, y) = 1/xy$, because

$$\frac{\partial}{\partial x} \left(\frac{100 - 4x - 2y}{y} \right) + \frac{\partial}{\partial y} \left(\frac{60 - x - y}{x} \right) = -\frac{4}{y} - \frac{1}{x} < 0.$$

Exercises

In each of Exercises 1 through 10, determine the qualitative behavior of solutions of the given system.

1. $x' = 3x(1 - \frac{x}{20}) - \frac{xy}{x+10}, y' = y(\frac{x}{x+10} - \frac{1}{2})$.
2. $x' = x(1 - \frac{x}{40}) - \frac{2xy}{x+15}, y' = 3y(\frac{2x}{x+15} - \frac{6}{5})$.
3. $x' = 3x(1 - \frac{x}{20}) - \frac{xy}{x+10}, y' = y(\frac{x}{x+10} - \frac{1}{6})$.
4. $x' = x(1 - \frac{x}{40}) - \frac{2xy}{x+15}, y' = 3y(\frac{2x}{x+15} - \frac{1}{2})$.
5. $x' = 3x(1 - \frac{x}{20}) - \frac{xy}{x+20}, y' = y(\frac{x}{x+10} - \frac{3}{4})$.
6. $x' = x(1 - \frac{x}{40}) - \frac{2xy}{x+15}, y' = 3y(\frac{2x}{x+15} - \frac{8}{5})$.
7. $x' = x(80 - x - y), y' = y(120 - x - 3y)$.
8. $x' = x(60 - 3x - y), y' = y(75 - 4x - y)$.
9. $x' = x(40 - x - y), y' = y(90 - x - 2y)$.
10. $x' = x(80 - 3x - 2y), y' = y(80 - x - y)$.
- 11.* Show that the equilibrium (x_∞, y_∞) with $x_\infty > 0, y_\infty > 0$ of the system

$$\begin{aligned}x' &= rx \left(1 - \frac{x}{K} \right) - \frac{axy}{x+A}, \\y' &= sy \left(\frac{ax}{x+A} - \frac{aJ}{J+A} \right),\end{aligned}$$

is unstable if $K > A + 2J$ and asymptotically stable if $J < K < A + 2J$.

- 12.* Determine the behavior as $t \rightarrow \infty$ of solutions of the system

$$\begin{aligned}x' &= rx \left(1 - \frac{x}{K+ay} \right) - \frac{axy}{x+A}, \\y' &= sy \left(1 - \frac{y}{M+bx} \right),\end{aligned}$$

with a, b, K, M, r, s positive constants. (*Warning:* The behavior if $ab < 1$ is different from the behavior if $ab > 1$.)

13. [Kaplan & Glass(1995)] A theoretical model for mutual inhibition is

$$\begin{aligned}\frac{d}{dt}x &= f(x) = \frac{\left(\frac{1}{2}\right)^n}{\left(\frac{1}{2}\right)^n + y^n} - x, \\ \frac{d}{dt}y &= g(y) = \frac{\left(\frac{1}{2}\right)^n}{\left(\frac{1}{2}\right)^n + x^n} - y,\end{aligned}$$

where x and y are positive variables and n is a positive constant greater than one. There is a steady state at $x^* = y^* = 1/2$. Discuss the bifurcation and sketch the flows in the (x, y) -plane as n varies.

14. Consider the system

$$\begin{aligned}x' &= x(ax + by), \\y' &= y(cx + dy).\end{aligned}$$

- (i) Show that every trajectory with $x(0) \geq 0$, $y(0) \geq 0$ satisfies $x(t) \geq 0$, $y(t) \geq 0$ for all $t \geq 0$ (i.e., trajectories starting in the first quadrant remain in the first quadrant forever).
- (ii) Use the Dulac criterion with $\beta(x, y) = 1/xy$ to show that there are no periodic orbits if $ad > 0$.

4.6 Appendix: Canonical Forms of 2×2 Matrices

In Section 4.4 we stated a theorem giving the possible real canonical forms of a real 2×2 matrix. This theorem identified six different cases relating both to the form of the matrix and to the signs of its eigenvalues. If only different canonical forms concern us, the cases reduce to three. Theorem 4.3 of Section 4.4 is a consequence

of the following result from linear algebra, whose proof requires an understanding of linear independence, eigenvalues, and eigenvectors but nothing more advanced. We include the proof because most linear algebra texts use a more general approach than ours, one that does require more advanced understanding.

Theorem 4.10. *If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is a real 2×2 matrix with $\det A = ad - bc \neq 0$ then there is a real nonsingular 2×2 matrix P such that $P^{-1}AP$ is one of the following:*

- (i). $\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$, $\lambda \neq 0$, $\mu \neq 0$, (Cases (i), (ii), (iv) of Theorem 4.3)
- (ii). $\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$, $\lambda \neq 0$, (Case (iii) of Theorem 4.3)
- (iii). $\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$, $\beta \neq 0$. (Cases (v), (vi) of Theorem 4.3)

Proof. First, consider the case that A has two linearly independent eigenvectors \mathbf{u}, \mathbf{v} corresponding to the eigenvalues λ, μ , respectively ($\lambda = \mu$ is permitted). Recall that an eigenvalue of a matrix A is a value of λ such that the linear homogeneous algebraic system $A\mathbf{u} = \lambda\mathbf{u}$ has a nontrivial solution and an eigenvector is a corresponding solution vector; λ is an eigenvalue if and only if $\det(A - \lambda I) = 0$. Let P be the matrix whose columns are \mathbf{u} and \mathbf{v} , $P = (\mathbf{u}, \mathbf{v})$. Since \mathbf{u} and \mathbf{v} are linearly independent, P is nonsingular. We now have

$$P^{-1}AP = P^{-1}(A\mathbf{u}, A\mathbf{v}) = P^{-1}(\lambda\mathbf{u}, \lambda\mathbf{v}).$$

Since $P^{-1}P = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $P^{-1}\mathbf{u}$ is the first column of I , namely $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and $P^{-1}\mathbf{v}$ is the second column of I , namely $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Thus

$$P^{-1}AP = P^{-1}(\lambda\mathbf{u}, \lambda\mathbf{v}) = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}.$$

Second, if λ is a double eigenvalue but there is only one linearly independent corresponding eigenvector, take \mathbf{v} to be any nonzero vector that is not an eigenvector, and let $\mathbf{u} = A\mathbf{v} - \lambda\mathbf{v} \neq \mathbf{0}$. If $c_1\mathbf{u} + c_2\mathbf{v} = \mathbf{0}$ with c_1 and c_2 not both zero, then c_1 and c_2 must both be different from zero because \mathbf{u} and \mathbf{v} are both different from zero. Then $c_1(A\mathbf{v} - \lambda\mathbf{v}) + c_2\mathbf{v} = \mathbf{0}$, which implies $A\mathbf{v} = (\lambda - c_2\mathbf{v}/c_1)\mathbf{v}$. But then $\lambda - c_2/c_1$ is an eigenvalue, contradicting the assumption that λ is the only eigenvalue. This shows that $c_1 = c_2 = 0$, and thus that \mathbf{u} and \mathbf{v} are linearly independent. If \mathbf{w} is any eigenvector we may write $\mathbf{w} = a\mathbf{u} + b\mathbf{v}$ with $a \neq 0$, because \mathbf{u} and \mathbf{v} span R^2 and \mathbf{v} is not an eigenvector. Then we have

$$\begin{aligned} \mathbf{0} &= A\mathbf{w} - \lambda\mathbf{w} = a(A\mathbf{u} - \lambda\mathbf{u}) + b(A\mathbf{v} - \lambda\mathbf{v}) \\ &= a(A\mathbf{u} - \lambda\mathbf{u}) + b\mathbf{u} = a\left(A\mathbf{u} - \left(\lambda - \frac{b}{a}\right)\mathbf{u}\right). \end{aligned}$$

This would mean that $\lambda - b/a$ is an eigenvalue, but since λ is the only eigenvalue $b = 0$ and therefore \mathbf{u} is an eigenvector. We now take $P = (\mathbf{u}, \mathbf{v})$, nonsingular because \mathbf{u} and \mathbf{v} are linearly independent, and calculate

$$\begin{aligned}
P^{-1}AP &= P^{-1}(A\mathbf{u}A\mathbf{v}) = P^{-1}(\lambda\mathbf{u}\mathbf{u} + \lambda\mathbf{v}) \\
&= \lambda P^{-1}(\mathbf{u}\mathbf{v}) + P^{-1}(0\mathbf{u}) \\
&= \lambda I + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.
\end{aligned}$$

Third, if there is a pair of complex conjugate eigenvalues $\alpha \pm i\beta$, with $\beta \neq 0$, let $\mathbf{u} + i\mathbf{v}$ be an eigenvector corresponding to the eigenvalue $\alpha + i\beta$, with \mathbf{u} and \mathbf{v} real. If $\mathbf{v} = \mathbf{0}$ then $A\mathbf{u} = (\alpha + i\beta)\mathbf{u}$, which is impossible because the left side is real and the right side is not; therefore, $\mathbf{v} \neq \mathbf{0}$. A similar argument shows that $\mathbf{u} \neq \mathbf{0}$. If $c_1\mathbf{u} + c_2\mathbf{v} = \mathbf{0}$, with $c_1 \neq 0$, $c_2 \neq 0$, we replace \mathbf{u} by $-c_2\mathbf{v}/c_1$ in $A(\mathbf{u} + c\mathbf{v}) = (\alpha + i\beta)(\mathbf{u} + i\mathbf{v})$, obtaining $A(-c_2/c_1 + i)\mathbf{v} = (\alpha + i\beta)(-c_2/c_1 + i)\mathbf{v}$, or $A = \mathbf{v} = (\alpha + i\beta)\mathbf{v}$. Then \mathbf{v} would be a real eigenvector, which is not possible, and therefore \mathbf{u} and \mathbf{v} are linearly independent. We now define the nonsingular matrix $P = (\mathbf{u}\mathbf{v})$, and $P^{-1}AP = P^{-1}(A\mathbf{u}A\mathbf{v})$. From $A(\mathbf{u} + i\mathbf{v}) = (\alpha + i\beta)(\mathbf{u} + i\mathbf{v}) = (\alpha\mathbf{u} - \beta\mathbf{v}) + i(\beta\mathbf{u} + \alpha\mathbf{v})$ we have

$$A\mathbf{u} = \alpha\mathbf{u} - \beta\mathbf{v}, \quad A\mathbf{v} = \beta\mathbf{u} + \alpha\mathbf{v},$$

and thus

$$\begin{aligned}
P^{-1}AP^{-1}(A\mathbf{u}A\mathbf{v}) &= P^{-1}(\alpha\mathbf{u} - \beta\mathbf{v} \quad \beta\mathbf{u} + \alpha\mathbf{v}) \\
&= \alpha P^{-1}(\mathbf{u}\mathbf{v}) + \beta P^{-1}(-\mathbf{v}\mathbf{u}) \\
&= \alpha I + \beta \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}.
\end{aligned}$$

We have now covered all the cases and have completed the proof of the theorem and of Theorem 4.3 as well.

4.7 Project: A Model for Giving Up Smoking

1. Let $S(t)$ denote the number of smokers in high school at time t ; $P(t)$ the number of potential smokers at time t ; $Q(t)$ the number of smokers who quit permanently.

c : average number of contacts per unit time

q : probability of becoming a smoker given that a P has a contact with a smoker

$\frac{1}{\mu}$: average time in the high school system

$\frac{1}{\gamma}$: average time as a smoker

$N(t)$: population size at time t ,

$\beta = qc$. The model is

$$\begin{aligned}\frac{dP}{dt} &= \mu N - \beta P \frac{S}{N} - \mu P, \\ \frac{dS}{dt} &= \beta P \frac{S}{N} - (\mu + \gamma) S, \\ \frac{dQ}{dt} &= \gamma S - \mu Q.\end{aligned}$$

- (i) Show that $N(t)$ is a constant.
- (ii) Explain each term and coefficient in the model.
- (iii) What are the assumptions of the model?
- (iv) Give the units of μ , β , and γ .
- (v) Give a criterion for invasion in terms of a basic reproductive number R_0 .
- (vi) Show that R_0 is a nondimensional quantity.
- (vii) Find all equilibria.
- (viii) Study the stability of these equilibria.
- (ix) Introduce the variables

$$x = \frac{P}{N}, \quad y = \frac{S}{N}, \quad z = \frac{Q}{N}$$

and write the system in terms of x , y , and z . Do you gain any understanding?

- (x) Alter the model by just one simple modification to make it more realistic.
- (x*) Perform some computer simulations to verify your stability results.

References: Castillo-Garsow, Jordan-Salivia, Rodriguez-Herrera (2000).

4.8 Project: A Model for Retraining of Workers by their Peers

Suppose we have a system with the following classes of worker at time t :

$R(t)$: reluctant workers

$P(t)$: positive workers

$M(t)$: master workers

$U(t)$: unchangeable (negative) workers

$I(t)$: inactive workers

Assume that $N(t) = R(t) + P(t) + M(t) + U(t) + I(t)$ and that the total number of workers is constant, that is, $N(t) = K$ for all t .

The model is

$$\begin{aligned}
\frac{dP}{dt} &= qK - \beta P \frac{M}{K} + \delta R - \mu P, \\
\frac{dR}{dt} &= (1-q)K - (\delta + \mu)R - \alpha R, \\
\frac{dM}{dt} &= \beta P \frac{M}{K} - (\gamma + \mu)M, \\
\frac{dU}{dt} &= -\mu U + \alpha R, \\
\frac{dI}{dt} &= \gamma M - \mu I,
\end{aligned}$$

where q , β , δ , μ , and α are constants and $0 \leq q \leq 1$.

1. Interpret the parameters.
2. Under what conditions will retraining of workers succeed?
3. Will every worker be retrained (e.g., become an M) over the long run?
4. What is the impact of changing $0 \leq q \leq 1$, $0 < \gamma < \infty$, and $0 < \delta < \infty$? (Compute R_0 and look at the role of these parameters.)
5. Look at the stability of the simplest equilibrium.
6. Do we have a nontrivial equilibrium? Use simulations to guess its stability.
7. What are your conclusions from this model?

4.9 Project: A Continuous Two-Sex Population Model

The two-sex population model of this project begins from a population of males and females that select their partners *more or less* at random. Individuals are either single or in temporarily monogamous heterosexual relationships. The recruitment and death per capita rates are constant but gender specific, while the divorce or separation constant per pair rate is assumed to be gender independent.

Definition of Variables and Parameters

State variables are denoted by the lowercase letters x , y , and p . Here $x(t)$ denotes single males at time t ; $y(t)$ denotes single females at time t ; and $p(t)$ denotes pairs or married couples at time t . Hence, the total population equals $x + y + 2p$. The constants β_x and β_y denote per capita birth or recruitment rates of females and males, respectively. Furthermore, the constants μ_x and μ_y denote the per capita death rates, while σ gives the constant per couple separation rate.

Using these variables and definitions, we arrive at the following two-sex model [Keyfitz (1949); Kendall (1949); Fredrickson (1971); McFarland (1972)]:

$$\begin{aligned}
\dot{x} &= (\beta_x + \mu_y + \sigma)p - \mu_x x - \varphi(x, y), \\
\dot{y} &= (\beta_y + \mu_x + \sigma)p - \mu_y y - \varphi(x, y), \\
\dot{p} &= -(\mu_x + \mu_y + \sigma)p + \varphi(x, y).
\end{aligned} \tag{4.19}$$

The population of singles increases because of recruitment, death of a partner, or divorce. The population of singles decreases because of death or marriage. The rate of marriage is *definitely* a nonlinear process. Here it is assumed that it depends only on the population of singles. It is modeled by a nonlinear function $\varphi(x, y)$ called the *marriage* or *mating* function. Hence, $\varphi(x, y)/x$ is the per capita marriage rate for males and $\varphi(x, y)/y$ is the per capita marriage rate for females.

The function $\varphi(x, y)$ is nonlinear, defined for $x, y \geq 0$, and it is commonly assumed to have the following properties [Keyfitz (1949); Kendall (1949); Fredrickson (1971); McFarland (1972)]:

1. $\varphi(x, 0) = \varphi(0, y) = 0$ for all $x, y \geq 0$, and $\varphi(x, y) \geq 0$.
2. $\varphi(\alpha x, \alpha y) = \alpha \varphi(x, y)$ for $\alpha \geq 0$, that is, it is homogeneous of degree one. In other words, if there are x male singles and y female singles, then there are $\varphi(x, y)$ marriages per unit of time, while if there are αx male singles and αy female singles, then there are $\alpha \varphi(x, y)$ marriages per unit of time.
3. If the number of single males and females increases, then the total number of marriages also increases.

Examples of φ 's. (Generalized means)

$$\varphi(a, b) = \rho(\theta x^\alpha + (1 - \theta)y^\alpha)^{1/\alpha}$$

for

$$\alpha \geq 0, \quad 0 \leq \theta \leq 1, \quad \rho > 0.$$

Special cases:

$$\begin{aligned}
\varphi(x, y) &= \rho \frac{xy}{\theta x + (1 - \theta)y} && \text{harmonic mean } (\alpha = -1), \\
\varphi(x, y) &= \rho x^\theta y^{1-\theta} && \text{geometric mean } (\alpha \rightarrow 0,) \\
\varphi(x, y) &= \rho \min\{x, y\} && \text{min function } (\alpha \rightarrow \infty).
\end{aligned}$$

This model was proposed by D.G. Kendall (1949), N. Keyfitz (1949), A.G. Fredrickson (1971), and D.McFarland (1972), but a complete mathematical analysis was only recently completed by K.P. Hadeler, R. Waldstätter, and A. Wörz-Busekros (1988).

- (a) Show that this model does not have (in general) time-independent steady states, that is, constant solutions.

Hint: Suppose such a steady state exists and show that it must then satisfy

$$\begin{aligned}
(\beta_x - \mu_x)p - \mu_x x &= 0, \\
(\beta_y - \mu_y)p - \mu_y y &= 0.
\end{aligned}$$

Then substitute the values for x and y in the p-equation of system (4.19). You should arrive at the condition

$$(\mu_x + \mu_y + \sigma)p = \varphi\left(\frac{\beta_x - \mu_x}{\mu_x}, \frac{\beta_y - \mu_y}{\mu_y}\right)p. \quad (4.20)$$

Hence, the parameters of the model have to satisfy equation (4.20). What do you conclude from this?

- (b) Since steady states are not generic, we look for exponential solutions, that is, for solutions of the form $ce^{\lambda t}$.

Look for solutions of the form (why is λ the same?)

$$\begin{aligned} x &= \bar{x}e^{\lambda t}, \\ y &= \bar{y}e^{\lambda t}, \\ p &= \bar{p}e^{\lambda t}, \end{aligned}$$

that is, substitute the above candidates for exponential solutions in the original system (4.19) to arrive at the following nonlinear eigenvalue problem:

$$\begin{aligned} (\beta_x + \mu_y + \sigma)\bar{p} - (\mu_x + \lambda)\bar{x} &= \varphi(\bar{x}, \bar{y}), \\ (\beta_y + \mu_x + \sigma)\bar{p} - (\mu_y + \lambda)\bar{y} &= \varphi(\bar{x}, \bar{y}), \\ (\mu_x + \mu_y + \sigma + \lambda)\bar{p} &= \varphi(\bar{x}, \bar{y}). \end{aligned} \quad (4.21)$$

- (c) Show that the above nonlinear eigenvalue problem has the trivial solutions

$$\begin{aligned} (1, 0, 0) \text{ with } \lambda &= -\mu_x, \\ (0, 1, 0) \text{ with } \lambda &= -\mu_y, \end{aligned}$$

and interpret them.

- (d) Look for nontrivial solutions. In fact, show that such solutions exist if and only if there exists a λ such that

$$\mu_x + \mu_y + \sigma + \lambda = \varphi\left(\frac{\bar{x}}{\bar{p}}, \frac{\bar{y}}{\bar{p}}\right),$$

or, by substituting \bar{x}/\bar{p} , \bar{y}/\bar{p} , show that such solutions exist if they satisfy the following nonlinear relationship:

$$\mu_x + \mu_y + \sigma + \lambda = \varphi\left(\frac{\beta_x}{\mu_x + \lambda} - 1, \frac{\beta_y}{\mu_y + \lambda} - 1\right). \quad (4.22)$$

Hence, we are looking for nonzero λ 's that satisfy (4.22). The equation (4.22) is called the *characteristic equation* of this two-sex population.

- (e) Assume that

$$\frac{\beta_x - \mu_x - \lambda}{\mu_x + \lambda} > 0, \quad \frac{\beta_y - \mu_y + \lambda}{\mu_y + \lambda} > 0,$$

Set $\underline{\lambda} = \max(-\mu_x, -\mu_y)$. Then $\mu_x + \lambda > 0$ and $\mu_y + \lambda > 0$ for $\lambda > \underline{\lambda}$. Set $\bar{\lambda} = \min(\beta_x - \mu_x, \beta_y - \mu_y)$. Then $\beta_x - \mu_x - \lambda > 0$ and $\beta_y - \mu_y - \lambda > 0$ for $\lambda < \bar{\lambda}$. Hence $\underline{\lambda} < \lambda < \bar{\lambda}$ (if it exists). Justify the above assumptions and definitions. Show that if $\underline{\lambda} > \bar{\lambda}$, then there is no solution, while if $\underline{\lambda} < \bar{\lambda}$ then there is a solution. In fact, show that if K is defined as

$$K = \varphi \cdot \left(\frac{\beta_x}{\mu_x + \lambda} - 1, \frac{\beta_y}{\mu_y + \lambda} - 1 \right) \leq \infty,$$

where the possibility that $K = \infty$ cannot be disregarded; this case must also be studied.

Then equation (4.22) has a unique solution if and only if

$$K > \mu_x + \mu_y + \sigma + \lambda.$$

- (f) Get a more detailed existence condition by *cleverly* expanding $\varphi(x, y)$ in a Taylor series (keeping just the first-order terms), that is, by using the fact that it is a homogeneous function of degree one and then letting $y \rightarrow \infty$ we have that

$$\varphi(x, y) = x\varphi\left(1, \frac{y}{x}\right) = x\left[\varphi(1, 0) + \varphi_y(1, 0)\frac{y}{x}\right] = y\varphi_y(1, 0).$$

Hence,

$$\begin{aligned}\varphi(x, y) &\approx y\varphi_y(1, 0) \text{ as } x \rightarrow \infty, \\ \varphi(x, y) &\approx x\varphi_x(1, 0) \text{ as } y \rightarrow \infty.\end{aligned}$$

Show that if $\underline{\lambda} = -\mu_x$, then as $x \rightarrow \infty$,

$$\varphi\left(\frac{\beta_x}{\mu_x + \lambda} - 1, \frac{\beta_y}{\mu_y + \lambda} - 1\right) \approx \left[\frac{\beta_y}{\mu_y - \mu_x} - 1\right] \varphi_y(1, 0).$$

Show that if $\underline{\lambda} = -\mu_y$, then as $y \rightarrow \infty$,

$$\varphi\left(\frac{\beta_x}{\mu_x + \lambda} - 1, \frac{\beta_y}{\mu_y + \lambda} - 1\right) \approx \left[\frac{\beta_x}{\mu_y - \mu_x} - 1\right] \varphi_x(1, 0).$$

Hence,

$$K = \begin{cases} \frac{\beta_y - \mu_y + \mu_x}{\mu_y - \mu_x} \varphi_y(1, 0) & \text{for } \underline{\lambda} = -\mu_x, \\ \frac{\beta_x - \mu_x + \mu_y}{\mu_y - \mu_x} \varphi_x(1, 0) & \text{for } \underline{\lambda} = -\mu_y. \end{cases}$$

Finally, verify that the conditions for the existence of nontrivial exponential solutions is

$$\begin{aligned}\mu_x &> \mu_y - \frac{\beta_y \varphi_y(1, 0)}{\mu_y + \sigma + \varphi_y(1, 0)}, \\ \mu_y &> \mu_x - \frac{\beta_x \varphi_x(1, 0)}{\mu_x + \sigma + \varphi_x(1, 0)}.\end{aligned}$$

- (g) Look at the proportions of each group (single males, single females, and pairs) and think of how one may study the stability of the exponential solutions that you have just shown to exist.

References: Castillo-Chavez and Huang (1995), Castillo-Chavez, Huang, and Li (1996), Castillo-Chavez and Hsu Schmitz (1993), Castillo-Chavez, Shyu, Rubin, and Umbauch (1992). (1992).

Chapter 5

Continuous Models for Two Interacting Populations

5.1 Species in Competition

In this chapter we will consider populations of two interacting species with population sizes $x(t)$ and $y(t)$, respectively, modeled by a system of two first order differential equations:

$$\begin{aligned}x' &= F(x, y), \\ y' &= G(x, y).\end{aligned}\tag{5.1}$$

There are different kinds of biological interaction represented mathematically by different signs for the rates of change of the growth rates $F(x, y)$ and $G(x, y)$ with respect to the population sizes x and y . We begin with the case of species in *competition*, perhaps competing for the same nutrient resource. Competition means that an increase in the size of either population tends to decrease the growth rate of the other population, so that

$$F_y(x, y) < 0, \quad G_x(x, y) < 0.\tag{5.2}$$

We also assume that an increase in the size of either population tends to inhibit the growth rate of that population, so that if x and y are not too close to zero,

$$F_x(x, y) < 0, \quad G_y(x, y) < 0.\tag{5.3}$$

The classical experimental result on species in competition is the *principle of competitive exclusion* of G.F. Gause (1934), that two species competing for the same resource cannot coexist. The experimental evidence is somewhat equivocal, and there is considerable doubt about the universality of the principle.

We will discuss primarily models in which the per capita growth rates are linear of the form

$$\begin{aligned}x' &= x(\lambda - ax - by), \\y' &= y(\mu - cx - dy),\end{aligned}\tag{5.4}$$

where λ, μ, a, b, c, d are positive constants, but we will also indicate extensions to models with nonlinear growth rates.

For competition models with linear per capita growth rates, there are four equilibria, as follows:

- I. $(0, 0)$, with community matrix

$$\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix},$$

necessarily an unstable node.

- II. $(K, 0)$, with $K = \lambda/a$. The community matrix is

$$\begin{bmatrix} -\lambda & -\frac{b\lambda}{a} \\ 0 & \mu - \frac{c\lambda}{a} \end{bmatrix},$$

with trace $\frac{1}{a}(a\mu - a\lambda - c\lambda)$ and determinant $\frac{\lambda}{a}(c\lambda - a\mu)$.

- III. $(0, M)$, with $M = \frac{\mu}{d}$. The community matrix is

$$\begin{bmatrix} \lambda - \frac{b\mu}{d} & 0 \\ -\frac{\mu c}{d} & -\mu \end{bmatrix},$$

with trace $\frac{1}{d}(d\lambda - d\mu - b\mu)$ and determinant $\frac{\mu}{d}(b\mu - d\lambda)$.

- IV. (x_∞, y_∞) , with $x_\infty > 0, y_\infty > 0, ax_\infty + by_\infty = \lambda, cx_\infty + dy_\infty = \mu$, or

$$x_\infty = \frac{d\lambda - b\mu}{ad - bc}, \quad y_\infty = \frac{a\mu - c\lambda}{ad - bc}.$$

The community matrix is

$$\begin{bmatrix} -ax_\infty & -bx_\infty \\ -cy_\infty & -dy_\infty \end{bmatrix},$$

with trace $-ax_\infty - dy_\infty < 0$ and determinant $x_\infty y_\infty(ad - bc)$.

We will distinguish four cases, corresponding to the four possible sign combinations for $b\mu - d\lambda$ and $a\mu - c\lambda$, but ignoring the possibilities $b\mu - d\lambda = 0$ (which would give $x_\infty = 0$) and $a\mu - c\lambda = 0$ (which would give $y_\infty = 0$). An equilibrium $(K, 0)$ or $(0, M)$ describes a situation in which one species survives but the other species loses the struggle for existence and becomes extinct. An equilibrium (x_∞, y_∞) with $x_\infty > 0, y_\infty > 0$ describes coexistence of the two species.

Because

$$\begin{aligned}
\Delta &= (-ax_\infty - dy_\infty)^2 - 4(ad - bc)x_\infty y_\infty \\
&= a^2 x_\infty^2 + 2ad x_\infty y_\infty + d^2 y_\infty^2 - 4ad x_\infty y_\infty + 4bc x_\infty y_\infty \\
&= a^2 x_\infty^2 - 2ad x_\infty y_\infty + d^2 y_\infty^2 + 4bc x_\infty y_\infty \\
&= (ax_\infty - dy_\infty)^2 + 4bc x_\infty y_\infty > 0,
\end{aligned}$$

and because it is easy to verify that $\Delta \geq 0$ for equilibria with either or both species absent, it is not possible for the system to have a spiral point or a center.

Case 1. $d\lambda - b\mu > 0, a\mu - c\lambda > 0$. In this case $d/b > \mu/\lambda > c/a$, and $ad - bc > 0$. Thus, there is an equilibrium (x_∞, y_∞) in the first quadrant of the phase plane. Because the determinant of the community matrix is positive while the trace is negative, this equilibrium is asymptotically stable and must be an asymptotically stable node. Since the community matrices for the equilibrium $(K, 0)$ and $(0, M)$ both have negative determinants, these equilibria are saddle points. The phase portrait is as shown in Figure 5.1: In this figure and in the other figures in the section, the symbol \swarrow means that orbits are directed downward and to the left, etc. Equilibria are labeled \bullet , with \textcircled{S} indicating that the equilibrium is asymptotically stable. In this case, every orbit tends to (x_∞, y_∞) as $t \rightarrow \infty$, indicating coexistence of the two species for all initial population sizes.

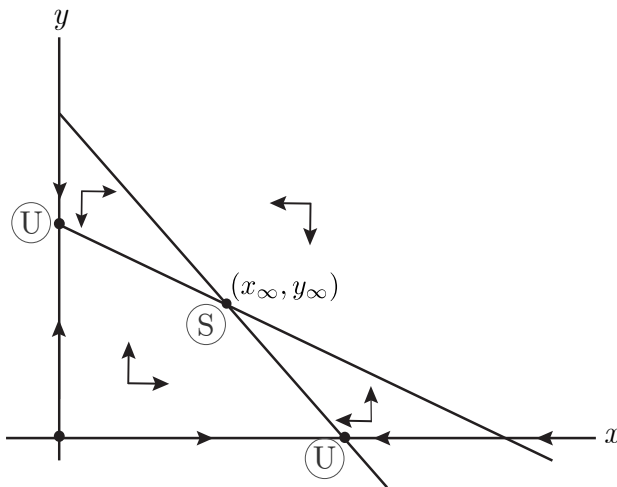


Fig. 5.1 Coexistence

Case 2. $c\lambda - a\mu > 0, b\mu - d\lambda > 0$. In this case $c/a > \mu/\lambda > d/b$, and $ad - bc < 0$. Again, there is an equilibrium (x_∞, y_∞) in the first quadrant, but because the determinant of the community matrix is negative, this equilibrium is a saddle point. The equilibria $(K, 0)$ and $(0, M)$ are both asymptotically stable nodes because their community matrices have positive determinant and negative trace.

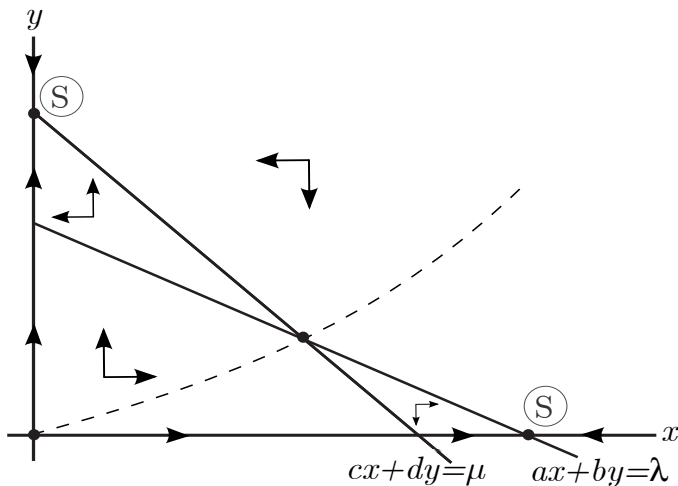


Fig. 5.2 Survival depending on initial state.

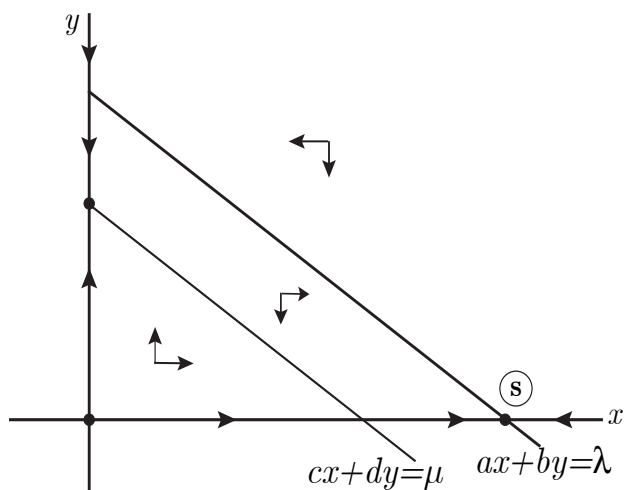
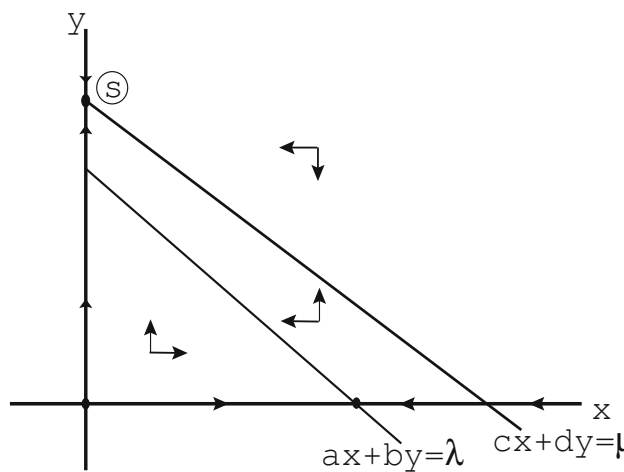
These separatrices divide the first quadrant into two regions, one containing initial points for which orbits tend to $(K, 0)$, and the other containing initial points for which orbits tend to $(0, M)$. In this case, coexistence of the two species is impossible, and one species will always win the competition for survival. The winner of this competition is determined by the initial population sizes. Although coexistence would be the result of initial population sizes on the separatrices, this possibility is too sensitive to disturbances to be biologically meaningful.

Case 3. $c\lambda - a\mu > 0, d\lambda - b\mu > 0$. In this case there is no equilibrium in the interior of the first quadrant. It is easy to verify by examination of the community matrices that the equilibrium at $(K, 0)$ is an asymptotically stable node, while the equilibrium at $(0, M)$ is a saddle point (Figure 5.3).

All orbits tend to $(K, 0)$ as $t \rightarrow \infty$, corresponding to extinction of the y species and survival of the x species for all initial population sizes.

Case 4. $a\mu - c\lambda > 0, b\mu - d\lambda > 0$. Again, there is no equilibrium in the interior of the first quadrant. It is again easy to determine the stability of the equilibria on the axes. Now $(K, 0)$ is a saddle point, while $(0, M)$ is an asymptotically stable node (Figure 5.4). All orbits tend to $(0, M)$ as $t \rightarrow \infty$, corresponding to extinction of the x species and survival of the y species for all initial population sizes.

The numbers K and M are the respective carrying capacities of the x and y species, the equilibrium population size that each species would reach in the absence of the other species. Sometimes, competition between species is described as “qualified” or “unqualified,” according to the inhibiting effect of each species on the other. We will say that the competition is *unqualified* if $x_\infty/K + y_\infty/M \leq 1$, that is, if competition tends to reduce the total population size (where the carrying capacities are the units of population size). This is intended to describe a situation in

Fig. 5.3 Survival of x -species.Fig. 5.4 Survival of y -species.

which both species depend on a common source of food and the competition makes survival more difficult for both species. Because

$$x_{\infty} = \frac{d\lambda - b\mu}{ad - bc}, \quad y_{\infty} = \frac{a\mu - c\lambda}{ad - bc},$$

the condition $x_{\infty}/K + y_{\infty}/M \leq 1$ is equivalent to

$$\frac{a}{\lambda} \frac{d\lambda - b\mu}{ad - bc} + \frac{d}{\mu} \frac{a\mu - c\lambda}{ad - bc} \leq 1,$$

which is equivalent to

$$a\mu(d\lambda - b\mu) + d\lambda(a\mu - c\lambda) \leq (ad - bc)\lambda\mu.$$

If $ad - bc > 0$, this reduces to $ab\mu^2 - (ad + bc)\lambda\mu + cd\lambda^2 \geq 0$, or $(c\lambda - a\mu)(d\lambda - b\mu) \geq 0$. Thus, the system is in Case 3 or 4. If $ad - bc < 0$ the inequalities are reversed, so that $(c\lambda - a\mu)(d\lambda - b\mu) < 0$, which shows that the system is in Case 2. This shows that the assumption of unqualified competition rules out the possibility of coexistence of the two species (Case 1). The *principle of competitive exclusion*, originally formulated by Gause in 1934 on the basis of experimental evidence, states that coexistence under unqualified competition is impossible. We have shown that the principle of competitive exclusion is a consequence of the assumption of linear per capita growth rates. The principle itself is quite controversial; there is experimental evidence suggesting that coexistence under unqualified competition is possible. Our argument shows only that the principle of competitive exclusion is a consequence of the specific model hypothesized, not that the principle has biological validity. It is important to remember that if a mathematical model predicts behavior that contradicts valid observations, one must conclude that there is a flaw in the model.

The condition $x_\infty/K + y_\infty/M \leq 1$ of unqualified competition says geometrically that the equilibrium (x_∞, y_∞) lies below the line $x/K + y/M = 1$ joining the equilibria $(K, 0)$ and $(0, M)$. The condition that the equilibrium (x_∞, y_∞) is asymptotically stable is that the crossing of the curves $x' = 0$ and $y' = 0$ at (x_∞, y_∞) must be as in Figure 5.5.

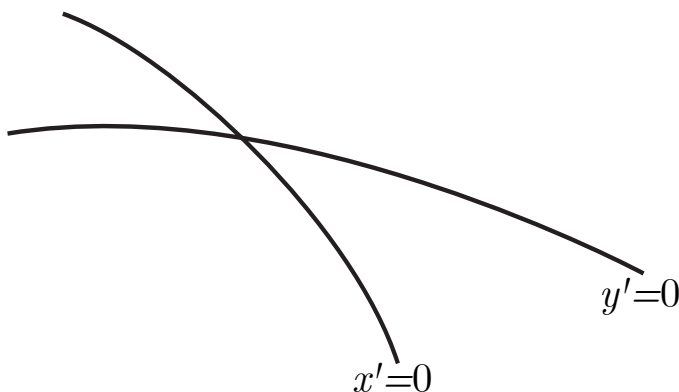


Fig. 5.5 Coexistence slope relation.

It may be seen by looking at the phase portraits in Cases 1 and 2 that if the *isoclines* $x' = 0$ and $y' = 0$ are straight lines, these two conditions are incompatible. However, it is possible to construct models of the form $x' = xf(x, y)$, $y' = yg(x, y)$ for species in competition for which coexistence and unqualified competition are compatible. An example of such a model, that has been proposed [Ayala, Gilpin,

and Ehrenfeld (1973)] is

$$\begin{aligned}x' &= \lambda x \left(1 - \left(\frac{x}{K_1} \right)^{\theta_1} - \frac{b}{\lambda} y \right), \\y' &= \mu y \left(1 - \left(\frac{y}{K_2} \right)^{\theta_2} - \frac{c}{\mu} x \right).\end{aligned}$$

In describing two species in competition by a model of the form

$$\begin{aligned}x' &= xf(x, y), \\y' &= yg(x, y),\end{aligned}$$

it is reasonable to assume that $f_x(x, y) < 0$ and $g_y(x, y) < 0$ for $x > 0, y > 0$, and also that $f_y(x, y) < 0, g_x(x, y) < 0$ for $x > 0, y > 0$. This assumption says that an increase in the population size of either species reduces the growth rate of the population size of both species. In order to locate equilibria, we examine the *isoclines* of the two species, namely the curves $f(x, y) = 0$ and $g(x, y) = 0$, respectively. Just as in the case of linear per capita growth rates, there are four types of equilibria: (I) $(0, 0)$; (II) $(K, 0)$, where K now is defined by $f(K, 0) = 0$; (III) $(0, M)$, where M now is defined by $g(0, M) = 0$; (IV) (x_∞, y_∞) , with $x_\infty > 0, y_\infty > 0$, where $f(x_\infty, y_\infty) = 0, g(x_\infty, y_\infty) = 0$.

At an intersection (x_∞, y_∞) of the isoclines $f(x, y) = 0$ and $g(x, y) = 0$, the slope of the x isocline is $-f_x(x_\infty, y_\infty)/f_y(x_\infty, y_\infty)$, and the slope of the y -isocline is $-g_x(x_\infty, y_\infty)/g_y(x_\infty, y_\infty)$. Because of the hypothesis $f_x < 0, f_y < 0, g_x < 0, g_y < 0$, each of these slopes is negative. The condition that the slope of the x isocline is less (more negative) than the slope of the y isocline is

$$-\frac{f_x(x_\infty, y_\infty)}{f_y(x_\infty, y_\infty)} < -\frac{g_x(x_\infty, y_\infty)}{g_y(x_\infty, y_\infty)},$$

or

$$f_x(x_\infty, y_\infty)g_y(x_\infty, y_\infty) - f_y(x_\infty, y_\infty)g_x(x_\infty, y_\infty) > 0.$$

The community matrix at the equilibrium is

$$\begin{bmatrix} x_\infty f_x(x_\infty, y_\infty) & x_\infty f_y(x_\infty, y_\infty) \\ y_\infty g_x(x_\infty, y_\infty) & y_\infty g_y(x_\infty, y_\infty) \end{bmatrix},$$

with trace

$$x_\infty f_x(x_\infty, y_\infty) + y_\infty g_y(x_\infty, y_\infty) < 0,$$

and determinant

$$x_\infty y_\infty (f_x(x_\infty, y_\infty)g_y(x_\infty, y_\infty) - f_y(x_\infty, y_\infty)g_x(x_\infty, y_\infty)).$$

Thus the determinant is positive, so that the equilibrium is asymptotically stable, if and only if the slope of the x -isocline is less than the slope of the y -isocline; otherwise the equilibrium is a saddle point. There are four cases, exactly as for

linear per capita growth rates, except that the isoclines are no longer necessarily straight lines. Otherwise, all the analysis carries over to this more general situation. Isoclines that are concave upward allow the possibility of an equilibrium below the line $x/K + y/M = 1$, so that the competition is unqualified. Such an equilibrium may be either asymptotically stable or unstable (Figure 5.6). Thus, the principle of competitive exclusion is not a consequence of these more general hypotheses.

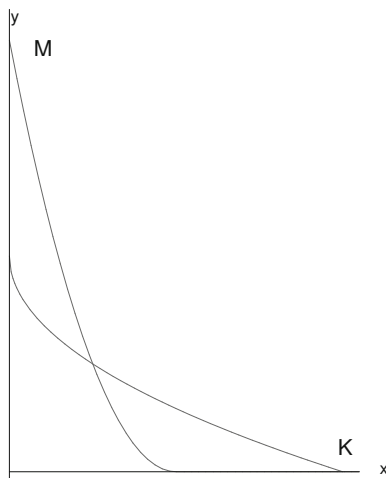


Fig. 5.6 Coexistence equilibrium.

Every orbit of a system $x' = xf(x,y)$, $y' = yg(x,y)$ for which $f(x,y) < 0$ when $x > K$, and $g(x,y) < 0$ when $y > M$ must be bounded because x and y must both be decreasing functions of t for x and y , respectively, large enough. According to the Poincaré–Bendixson theorem, every orbit then must tend either to an equilibrium or to a limit cycle. For competition models with $f_x(x,y) < 0$, $g_y(x,y) < 0$ for $x > 0, y > 0$, we have seen in Section 4.5 that the Dulac criterion rules out the possibility of periodic solutions. Therefore every orbit must tend to an equilibrium, and by analyzing the nature of each equilibrium we have actually described the global behavior of solutions, not merely the behavior of solutions starting near an equilibrium.

Example 1. Determine the outcome of a competition modeled by the system

$$\begin{aligned}x' &= x(100 - 4x - y), \\y' &= y(60 - x - 2y).\end{aligned}$$

Solution: A coexistence equilibrium is found by solving the system

$$\begin{aligned}4x + y &= 100, \\x + 2y &= 60\end{aligned}$$

By eliminating one of the variables, we can obtain the equilibrium $(20, 20)$. The community matrix at an equilibrium (x_∞, y_∞) is

$$\begin{bmatrix} -4x_\infty & -y_\infty \\ -x_\infty & -2y_\infty \end{bmatrix},$$

which in this case is

$$\begin{bmatrix} -80 & -20 \\ -20 & -40 \end{bmatrix},$$

with negative trace and positive determinant. Thus, the coexistence equilibrium is asymptotically stable and every orbit approaches this equilibrium. The two species coexist.

Exercises

In each of Exercises 1 through 4, determine the outcome of the competition modeled by the given system.

1. $x' = x(80 - x - y)$, $y' = y(120 - x - 3y)$ [see Exercise 7, Section 4.5].
2. $x' = x(60 - 3x - y)$, $y' = y(75 - 4x - y)$ [see Exercise 8, Section 4.5].
3. $x' = x(40 - x - y)$, $y' = y(90 - x - 2y)$ [see Exercise 9, Section 4.5].
4. $x' = x(80 - 3x - 2y)$, $y' = y(80 - x - y)$ [see Exercise 10, Section 4.5].
5. What is the outcome of a competition modeled by the system

$$\begin{aligned} x' &= x(2 - x - x^2 - y) \\ y' &= y(16 - 2x - x^2 - y)? \end{aligned}$$

- 6.* What is the outcome of a competition modelled by the system

$$\begin{aligned} x' &= x(80 - x - y), \\ y' &= y(120 - x - 3y - 2y^2)? \end{aligned}$$

5.2 Predator–Prey Systems

We have already discussed the Lotka–Volterra system

$$x' = x(\lambda - by), \quad y' = y(-\mu + cx), \quad (5.5)$$

which unrealistically predicts population oscillations that have been observed in real populations (see Section 4.1). The reason for describing this prediction as “unrealistic” is that the model is extremely sensitive to perturbations. A change in initial population size would produce a change to a different periodic orbit, while the ad-

dition of a perturbing term to the system of differential equations could produce the same type of change or could produce a *qualitative* change in the behavior of orbits, which might either spiral in to an equilibrium or spiral out from an equilibrium.

Let us try to construct a more realistic model by assuming that in the absence of predators, the prey species would obey a logistic model. This would suggest a system

$$x' = x(\lambda - ax - by), \quad y' = y(-\mu + cx). \quad (5.6)$$

There are two possible direction fields, depending on the sign of $c\lambda - a\mu$, as shown in Figure 5.7.

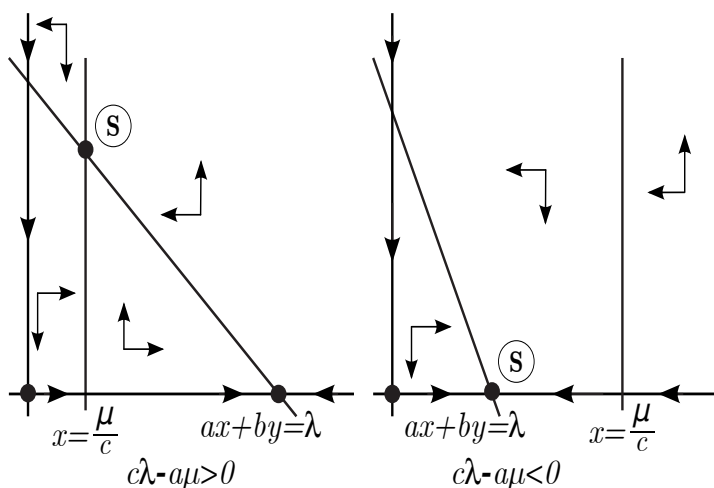


Fig. 5.7 Predator-prey model alternatives.

The community matrix at the interior equilibrium (x_∞, y_∞) (if there is one) is

$$\begin{bmatrix} -ax_\infty & -bx_\infty \\ cy_\infty & 0 \end{bmatrix},$$

with determinant $bcx_\infty y_\infty > 0$ and trace $-ax_\infty < 0$; thus (x_∞, y_∞) is asymptotically stable. The equilibrium (x_∞, y_∞) may be a spiral point, so that oscillations are possible, but only damped oscillations. The Dulac criterion shows that there cannot be a periodic orbit. It is easy to verify that if $c\lambda - a\mu > 0$ the equilibrium $(\lambda/a, 0)$ is a saddle point, while if $c\lambda - a\mu < 0$, the equilibrium $(\lambda/a, 0)$ is an asymptotically stable node. Thus in each case, there is exactly one asymptotically stable equilibrium to which every orbit tends.

Another plausible modification of the Lotka–Volterra system is

$$x' = x(\lambda - ax - by), \quad y' = y(\mu + cx - dy), \quad (5.7)$$

with μ allowed to be either positive or negative. The direction field is as in Figure 5.8.

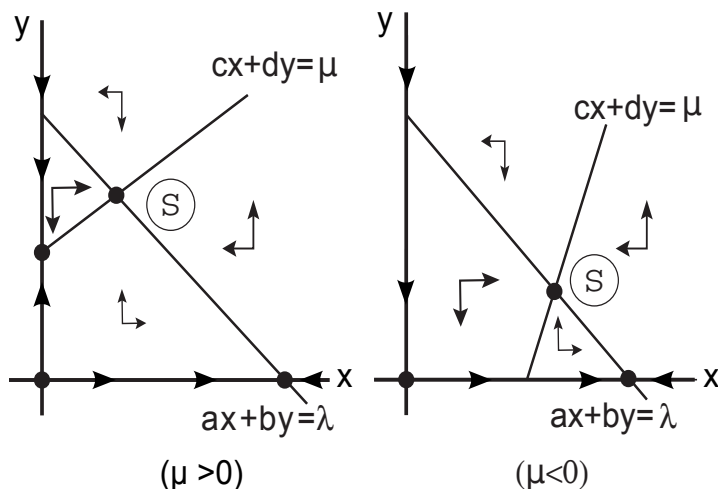


Fig. 5.8 Coexistence of predators and prey.

(We shall omit the cases in which there is no interior equilibrium, but the reader may easily sketch these also.) At the interior equilibrium (x_∞, y_∞) the community matrix is

$$\begin{bmatrix} -ax_\infty & -bx_\infty \\ cy_\infty & -dy_\infty \end{bmatrix},$$

with determinant $(ad + bc)x_\infty y_\infty > 0$ and trace $-ax - dy < 0$; therefore, (x_∞, y_∞) is asymptotically stable. Again, damped oscillations are possible but there cannot be a periodic orbit.

In order to obtain a possibility of periodic orbits without the extreme sensitivity of the Lotka–Volterra model, we must assume nonlinear per capita growth rates. We consider models of the form

$$\begin{aligned} x' &= xf(x) - xy\phi(x), \\ y' &= y(cx\phi(x) - e). \end{aligned} \tag{5.8}$$

In the absence of predators ($y \equiv 0$), this model would reduce to $x' = xf(x)$. Thus, $xf(x)$ represents the growth rate of the prey species by itself. The term $xy\phi(x)$ is called the *predator functional response*; $x\phi(x)$ is the number of prey consumed per predator in unit time. The constant c is the conversion efficiency of prey into predator food, and the term $cxy\phi(x)$ is called the *predator numerical response*. The constant e is the predator mortality rate. Models of the form (5.8) are known as Rosenzweig–MacArthur models, having been proposed by M.L. Rosenzweig and R.H. MacArthur (1963).

In (5.8) it is reasonable to assume

$$\phi(x) \geq 0, \quad \phi'(x) \leq 0, \quad [x\phi(x)]' \geq 0, \quad (5.9)$$

and that $x\phi(x)$ is bounded as $x \rightarrow \infty$. These assumptions express the idea that, as prey population increases, the consumption rate of prey per predator increases but that the fraction of the total prey population consumed per predator decreases. Some explicit forms for the predator functional response that have been used are

$$\begin{aligned} x\phi(x) &= \frac{\alpha x}{x+A} && [\text{Holling (1965)}], \\ x\phi(x) &= a(1 - e^{-cx}) && [\text{Ivlev (1961)}], \\ x\phi(x) &= ax^q \quad (q < 1) && [\text{Rosenzweig (1971)}]. \end{aligned}$$

In order to study the system (5.8), we plot the *isoclines*, which are the curves in the (x, y) -plane, or phase plane, along which $x' = 0$ and $y' = 0$. These curves should properly be called *nullclines*. The predator isocline is $cx\phi(x) = e$, which is a vertical straight line. To plot the prey isocline $y = f(x)/\phi(x)$, we proceed as follows: Sketch the graph of $xf(x)$ and of $y_i x\phi(x)$ for a sequence of values y_i . If $y_i x\phi(x)$ intersects $xf(x)$ for $x = x_i$, then (x_i, y_i) is a point on the prey isocline. The prey isocline may then be sketched by projecting to another graph in the (x, y) -plane (Figure 5.9).

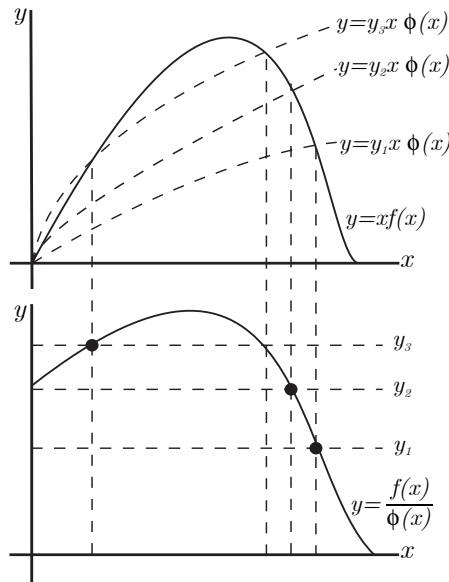


Fig. 5.9 Prey isocline.

Observe that the prey isocline may have a maximum if the graph of $x\phi(x)$ is concave down, so that $y_i x\phi(x)$ and $xf(x)$ may intersect in more than one point for some

values of y_i . There are, in fact, biological reasons for expecting the prey isocline to have a maximum [Rosenzweig (1969)]. It is reasonable for the function $x\phi(x)$ to increase with x but approach a maximum when the number of prey is large. As the prey population increases it should be able to survive in the face of an increase in the predator population because of this saturation, but when the prey population gets still larger, self-limiting effects tend to take over to bring the prey isocline down. It is possible to incorporate an *Allee effect*, in which the prey population is required to maintain a minimum size for survival, by having the prey isocline intersect the x -axis at two points (Figure 5.10).

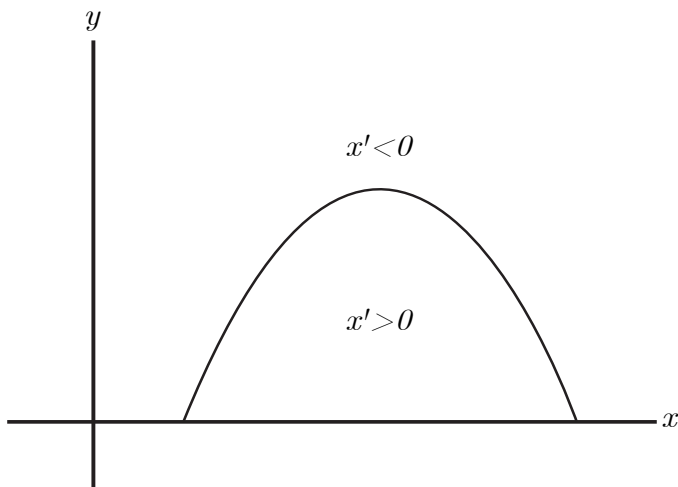


Fig. 5.10 Allee effect.

The Lotka–Volterra model (5.5) and the refinements (5.6), (5.7) with linear per capita growth rates, do not allow any bending of the isoclines. Also, the refinements (5.6), (5.7) satisfy $f_x(x, y) < 0$, $g_y(x, y) \leq 0$ in the first quadrant of the phase plane. Thus, by the Dulac criterion they cannot have periodic orbits.

For a Rosenzweig–MacArthur model (5.8) with an equilibrium (x_∞, y_∞) given by $x_\infty\phi(x_\infty) = e/c$, $y_\infty = f(x_\infty)/\phi(x_\infty) = cf(x_\infty)/e$, the community matrix at equilibrium is

$$\begin{bmatrix} x_\infty f'(x_\infty) - x_\infty y_\infty \phi'(x_\infty) & -x_\infty \phi(x_\infty) \\ cy_\infty [x\phi(x)]'_{x_\infty} & 0 \end{bmatrix},$$

whose determinant is $cx_\infty y_\infty \phi(x_\infty) [x\phi(x)]'_{x_\infty} > 0$, and whose trace is $x_\infty (f'(x_\infty) - y_\infty \phi'(x_\infty))$.

The slope of the prey isocline at the equilibrium is

$$\frac{\phi(x_\infty)f'(x_\infty) - f(x_\infty)\phi'(x_\infty)}{(\phi(x_\infty))^2} = \frac{f'(x_\infty) - y_\infty\phi'(x_\infty)}{\phi(x_\infty)},$$

which has the same sign as the trace of the community matrix. Thus, the equilibrium is asymptotically stable if the prey isocline has negative slope at equilibrium and unstable if the prey isocline has positive slope at equilibrium (Figure 5.11).

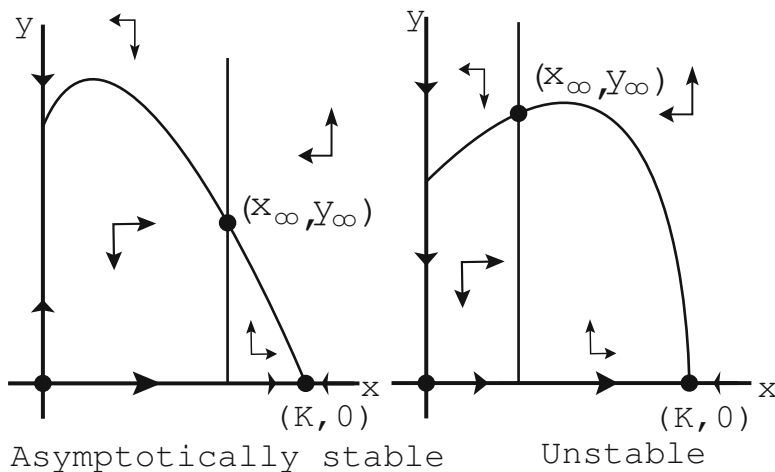


Fig. 5.11

No orbit can cross either the positive x axis or the positive y axis. An orbit starting on the y -axis tends to the origin, while an orbit starting on the x -axis tends to the equilibrium on the positive x -axis. Every orbit starting in the interior of the first quadrant, if bounded, remains in the interior of the first quadrant and must, by the Poincaré–Bendixson theorem, tend to the equilibrium (x_∞, y_∞) or to the equilibrium $(K, 0)$ on the positive x -axis or to a limit cycle around the equilibrium (x_∞, y_∞) . We will prove shortly in a more general setting that every orbit must remain bounded. If there is no equilibrium (x_∞, y_∞) , that is, if $c\phi(K) < e$, where $f(K) = 0$, then the equilibrium $(K, 0)$ is asymptotically stable and every orbit tends to it (Figure 5.12).

If the prey isocline has negative slope at (x_∞, y_∞) , then (x_∞, y_∞) is asymptotically stable and every orbit tends to it; $(K, 0)$ is a saddle point (Figure 5.13). If the prey isocline has positive slope at (x_∞, y_∞) , then (x_∞, y_∞) is unstable and there must be a limit cycle around (x_∞, y_∞) to which every orbit tends (Figure 5.14). We ignore the pathological (and very sensitive to perturbations) possibility of an infinite sequence of periodic orbits around (x_∞, y_∞) and assume that there can be only a single limit cycle. Then every orbit spirals either outward or inward to this limit cycle. Thus we have a more plausible model than the original Lotka–Volterra model to explain the observed periodic behavior.

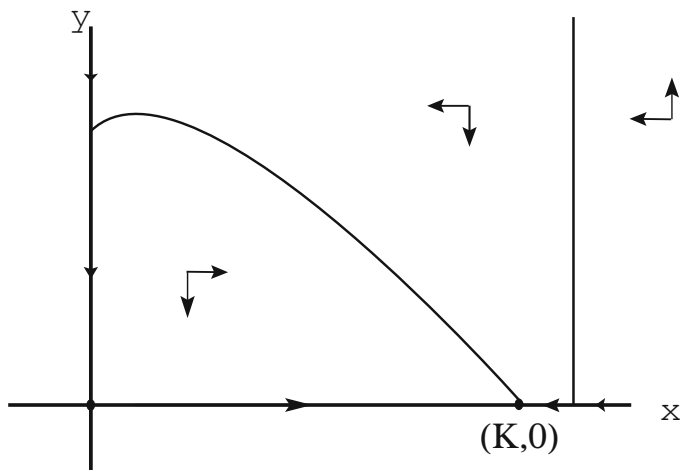


Fig. 5.12

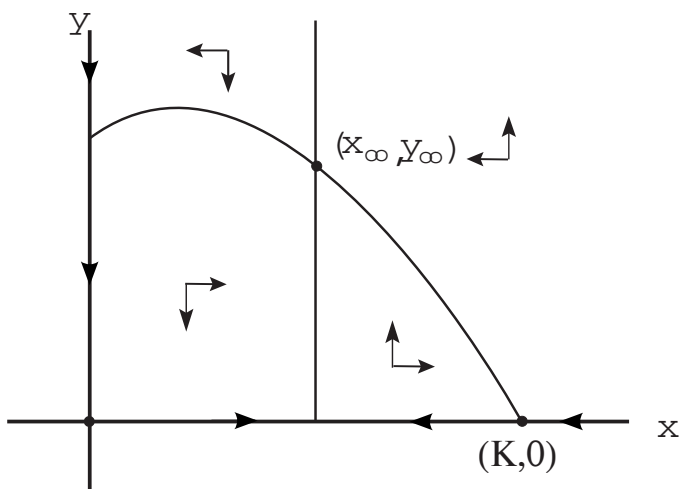


Fig. 5.13 Saddle point.

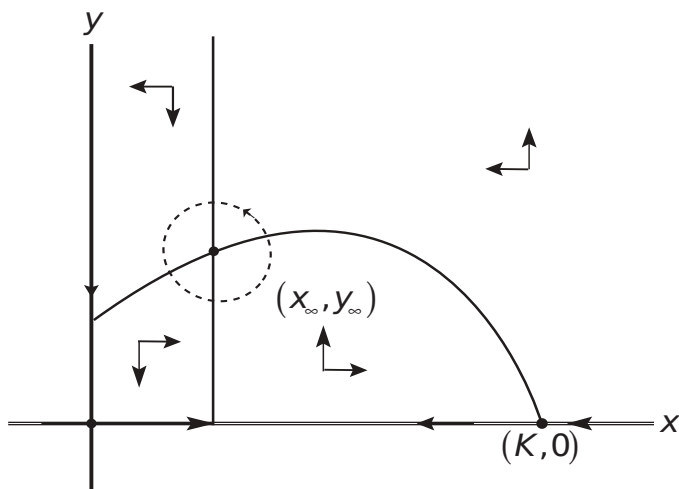
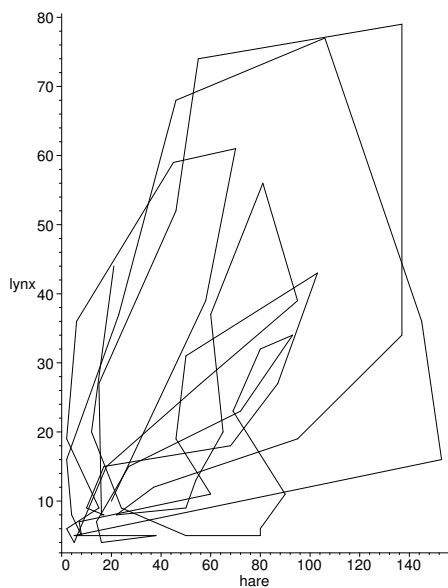
Another classical example of interacting populations in which oscillations have been observed is the data collected by the Hudson's Bay Company in Canada during the period 1821–1940 on furs of the snowshoe hare (*Lepus Americanus*) and Canadian lynx (*Lynx Canadensis*) brought to the company by trappers. Table 5.1 shows some data for the period 1848 to 1907, with the hare data estimated from the data shown in graph form in [MacLulich (1937)] and the lynx data taken from [Elton and Nicholson (1942)], and with population numbers given in thousands. In each case, the data were obtained by analyzing a variety of reports. This data is also depicted in Figure 5.15 in the phase (hare–lynx) plane.

Year	Hares	Lynx
1848	21	44
1849	12	20
1850	24	9
1851	50	5
1852	80	5
1853	80	6
1854	90	11
1855	69	23
1856	80	32
1857	93	34
1858	72	23
1859	27	15
1860	14	7
1861	16	4
1862	38	5
1863	5	5
1864	153	16
1865	145	36
1866	106	77
1867	46	68
1868	23	37
1869	2	16
1870	4	8
1871	8	5
1872	7	7
1873	60	11
1874	46	19
1875	50	31
1876	103	43
1877	87	27

Year	Hares	Lynx
1878	68	18
1879	17	15
1880	10	9
1881	17	8
1882	16	8
1883	15	27
1884	46	52
1885	55	74
1886	137	79
1887	137	34
1888	95	19
1889	37	12
1890	22	8
1891	50	9
1892	54	13
1893	65	20
1894	60	37
1895	81	56
1896	95	39
1897	56	27
1898	18	15
1899	5	4
1900	2	6
1901	15	9
1902	2	19
1903	6	36
1904	45	59
1905	50	61
1906	58	39
1907	20	10

Table 5.1 Hare and Lynx population sizes.

There are many problems with this data as a measure of population sizes. One problem is that trapping data may not accurately describe the total population sizes, and another problem is the varying delays between catching animals and bringing their skins to a trading post. For example, it is pointed out in [Elton & Nicholson (1942)] that the data in [MacLulich (1937)] should be shifted forward by a year. In the earlier studies, it was found that the population sizes of hares, lynx, and also several other species, fluctuated in time with a period of approximately 10 years. The lynx oscillations were ascribed to the fact that hare is the main food for lynx. However, the relation was not viewed as a predator–prey interaction until the systems view was introduced into theoretical ecology, notably in the text by Odum (1953), which was the first to display the lynx and hare data in the same graph. This appears to be another example of a predator–prey system with the oscillations being explained as a limit cycle. However, a graph of the hare and lynx populations in the hare (prey)–lynx (predator) plane shows orbits going clockwise in some

**Fig. 5.14** Limit cycle.**Fig. 5.15** The hare-lynx phase plane

cycles, whereas the limit cycle predicted by a Rosenzweig–MacArthur model is always counterclockwise. This does not show that hares eat lynx [Gilpin (1973)]! It indicates that something is wrong with the model. Various suggestions have been made to explain the anomaly. One possibility is that the numbers of animals of the two species caught are not necessarily proportional to the actual population sizes. The trappers should be viewed as predators who choose which type of animal to

pursue rather than as external experimenters who measure population size. Lynx are useful for their fur, which is more valuable than rabbit fur, but rabbits are also useful as food for trappers. Thus, when rabbits are abundant and food is plentiful, hunters can conveniently pursue both hares and lynx, since lynx tend to go where there are rabbits. When rabbits are scarce, trappers must concentrate on rabbits for food and therefore will catch fewer lynx. An important lesson to be learned is that population models outside the laboratory are seldom simple and comprehensive, while experimental data are also subject to substantial errors. Predictions of qualitative behavior may be possible, but quantitative data are unreliable, especially over long periods of time.

There is an even more basic flaw in the predator–prey model as an explanation for the lynx-hare cycle: Hare population size oscillates even in the absence of lynx, probably due to climatic variations and to epidemics that recur, killing most of the hares. Also, hares could have other predators. A two-species predator–prey model, whether of Lotka–Volterra or Rosenzweig–MacArthur type cannot exhibit oscillations when only one species is present.

Example 1. In Section 4.2 we formulated the model for the simple chemostat

$$\begin{aligned} C' &= q(C^{(0)} - C) - \frac{\beta a C y}{C + A}, \\ y' &= \frac{a C y}{C + A} - q y, \end{aligned} \tag{5.10}$$

with a, A, q , and β constants. This is a Rosenzweig–MacArthur model with the nutrient concentration C playing the role of the prey population and the bacteria population size y playing the role of the predator population. In the model (5.8) we have

$$\begin{aligned} x f(x) &= q(C^{(0)} - C), \\ x \phi(x) &= \frac{\beta a C}{C + A}, \\ c &= \frac{1}{\beta}, \\ e &= q. \end{aligned}$$

For (5.10), the prey isocline is the curve

$$y = \frac{q(C + A)(C^{(0)} - C)}{a\beta C},$$

which is unbounded as $C \rightarrow 0$, and it is not difficult to verify that $dy/dC < 0$ at every point on the prey isocline, and that the prey isocline meets the C -axis for $C = C^{(0)}$, while the predator isocline is the line $C = Aq/(a - q)$. Because the prey isocline is monotone decreasing, an intersection of the two isoclines with $C > 0$, $y > 0$ is asymptotically stable. There are two possibilities for this behavior. Either

the two isoclines do not intersect, and the only equilibrium is $(C^{(0)}, 0)$, which is asymptotically stable, or there is an asymptotically stable equilibrium with $C > 0$, $y > 0$, and the equilibrium $(C^{(0)}, 0)$ is unstable, as in Figure 5.16.

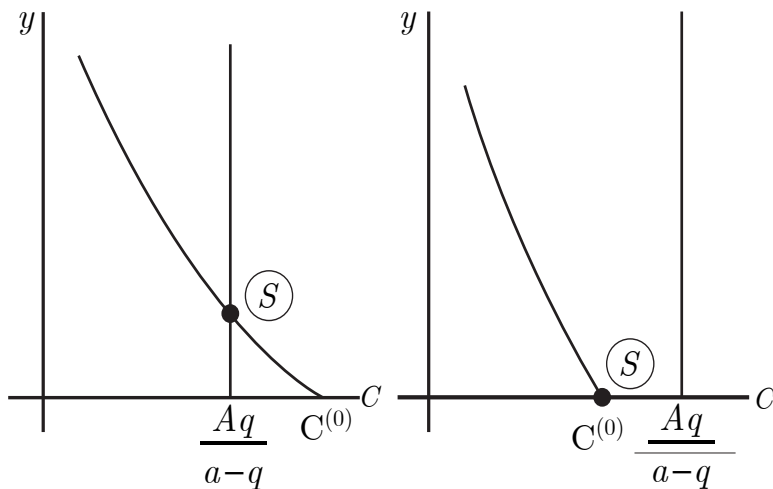


Fig. 5.16 Chemostat equilibrium possibilities.

The condition that the chemostat have an equilibrium with a positive bacteria population is

$$C^{(0)} > \frac{Aq}{a-q}.$$

In terms of the parameter q , which depends on the flow rate and is thus adjustable, this coexistence condition is

$$q < \frac{aC^{(0)}}{A + C^{(0)}}.$$

Example 2. Determine the qualitative behavior of a predator–prey system modeled by

$$\begin{aligned} x' &= x \left(1 - \frac{x}{30} \right) - \frac{xy}{x+10}, \\ y' &= y \left(\frac{x}{x+10} - \frac{1}{3} \right). \end{aligned}$$

Solution. We have studied this system in Example 1, Section 4.5, and have shown that every orbit approaches a periodic orbit around the (unstable) equilibrium $(5, 12.5)$. Thus the two species coexist with oscillations.

Example 3. Determine the equilibrium behavior of a predator–prey system modeled by

$$\begin{aligned}x' &= x \left(1 - \frac{x}{30} \right) - \frac{xy}{x+10}, \\y' &= y \left(\frac{x}{x+10} - \frac{3}{5} \right).\end{aligned}$$

Solution. Equilibria are solutions of the pair of equations

$$x \left(1 - \frac{x}{30} - \frac{y}{x+10} \right) = 0, \quad y \left(\frac{x}{x+10} - \frac{3}{5} \right) = 0.$$

There is an equilibrium $(0,0)$, and if $y = 0$ we must have $x = 30$. There is a coexistence equilibrium obtained by solving

$$\frac{x}{x+10} - \frac{3}{5} = 0, \quad 1 - \frac{x}{30} - \frac{y}{x+10} = 0.$$

The first of these equations gives $x = 15$, and then the second equation gives $y = 12.5$. Thus, there are three equilibria: $(0,0)$, $(30,0)$, and $(15,12.5)$. The community matrix at an equilibrium (x_∞, y_∞) is

$$\begin{bmatrix} 1 - \frac{x_\infty}{15} - \frac{10y_\infty}{(x_\infty+10)^2} & -\frac{x_\infty}{x_\infty+10} \\ \frac{10y_\infty}{(x_\infty+10)^2} & \frac{x_\infty}{x_\infty+10} - \frac{3}{5} \end{bmatrix}.$$

At $(0,0)$, the community matrix is

$$\begin{bmatrix} 1 & 0 \\ 0 & -\frac{3}{5} \end{bmatrix}.$$

and thus $(0,0)$ is unstable. At $(30,0)$ the community matrix is

$$\begin{bmatrix} -1 & -\frac{3}{4} \\ 0 & \frac{3}{20} \end{bmatrix}.$$

Since the determinant of this matrix is negative, $(30,0)$ is also unstable. The community matrix at $(15,12.5)$ is

$$\begin{bmatrix} -\frac{1}{5} & -\frac{3}{5} \\ \frac{1}{5} & 0 \end{bmatrix}$$

with negative trace and positive determinant. Thus $(15,12.5)$ is asymptotically stable, and in fact, every orbit approaches this equilibrium.

Exercises

In each of Exercises 1 through 6, determine the outcome of the predator–prey relationship modeled by the given system

1. $x' = 3x(1 - \frac{x}{20}) - \frac{xy}{x+10}$, $y' = y(\frac{x}{x+10} - \frac{1}{2})$ [see Exercise 1, Section 4.5].
2. $x' = x(1 - \frac{x}{40}) - \frac{2xy}{x+15}$, $y' = 3y(\frac{2x}{x+15} - \frac{6}{5})$ [see Exercise 2, Section 4.5].
3. $x' = 3x(1 - \frac{x}{20}) - \frac{xy}{x+10}$, $y' = y(\frac{x}{x+10} - \frac{1}{6})$ [see Exercise 3, Section 4.5].
4. $x' = x(1 - \frac{x}{40}) - \frac{2xy}{x+15}$, $y' = 3y(\frac{2x}{x+15} - \frac{1}{2})$ [see Exercise 4, Section 4.5].
5. $x' = 3x(1 - \frac{x}{20}) - \frac{xy}{x+10}$, $y' = y(\frac{x}{x+10} - \frac{3}{4})$ [see Exercise 5, Section 4.5].
6. $x' = x(1 - \frac{x}{40}) - \frac{2xy}{x+15}$, $y' = 3y(\frac{2x}{x+15} - \frac{8}{5})$ [see Exercise 6, Section 4.5].
7. Show that the equilibrium (x_∞, y_∞) with $x_\infty > 0$, $y_\infty > 0$ of the predator–prey system modeled by

$$\begin{aligned}x' &= rx \left(1 - \frac{x}{K}\right) - \frac{axy}{x+A}, \\y' &= sy \left(\frac{ax}{x+A} - \frac{aJ}{J+A}\right),\end{aligned}$$

is unstable if $K > A + 2J$, and asymptotically stable if $J < K < A + 2J$ [see Exercise 11, Section 4.5].

8. Use a computer algebra system to plot the hare and lynx populations given in [Table 5.1](#) as functions of time separately but in the same graph.
9. Use a computer algebra system to plot the hare and lynx data given in [Table 5.1](#) over some time intervals of 10–12 years in the phase plane.

5.3 Laboratory Populations: Two Case Studies

Such ideas as competitive exclusion and oscillations in predator–prey systems have been suggested by real-world observations, but the data is never completely unequivocal. There are too many factors that are ignored in simple models for a close fit between model and data. For example, it is rare for a population system to involve only two species without any interaction with other species. However, laboratory experiments may allow more control and give better opportunities to test the validity of models. In this section, we describe two experiments conducted by G. F. Gause (1934a, 1934b), one on two species in competition and one on a predator–prey system.

The possibility of competitive exclusion is illustrated by some experiments conducted by Gause (1934a) on *Paramecium aurelia* and *Paramecium caudatum* (two kinds of protozoa). First, Gause measured the population sizes of each separately and fit the results to logistic models. Then he measured the sizes of the two popula-

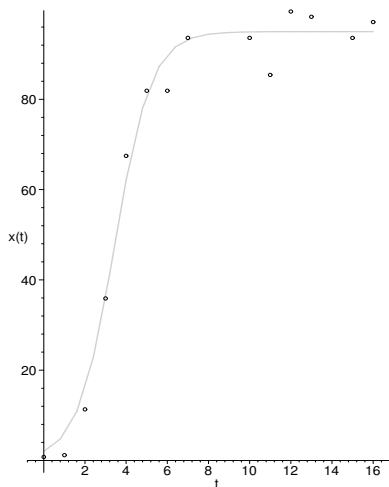
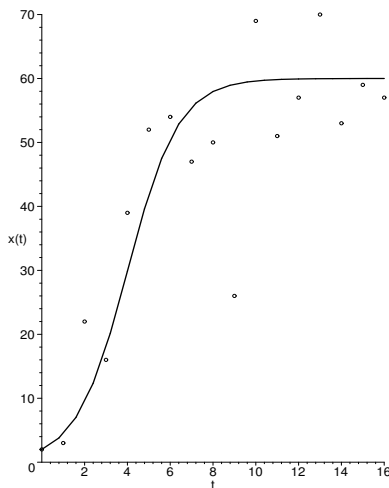
tions together. The results are shown in Table 5.2, with time measured in days and population sizes measured in numbers of individuals per 0.5 cc.

Time	P. aurelia	P. caudatum	Competition	
			P. aurelia	P. caudatum
0	2	2	2	2
1	3	5	4	8
2	29	22	29	20
3	92	16	66	25
4	173	39	141	24
5	210	52	162	-
6	210	54	219	-
7	240	47	153	-
8	-	50	162	21
9	-	26	150	15
10	240	69	175	12
11	219	51	260	9
12	255	57	276	12
13	252	70	285	6
14	270	53	225	9
15	240	59	222	3
16	249	57	220	0

Table 5.2 Paramecium population sizes.

For comparison of the population sizes of the two species, the volumes are used rather than the numbers of individuals. Because *P. caudatum* has larger volume than *P. aurelia*, the volume of *P. caudatum* is taken as the unit and the number of *P. aurelia* is multiplied by 0.39. The general shape of the individual growth curves looks rather like logistic growth, although there appear to be oscillations around the carrying capacity, which are probably caused by small variations in temperature and the composition of the growth medium. Gause therefore fit the single-species data to logistic models. We use slightly different values for the parameters from those obtained by Gause in order to fit his data better. We use intrinsic growth rate 1.12 and carrying capacity 95 for *P. aurelia* and intrinsic growth rate 0.84 and carrying capacity 60 for *P. caudatum*. Figure 5.17 shows the data and model curve for *P. aurelia* and Figure 5.18 shows the data and model curve for *P. caudatum*.

The data for the mixed populations, shown as a phase portrait in Figure 5.19, indicates clearly a *P. caudatum* population dying out and a *P. aurelia* population approaching its single-species carrying capacity. We may fit the data to a model of the form (5.4) with x the density of *P. aurelia* and y the density of *P. caudatum* and parameter values $\lambda = 1.12$, $\mu = 0.84$, $a = 1.12/95$, $d = 0.84/60$, and b and c to be determined. By comparing the conversion of food to biomass for each species, Gause estimated values for $b = 1.8103/105$, $c = 0.4846/64$ in the initial stages, but he considered these parameters to be functions of time. In the model (5.4), these coefficients are constants, and a reasonable fit with the data is obtained with the

Fig. 5.17 *P. aurelia*.Fig. 5.18 *P. caudatum*.

parameter values $b = (1.55)(1.12)/95$, $c = (0.65)(0.84)/60$. Using these parameter values, we may calculate that $c\lambda - a\mu > 0$, $d\lambda - b\mu > 0$. This suggests that the system is in what we called Case 3 in Section 5.1, and thus that *P. aurelia* should survive and *P. caudatum* should become extinct. The behavior of the model (5.4) agrees quite well with the experimental data but the approach to extinction of *P. caudatum* is considerably slower than the experimental data indicates. The experiment is consistent with the principle of competitive exclusion. More elaborate experiments have been carried out [Park (1948)], and these also support the principle of competitive

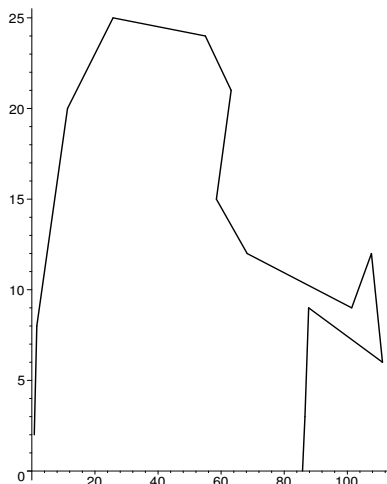


Fig. 5.19 *P. aurelium* and *P. caudatum* in competition

exclusion. However, other experiments, such as those reported in [Ayala, Gilpin, and Ehrenfeld (1973)] cast doubt on the universality of the principle.

In the real world, there are many situations in which species avoid competition by using somewhat different food supplies or by searching for food in different locations. This suggests that the principle of competitive exclusion may be broadly applicable but that species may arrange to coexist in spite of it.

Another set of experiments performed by Gause (1934a, 1934b) involved *Paramecium aurelia* (protozoa) and *Saccharomyces exiguus* (a yeast on which the protozoa feed). Gause obtained the data given in Table 5.3, with time measured in days. Because the original data was displayed only in a graph, the population sizes are estimates read from the graph.

A phase portrait display of this data is shown in Figure 5.20. The phase portrait looks more like a limit cycle than a Lotka–Volterra model, as there is some decrease in amplitude from the initial state, and Gause recognized that the Lotka–Volterra model was not a perfect fit. However, if we try to fit the data to a Lotka–Volterra model (5.5), we may use the estimates $\lambda = 0.65, \mu = 0.32$ obtained by Gause. In order to estimate the other two parameters b and c , we may use the fact that in the Lotka–Volterra model the center of the closed orbits is the equilibrium point $(\mu/c, \lambda/b)$ and then we may use the values $b = 0.0108$ and $c = 0.0058$ obtained by estimating that the equilibrium point is $(55, 60)$. Perhaps a model of Rosenzweig–MacArthur type would give a better fit, but to fit data to such a model is a much more difficult matter.

Some of Gause's experiments led to neither coexistence in equilibrium nor coexistence with oscillations of the predator and prey species. Instead, the predators consumed all the prey and then died from lack of food. This problem also arose in more elaborate experiments conducted by Huffaker (1958); in some cases the sys-

Time	<i>S. exiguus</i>	<i>P. aurelia</i>
0	155	90
1	40	175
2	20	120
3	10	60
4	25	10
5	55	20
6	120	15
7	110	55
8	50	130
9	20	70
10	15	30
11	20	15
13	70	20
15	135	30
16	135	80
17	50	170
18	15	90
19	20	30

Table 5.3 Predator–prey population sizes.

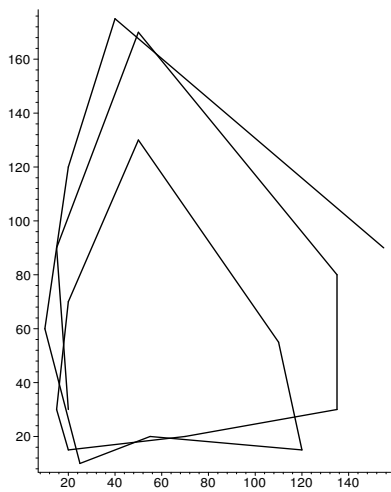


Fig. 5.20 *S. exiguus* and *P. aurelia* phase portrait.

tem went through a few cycles and then collapsed. In experiments on predator–prey systems, it may be necessary to perturb the system by providing a refuge for prey or allowing some immigration to make coexistence possible.

Exercises

1. Use a computer algebra system to plot solutions of the system (5.4) with the parameters derived here and initial data $x(0) = 2$, $y(0) = 2$.
2. Use a computer algebra system to plot solutions of the system (5.5) with the parameters derived here and an initial point on the periodic orbit indicated by Figure 5.20.

5.4 Kolmogorov Models

Two-species models whose per capita growth rates are functions of the population size are described generally by models of the form

$$\begin{aligned}x' &= xf(x, y), \\ y' &= yg(x, y),\end{aligned}\tag{5.11}$$

sometimes called *Kolmogorov* models. We model predator–prey behavior by the following assumptions:

- (i) $f_y(x, y) < 0$, $g_x(x, y) > 0$, $g_y(x, y) \leq 0$.
- (ii) For some (prey carrying capacity) $K > 0$, $f(K, 0) = 0$ and $f(x, y) < 0$ if $x > K$; for some (minimum prey population to support predators) $J > 0$, $g(J, 0) = 0$.

Because $g_x(x, y) > 0$, $g_y(x, y) \leq 0$, the predator isocline $g(x, y) = 0$ slopes upward to the right from $(J, 0)$ (or is vertical if $g_y(x, y) = 0$), as in the Rosenzweig–MacArthur model). The interpretation of a condition $g_y(x, y) \equiv 0$ is that the predators do not interfere with one another in searching for prey. The Rosenzweig–MacArthur model is a special case of the Kolmogorov model with $f(x, y) = f(x) - y\phi(x)$, $g(x, y) = cx\phi(x) - e$. Then $f_y(x, y) = -\phi(x) < 0$, $g_x(x, y) = [x\phi(x)]' > 0$, K is determined by $f(K) = 0$, and J by $cJ\phi(J) = e$. An example of a Kolmogorov model that is not a Rosenzweig–MacArthur model is the Leslie model (1948) with a predator “carrying capacity” proportional to prey population size, that is,

$$g(x, y) = b\left(1 - \frac{y}{ax}\right).$$

Theorem 5.1. *Under the hypotheses (i), (ii) above, every solution of a Kolmogorov model with $x(0) > 0$, $y(0) > 0$ remains bounded for $0 \leq t < \infty$.*

Proof. The proof of this theorem is quite difficult and may be omitted without seriously affecting comprehension. Since $x'(t) < 0$ whenever $x > K$, it is impossible to have $x(t) \rightarrow \infty$. A solution would become unbounded only if $y(t) \rightarrow \infty$ in the region where $x' < 0$ or $f(x, y) < 0$, and $y' > 0$ or $g(x, y) > 0$. Pick a constant $\alpha > 0$ and choose any $x_0 > K$, y_0 above the maximum height of the curve $f(x, y) = -\alpha$. Let $(x(t), y(t))$ be the solution with $x(0) = x_0$, $y(0) = y_0$. We will

show that this solution crosses $g(x, y) = 0$ and enters the region where $x' < 0$, $y' < 0$, and thus is not unbounded. Any solution with $x(0) < x_0, y(0) < y_0$ must also remain bounded because it cannot cross the orbit of $(x(t), y(t))$. So long as this orbit remains in the region in which $x' < 0, y' > 0$, $(x(t), y(t))$ remains above the curve $f(x, y) = -\alpha$, and therefore we have $f(x(t), y(t)) \leq -\alpha$. Also, $(d/dt)g(x(t), y(t)) = g_x(x(t), y(t))x'(t) + g_y(x(t), y(t))y'(t) \leq 0$, since $g_x(x, y) > 0, x' < 0, g_y(x, y) \leq 0, y' > 0$. Thus, $g(x(t), y(t)) \leq g(x_0, y_0)$ for $t \geq 0$. We now let $\beta = g(x_0, y_0)/\alpha > 0$ and define $V(x, y) = x^\beta y$. Then

$$\begin{aligned} \frac{d}{dt}V(x(t), y(t)) &= \beta x^{\beta-1} y x'(t) + x^\beta y'(t) \\ &= \beta x^{\beta-1} y x f(x, y) + x^\beta y g(x, y) \\ &= x^\beta y (f(x, y) + g(x, y)) \\ &\leq x^\beta y (-\beta \alpha + g(x_0, y_0)) \leq 0. \end{aligned}$$

Thus $V(x(t), y(t))$ is a decreasing function, and $V(x(t), y(t)) \leq V(x_0, y_0)$ for $t \geq 0$. We now have

$$y(t) \leq \frac{x_0^\beta y_0}{[x(t)]^\beta} \leq \left(\frac{x_0}{J}\right)^\beta y_0$$

in the region $x \geq J$, which contains the region where $x' < 0, y' > 0$. Thus the solution $(x(t), y(t))$ cannot have $y(t) \rightarrow \infty$ but must proceed to the region where $x' < 0, y' < 0$ and remain bounded.

An immediate consequence of Theorem 5.1, together with the Poincaré–Bendixson theorem, is the theorem of Kolmogorov (1936): Every orbit of a predator–prey system of Kolmogorov type tends either to a stable equilibrium or to a stable limit cycle as $t \rightarrow \infty$. We may now return to the problem of explaining the periodic oscillations that originally led to the Lotka–Volterra model (Section 4.1) and that suggest that a predator–prey model with nonlinear per capita growth rates is a more suitable description.

Models of Kolmogorov type are the most general models for situations in which the per capita growth rate of each species depends only on the population sizes of both species. The effect of the factors x and y in the equations for x' and y' , respectively, is to ensure that neither species generates spontaneously; an orbit for which $x = 0$ at some time must remain on the y -axis of the phase plane for all later times. We are also assured that every orbit starting in the first quadrant ($x > 0, y > 0$) of the phase plane remains in the first quadrant.

5.5 Mutualism

There are situations in which the interaction of two species is mutually beneficial, for example, plant–pollinator systems. The interaction may be *facultative*, mean-

ing that the two species could survive separately, or *obligatory*, meaning that each species will become extinct without the assistance of the other.

If we model a mutualistic system by a pair of differential equations with linear per capita growth rates

$$\begin{aligned}x' &= x(\lambda - ax + by), \\y' &= y(\mu + cx - dy),\end{aligned}$$

the mutualism of the interaction is modeled by the positive nature of the interaction terms cx and by . In a facultative interaction, the constants λ and μ are positive, while in an obligatory relation the constants λ and μ are negative. In each type of interaction there are two possibilities, depending on the relationship between the slope a/b of the x isocline and the slope c/d of the y -isocline (Figure 5.21). If $ad > bc$, the mutualistic effects are smaller than the self-limiting terms in the per capita growth rates, and the slope of the x -isocline is greater than the slope of the y -isocline.

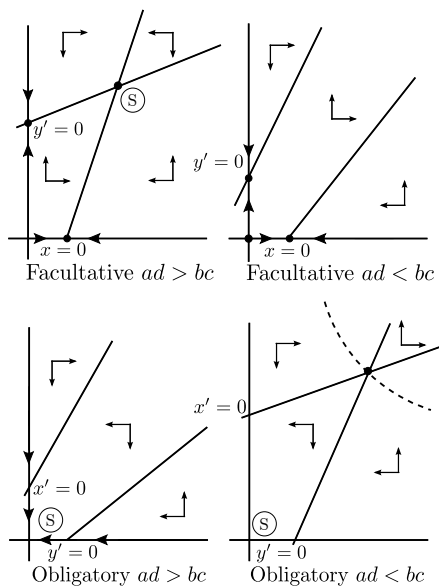


Fig. 5.21 Types of mutualism.

In both facultative and obligatory interactions, if $ad < bc$ there is a region of the phase plane in which solutions become unbounded, and this suggests that either we must restrict models of this form by requiring $ad > bc$, or we must consider models with nonlinear per capita growth rates. For models with linear per capita growth rates and $ad > bc$, it is easy to verify that in the facultative case the only asymptotically stable equilibrium is the intersection (x_∞, y_∞) of the lines $ax - by = \lambda$, $-cx + dy = \mu$ with $x_\infty > 0, y_\infty > 0$, and every orbit tends to this equilibrium. In the

obligatory case, the only asymptotically stable equilibrium is the origin, and every orbit tends to the origin. Thus, in the obligatory case neither species survives. While our model may be acceptable in the facultative case, it is clear that the possibility of obligatory mutualism is not described by this model. If we consider the obligatory case with $ad < bc$, there is an equilibrium (x_∞, y_∞) with $x_\infty > 0$, $y_\infty > 0$, which may be shown to be a saddle point whose stable separatrices separate the phase plane into a region of mutual extinction and a region of unbounded growth. Such a separation is plausible biologically, but we must alter the model so as to rule out the possibility of unbounded growth in order to give a more realistic model.

We shall describe a model of Kolmogorov type to describe mutualistic systems, either facultative or obligatory, and show how general hypotheses that are biologically reasonable lead to qualitative predictions about the behavior of solutions. We assume a model of the form

$$\begin{aligned}x' &= xf(x, y), \\y' &= yg(x, y).\end{aligned}$$

To describe the mutualistic effect of the y species on the x species, we assume

$$f_x(x, y) < 0, \quad f_y(x, y) \geq 0,$$

for $x \geq 0, y \geq 0$. The case $f_y(x, y) \equiv 0$ is known as *commensalism*. Under commensalism the population size of the x species is independent of the population size of the y species. Our treatment will include the possibility of commensalism, but obviously we will not assume that the model is commensal for both species.

Because $f_x(x, y) < 0$, the x -isocline $f(x, y) = 0$ may be written in the form $x = \phi(y) \geq 0$ with $\phi'(y) \geq 0$ on some interval $\alpha \leq y < \infty$. If $\alpha = 0$ and $\phi(0) = K > 0$ so that $f(K, 0) = 0$, the x species is facultative and K is the carrying capacity. If $\alpha > 0$ or $\alpha = 0$ but $\phi(0) = 0$, the x species is obligatory. In this case the x species would die out in the absence of the y species.

We assume also that

$$\phi(\infty) = K^* < \infty,$$

so that the mutualistic effect of the y population cannot be so strong as to allow the x species population size to become unbounded. Since ϕ is an increasing function, $K^* \geq K$; the number K^* may be viewed as an increased carrying capacity for the x species produced by the mutualistic effect of the y species. Specifically, we assume

$$g_x(x, y) \geq 0, \quad g_y(x, y) < 0,$$

for $x \geq 0, y \geq 0$, so that the y isocline $g(x, y) = 0$ may be written in the form $y = \psi(x) \geq 0$ with $\psi'(x) \geq 0$ for $\beta \leq x < \infty$ ($\beta \geq 0$). The y species is facultative if $\beta = 0$ and $\psi(0) = M > 0$, so that $g(0, M) = 0$. If $\beta > 0$ or $\beta = 0$ but $\psi(0) = 0$, the y species is obligatory. We also assume

$$\psi(\infty) = M^* < \infty.$$

At an interior equilibrium (x_∞, y_∞) , the community matrix is

$$\begin{bmatrix} x_\infty f_x(x_\infty, y_\infty) & x_\infty f_y(x_\infty, y_\infty) \\ y_\infty g_x(x_\infty, y_\infty) & y_\infty g_y(x_\infty, y_\infty) \end{bmatrix},$$

with trace

$$x_\infty f_x(x_\infty, y_\infty) + y_\infty g_y(x_\infty, y_\infty) < 0$$

and determinant

$$x_\infty y_\infty (f_x(x_\infty, y_\infty) g_y(x_\infty, y_\infty) - f_y(x_\infty, y_\infty) g_x(x_\infty, y_\infty)),$$

which is positive if and only if the crossing of the isoclines at (x_∞, y_∞) is such that the x -isocline is above the y -isocline to the right of x_∞ . Also the quantity Δ satisfies

$$\Delta = (x_\infty f_x(x_\infty, y_\infty) - y_\infty g_y(x_\infty, y_\infty))^2 + 4x_\infty y_\infty f_y(x_\infty, y_\infty) g_x(x_\infty, y_\infty) \geq 0.$$

Thus, the equilibrium (x_∞, y_∞) is either a saddle point or an asymptotically stable node, depending on the crossing of the isoclines. If we assume that the isoclines are not tangent at any equilibrium, then interior equilibria must alternate between saddle points and asymptotically stable nodes. The stable separatrices at a saddle point separate the domains of attractions of the adjacent nodes.

Because the x -isocline approaches the vertical line $x = K^*$ asymptotically and the y -isocline approaches the horizontal line $y = M^*$ asymptotically, there must be a “last” equilibrium—the one with the largest values of x and y ,—and this equilibrium must be an asymptotically stable node. Also, if the y -isocline is above the x -isocline for small x , there must be at least one interior equilibrium.

We can now describe the qualitative behavior of mutualistic systems of Kolmogorov type for both facultative and obligatory interactions. In the accompanying figures we will use **(SP)** to describe a saddle point, **(U)** to designate an unstable node, and **(AS)** to designate an asymptotically stable node.

Case I. Both species facultative. In this case $f(K, 0) = 0$, $g(0, M) = 0$, and $f(0, 0) > 0$. Thus, the origin is an unstable node and $(K, 0)$, $(0, M)$ are saddle points. Since the y -isocline is above the x -isocline for small x , the first interior equilibrium is an asymptotically stable node, and since the last equilibrium is an asymptotically stable node, there must be an odd number (at least one) of interior equilibria. Every orbit starting in the interior of the first quadrant tends to an interior node (Figure 5.22).

Case II. Both species obligatory. In this case the x -isocline meets the y -axis at $(0, \alpha)$ and the y -isocline meets the x -axis at $(\beta, 0)$. Since $f(0, 0) < 0$, $g(0, 0) < 0$, the origin is an asymptotically stable node. Since the x -isocline is above the y -isocline for small x , the first interior equilibrium is a saddle point. Thus there is an even number, possibly zero, of interior equilibria. If there are two interior equilibria, the first one is a saddle point whose stable separatrices separate the region of initial states for which orbits tend to the origin (extinction of both species) from the region of initial states for which orbits tend to the asymptotically stable interior node (coexistence) (Figure

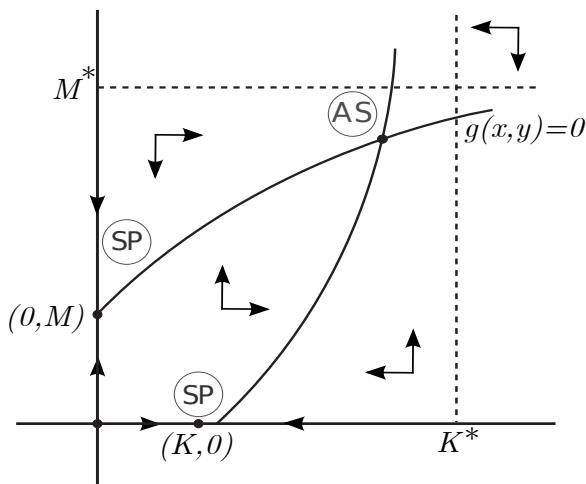


Fig. 5.22 Both species facultative.

5.23). If there is no interior equilibrium, the origin is an asymptotically stable node and all orbits tend to the origin (extinction of both species) (Figure 5.24).

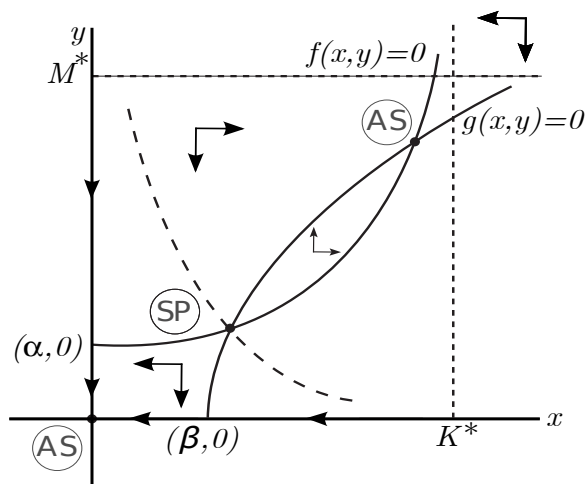


Fig. 5.23 Coexistence.

Which of the two situations occurs depends on the parameters of the model. Changes in external conditions may alter the parameters of the system and cause a transition from one state to another. Thus coexistence may be destroyed, either by such a transition or by a perturbation that moves an orbit across the separatrix into the extinction region.

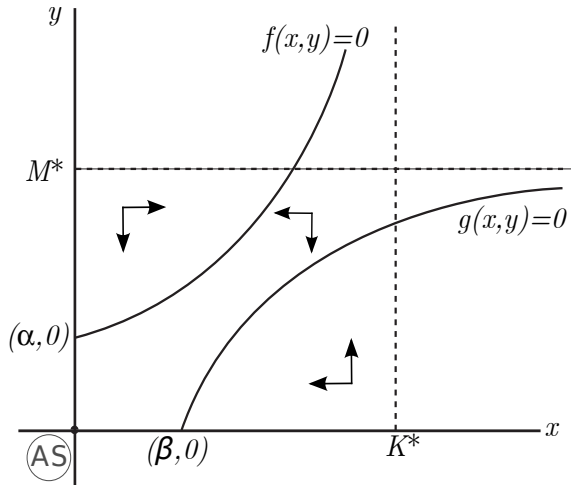


Fig. 5.24 Extinction.

Case III. Facultative–obligatory mutualism. We take the x -species to be obligatory, so that the x -isocline meets the y -axis at $(0, \alpha)$, and the y -species to be facultative, so that the y -isocline meets the y -axis at $(0, M)$. There are two distinct subcases, depending on the relative sizes of α and M , but in each of these subcases, $f(0, 0) < 0$, $g(0, 0) > 0$, and the origin is a saddle point.

If $\alpha > M$, then $f(0, M) < 0$, and $(0, M)$ is an asymptotically stable node. Since the x -isocline is above the y -isocline for small x , the first interior equilibrium is a saddle point. As in the obligatory–obligatory case (Case II), there must be an even number (possibly zero) of interior equilibria. If there are two interior equilibria, the first is a saddle point whose stable separatrices separate the region of initial states for which orbits tend to the node $(0, M)$ (extinction of the x species) from the region of initial states for which orbits tend to the interior stable node (coexistence) (Figure 5.25). If there are no interior equilibria all orbits tend to the node $(0, M)$ (Figure 5.26).

If $\alpha < M$, then $f(0, M) > 0$, and $(0, M)$ is a saddle point. Since the y -isocline is above the x -isocline for small x , the first interior equilibrium is an asymptotically stable node. As in the facultative–facultative case (Case I) there must be an odd number of interior equilibria, and every orbit starting in the interior of the first quadrant of the phase plane must tend to an asymptotically stable interior node (Figure 5.27).

We now have a general description of what a two-species mutualistic model should be and how its solutions should behave. There have been relatively few attempts to suggest explicit models for mutualism, possibly because experimental data to fit a model are difficult to obtain. While in the model we have proposed both isoclines be curved, all the results that we have deduced on behavior of solutions would remain true if one of the isoclines were a straight line.

Some forms that have been used include:

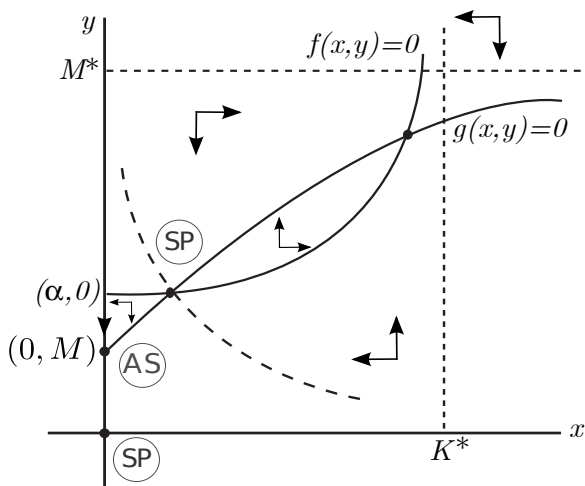


Fig. 5.25 Coexistence.

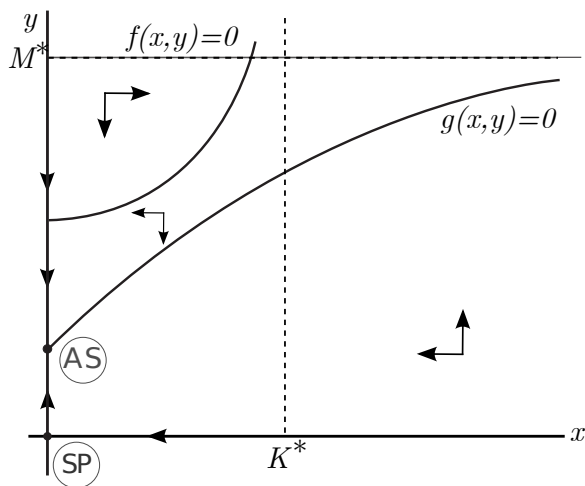


Fig. 5.26 Stable node.

$$\begin{aligned}
 f(x, y) &= \frac{a_1 y}{1 + b_1 x} - c, \\
 g(x, y) &= d_2(K - y) + \frac{a_2 x}{1 + b_2 x} \quad [\text{Soberon and DelRio (1981)}], \\
 f(x, y) &= r \left(1 - \frac{x}{K(1 - e^{x(1-y/M)})} \right), \\
 g(x, y) &= s \left(1 - \frac{y}{L(1 - e^{y(1-x/N)})} \right) \quad [\text{Dean (1983)}].
 \end{aligned}$$

If we think of a logistic model with parameters r and K , we could use

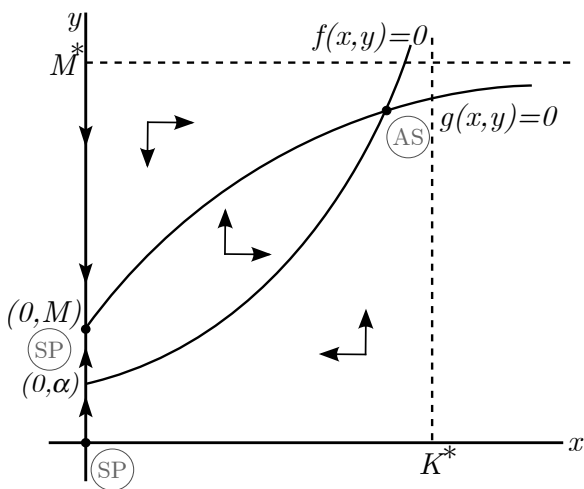


Fig. 5.27 Stable node.

$$f(x, y) = r \left(1 - \frac{x(y+A)}{KA + K^*y} \right)$$

to model enhancement of K , or

$$f(x, y) = \left(\frac{rA + r^*y}{y+A} \right) \left(1 - \frac{x}{K} \right)$$

to model enhancement of r , or

$$f(x, y) = \left(\frac{rA + r^*y}{y+A} \right) \left(1 - \frac{x(y+B)}{KB + K^*y} \right)$$

to model simultaneous enhancement of r and K through mutualism, with similar forms for $g(x, y)$.

Exercises

In Exercises 1 through 4, find all equilibria of the given mutualistic system and determine their stability. Are there any unbounded orbits?

1. $x' = x(-20 - x + 2y)$, $y' = y(-50 + x - y)$.
2. $x' = x(-20 - x + 2y)$, $y' = y(-50 + x - y)$.
3. $x' = x(y/(1+x) - 1)$, $y' = y(10 - y + x/(1+y))$.
4. $x' = x(1 - x/10) \cdot ((2y+1)/(1+y))$, $y' = y(1 - y/20)$.
- 5.* Determine the behavior as $t \rightarrow \infty$ of solutions of the system

$$x' = rx \left(1 - \frac{x}{K + ay} \right), \quad y' = sy \left(1 - \frac{y}{M + bx} \right)$$

with a , b , K , M , r , and s positive constants. [see Exercise 12, Section 4.5].
[Warning: The behavior if $ab < 1$ is different from the behavior if $ab > 1$].

5.6 The Spruce Budworm: A Case Study

The spruce budworm is an insect that inhabits spruce and fir trees in eastern North America. In most years it is relatively scarce, but it exhibits outbreaks in which the population grows by a factor of as much as 1000. In an outbreak year, budworms may devour the new needles produced by an evergreen forest, ultimately killing 80 percent of the mature trees in the forest. Such outbreaks may destroy forests. When the food supply is destroyed (this may take 7 to 14 years from the start of the outbreak) the budworm population declines rapidly to a low level, and the forest may then begin to recover. We have described a simple model for the budworm population in Exercise 13, Section 1.4 and Project 1.1. Here we will recall this model and refine it to incorporate the forest as well. The dynamics of the spruce budworm proceed on a much faster time scale than the dynamics of the forest, and our model will incorporate two time scales. The process is to begin with the fast (budworm) variable, assuming that the slow (forest) variables remain constant. The next step is to formulate a slow-variable model and analyze it in order to decide how to refine the fast variable model to incorporate dependence of its parameters on the slow variables. This will allow us to simulate outbreaks caused by a change in the behavior of the fast -variable system caused by parameter changes. Our description is based on the paper [Ludwig, Jones, and Holling (1978)].

We let B denote the spruce budworm population size and assume the fast variable model

$$B' = r_B \left(1 - \frac{B}{K_B} \right) - \beta \frac{B^2}{\alpha^2 + \beta^2}. \quad (5.12)$$

This is the model of Exercise 13, Section 1.4, with different notation. Here r_B is the intrinsic growth rate and K_B is the carrying capacity in a logistic growth model. In addition, there is predation by birds and parasites that saturates for high budworm populations at a level β , and α is the budworm population at which predation is half the maximum. The model involves the four parameters r_B , K_B , β , α . Later, when we incorporate the slow variables, we will allow K_B and α to depend on these slow variables.

It is convenient to scale equation (5.12) to reduce the number of parameters. We use α as the unit of budworm population size and introduce $u = B/\alpha$ as a new dependent variable. This transforms (5.12) to the equation

$$\alpha u' = r_B u \alpha \left(1 - \frac{u}{K_B} \right) - \beta \frac{u^2}{1 + u^2},$$

or

$$\frac{\alpha}{\beta} u' = r_B \frac{\alpha}{\beta} u \left(a - \frac{\alpha u}{K_B} \right) - \frac{u^2}{1+u^2}.$$

Now we define the new parameters

$$R = \frac{\alpha}{\beta}, \quad Q = \frac{K_B}{\alpha},$$

and rewrite the differential equation as

$$\frac{\alpha}{\beta} u' = Ru \left(1 - \frac{u}{Q} \right) - \frac{u^2}{1+u^2}. \quad (5.13)$$

Equilibria of (5.13) are $u = 0$ and the intersections of the line $y = R(1 - u/Q)$ and the curve $y = u/(1+u^2)$. It will turn out that during an outbreak Q remains nearly constant but R may vary considerably. Thus, in order to understand the dynamics we will need to investigate how the equilibria depend on R . The curve $y = u/(1+u^2)$ starts at the origin, increases to a maximum at $(1, 1/2)$, and then decreases with an inflection point at $(\sqrt{3}, \sqrt{3}/4)$, and approaches zero as $u \rightarrow \infty$. The line $y = R(1 - u/Q)$ passes through the points $(0, R)$ and $(Q, 0)$ in the (u, y) -plane. A “typical” situation is as shown in [Figure 5.28](#).

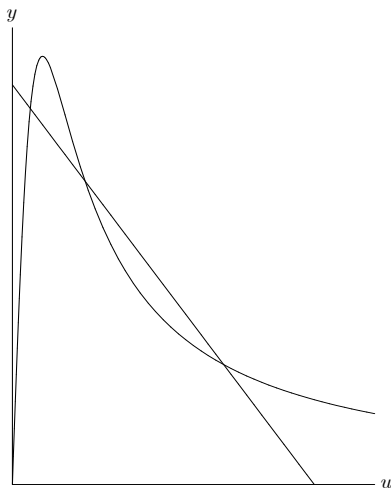


Fig. 5.28 Equilibria of fast-variable model.

There are four equilibria: an unstable equilibrium at $u = 0$, two asymptotically stable equilibria u_- and u_+ , separated by an unstable equilibrium $u = u_c$. The domain of attraction of the small equilibrium u_- is the interval $(0, u_c)$, and the domain of attraction of the small equilibrium u_+ is the interval (u_c, ∞) . If R increases until u_- and u_c coalesce, as in [Figure 5.29](#), then only the equilibrium u_+ remains. Thus,

if a system is in equilibrium at u_- and R is increased, there could be a jump to u_+ . Conversely, if R is decreased until u_+ and u_c coalesce, as in Figure 5.30, then only the equilibrium u_- remains. Thus, if a system is in equilibrium at u_+ and R is decreased, there could be a crash to u_- .

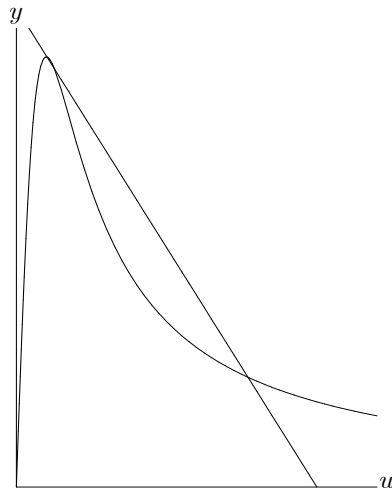


Fig. 5.29 Coalescence of equilibria.

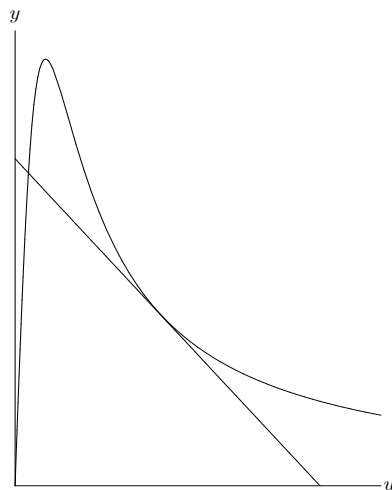


Fig. 5.30 Coalescence of equilibria.

We think of the parameter R as representing the resources of the forest supply as food for budworms, and the equilibria u_- and u_+ as corresponding to budworm

limitation by predators and food supply, respectively. As forest conditions improve, budworm growth exceeds control by predators and there is an outbreak. If this outbreak destroys the forest, the predators may regain control and cause a budworm crash. The outbreak and crash will generally occur at different population levels.

In order to build in the possibility of forest destruction, we need to form a model to describe the state of the forest assuming a fixed budworm population. We will use as forest variables the *total surface area of branches* S and a variable E , which may be viewed as an *energy reserve*, to describe the foliage and health of the trees. We choose a logistic form for the equation governing S ,

$$S' = r_S S \left(1 - \frac{SK_E}{K_S E} \right). \quad (5.14)$$

The factor K_E/E is included because S may decrease under stress through the death of branches or whole trees. Normally, however, E will be close to its carrying capacity K_E and S will approach its carrying capacity K_S . We assume that the energy reserve E also satisfies an equation of logistic type

$$E' = r_E E \left(1 - \frac{E}{K_E} \right) - p \frac{B}{S}. \quad (5.15)$$

The term pB/S describes the stress on the trees excited by budworm consumption of foliage. Since B has units of number of budworms per acre and S has units of branch surface area per acre, B/S is the number of budworms per unit of branch surface area.

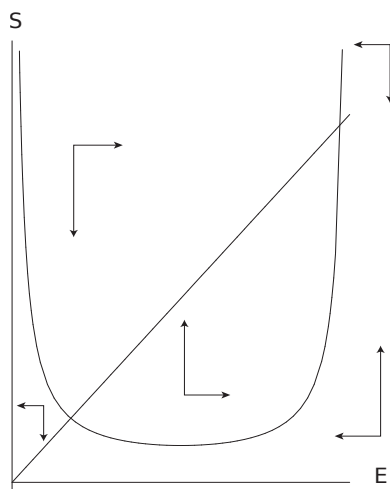
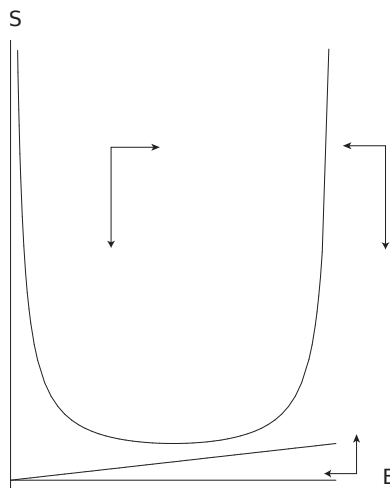
The next step is to determine the qualitative behavior of the slow-variable system (5.14), (5.15), treating B as constant, by analyzing its equilibria. The equilibrium conditions are $S = 0$ or $S = \frac{K_S}{k_E} E$ together with

$$S = \frac{pBK_E}{r_E} \frac{1}{E(K_E - E)}. \quad (5.16)$$

The curve (5.16) has vertical asymptotes at $E = 0$ and $E = K_E$ and has a minimum when $E = K_E/2$. Equilibria are intersections of this curve and the line $S = EK_S/K_E$ through the origin with positive slope. If B is small, there are two equilibria (Figure 5.31), while if B is large, as during a budworm outbreak, the curve (5.16) is elevated and there are no equilibria (Figure 5.32).

We may see from the flow arrows that in Figure 5.31, the equilibrium on the left is a saddle point and the equilibrium on the right is asymptotically stable. The stable separatrices at the saddle point divide the plane into the domain of attraction of the stable equilibrium (to the right of the separatrices) and the region for which $E \rightarrow -\infty$ (to the left of the separatrices). (In our model, $E' < 0$ when $E = 0$, and this is unrealistic; later we shall describe how to modify the model to make $E' = 0$ when $E = 0$ and allow the forest to recycle.)

When the budworm population is small enough that Figure 5.31 describes the slow-time situation, S and E increase slowly toward the stable equilibrium. As B

**Fig. 5.31** Two equilibria.**Fig. 5.32** No equilibrium.

increases, the curve (5.16) rises, causing the equilibrium values of S and E to decrease somewhat, indicating a deterioration of the health of the forest. When there is a catastrophic outbreak of budworms, the equilibria coalesce and disappear, shifting the situation to that of [Figure 5.32](#). Then the forest collapses. There is another possibility: The budworm population level, though high, may be held down enough to avoid collapse of the forest. This is called “perpetual outbreak” and has been observed as a consequence of the spraying done to try to control the budworm.

We now have two separate models for budworm and forest. In order to combine them into one model, we need to reconsider the two models by coupling the param-

eters appropriately. We would expect r_B , the intrinsic growth rate of the budworm population, to be independent of the forest parameters, but the carrying capacity should be proportional to the foliage surface area S . Thus, we replace K_B by a new parameter K with

$$K_B = KS.$$

The maximum predation rate β is not affected by the forest state, but the budworm population level α at which predation is half its maximum rate will also increase proportionally to S ; as S increases predators will have to search for more foliage, and a higher budworm population can be supported at the same predation rate. Thus, we replace α by a new parameter α' with $\alpha = \alpha'S$. We now have a coupled model

$$\begin{aligned} B' &= r_B B \left(1 - \frac{B}{KS}\right) - \beta \frac{B^2}{(\alpha'S)^2 + B^2}, \\ S' &= r_S S \left(1 - \frac{S}{K_S} \frac{K_E}{E}\right), \\ E' &= r_E E \left(1 - \frac{E}{K_E}\right) - p \frac{B}{S}. \end{aligned} \tag{5.17}$$

The parameters R and Q introduced in the analysis of (5.12) now become

$$R = \frac{\alpha' r_B}{\beta} S, \quad Q = \frac{K}{\alpha}.$$

Q remains constant but R increases proportionally to S as S increases. As a young forest matures, S increases, and the consequent increase in R in the analysis of the fast variable will lead to a budworm outbreak. Whether there will be perpetual outbreak or forest collapse depends on parameter values.

Appropriate values for the parameters of the model (5.17) have been estimated [Ludwig, Jones, & Holling (1978)], first on the basis of general knowledge of the biology of the situation and then in more refinement from an extensive field study of the forests of New Brunswick. These studies led to the following choices for parameter values:

$$\begin{aligned} r_B &= 1.52, & K &= 335, & \beta &= 43,200, \\ \alpha' &= 1.11, & r_S &= 0.095, & K_S &= 25,440, \\ K_E &= 1.0, & r_E &= 0.92, & p &= 0.00195. \end{aligned}$$

These led to the values $R = 3.91 \times 10^{-5} S$, $Q = 302$. We could then use these parameter values in (5.17) to simulate the behavior of the spruce budworm-forest system, and we would find outbreak followed by collapse of the forest. This conforms to observed behavior in part, but does not exhibit the cycles of budworm outbreak, forest collapse, and budworm collapse, followed by forest regeneration.

This shortcoming of our model is reflected in the fact that E , a variable meant to describe a general status and considered to vary from zero to one, may become negative. We may improve the model in an ecological sense by observing that budworm carrying capacity should depend on E as well as S and by incorporating dependence on E in the stress due to defoliation by budworms. We replace $K_B = KS$ by

$$K_B = KS \frac{E^2}{E^2 + T_E^2},$$

where T_E is a constant threshold value. The quantity $E^2/(E^2 + T_E^2)$ is near one in a healthy forest but decreases sharply as $E \rightarrow 0$. We also replace the stress term pB/S by

$$p \frac{B}{S} \frac{E^2}{T_E^2 + E^2}.$$

This means that the stress term will decrease rapidly as the forest collapses. The refined model would be

$$\begin{aligned} B' &= r_B B \left(1 - \frac{B}{KS} \frac{E^2 + T_E^2}{E^2} \right) - \beta \frac{B^2}{(\alpha' S)^2 + B^2}, \\ S' &= r_S S \left(1 - \frac{S}{K_S} \frac{K_E}{E} \right), \\ E' &= r_E E \left(1 - \frac{E}{K_E} \right) - P \frac{B}{S} \frac{E^2}{E^2 + T_E^2}. \end{aligned} \tag{5.18}$$

With suitably chosen T_E , this model will behave like the model (5.17) except that when B , S , and E are small it will allow regeneration of the forest and a new cycle.

The models (5.17) and (5.18) give a qualitative picture similar to what has been observed. For short term predictions and actual forest management it may be preferable to use a more detailed model that does not attempt to describe a whole forest by two variables. Such a model would not be analytically tractable but could be used for numerical simulations. However, there are, necessarily, approximations in the model and errors in the parameters, which could give accumulating errors in a simulation. Thus, such a model is best suited for short-term predictions. Our less-detailed model is better suited for explaining *why things happen*.

References: Arreola, Mijares-Bernal, Ortiz-Navarro, and Saenz (2000) have looked at the effects of insecticides on a two-dimensional version of the above model that supports oscillations. They show that insecticides can increase the frequency of these oscillations, a situation that in the long run may not be good for the trees.

5.7 The Community Matrix

In this chapter, we are concerned primarily with models for two interacting species. However, in most real-life situations there are more than two species involved. For example, the spruce budworm model considered in Section 5.5 involves three species. In the remainder of this chapter, we shall give a very brief sketch of how one may use two-species models (competition, predator–prey, and mutualism) as building blocks to study some larger systems. In general, the study of multispecies models is very complicated, and we shall confine ourselves to some remarks about their general structure and two three-species examples.

If x_1, x_2, \dots, x_n represent the population sizes of n different interacting species in a system for which it is assumed that the growth rate of each population size at any time depends only on the various population sizes at that time, then the system is modeled by an autonomous system of n first-order differential equations

$$\begin{aligned} x_1' &= F_1(x_1, x_2, \dots, x_n), \\ x_2' &= F_2(x_1, x_2, \dots, x_n), \\ &\vdots \\ x_n' &= F_n(x_1, x_2, \dots, x_n). \end{aligned} \tag{5.19}$$

We are interested only in nonnegative population sizes. Since it is biologically reasonable to require that no species can generate spontaneously, it is reasonable to require that a solution for which some x_j is zero at any time should continue to have $x_j = 0$ for all time. Thus, it is natural to write the system (5.19) in the Kolmogorov form

$$\begin{aligned} x_1' &= x_1 r_1(x_1, x_2, \dots, x_n), \\ x_2' &= x_2 r_2(x_1, x_2, \dots, x_n), \\ &\vdots \\ x_n' &= x_n r_n(x_1, x_2, \dots, x_n). \end{aligned} \tag{5.20}$$

Exactly as for single-species models (Chapter 1) and two-species models (Chapter 4) we define an *equilibrium* of the system (5.19) to be a solution $(\xi_1, \xi_2, \dots, \xi_n)$ of the system of equations

$$\begin{aligned} F_1(\xi_1, \xi_2, \dots, \xi_n) &= 0, \\ F_2(\xi_1, \xi_2, \dots, \xi_n) &= 0, \\ &\vdots \\ F_n(\xi_1, \xi_2, \dots, \xi_n) &= 0. \end{aligned} \tag{5.21}$$

An equilibrium is a constant solution of the system of differential equations (5.19). The *linearization* of the system (5.19) at an equilibrium $(\xi_1, \xi_2, \dots, \xi_n)$ is defined to be the *linear* system of differential equations

$$\begin{aligned}
u'_1 &= \frac{\partial F_1}{\partial x_1}(\xi_1, \dots, \xi_n)u_1 + \frac{\partial F_1}{\partial x_2}(\xi_1, \dots, \xi_n)u_2 + \dots + \frac{\partial F_1}{\partial x_n}(\xi_1, \dots, \xi_n)u_n, \\
u'_2 &= \frac{\partial F_2}{\partial x_1}(\xi_1, \dots, \xi_n)u_1 + \frac{\partial F_2}{\partial x_2}(\xi_1, \dots, \xi_n)u_2 + \dots + \frac{\partial F_2}{\partial x_n}(\xi_1, \dots, \xi_n)u_n, \\
&\vdots \\
u'_n &= \frac{\partial F_n}{\partial x_1}(\xi_1, \dots, \xi_n)u_1 + \frac{\partial F_n}{\partial x_2}(\xi_1, \dots, \xi_n)u_2 + \dots + \frac{\partial F_n}{\partial x_n}(\xi_1, \dots, \xi_n)u_n.
\end{aligned}$$

It is convenient to use vector–matrix notation to define \mathbf{x} to be the *column vector* with components (x_1, \dots, x_n) , ξ to be the column vector with components (ξ_1, \dots, ξ_n) , \mathbf{u} to be the column vector with components (u_1, \dots, u_n) , $\mathbf{F}(\mathbf{x})$ to be the column vector function with components $(F_1(x_1, \dots, x_n), \dots, F_n(x_1, \dots, x_n))$, and \mathbf{A} to be the matrix with element

$$\frac{\partial F_i}{\partial x_j}(\xi_1, \dots, \xi_n)$$

in the i th row, j th column. Then the system (5.19) can be written in the vector form

$$\mathbf{x}' = \mathbf{F}(\mathbf{x}). \quad (5.22)$$

An equilibrium is a vector ξ satisfying

$$\mathbf{F}(\xi) = \mathbf{0}, \quad (5.23)$$

and the linearization of the system (5.19) or (5.22) at an equilibrium is the linear system

$$\mathbf{u}' = \mathbf{A}\mathbf{u}. \quad (5.24)$$

The matrix

$$\mathbf{A} = \left(\frac{\partial F_i}{\partial x_j}(\xi_1, \dots, \xi_n) \right)$$

is called the *community matrix* of the system (5.19) or (5.22) at the equilibrium ξ .

There is a general theorem, analogous to Theorem 4.1, Section 4.3, that says that an equilibrium of the system (5.22) is asymptotically stable if all solutions of the linearization at this equilibrium tend to zero as $t \rightarrow \infty$, while an equilibrium ξ is unstable if the linearization has any solution that grows exponentially. It is also true that all solutions of the linearization tend to zero if all roots of the *characteristic equation*

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \quad (5.25)$$

have negative real part and that there are solutions of the linearization that grow exponentially if the characteristic equation (5.25) has any roots with positive real part. The roots of the characteristic equation (5.25) are the eigenvalues of the matrix \mathbf{A} . The characteristic equation has the property that its roots are the values of λ such that the linearization at the equilibrium has a solution $e^{\lambda t} \mathbf{c}$ for some constant column vector \mathbf{c} .

Thus, the stability of an equilibrium ξ can be determined from the eigenvalues of the community matrix at the equilibrium.

Theorem 5.2. *If all eigenvalues of the community matrix of a system (5.19) or (5.22) at an equilibrium ξ have negative real part then the equilibrium is asymptotically stable.*

This is the generalization to n dimensions of Theorem 4.4 for two-dimensional systems, but we no longer have such a simple set of conditions on the trace and determinant of the community matrix to determine the stability of an equilibrium.

The characteristic equation for an n -dimensional system is a polynomial equation of degree n for which it may be difficult or impossible to find all roots explicitly. There is, however, a general criterion for determining whether all roots of a polynomial equation have negative real part known as the *Routh–Hurwitz criterion*. This gives conditions on the coefficients of a polynomial equation

$$\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_{n-1}\lambda + a_n = 0$$

under which all roots have negative real part. For $n = 2$, the Routh–Hurwitz conditions are

$$a_1 > 0, \quad a_2 > 0$$

(equivalent to the conditions that the trace of the matrix \mathbf{A} be negative and the determinant of the matrix \mathbf{A} be positive). For $n = 3$, the Routh–Hurwitz conditions are

$$a_3 > 0, \quad a_1 > 0, \quad a_1a_2 > a_3.$$

For $n = 4$, the Routh–Hurwitz conditions are

$$a_4 > 0, \quad a_2 > 0, \quad a_1 > 0, \quad a_3(a_1a_2 - a_3) > a_1^2a_4.$$

For a polynomial of degree n , there are n conditions. While the Routh–Hurwitz criterion may be useful on occasion, it is complicated to apply in problems of many dimensions.

For autonomous systems of two differential equations, the Poincaré–Bendixson theorem (Section 4.5) makes it possible to analyze the qualitative behavior of a system under very general conditions. Essentially, we know that a bounded orbit must approach either an equilibrium point or a limit cycle. For autonomous systems of more than two differential equations, a much greater range of behavior is possible. An example is the *Lorenz equation*

$$\begin{aligned} x' &= \sigma(y - x), \\ y' &= rx - y - xz, \\ z' &= xy - bz, \end{aligned}$$

with σ , r , and b three positive parameters. This equation, which arose originally as a highly simplified model of fluid convection in a meteorological system, can exhibit chaotic behavior for some ranges of the parameters [Lorenz (1963)].

There is some biological evidence to suggest that complicated population systems have a tendency to be more stable than simple systems. For example, a predator species that can switch between different prey species for its food supply may be less sensitive to disturbances than if it were dependent on a single food supply. On the other hand, removal of one species can lead to a collapse of population systems; an example was observed in which the removal of one species from a 15-species system led to its collapse to an 8-species system in less than two years [Paine (1966)]. The relationship between stability and complexity in population systems, a question raised by May (1974), is not well understood. One aspect of this question that can provide some useful information is the vulnerability of a system to invasion by a new species (Section 5.9).

5.8 The Nature of Interactions Between Species

If we model a two-species interaction by a system of Kolmogorov type

$$\begin{aligned}x' &= xf(x, y), \\y' &= yg(x, y),\end{aligned}\tag{5.26}$$

it may appear reasonable to require

$$f_x(x, y) < 0, \quad g_y(x, y) < 0,\tag{5.27}$$

to express the idea that each species tends to limit its own growth. However, the condition (5.27) is not satisfied for all $x > 0, y > 0$ in the standard predator-prey models; we should require only that each species tend to limit its own growth if the other species is not present. The assumption $f_x(x, y) < 0$ is sometimes replaced by the requirement

$$xf_x(x, y) + yf_y(x, y) < 0$$

for $x > 0, y > 0$ [Kolmogorov (1936)], which is satisfied by models of Rosenzweig-MacArthur type. Another approach [Bulmer(1976)] is to require $f_x(x, 0) < 0$ together with a condition modeling the idea that predators increase only through the consumption and conversion of prey, such as

$$yg(x, y) \leq \alpha x(f(x, 0) - f(x, y)) - \mu y$$

for some positive constants α, μ .

In order to separate the growth rates in the system (5.26) into self-limiting and interaction terms, we write (5.26) in the form

$$\begin{aligned}x' &= xf(x, 0) + x(f(x, y) - f(x, 0)), \\y' &= yg(0, y) + y(g(x, y) - g(0, y)).\end{aligned}\tag{5.28}$$

We then require that the respective single-species dynamics be self-limiting,

$$f_x(x, 0) < 0, \quad g_y(0, y) < 0, \quad (5.29)$$

except that we will permit a species that is dependent on the interaction for survival (predator or obligate mutualist) to have, for example, $g_y(0, y) = 0$. This would mean that $g(0, y)$ is a constant, and we would require this constant to be negative so that $y(t) \rightarrow 0$ as $t \rightarrow \infty$ in the absence of the x species. We may then characterize the interaction as predator-prey, competitive, or mutualistic, depending on the signs of

$$\begin{aligned} \frac{\partial}{\partial y}(f(x, y) - f(x, 0)) &= f_y(x, y), \\ \frac{\partial}{\partial x}(g(x, y) - g(0, y)) &= g_x(x, y). \end{aligned}$$

The terms $x(f(x, y) - f(x, 0))$ and $y(g(x, y) - g(0, y))$ represent the interaction terms in (5.28). Thus, the system (5.26) or (5.28) is a predator-prey system if

$$f_y(x, y) < 0, \quad g_x(x, y) > 0,$$

a competitive system if

$$f_y(x, y) < 0, \quad g_x(x, y) < 0,$$

and a mutualistic system if

$$f_y(x, y) > 0, \quad g_x(x, y) > 0.$$

If we attempt to decompose a three-species Kolmogorov model

$$\begin{aligned} x' &= xf(x, y, z), \\ y' &= yg(x, y, z), \\ z' &= zh(x, y, z), \end{aligned} \quad (5.30)$$

into self-limiting terms and two-species interaction terms, we may write the first equation of (5.30) as either

$$x' = xf(x, 0, 0) + x(f(x, y, 0) - f(x, 0, 0)) + x(f(x, y, z) - f(x, y, 0))$$

or

$$x' = xf(x, 0, 0) + x(f(x, y, z) - f(x, 0, z)) + x(f(x, 0, z) - f(x, 0, 0)).$$

In characterizing the x - y interaction we would consider the sign of the partial derivative with respect to y of the effect of y on the per capita growth rate of x , which would be either

$$\frac{\partial}{\partial y}(f(x, y, 0) - f(x, 0, 0)) = f_y(x, y, 0)$$

or

$$\frac{\partial}{\partial y}(f(x, y, z) - f(x, 0, z)) = f_y(x, y, z),$$

and these do not necessarily have the same sign.

Consider, for example, a population system consisting of two competing species x and y and a third species z acting as a predator on both. We may model such a system in more specific form by a system of Rosenzweig–MacArthur type:

$$\begin{aligned}x' &= rxf(x, y) - xz\phi(x, y, z) \\y' &= syg(x, y) - yz\psi(x, y, z) \\z' &= z(\lambda x\phi(x, y, z) - \mu y\psi(x, y, z) - c).\end{aligned}\tag{5.31}$$

This is of the form (5.30) with

$$\begin{aligned}f(x, y, z) &= rxf(x, y) - xz\phi(x, y, z), \\g(x, y, z) &= syg(x, y) - yz\psi(x, y, z), \\h(x, y, z) &= \lambda x\phi(x, y, z) + \mu y\psi(x, y, z) - c,\end{aligned}$$

so that

$$f_y(x, y, 0) = rf_y(x, y), \quad f_y(x, y, z) = rf_y(x, y) - z\phi_y(x, y, z).$$

We model the competition between x and y by the assumptions

$$f_y(x, y) < 0, \quad g_x(x, y) < 0.$$

If $\phi(x, y)$ and $\psi(x, y)$ are increasing functions of the ratios $x/(x+y)$ and $y/(x+y)$, respectively, then z is a predator that switches between x and y and $\phi_y(x, y) < 0$, $\psi_x(x, y) < 0$. This would imply that $f_y(x, y, 0) < 0$ but $f_y(x, y, z) > 0$ for sufficiently large z . The same kind of sign change would occur for the partial derivative with respect to x of the effect of x on the per capita growth rate of y . If we try to describe the x - y interaction in this three-species model in terms of the signs of $f_y(x, y, z)$ and $g_x(x, y, z)$ in analogy to the two-species model, we would be led to say that this interaction is competitive when few predators are present, and mutualistic in the presence of many predators, with an intermediate stage in which the interaction is predator–prey. Such an interpretation suggests that the interpretation of the interaction between two species in a three-species model is more complicated. Presumably, a proper interpretation should be in terms of the signs of $f_y(x, y, 0)$ and $g_x(x, y, 0)$ together with some additional conditions whose form is not clear. For this reason, systems of more than two species are usually modeled by systems of differential equations less general than a Kolmogorov model. Frequently, a system of Lotka–Volterra type is used to give some indication of the type of behavior one might expect, even though the Lotka–Volterra model is a specific form not necessarily representative of general behavior, as we have seen for two-species models.

5.9 Invading Species and Coexistence

In analyzing population systems with more than two species there are many different cases that must be considered. For three species, even if we eliminate mutualism from consideration, there are six different cases: (i) three species in competition, (ii) one predator and two competing prey, (iii) two predators with a common prey, (iv) a prey whose predator is the prey of a higher-level predator, (v) two species in competition and a third species that is the prey of one of them, (vi) two species in competition and a third species that is a predator of one of them. Since general Kolmogorov-type models are not suitable for distinguishing these types, as we have seen in Section 5.7, each case must be analyzed separately.

A simplification that may be helpful is to ask for less information. Rather than trying to describe the qualitative behavior of a system completely, we might try to determine only the survival or extinction of each species. From a biological point of view, this may be the most significant aspect of the problem. A species with population size $x(t)$ is said to be persistent [Freedman and Waltman (1984)] if

$$\liminf_{t \rightarrow \infty} x(t) > 0$$

provided $x(0) > 0$. Persistence, however, does not ensure survival of a species in a biological sense. Examples have been given of three-species systems that oscillate wildly so that each population size repeatedly comes arbitrarily close to zero (at which time a small perturbation could drive the population size to zero) and then recovers [Armstrong and McGehee (1980); May and Leonard (1975)]. A more realistic requirement would be uniform persistence [Butler and Waltman (1986)] or permanent coexistence [Hofbauer (1981)], a positive lower bound for population size. Conditions have been given, usually involving the behavior of lower-dimensional systems, under which persistence implies uniform persistence [Butler, Freedman, and Waltman (1986); Butler and Waltman (1986)]. We shall examine only persistence, avoiding the more delicate problem of uniform persistence.

In order to obtain some insight into the behavior of multispecies systems it may be helpful to consider the effect on an n -species system of introducing another species to give an $(n + 1)$ -species system. This additional species may be viewed biologically as an invading species disrupting an existing population system. Such situations have been observed in nature. For example, in systems consisting of two species in competition that are the prey of a common predator (to be examined in more detail later) the introduction of an alternative prey may lead to the extinction of the original prey [Holt (1977)]. Such a situation may also be viewed as the introduction of a predator into a system of coexisting competing species, leading to the exclusion of one of the prey species [Poole (1974); Dodson (1974); Holt (1977)].

We consider an n -species model

$$\begin{aligned}
x'_1 &= x_1 F_1(x_1, x_2, \dots, x_n, 0), \\
x'_2 &= x_2 F_2(x_1, x_2, \dots, x_n, 0), \\
&\vdots \\
x'_n &= x_n F_n(x_1, x_2, \dots, x_n, 0),
\end{aligned} \tag{5.32}$$

with an equilibrium $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ and having community matrix \mathbf{A} at this equilibrium. We also consider the $(n+1)$ -species model

$$\begin{aligned}
x'_1 &= x_1 F_1(x_1, x_2, \dots, x_n, y), \\
x'_2 &= x_2 F_2(x_1, x_2, \dots, x_n, y), \\
&\vdots \\
x'_n &= x_n F_n(x_1, x_2, \dots, x_n, y), \\
y' &= y G(x_1, x_2, \dots, x_n, y),
\end{aligned} \tag{5.33}$$

representing the introduction of an invading species y into the n -species system. The $(n+1)$ -species system (5.33) has an equilibrium $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0)$, and the community matrix at this equilibrium is

$$\begin{bmatrix}
\bar{x}_1 \frac{\partial F_1}{\partial y}(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0) \\
\bar{x}_2 \frac{\partial F_2}{\partial y}(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0) \\
\vdots \\
\bar{x}_n \frac{\partial F_n}{\partial y}(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0) \\
\mathbf{0} \quad G(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0)
\end{bmatrix}.$$

The eigenvalues of this matrix are the eigenvalues of \mathbf{A} together with $G(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0)$.

We assume that the equilibrium $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0)$ of the n -species model (5.32) is asymptotically stable, and thus that the eigenvalues of \mathbf{A} have negative real part. Then if $G(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0) < 0$ the equilibrium $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0)$ of the $(n+1)$ -species system (5.33) is also asymptotically stable, and orbits starting near this equilibrium, corresponding to the introduction of a small number of members of the invading species, tend to this equilibrium. Thus, the invading species fails to survive.

On the other hand, if $G(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0) > 0$, the equilibrium $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0)$ of (5.33) is unstable, because the community matrix at this equilibrium has one positive eigenvalue. The hyperplane $y = 0$ is the stable manifold, but orbits starting near the equilibrium with $y(0) > 0$ tend away from the stable manifold. This does not ensure survival of the invading species because the orbit could tend to another equilibrium on $y = 0$. However, if the per capita growth rate $G(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0)$ of the invading species is positive at every asymptotically stable equilibrium $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, 0)$ of the n -species system (5.32) then the invading species survives. In this case, we will say that y is a survivor species. A survivor species is persistent, but as we have pointed out, it may not survive in the biological sense unless it is actually uniformly persistent.

In addition to using the idea of survivor species to model the effect of introducing an invading species into an existing population system, we may also try to analyze the behavior of an $(n + 1)$ -species system by considering each species in turn as an invader of the system consisting of the remaining n species. The basic result is that there is coexistence of the $(n + 1)$ -species system for all initial states if and only if every species is a survivor species. Two examples for $n = 1$ may be illuminating.

- (i) Consider two species in competition with an unstable equilibrium, so that either species may win the competition depending on the initial state (Figure 5.33). Neither species is a survivor; the existence of an equilibrium with both species present is not sufficient to ensure coexistence.

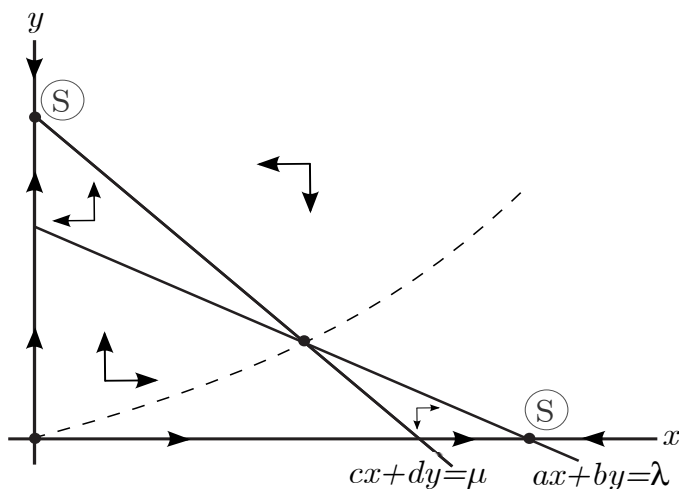


Fig. 5.33 Unstable coexistence equilibrium.

- (ii) Consider a predator–prey system with an Allee effect for the prey. There is coexistence for some but not all initial states, but neither species is a survivor species (the predator species is a survivor at one prey equilibrium but not at the other) (Figure 5.34).

5.10 Example: A Predator and Two Competing Prey

As an example we shall now apply the ideas of Section 5.8 to a three species system consisting of a predator species z and two prey species x and y in competition with each other. Phenomena that have been observed for such systems include predator-induced exclusion of one of the prey species [Poole (1974); Dodson (1974); Holt (1977)], exclusion of prey by alternative prey [Holt (1977)], and predator-induced coexistence [Oaten and Murdoch (1975); Paine (1966)].

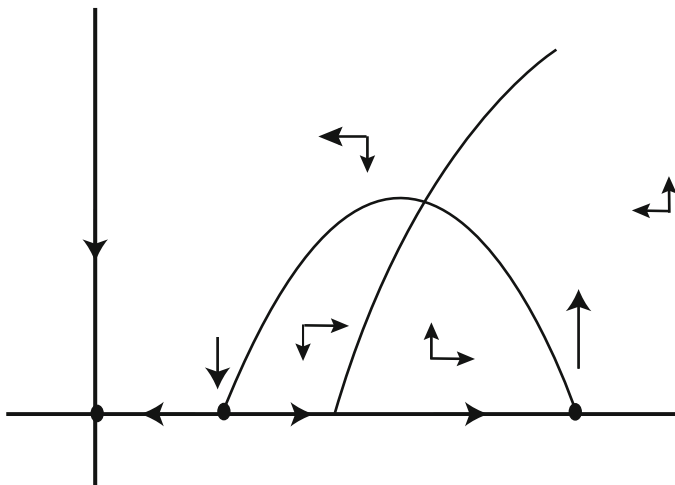


Fig. 5.34 Coexistence for some initial states.

In the Kolmogorov model

$$\begin{aligned} x' &= xf(x, y, z), \\ y' &= yg(x, y, z), \\ z' &= zh(x, y, z), \end{aligned} \quad (5.34)$$

we define the numbers $J_1, K_1, L_1, J_2, K_2, L_2$ by

$$\begin{aligned} f(K_1, 0, 0) &= 0, & f(0, L_2, 0) &= 0, \\ g(0, K_2, 0) &= 0, & f(L_1, 0, 0) &= 0, \\ h(J_1, 0, 0) &= 0, & h(0, J_2, 0) &= 0. \end{aligned}$$

Thus, K_1 is the natural x population size in the absence of the other two species, J_1 is the minimum x population size that allows z to survive in the absence of y , and L_1 is the maximum x population size that allows y to survive unconditionally (for all initial states) in the absence of z . The numbers K_2, J_2, L_2 have analogous interpretations for the y population size. We may also define these quantities for the more explicit Rosenzweig–MacArthur model.

$$\begin{aligned} x' &= rxf(x, y) - xz\phi(x, y), \\ y' &= syg(x, y) - yz\psi(x, y), \\ z' &= z(\lambda x\phi(x, y) + \mu y\psi(x, y) - c). \end{aligned} \quad (5.35)$$

Before looking at the conditions for each species to be a survivor we must recall the properties of the three two-species subsystems.

- (i) If $K_1 < L_1$ and $K_2 < L_2$, then x and y coexist unconditionally (written $x \sim y$) with equilibrium $(\bar{x}, \bar{y}, 0)$. If $K_1 > L_1$ and $K_2 < L_2$, we have the bistable case ($x \oplus y$)

- with two possible equilibria, $(K_1, 0, 0)$ and $(0, K_2, 0)$. If $K_1 > L_1$ and $K_2 < L_2$, then x always wins the competition ($x \gg y$) with equilibrium $(K_1, 0, 0)$.
- (ii) If $J_1 < K_1$, then x and z coexist ($x \sim y$) with equilibrium $(x^*, 0, z^*)$. If $J_1 > K_1$, then only x survives ($x \gg z$) with equilibrium $(K_1, 0, 0)$.
- (iii) If $J_2 < K_2$, then $y \sim z$, with equilibrium $(0, y^*, z^*)$. If $J_2 > K_2$, then $y \gg z$, with equilibrium $(0, K_2, 0)$.

The survivor conditions for (5.34) are as follows:

$$\begin{aligned}
 x \text{ is a survivor if } & \begin{cases} f(0, y^*, z^*) > 0 & (y \sim z), \\ f(0, K_2, 0) > 0 \text{ or } K_2 < L_2 & (y \gg z), \end{cases} \\
 y \text{ is a survivor if } & \begin{cases} g(x^*, 0, z_1^*) > 0 & (x \sim z), \\ g(K_1, 0, 0) > 0 \text{ or } K_1 < L_1 & (x \gg z), \end{cases} \\
 z \text{ is a survivor if } & \begin{cases} h(\bar{x}, \bar{y}, 0) > 0, & (x \sim z) \\ h(K_1, 0, 0) > 0, h(0, K_2, 0) > 0, \\ K_1 < J_1, K_2 > J_2, & (x \oplus y) \\ h(K_1, 0, 0) > 0 \text{ or } K_1 > J_1 & (x \gg y). \end{cases}
 \end{aligned}$$

For the Rosenzweig–MacArthur model (5.35) some of these conditions may be described more explicitly. For example, if $y \sim z$, we have $y^* = J_2$, $z_1^* = sg(0, J_2)/\psi(0, J_2)$, and thus x is a survivor if $rf(0, J_2)\psi(0, J_2) - sg(0, J_2)\phi(0, J_2) > 0$. Similarly, if $x \sim z$, y is a survivor if $-rf(J_1, 0)\psi(J_1, 0) + sg(J_1, 0)\phi(J_1, 0) > 0$.

Next we use these conditions to determine whether our model will support the possibility of predator-induced exclusion, which in our terminology would be $x \sim y$ but x not a survivor species. The assumption that $x \sim y$ is equivalent to $K_1 < L_1$ and $K_2 < L_2$. If $y \sim y$ so that $J_2 < K_2$, the condition that x not be a survivor is $f(0, y^*, z^*) < 0$. In the Rosenzweig–MacArthur model, $y^* = J_2$ and $z_1^* = sg(0, J_2)/\psi(0, J_2)$, so that this condition becomes $rf(0, J_2)\psi(0, J_2) - sg(0, J_2)\psi(0, J_2) < 0$. Since $J_2 < L_2 < K_2$, we have $f(0, J_2) > 0$ and $g(0, J_2) > 0$. Thus if s/r is sufficiently large, x is not a survivor. If $y \gg z$, so that $J_2 > K_2$, the condition for x to be a survivor is $K_2 < L_2$, which is satisfied. Thus, it is possible to have $x \sim y, y \sim z$ with x not a survivor, which can be viewed as predator-induced exclusion. It may also be viewed as exclusion of the prey x by the alternative prey.

A more complicated analysis is needed to examine the possibility of predator-induced coexistence. In our terminology this means that x and y do not coexist unconditionally, either $x \gg y$ or $x \oplus y$, but all three species are survivors. Let us begin with the possibilities when $x \gg y$ or $x \oplus y$, but all three species are survivors. Let us begin with the possibilities when $x \gg y$ or $L_1 < K_1$ and $K_2 < L_2$.

- (i) If $x \sim z$ and $y \gg z$, or $J_1 < K_1$ and $J_2 > K_2$, then x and z are survivors, and y is a survivor if $g(x^*, 0, z_1^*) > 0$, which is possible if $J_1 < L_1$. In the Rosenzweig–MacArthur case y will be a survivor if $J_1 < L_1$ and s/r are large.
- (ii) If $x \sim y, y \sim z$, or $J_1 < K_1$ and $J_2 < K_2$, z is a survivor, while x is a survivor if $f(0, y^*, z_2^*) > 0$, and y is a survivor if $g(x^*, 0, z_1^*) > 0$. In the Rosenzweig–MacArthur case x is a survivor if s/r is small enough, and y is a survivor if s/r

is large enough. Thus, either or both competitors may survive, but for both to be survivors two conditions that may be incompatible must be satisfied.

- (iii) If $x \gg z$, $y \gg z$, or $J_1 > K_1$ and $J_2 > K_2$, x is a survivor but y and z are not survivors.
- (iv) If $x \gg z$, $y \sim z$, or $J_1 > K_1$ and $J_2 > K_2$, x is a survivor if $f(0, y^* z_2^*) > 0$ but y and z are not survivors.

From this we conclude that predator-induced coexistence is impossible if $x \gg y$ and $x \gg z$, but is possible if $x \gg y, x \sim z, y \gg z$, and may be possible if $x \gg y, x \sim z, y \sim z$. Further, if $x \sim z, y \sim z, x \gg y$, it is possible for the predator to reverse the outcome of the competition.

If $x \oplus y$, we have $K_2 > L_2$ in place of $K_2 < L_2$. In this case x cannot be a survivor if $x \sim z, y \gg z$, or if $x \gg z, y \gg z$ but otherwise the results are the same. The only case in which both x and y can be survivors is the case $x \oplus y, x \sim z$, and the survivor conditions for x and y could be incompatible. In the Lotka–Volterra model, $f(x, y) = 1 - x/K_1 - y/L_2$, $g(x, y) = 1 - x/L_1 - y/K_2$, $\phi(x, y) = c/\lambda J_1$, $\psi(x, y) = c/\mu J_2$, it is not difficult to show that the two survivor conditions are indeed incompatible.

Three-species coexistence is not possible in the bi-stable case [Hutson and Vickers (1983)]. However, with more complicated predator response terms it is possible to give examples in which three-species coexistence may or may not be possible, depending on the parameters of the model.

The possibilities of predator-induced exclusion of one of the prey species and predator-induced coexistence are both supported by the Rosenzweig–MacArthur model (5.35). This both vindicates the Rosenzweig–MacArthur model as an approximation to biological reality and illustrates the use of the survivor species idea to study three-dimensional systems.

5.11 Example: Two Predators Competing for Prey

We discussed the principle of competitive exclusion of Gause (1934) in Section 5.1. There we analyzed a two-dimensional model for species in competition, without going into detail as to the nature of the competition. The principle of competitive exclusion is sometimes formulated as saying that it is not possible for n species to coexist when competing for fewer than n different resources as food supply. If we consider the resources as prey species, the simplest case would be two species each acting as predator on a third species. Let us consider a three-species Rosenzweig–MacArthur model

$$\begin{aligned} x' &= sf(x) - xy_1\phi_1(x) - xy_2\phi_2(x), \\ y_1' &= y_1(c_1x\phi_1(x) - e_1), \\ y_2' &= y_2(c_1x\phi_2(x) - e_2), \end{aligned} \tag{5.36}$$

to describe such a situation. Here x is the population size of the common prey species, and y_1 and y_2 are the population sizes of the competing predator species. For three-species coexistence to be possible, the two-species systems describing x

and y_1 in the absence of y_2 , and x and y_2 in the absence of y_1 , must both admit coexistence. The x - y_1 system model is

$$x' = xf(x) - xy_1\phi_1(x) \quad \text{and} \quad y_1' = y_1(c_1x\phi_1(x) - e_1), \quad (5.37)$$

and coexistence requires an equilibrium with

$$x = J_1, \text{ where } c_1J_1\phi_1(J_1) = e_1 \quad (5.38)$$

and $y_1 = f(J_1)/\phi_1(J_1)$. Then y_2 , considered as an invading species, is a survivor species if

$$c_2J_1\phi_2(J_1) > e_2. \quad (5.39)$$

Similarly, the x - y_2 system

$$x' = xf(x) - xy_2\phi_2(x) \quad \text{and} \quad y_2' = y_2(c_2x\phi_2(x) - e_2) \quad (5.40)$$

has an equilibrium with $x = J_2$, where

$$c_2J_2\phi_2(J_2) > e_2 \quad (5.41)$$

and $y_2 = f(J_2)/\phi_2(J_2)$. Then y_1 is a survivor species if

$$c_1J_2\phi_1(J_2) > e_1. \quad (5.42)$$

Because of the standard assumption that the predator functional responses $x\phi_1(x)$ and $x\phi_2(x)$ are increasing functions, the relationships (5.38) and (5.42) imply $J_2 > J_1$, while the relationships (5.39) and (5.41) imply $J_1 > J_2$. Thus, y_1 and y_2 cannot both be survivor species, and this at least suggests the possibility that y_1 and y_2 cannot coexist.

However, an example has been constructed by R.A. Armstrong and R. McGehee (1980) of a system of the form (5.36) that does admit coexistence. The construction is indirect and quite complicated, and it is not known whether the asymptotically stable orbit of the example is periodic or a recurrent orbit that is not periodic. In any case, the example violates the competitive exclusion principle.

5.12 Project: A Simple Neuron Model

Neurons are cells in the body that transmit information to the brain and the body by amplifying an incoming stimulus (electrical charge input) and transmitting it to neighboring neurons, then turning off to be ready for the next stimulus. Neurons have fast and slow mechanisms to open ion channels in response to electrical charges. The key quantities are the concentrations of sodium ions and potassium ions (both positively charged). A resting neuron has an excess of potassium and a deficit of sodium and a negative resting potential (an excess of negative ions).

Neurons use changes of sodium and potassium ions across the cell membrane to amplify and transmit information. *Voltage-gated channels* exist for each kind of ion, which open and close in response to voltage differences, which are closed in a resting neuron. When a burst of positive charge enters the cell, making the potential less negative, the voltage-gated sodium channels open. Since there is an excess of sodium ions outside the cell, more sodium ions enter, increasing the potential until it eventually becomes positive. Next a slow mechanism acts to block the voltage-gated sodium channels, and another slow process begins to open voltage-gated potassium channels. Both of these diminish the buildup of positive charge by blocking sodium ions from entering and by allowing excess potassium ions to leave. When the potential decreases to or below the resting potential these slow mechanisms turn off, and then the process can start over. If the electrical excitation reaches a sufficiently high level, called an *action potential*, the neuron fires and transmits the excitations to other neurons.

In order to describe a simple model for this process, we let the potential be v , scaled so that $v = 0$ is the resting potential. We let $v = a$ be the potential above which the neuron fires and $v = 1$ the potential at which sodium channels open ($0 < a < 1$). A model of the form

$$v' = -v(v - a)(v - 1)$$

with asymptotically stable equilibria at $v = 0$ and $v = 1$ and unstable equilibrium at $v = a$ would explain part of the observed behavior. If the initial potential is above a , the potential increases to one and if the initial potential is below a , it decreases to zero. Thus, the model allows the signal amplification of the neuron but stops at $v = 1$. We must also build in a blocking mechanism.

Let w denote the strength of the blocking mechanism with $w = 0$ (turned off) when $v = 0$. As v approaches one, the blocking mechanism becomes stronger but remains bounded, and we assume an equation of the form

$$w' = \varepsilon(v - \xi w)$$

with a limiting value v/ξ for w if v is fixed. If $v = 0$, then $w \rightarrow 0$, and if $v = 1$, $w \rightarrow 1/\xi$ (the maximum strength of the blocking mechanism). The parameter ε influences the rate of approach to equilibrium but does not affect the equilibrium value. We use a small value of ε to indicate a slow-acting mechanism.

In order to formulate a model that includes both v and w , we must also take account of the effect of the blocking mechanism on v . The model we shall examine is the two-dimensional system

$$\begin{aligned} v' &= -v(v - a)(v - 1) - w \\ w' &= \varepsilon(v - \xi w) \end{aligned} \tag{5.43}$$

known as the *Fitzhugh–Nagumo* system. [Fitzhugh (1961); Nagumo, Arimoto, & Yoshizawa (1962)] It is a simplification of the four-dimensional *Hodgkin–Huxley* model proposed in the early 1950s by Sir Alan Hodgkin and Sir Andrew Huxley [Hodgkin and Huxley (1952)], for which they received the 1963 Nobel Prize in

Physiology and Medicine, and which is still being used in the study of neurons and other kinds of cells. Excitable systems occur in a large variety of biological systems, and the Fitzhugh–Nagumo and Hodgkin–Huxley models are prototypical models of excitable systems.

Question 1

Show that the only equilibrium of the Fitzhugh–Nagumo system (5.43) is $v = 0$, $w = 0$, and that this equilibrium is asymptotically stable.

Question 2

Use a computer algebra system to draw the orbit of the system (5.43) with $a = 0.3$, $\xi = 1$, $\varepsilon = 0.01$ and a starting point $(v_0, 0)$ with $v_0 > a$, say $v_0 = 0.4$. You should observe that v grows quickly and then returns to zero. You should also note that after the neuron fires and the potential drops, it overshoots zero (this can also be observed experimentally).

Another experiment gives the cell a constant input of positive ions instead of a single pulse. If we apply a constant electrical current J , we add J to the rate of change of potential (taking units of current so that one unit of current raises the potential by one unit in unit time). Thus, we replace the system (5.43) by

$$\begin{aligned} v' &= -v(v-a)(v-1) - w + J, \\ w' &= \varepsilon(v - \xi w). \end{aligned} \tag{5.44}$$

Question 3

Show that increasing J moves the equilibrium from $(0, 0)$ into the first quadrant of the (v, w) -plane. Show that this equilibrium is asymptotically stable for small values of J but becomes unstable for larger values of J .

Question 4

Use a computer algebra system to experiment with different values of J , taking $a = 0.3$, $\xi = 1$, $\varepsilon = 0.01$, and find a value for which the system (5.44) has a periodic orbit.

Question 5

Use a computer algebra system to graph v as a function of t for the periodic orbit found in Question 4. You should observe a “bursting” behavior, with potential rising close to one and dropping below zero (similar to observations in real neurons).

Question 6

Add a periodic term to the equation for u in the model and experiment using a computer algebra system with different amplitudes and periods to see the effects on the behavior of the model.

5.13 Project: A Plant–Herbivore Model

This project is taken from Edelstein-Keshet (1988).

Assume that a population of herbivores of density y causes changes in the vegetation on which it preys. An internal variable x reflects some physical or chemical property of the plant, which undergoes changes in response to herbivory. We refer to this attribute as the *plant quality* of the vegetation and assume that it may in turn affect the fitness or survivorship of the herbivores. If this happens in a graded, continuous interaction, plant quality may be modeled by a system

$$x' = f(x, y), \quad y' = yg(x, y).$$

1. In one case, the function $f(x, y)$ is assumed to be

$$f(x, y) = x(1 - x)[\alpha(1 - y) + x], \quad 0 \leq x \leq 1.$$

Sketch this as a function of x and reason that plant quality x always remains in the interval $(0, 1)$ if $x(0)$ is in this interval. Show that plant quality may either increase or decrease, depending on the initial value of x and the population of herbivores. For a given herbivore population density \hat{y} , what is the “break-even” point (the level of x for which the rate of change of x is zero)?

2. It is assumed that the herbivore population undergoes logistic growth with reproductive rate β and a carrying capacity that is directly proportional to current plant quality. What is the function g ?
3. Draw the nullclines of the system you have obtained. (There is more than one possibility, depending on the parameter values.)
4. Find the direction of motion along all nullclines obtained in part (c).
5. Show that there is a particular configuration with a set in the (x, y) plane that “traps” trajectories.
6. Define $\gamma = \alpha/(\alpha K - 1)$. Interpret the meaning of this parameter. Show that $(\gamma, K\gamma)$ is a steady state of your equations, and locate it on your phase plot. Find the other steady state.
7. Show that $(\gamma, K\gamma)$ is a saddle point if $\gamma > 1$ and a focus if $\gamma < 1$.
8. Now show that as β decreases from large to small values, the steady state for $\gamma < 1$ undergoes a transition from a stable focus to an unstable focus.
9. Use the results you have obtained to comment on the existence of periodic solutions. Interpret your answer in biological terms.

[Reference: Edelstein-Keshet (1988)]

Chapter 6

Harvesting in Two-species Models

6.1 Harvesting of Species in Competition

The topics in this chapter are part of the subject of natural resource management and bioeconomics. This is an important subject that is developing rapidly. The classical reference is the book by Clark (1990), where additional references may be found.

In Section 1.5 we examined the effects of both constant-yield harvesting and constant-effort harvesting on single-species populations modeled by an ordinary differential equation. In this section we shall study the effects of harvesting one of a pair of interacting species in competition.

In studying models for competition between two species (Section 5.1), we began by assuming linear per capita growth rates, obtaining a system

$$\begin{aligned}x' &= x(\lambda - ax - by), \\ y' &= y(\mu - cx - dy).\end{aligned}\tag{6.1}$$

We observed that there are four cases depending on the relationship between the x -isocline $ax + by = \lambda$, on which $x' = 0$ and the y -isocline $cx + dy = \mu$, on which $y' = 0$. These cases may be depicted graphically in [Figure 6.1](#), with equilibria marked \times and asymptotically stable equilibria marked \bullet .

We will consider only the harvesting of one of the two species, say the x -species. For constant-effort harvesting, the model would be

$$\begin{aligned}x' &= x(\lambda - ax - by) - Ex, \\ y' &= y(\mu - cx - dy).\end{aligned}$$

This is the same as the unharvested model but with λ replaced by $\lambda - E$, or with the x -isocline moved parallel to itself toward the origin. It has boundary equilibria $(K(E), 0)$ and $(0, M)$ with $M = \frac{\mu}{d}$ and $K(E) = \frac{\lambda}{a+E}$. Examination of the four cases shows that it is possible by harvesting to change Case 1 to Case 4 (coexistence to x -extinction), Case 2 to Case 4 (survival of either species alone to x -extinction), or Case 3 to Case 1 or Case 2 and then to Case 4 (y -extinction to x -extinction via

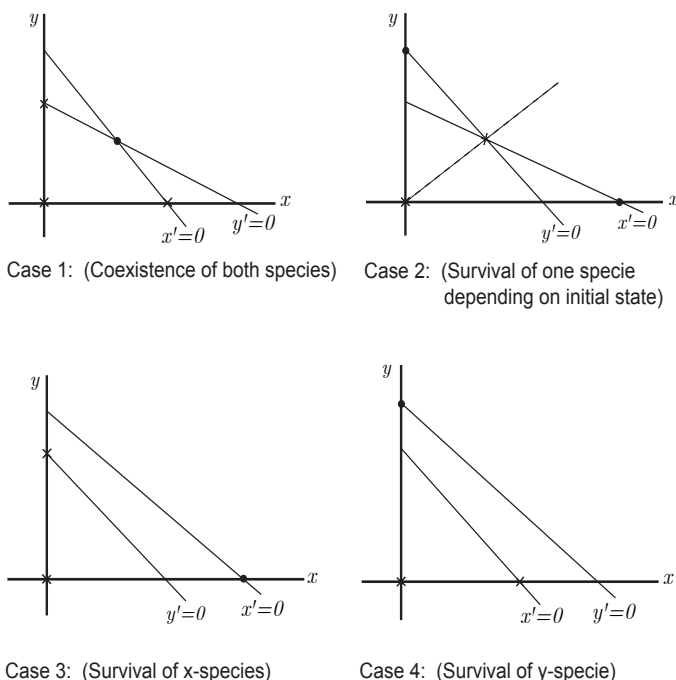


Fig. 6.1 Species in competition cases.

coexistence). If the growth rates are not linear, the same type of argument can be used to see how the isoclines move under harvesting and to obtain the same results, all of which are in accordance with our intuition.

With constant-yield harvesting, the model is

$$x' = x(\lambda - ax - by) - H, y' = y(\mu - cx - dy).$$

The x -isocline, instead of being the pair of lines $x = 0$, $ax + by = \lambda$, is now the curve $x(\lambda - ax - by) = H$, which is a hyperbola having the lines $x = 0$ and $ax + by = \lambda$ as asymptotes and which moves away from these asymptotes as H increases. The effect of harvesting on the system can be depicted graphically by sketching these hyperbolas; the results are somewhat different in the four cases.

Case 1, the asymptotically stable (coexistence) equilibrium moves up and to the left and the saddle point moves from $(0, M)$ down and to the right (see [Figure 6.2](#)). The stable separatrices at this saddle point move from the y -axis into the first quadrant, and divide this quadrant into a coexistence region (initial states that tend to the asymptotically stable equilibrium) and a region for which the x -species becomes extinct in finite time. Since $x = 0$ is not a solution of the harvested system, it is possible for orbits to cross the y -axis, corresponding to extinction of the x species in finite time. As the hyperbola moves down, the saddle point and asymptotically stable equilibrium coalesce and disappear, producing a catastrophe. For harvest rates

larger than the critical rate at which this coalescence occurs, the x species becomes extinct in finite time for all initial states.

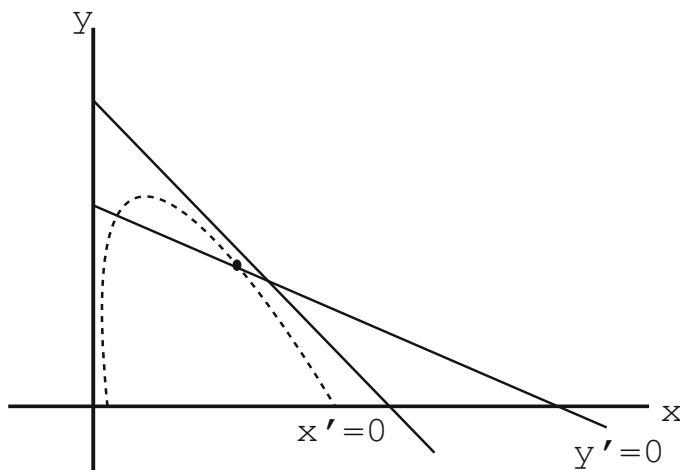


Fig. 6.2 Case 1.

Case 2, the asymptotically stable equilibrium at $(K, 0)$ moves to the left along the x -axis and the asymptotically stable equilibrium at $(0, M)$ moves into the second quadrant, corresponding to extinction of the x species in finite time (see Figure 6.3). The saddle point at (x_∞, y_∞) moves down and to the right, remaining a saddle point, and its stable separatrices divide the first quadrant into a region of x -extinction in finite time and a region of ultimate y -extinction. As the hyperbola moves down with increasing H , the equilibrium at (x_∞, y_∞) coalesces either with the equilibrium at $(0, 0)$ or the equilibrium at $(K, 0)$ and disappears; then the two remaining equilibria on the x -axis coalesce and disappear. The second coalescence signifies x -extinction for all initial states.

Case 3, the asymptotically stable equilibrium moves from $(K, 0)$ along the x -axis. If $ad - bc > 0$, when the hyperbola meets the line $cx + dy = \mu$ this equilibrium moves into the interior of the first quadrant, giving coexistence (see Figure 6.4(A)). The other equilibrium is a saddle point whose stable separatrices give a division into a coexistence region and a region of extinction of the x -species in finite time (see Figure 6.4(B)). When the two equilibria coalesce and disappear, we have extinction of the x -species in finite time for all initial states.

Case 4, the asymptotically stable equilibrium moves from $(0, M)$ into the second quadrant, giving extinction of the x species in finite time for all initial states (see Figure 6.5).

If we do not assume that the per capita growth rates are linear, but instead require only a model of Kolmogorov type,

$$x' = xf(x, y), \quad y' = yg(x, y),$$

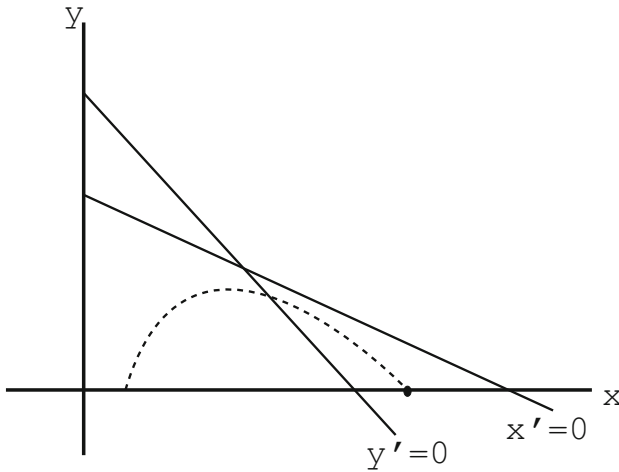


Fig. 6.3 Case 2.

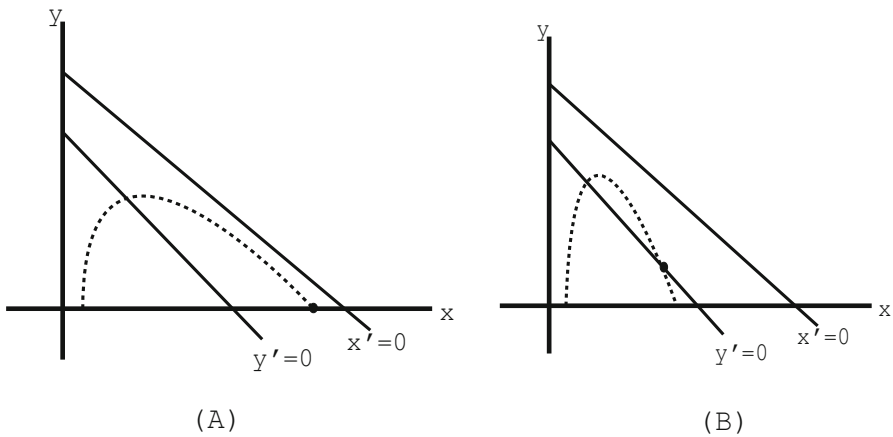


Fig. 6.4 Case 3.

with $f_x(x, y) < 0$, $f_y(x, y) < 0$, $g_x(x, y) < 0$, $g_y(x, y) < 0$ for $x > 0, y > 0$, then the isoclines are curves with negative slope but are not necessarily straight lines. The x -isocline under constant-yield harvesting is no longer necessarily a hyperbola, but it does approach the line $x = 0$ and the curve $f(x, y) = 0$ asymptotically. The same cases as those described for linear growth rates may occur, and the qualitative responses to harvesting are the same.

Example 1: Determine the response of the system

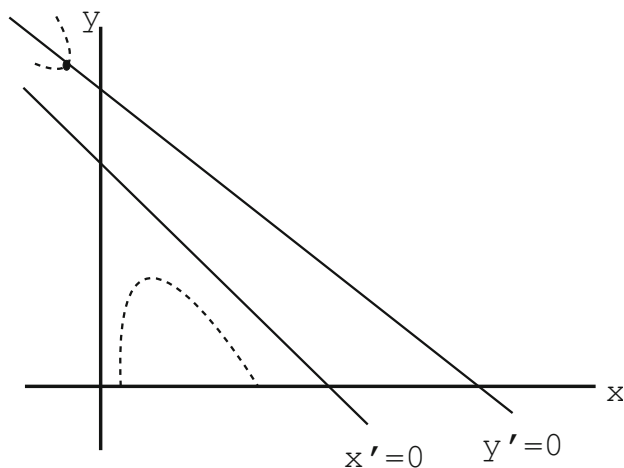


Fig. 6.5 Case 4.

$$\begin{aligned}x' &= x(100 - 4x - y), \\y' &= y(60 - x - 2y),\end{aligned}$$

to constant-effort harvesting of the x species.

Solution: With no harvesting, there is an asymptotically stable equilibrium at $(20, 20)$ [Example 1, Section 5.1]. A harvested system has the form

$$\begin{aligned}x' &= x[(100 - E) - 4x - y], \\y' &= y(60 - x - 2y).\end{aligned}$$

Equilibria are given by the pair of equations

$$4x + y = 100 - E, \quad x + 2y = 60,$$

and elimination gives the solution

$$x = 20 - \frac{2}{7}E, \quad y = 20 + \frac{E}{7}.$$

When $E = 70$, x reaches zero, and $y = 30$. Thus harvesting decreases the x -population size and increases the y -population size, eventually moving coexistence to x -extinction.

Example 2: Determine the response of the system

$$\begin{aligned}x' &= x(100 - 4x - y), \\y' &= y(60 - x - 2y),\end{aligned}$$

to constant-yield harvesting of the x species.

Solution: With no harvesting, there is an asymptotically stable equilibrium at $(20, 20)$ [Example 1, Section 5.1]. A harvested system has the form

$$\begin{aligned}x' &= x(100 - 4x - y) - H, \\y' &= y(60 - x - 2y).\end{aligned}$$

Equilibria are given by the pair of equations

$$x(100 - 4x - y) = H, \quad x + 2y = 60.$$

Replacing x by $60 - 2y$ in the first of these equations, we obtain a quadratic equation for y

$$\begin{aligned}(60 - 2y)[100 - 4(60 - 2y) - 2y] &= H, \\14(y - 30)(y - 20) + H &= 0, \\14y^2 - 700y + (8400 + H) &= 0.\end{aligned}$$

The roots of this equation may be found using the quadratic formula:

$$y = 25 \pm \frac{1}{2} \sqrt{100 - \frac{2H}{7}}.$$

For $H = 0$, these roots are $y = 20$ (which gives $x = 20$) and $y = 30$ (which gives $x = 0$). There is an asymptotically stable equilibrium at $(20, 20)$ and a saddle point at $(0, 30)$. As H increases, these equilibria move along the line $x + 2y = 60$ until they coalesce at $(10, 25)$ for $H = 350$.

Exercises

In each of Exercises 1–4, determine the response of the system to constant-effort harvesting of the x species.

1. $x' = x(80 - x - y)$, $y' = y(120 - x - 3y)$ [cf. Exercise 1, Section 5.1].
2. $x' = x(60 - 3x - y)$, $y' = y(75 - 4x - y)$ [cf. Exercise 2, Section 5.1].
3. $x' = x(40 - x - y)$, $y' = y(90 - x - 2y)$ [cf. Exercise 3, Section 5.1].
4. $x' = x(80 - 3x - 2y)$, $y' = y(80 - x - y)$ [cf. Exercise 4, Section 5.1].

In each of Exercises 5–8, determine the response of the system to constant-yield harvesting of the x species.

5. $x' = x(80 - x - y)$, $y' = y(120 - x - 3y)$ [cf. Exercise 1, Section 5.1].
6. $x' = x(60 - 3x - y)$, $y' = y(75 - 4x - y)$ [cf. Exercise 2, Section 5.1].
7. $x' = x(40 - x - y)$, $y' = y(90 - x - 2y)$ [cf. Exercise 3, Section 5.1].
8. $x' = x(80 - 3x - 2y)$, $y' = y(80 - x - y)$ [cf. Exercise 4, Section 5.1].

6.2 Harvesting of Predator–Prey Systems

Predator–prey systems are most commonly used to describe the interaction of a species with its food supply. In other words, the primary interest is in the predator population size, and the prey population size is of interest only because of its effect on the predator growth rate. A predator–prey model may be viewed as a refinement of a single-species model through more realistic description of the limitations on the growth rate imposed by resource limitations. Thus we will study only the harvesting of predators. It would be possible, however, to use an analogous approach to study the harvesting of prey and this would be appropriate in an examination of the extent to which one can control a population by tampering with its food supply.

Constant-effort harvesting: For a system modeled by

$$x' = xf(x) - xy\phi(x), \quad y' = y(cx\phi(x) - e),$$

(Rosenzweig–MacArthur model), the harvested system is modeled by

$$x' = xf(x) - yx\phi(x), \quad y' = y(cx\phi(x) - e - E).$$

The effect of harvesting is to move the predator isocline $x = J$, where $cJ\phi(J) = e$, to the right. This applies to the more general (Kolmogorov type) model

$$x' = xf(x, y), \quad y' = y(g(x) - E),$$

with the predator per capita growth rate independent of predator population size. If the prey isocline $y = \frac{f(x)}{\phi(x)}$ (as in the Rosenzweig–MacArthur model) or $f(x, y) = 0$ (as in the Kolmogorov model) is concave downwards and has a maximum, an equilibrium that is unstable for $E = 0$ is stabilized by harvesting as the harvested predator isocline moves past the maximum of the prey isocline (see Figure 6.6).

For a given model, the effort that maximizes yield may be calculated by solving

$$\frac{d}{dE}(Ey_\infty) = E \frac{dy_\infty}{dE} + y_\infty = 0$$

for E and then calculating $Ey_\infty(E)$ for this value of E . Alternatively, one would proceed as follows for a Kolmogorov model. For a given effort E , the equilibrium conditions are $f(x_\infty(E), y_\infty(E)) = 0$, $g(x_\infty(E)) = E$, and the yield is $Ey_\infty(E) = y_\infty(E)g(x_\infty(E))$. The maximum yield is thus the maximum of the function $yg(x)$ subject to the constraint $f(x, y) = 0$. So long as $E < cK\phi(K) - e$, where $f(K) = 0$ (Rosenzweig–MacArthur model) or $E < g(K)$, where $f(K, 0) = 0$ (Kolmogorov model), $y_\infty(E) > 0$ and the equilibrium depends continuously on E . For $E \geq cK\phi(K) - e$ or $E \geq g(K)$ in the respective models, we have $y_\infty(E) = 0$, corresponding to the ultimate extinction of predators.

Constant-yield harvesting: For a Rosenzweig–MacArthur model the harvested system is

$$x' = xf(x) - xy\phi(x), \quad y' = y(cx\phi(x) - e) - H.$$

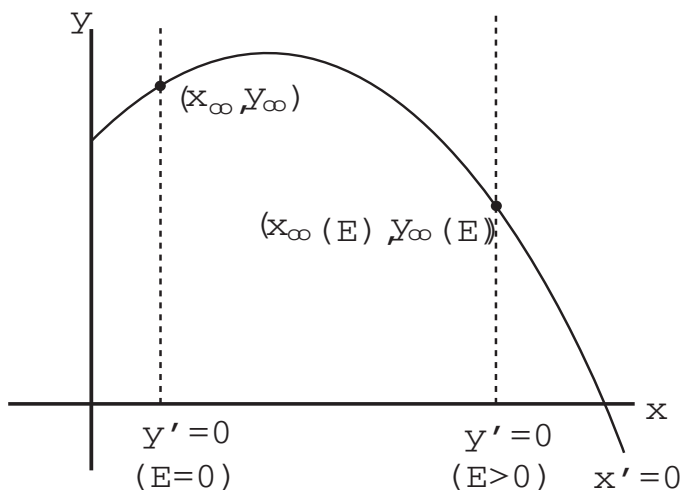


Fig. 6.6 Stabilization under harvesting.

The predator isocline $cx\phi(x) = e$ is replaced by the curve $y(cx\phi(x) - e) = H$ to the right of $cx\phi(x) = e$ and having $cx\phi(x) = e$ as a vertical asymptote (since $cx\phi(x) - e \rightarrow 0$ as $y \rightarrow \infty$) and a horizontal asymptote $y = \frac{H}{c \lim_{x \rightarrow \infty} x\phi(x) - e}$ (since $y \rightarrow \frac{H}{c \lim_{x \rightarrow \infty} x\phi(x) - e}$ as $x \rightarrow \infty$). It is no more difficult to study the more general system

$$x' = xf(x, y), \quad y' = yg(x),$$

or indeed the general Kolmogorov model

$$x' = xf(x, y), \quad y' = yg(x, y),$$

with harvested system

$$x' = xf(x, y), \quad y' = yg(x, y) - H,$$

in which $f_y(x, y) < 0$, $g_x(x, y) > 0$, $g_y(x, y) \leq 0$, $f(K, 0) = 0$, $g(J, 0) = 0$; we will usually think in terms of the special case where g is independent of y . As H increases, the “hyperbola-like” curve $y(cx\phi(x) - e) = H$ or $yg(x, y) = H$ moves away from the line $x = J$ or the curve $g(x, y) = 0$ and upward to the right.

It intersects the prey isocline $f(x, y) = 0$ at two points: $(x_\infty(H), y_\infty(H))$ coming from (x_∞, y_∞) when $H = 0$, and $(\xi(H), y(H))$ coming from $(K, 0)$ when $H = 0$ (Figure 6.7). As H increases, these two equilibria move together, coalesce, and disappear. The value of H for which they coalesce is the value for which the level curve $yg(x, y) = H$ of the function $yg(x, y)$ is tangent to the prey isocline. This is the largest value of the function $yg(x, y)$ attained on the curve $f(x, y) = 0$; we define

$$H_c = \max_{f(x, y)=0} yg(x, y).$$

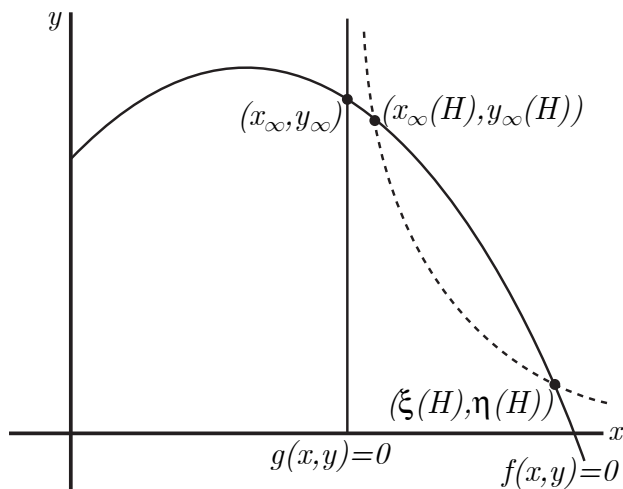


Fig. 6.7 Two equilibria.

Then for $H > H_c$ there are no equilibria in the first quadrant (Figure 6.8).

If $H > 0$, then $y' < 0$ on the x -axis. This means that solutions can reach and cross the x -axis in finite time (predator extinction). Indeed, if $H > H_c$, this occurs for all solutions, independent of initial values.

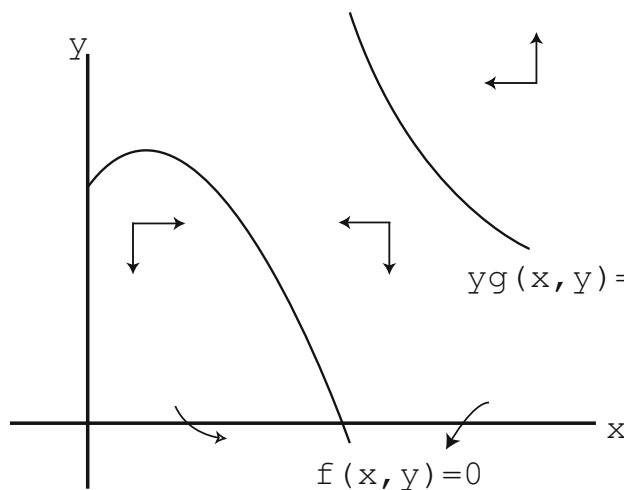


Fig. 6.8 No equilibrium.

It is possible to prove that no solution can be unbounded; the proof is similar to the proof in the case $H = 0$ but the details are more complicated. Thus a solution either tends to an equilibrium, or tends to a limit cycle, or reaches the x -axis in finite

time. If a solution reaches the x -axis in finite time, we consider the system to have collapsed.

To study the stability of any equilibrium (x_0, y_0) we look at the community matrix

$$A(x_0, y_0) = \begin{pmatrix} x_0 f_x(x_0, y_0) & x_0 f_y(x_0, y_0) \\ y_0 g_x(x_0, y_0) & g(x_0, y_0) + y_0 g_y(x_0, y_0) \end{pmatrix}.$$

If $\det A(x_0, y_0) < 0$, then (x_0, y_0) is a saddle point and if $\det A(x_0, y_0) > 0$, then (x_0, y_0) is a node or spiral point that is unstable if $\text{tr} A(x_0, y_0) > 0$ and asymptotically stable if $\text{tr} A(x_0, y_0) < 0$.

The slope of the prey isocline $f(x, y) = 0$ at (x_0, y_0) is $-\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)}$, and the slope of the (harvested) predator isocline $yg(x, y) = H$ at (x_0, y_0) is $-\frac{y_0 g_x(x_0, y_0)}{y_0 g_y(x_0, y_0) + g(x_0, y_0)}$. If we assume $g(x_0, y_0) + y_0 g_y(x_0, y_0) \geq 0$, which is certainly true if $g(x, y)$ is independent of y , and use the hypothesis $f_y(x_0, y_0) < 0$, we see that the slope of the prey isocline is less than the slope of the predator isocline if and only if $f_x(yg_y + g) < yg_x f_y$, or

$$y_0(f_x(x_0, y_0)g_y(x_0, y_0) - f_y(x_0, y_0)g_x(x_0, y_0)) + g(x_0, y_0)f_x(x_0, y_0) < 0.$$

Because

$$\det A(x_0, y_0) = x_0 y_0(f_x(x_0, y_0)g_y(x_0, y_0) - f_y(x_0, y_0)g_x(x_0, y_0)) + g(x_0, y_0)f_x(x_0, y_0),$$

(x_0, y_0) is a saddle point if and only if the slope of the prey isocline at (x_0, y_0) is less than the slope of the predator isocline at (x_0, y_0) . Thus $(\xi(H), y(H))$ is a saddle point and $(x_\infty(H), y_\infty(H))$ is not a saddle point. If $(x_\infty(H), y_\infty(H))$ is asymptotically stable, some orbits tend to it, and if $(x_\infty(H), y_\infty(H))$ is unstable, then either there is a limit cycle to which some orbits tend, or every orbit reaches the x -axis in finite time. We define the region of coexistence to be the set of initial population sizes $(x(0), y(0))$ for which the orbit tends either to the equilibrium $(x_\infty(H), y_\infty(H))$ or to a limit cycle around this equilibrium as $E \rightarrow \infty$. We know that if $H = 0$, the region of coexistence is the whole first quadrant of the x - y plane. If $H > 0$, the region of coexistence is smaller and depends on the separating orbits at the saddle point $(\xi(H), \eta(H))$. There is one stable separatrix in the region $x' < 0, y' > 0$, and there is one unstable separatrix in the region $x' > 0, y' < 0$ (Figure 6.9).

There are two possible cases for the other two separatrices:

“Good case”: There is a stable separatrix running from $x \rightarrow +\infty$ around $(x_\infty(H), y_\infty(H))$ to $(\xi(H), \eta(H))$ and an unstable separatrix running from $(\xi(H), \eta(H))$ to $(x_\infty(H), y_\infty(H))$ or a limit cycle around $(x_\infty(H), y_\infty(H))$. The coexistence region is bounded by the two stable separatrices, and orbits starting in this region tend to $(x_\infty(H), y_\infty(H))$ or to a limit cycle around $(x_\infty(H), y_\infty(H))$ (Figure 6.10).

“Bad case”: There is a stable separatrix running from the equilibrium $(x_\infty(H), y_\infty(H))$ or from a periodic solution around $(x_\infty(H), y_\infty(H))$ and an unstable separatrix around $(x_\infty(H), y_\infty(H))$ and then down to the x -axis. Either there is no coexistence region at all, or there is an unstable periodic orbit around the asymptotically sta-

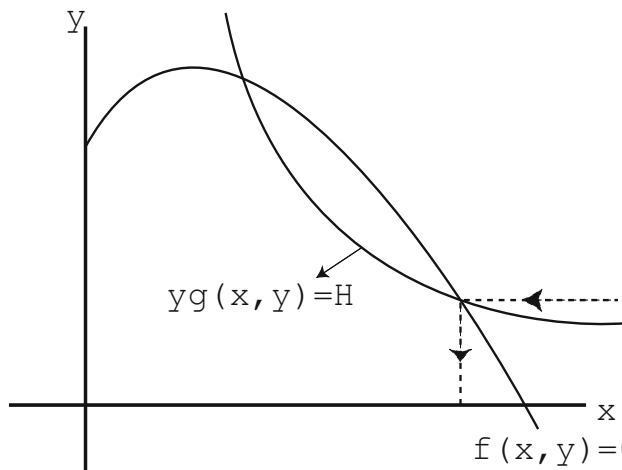


Fig. 6.9 Stable and unstable separatrices.

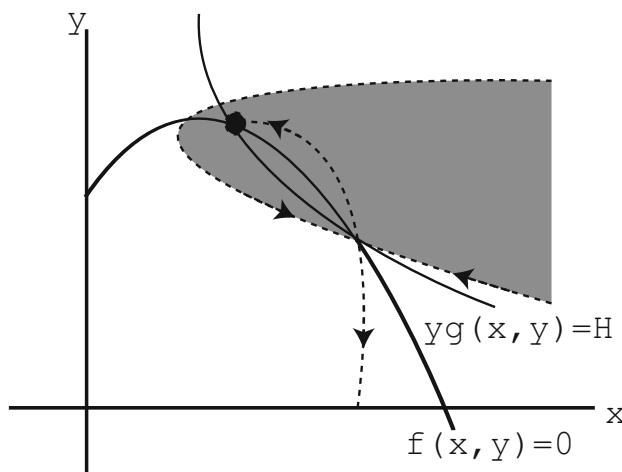


Fig. 6.10 “Good case”.

ble equilibrium $(x_\infty(H), y_\infty(H))$ and the coexistence region is the interior of this periodic orbit (Figure 6.11).

It is possible to have a transition from the “good” to the “bad” case as H increases. If this occurs, then at the transition there must be a homoclinic orbit from saddle point to saddle point (Figure 6.12).

For $H = 0$, the system must be in the “good” case since the unstable separatrix at $(K, 0)$ cannot return to the x -axis and must therefore go to (x_∞, y_∞) or to a limit cycle around (x_∞, y_∞) . Since the orbits depend continuously on the system, the system must be in the “good” case if H is sufficiently small.

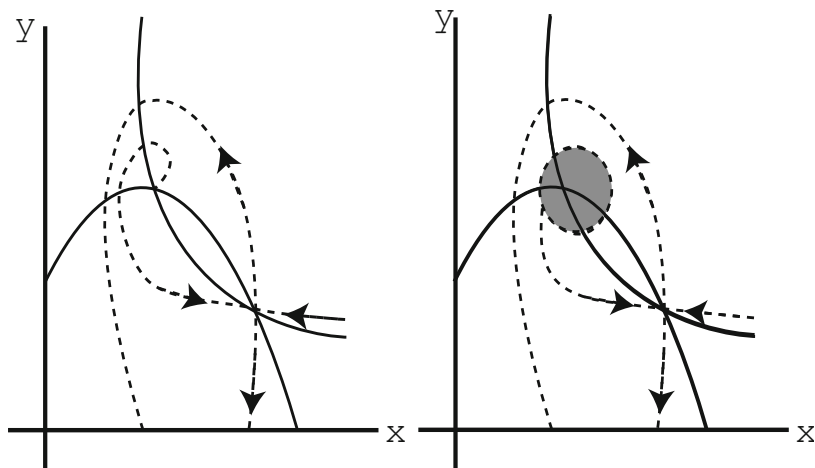


Fig. 6.11 “Bad case”.

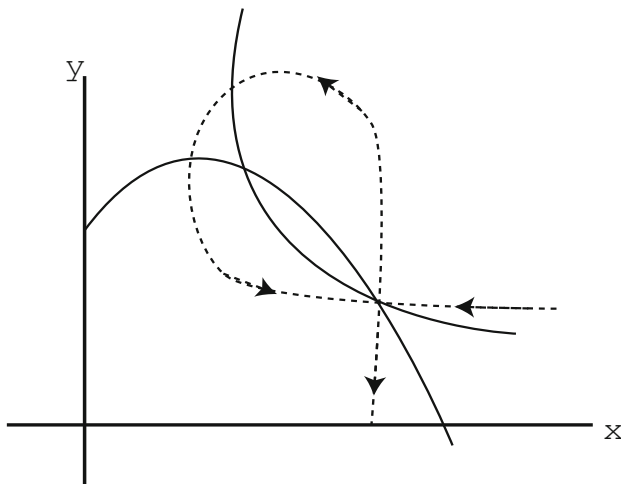


Fig. 6.12 Homoclinic orbit.

For H close to H_c we can decide whether the system is in the “good” or “bad” case by looking at the local stability of the equilibrium $(x_\infty(H), y_\infty(H))$, as is shown by the following theorem.

Theorem 6.1. Suppose that $\text{tr}A(x_\infty(H_c), y_\infty(H_c)) \neq 0$. Then for $H < H_c$ but H sufficiently close to H_c , the system is in the “good” case if $(x_\infty(H_c), y_\infty(H_c))$ is asymptotically stable, that is, if $\text{tr}A(x_\infty(H_c), y_\infty(H_c)) < 0$, and it is in the “bad” case if $(x_\infty(H_c), y_\infty(H_c))$ is unstable, that is, if $\text{tr}A(x_\infty(H_c), y_\infty(H_c)) > 0$.

Proof. The proof depends on the fact that for H close to H_c there must be a (full) orbit connecting $(x_\infty(H), y_\infty(H))$ and $(\xi(H), \eta(H))$ as t runs from $-\infty$ to $+\infty$. This implies that there cannot be a closed orbit around $(x_\infty(H), y_\infty(H))$ and thus rules out the “good” unstable and “bad” asymptotically stable cases.

The details of the proof of the existence of a connecting orbit may be omitted without serious loss. To prove the existence of a connecting orbit, we proceed as follows. For $H < H_c$ but H sufficiently close to H_c , the continuity of $\text{tr}A(x_0, y_0)$ in (x_0, y_0) and of $(x_\infty(H), y_\infty(H))$ in H implies that $\text{tr}A(x_\infty(H), y_\infty(H)) \neq 0$. Again by continuity, there is a neighborhood of $(x_\infty(H), y_\infty(H))$ in which $\text{tr}A(x, y) \neq 0$. We take a point C on $f(x, y) = 0$ above $(x_\infty(H), y_\infty(H))$ but close to $(x_\infty(H), y_\infty(H))$ and construct the orbit of $x' = xf(x, y), y' = yg(x, y) - H$ backward in time until it meets $yg(x, y) = H$ in a point B . We may assume that B is close to $(x_\infty(H), y_\infty(H))$, for if it is not we can take a point B' on $yg(x, y) = H$ close to $(x_\infty(H), y_\infty(H))$ and construct the orbit forward in time until it meets $f(x, y) = 0$ closer to $(x_\infty(H), y_\infty(H))$ than C , and use B' and C' in place of B and C . Now continue the orbit through B backward in time to a point A below and to the right of the saddle point $(\xi(H), \eta(H))$ and above the curve $yg(x, y) = H$. This is possible unless the orbit meets $f(x, y) = 0$ between $(x_\infty(H), y_\infty(H))$ and $(\xi(H), \eta(H))$, in which case we choose H larger so that $(\xi(H), \eta(H))$ moves up to the point on $f(x, y) = 0$ where the orbit meets it; increasing H reduces y' , and since $x' < 0$, this flattens the orbit and moves the intersection of the orbit with $f(x, y) = 0$ down below $(\xi(H), \eta(H))$. Thus we may assume that A is below and to the right of the saddle point. By an argument similar to that used to show that B may be assumed close to $(x_\infty(H), y_\infty(H))$, we may assume that A is close to $(\xi(H), \eta(H))$. Now draw a vertical line through C and a horizontal line through A , meeting in D . For H sufficiently close to H_c , the points $(x_\infty(H), y_\infty(H))$ and $(\xi(H), \eta(H))$ are close together, and we may assume that in the region \mathcal{N} bounded by the arc ABC , the vertical line segment CD , and the horizontal line segment DA , we have $\text{tr}A(x, y) \neq 0$. By the Bendixson criterion there can be no closed orbit contained in \mathcal{N} , and no homoclinic orbit from the saddle point $(\xi(H), \eta(H))$ contained in \mathcal{N} (Figure 6.13).

By the Poincaré–Bendixson theorem, a nonconstant orbit that remains in \mathcal{N} for $-\infty < t < \infty$ must have the equilibrium $(x_\infty(H), y_\infty(H))$ and the saddle point $(\xi(H), \eta(H))$ as its limiting sets for $t \rightarrow \pm\infty$ and must therefore be a connecting orbit. Thus if there is no connecting orbit, then every orbit starting in \mathcal{N} must leave \mathcal{N} . We know that one of the unstable separatrices at the saddle point leaves \mathcal{N} (downward). If the other one also leaves \mathcal{N} as $t \rightarrow \infty$, it must do so across the line segment DA . But then there is an asymptotically stable separatrix at the saddle point between these two unstable separatrices that cannot leave \mathcal{N} as $t \rightarrow -\infty$ because it could leave \mathcal{N} only across the line segment DA and such a crossing would be in the wrong direction. Thus there must be a separatrix that remains in \mathcal{N} for $-\infty < t < \infty$, and this is the desired connecting orbit. We have now established Theorem 6.1.

The value H_c is the maximum harvest rate for which coexistence is possible. However, if the system is in the “bad” case (unstable) at H_c , there must in fact be a value $H^* < H_c$ such that coexistence is possible only for $H \leq H^*$. Then H^* is the true maximum harvest rate. The value H_c may be calculated by looking for equilibria,

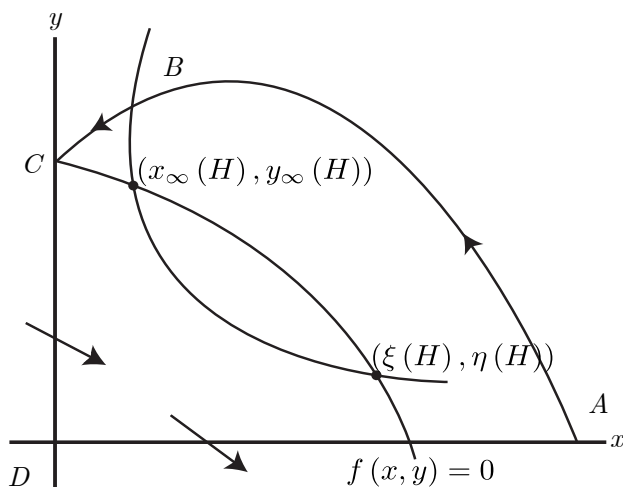


Fig. 6.13 No homoclinic orbit.

but to find H^* it is necessary to compute orbits near the saddle point and determine whether as $t \rightarrow \infty$ they are unbounded (“good”) or go to $(x_\infty(H), y_\infty(H))$ (“bad”).

Example 1: Consider the model of Rosenzweig–MacArthur type

$$x' = rx \left(1 - \frac{x}{K}\right) - \frac{cxy}{x+A}, \quad (6.2)$$

$$y' = sy \left(\frac{x}{x+A} - \frac{J}{J+A}\right) - H = \frac{sA}{J+A} \cdot \frac{x-J}{x+A} - H, \quad (6.3)$$

in which the prey satisfies a logistic equation in the absence of predators and $\frac{cx}{x+A}$ is the *predator functional response*—the rate at which predators remove prey. Here, A represents a prey population level at which the predators’ attack capability begins to saturate, K is the equilibrium population size for prey in the absence of predators, and J is the prey population size in equilibrium with predators without harvesting. For this model we may calculate that

$$x_\infty(H) = \frac{1}{2} \left((K+J) - \sqrt{(K-J)^2 - \frac{4K(J+A)}{rsA} H} \right),$$

$$y_\infty(H) = \frac{r}{K} (x_\infty(H) + A) (K - x_\infty(H)).$$

In particular, $x_\infty(0) = J$ and $x_\infty(H_c) = \frac{1}{2}(K+J)$, with $H_c = \frac{rsA(K-J)^2}{4K(J+A)}$. An equilibrium at $(x_\infty(H), y_\infty(H))$ is asymptotically stable if and only if

$$sAK(x_\infty(H) - J) < rx_\infty(H)(J+A)(2x_\infty(H) + A - K).$$

For example, if $A = 10$, $J = 20$, $K = 40$, $c = 1$, $r = 1$, $s = 1$, we have $H_c = 0.833$. Calculation shows that the equilibrium is asymptotically stable for $0 \leq H < H_c$, and numerical simulation of orbits shows that the system is in the “good” case for $0 \leq H < H_c$.

If $A = 10$, $J = 20$, $K = 40$, $c = 1$, $r = 1$, $s = 7$, we have $H_c = 5.833$. Calculation shows that the equilibrium is asymptotically stable for $0 \leq H < 5.185$, and simulation shows a transition from “good” to “bad” case at $H = 5.185$. Thus $H^* = 5.185$, the effective maximum harvest.

If $A = 10$, $J = 20$, $K = 45$, $c = 1$, $r = 1$, $s = 5$, we have $H_c = 5.787$. Calculation shows that the equilibrium is asymptotically stable if $0 \leq H < 3.704$ or $5.556 < H < H_c$. Thus if $3.704 < H < 5.556$ there is an asymptotically stable limit cycle; otherwise, there is an asymptotically stable equilibrium.

If $A = 10$, $J = 20$, $K = 60$, $c = 1$, $r = 2$, $s = 1$, we have $H_c = 4.444$. Calculation shows that the equilibrium is unstable for $0 \leq H < 2.322$, and asymptotically stable for $2.322 < H < H_c$. Thus there is an asymptotically stable limit cycle for small H that disappears when $H = 2.322$ and is replaced by an asymptotically stable equilibrium.

A final example is $A = 10$, $J = 20$, $K = 60$, $c = 1$, $r = 1$, $s = 4$, for which $H_c = 8.889$. Calculation shows that the equilibrium is unstable for all H and simulation indicates a transition from “good” to “bad” cases for $H = 0.785$. Thus there is an asymptotically stable limit cycle for $0 \leq H < 0.785$ and predator extinction for all harvest rates greater than $H^* = 0.785$. Observe that the true maximum harvest rate is much less than the critical harvest rate H_c , which can be calculated from equilibrium analysis.

Exercises

In each of Exercises 1–5, find the critical harvest rate of the system (6.2) with the given parameter values, find the range of harvest rates for which there is an asymptotically stable equilibrium, and carry out numerical simulations using a computer algebra system to describe the behavior of the system.

1. $A = 10$, $J = 20$, $K = 40$, $c = 1$, $r = 1$, $s = 1$.
2. $A = 10$, $J = 20$, $K = 40$, $c = 1$, $r = 1$, $s = 7$.
3. $A = 10$, $J = 20$, $K = 45$, $c = 1$, $r = 1$, $s = 5$.
4. $A = 10$, $J = 20$, $K = 60$, $c = 1$, $r = 2$, $s = 1$.
5. $A = 10$, $J = 20$, $K = 60$, $c = 1$, $r = 1$, $s = 4$.

6.3 Intermittent Harvesting of Predator–Prey Systems

The harvesting of predators in a predator–prey system at a rate u that may depend on time is modeled by the system

$$x' = xf(x, y), \quad y' = yg(x, y) - u(t).$$

Let us assume that the harvest rate is constrained by a condition $0 \leq u(t) \leq H$. Our goal is to determine a harvesting policy, or choice of u as a function of t , that will maximize the long-term average yield from harvesting,

$$Y = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t u(\tau) d\tau,$$

but we will also require that u be chosen in such a way as to avoid the collapse of the system by demanding that $x(t) > 0$, $y(t) > 0$ for $0 \leq t < \infty$.

The problem that we have formulated is an *optimal control* problem with an objective function $Y[u]$ that is linear in the control variable. It is an immediate consequence of the general theory of , specifically the Pontryagin maximum principle [Pontryagin, et al, (1962)] that the optimal control must be a “bang-bang” control, that is, a control that switches between the values $u = 0$ and $u = H$, possibly combined with some time intervals on which the control is “singular.”

If an optimal control makes only a finite number of switches between the values $u = 0$ and $u = H$, then we must have $u = H$ for all large t , and this implies that the yield Y is equal to H . As we have seen in the preceding section, there is a value $H_c = \max_{f(x,y)=0} yg(x, y)$ with the property that if $H > H_c$, then every solution of $x' = xf(x, y)$, $y' = yg(x, y) - H$ reaches the x -axis in finite time. Thus an optimal control can make a finite number of switches only if $H \leq H_c$. Even for $H < H_c$, the system may collapse if it is in the “bad” case, which occurs if its equilibrium is unstable as $H \rightarrow H_c$, and then there is a value $H^* < H_c$ such that the system actually collapses for $H > H^*$. If the equilibrium is asymptotically stable as $H \rightarrow H_c$ then for every $H < H_c$ it is possible to bring the solution of $x' = xf(x, y)$, $y' = yg(x, y) - u(t)$ into the set of initial conditions for which the solution of $x' = xf(x, y)$, $y' = yg(x, y) - H$ tends to the equilibrium by making a finite number of switches between $u = 0$ and $u = H$. Thus for such a system we can obtain a yield H for every $H \leq H_c$ with a finite number of switches.

In considering controls with an infinite number of switches between $u = 0$ and $u = H$, it would be desirable for practical applications to have a reasonably simple criterion for determining switching points. One very simple control strategy would be to assign upper and lower thresholds y_U and y_L for predator population size, start the system at $t = 0$ with $y = y_L$ and $u = 0$, and switch from $u = 0$ to $u = H$ whenever y increases through y_U and switch from $u = H$ to $u = 0$ whenever y decreases through y_L . This procedure is safe, in the sense that it prevents predator extinction. Numerical simulations indicate that it is always possible to choose the parameters H , y_U , and y_L so as to make the yields Y arbitrarily close to H_c , even for systems that are unstable at H_c and therefore collapse for smaller yields under constant yield harvesting. For such systems there is a tendency for the orbit to approach a periodic orbit with two switch points per cycle (Figure 6.14).

There may be an interval of values of H for given y_U and y_L on which the behavior is more complicated, with period-doubling indicated by periodic orbits having four or eight switch points per cycle, and apparent chaotic behavior (Figure 6.15).

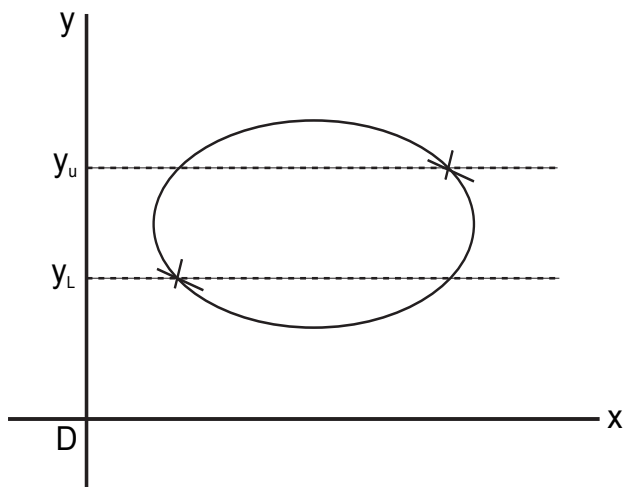


Fig. 6.14 Periodic orbit with switching.

Such behavior appears to occur only for $H < H_c$ and with yields that are not close to optimal.

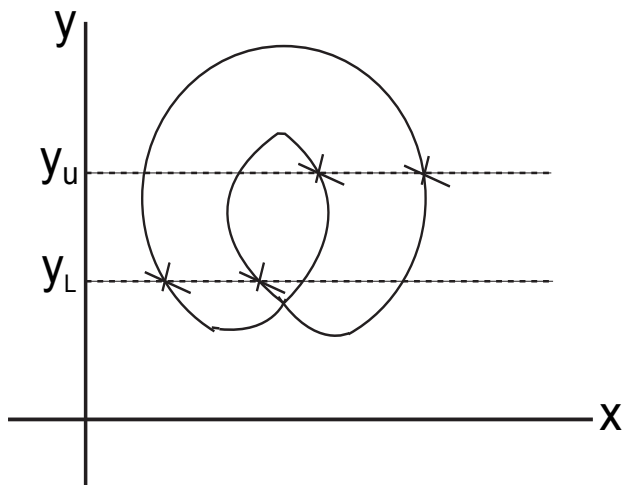


Fig. 6.15 Multiple switch points.

It is possible to prove that if the functions $f(x, y)$ and $g(x, y)$ satisfy some additional concavity conditions (always satisfied by Rosenzweig–MacArthur models and by Kolmogorov models with $g(x, y)$ independent of y and $f(x, y)$ linear in y), then if the control that maximizes the yield gives either a periodic solution of the controlled system or a solution that tends to an equilibrium, then the yield cannot be

greater than H_c for any H . If $H < H_c$, it is clear that the yield cannot be greater than H_c . For $H > H_c$ it appears plausible that optimal controls lead to periodic orbits and therefore that the yield cannot be greater than H_c .

If no bound $u(t) \leq H$ is imposed on the harvest rate, then the analogue of a bang-bang control that switches between the value $u = 0$ and the maximum permissible value of u would be an *instantaneous harvest*. We may model this procedure by setting an upper threshold y_U and a lower threshold y_L and considering only the unharvested system $x' = xf(x, y)$, $y' = yg(x, y)$; whenever $y(t)$ reaches y_U we remove $(y_U - y_L)$ predators to make $y(t) = y_L$ and restart the system with the same x -value.

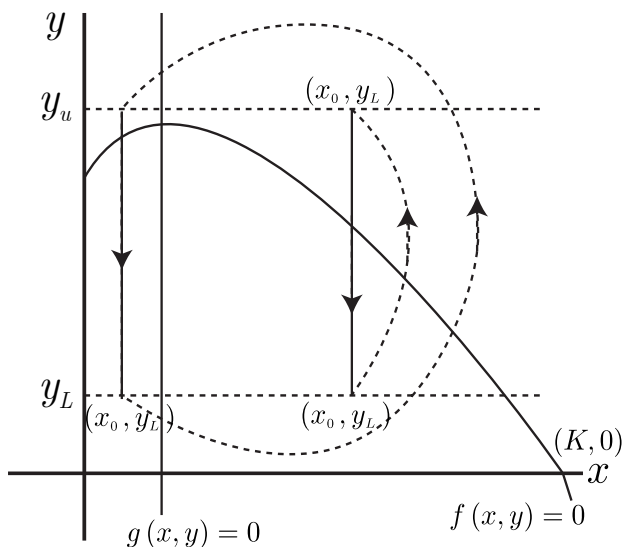
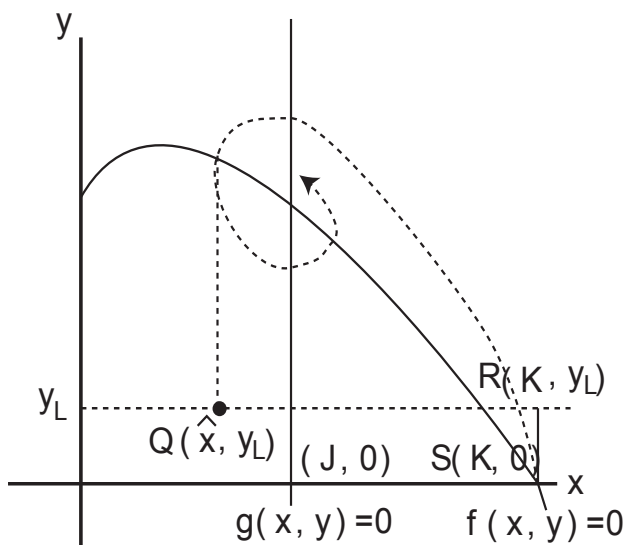
Another instantaneous harvesting policy that occurs more commonly in practice is to set a fixed “closed season” of length T at the end of which the predator population is harvested instantaneously down to a set escapement level y_L . These two policies are equivalent mathematically; the only difference is that in the first, y_U and y_L are assigned and T is to be determined, while in the second, y_L and T are assigned and y_U is determined from the orbit. Another equivalent variant would assign T and $(y_U - y_L)$ with y_L or y_U to be determined from the orbit. In practice, there may be a very short hunting season and a relatively long closed season, and instantaneous harvesting approximates this procedure.

For instantaneous harvesting, the following convergence result holds.

Theorem 6.2. *Let the value y_L be small enough that the line $y = y_L$ and the curve $f(x, y) = 0$ intersect. If $y_U > y_L$ is sufficiently close to y_L , then the instantaneous harvesting procedure converges to a periodic orbit consisting of a solution of $x' = xf(x, y)$, $y' = yg(x, y)$ from a point (x_0, y_L) to a point (x_0, y_U) and a vertical line segment from (x_0, y_U) to (x_0, y_L) (see Figure 6.16).*

Proof. The unharvested system has a unique equilibrium P_∞ in the interior of the first quadrant which may be asymptotically stable or unstable, and a saddle point $S(K, 0)$. There is a full orbit from S to P_∞ or to a limit cycle around P_∞ for $-\infty < t < \infty$. Let \hat{x} be the minimum value of x on this orbit and let Q be the point (\hat{x}, y_L) , R the point (K, y_L) (Figure 6.17).

For any solution $(x(t), y(t))$, if $x(0) < x_L, y(0) = y_L$, then $x(t) > x(0)$ for $0 < t < \infty$, and if $x(0) > K, y(0) = y_L$, then $x(t) < x(0)$ for $0 < t < \infty$. The verification of these facts depends on the observation that no other orbit can cross the orbit from S to P_∞ or a limit cycle around P_∞ . It now follows that for each $T > 0$ there is an x_0 with $\hat{x} < x_0 < K$ such that if $x(0) = x_0, y(0) = y_L$, then $x(T) = x(0) = x_0$. Further, the policy of starting the system at a point (x^*, y_L) at $t = 0$, letting it run to $(x(T), y(T))$, and then harvesting $y(T) - y_L$ predators to bring the system to the point $(x(T), y_L)$ will bring $x(T)$ closer to x_0 than x^* is. Thus repetition of the process must give convergence to a periodic orbit starting at (x_0, y_L) , following the orbit of the system to $(x_0, y(T))$ at time T , and then going vertically to (x_0, y_L) . The procedure of setting $y_U = y(T)$ rather than specifying T gives the same periodic orbit, which then determines T , provided y_U is close enough to y_L that the orbit starting at (x_0, y_L) reaches the horizontal line $y = y_U$. This completes the proof of the theorem.

**Fig. 6.16** Instantaneous harvesting.**Fig. 6.17** Instantaneous harvesting.

The yield Y under instantaneous harvesting is given by $Y = \frac{y_U - y_L}{T}$. Since the solution of the controlled system is periodic, it follows from the results cited earlier in this section that this yield is at most H_c . Obviously, it is possible to make the yield arbitrarily close to H_c by making T very small and arranging that the orbit lie very close to $(x_\infty(H_c), y_\infty(H_c))$ so that $y(t)g(x(t), y(t))$ is very close to H_c for all

points of the orbit. Since the yield is the average of $y'(t) = y(t)g(x(t), y(t))$ over $0 \leq t \leq T$, this yield is close to H_c . Of more interest is the fact that numerical simulations indicate that yields arbitrarily close to H_c can always be obtained for a range of values of T not confined to very small values. This is true in the “bad” case, in which such yields cannot be achieved by constant-yield harvesting.

The material in this section is from [Brauer & Soudack (1979)], where additional discussion may be found.

Exercises

1. Consider the system

$$\begin{aligned}x' &= x\left(1 - \frac{x}{60}\right) - \frac{xy}{x+10}, \\y' &= 4y\left(\frac{x}{x+10} - \frac{2}{3}\right) = \frac{4y(x-20)}{3(x+10)}.\end{aligned}$$

We will choose $y_L = 5$, $T = 4$. Use a computer algebra system to simulate this system with $x(0) = 40$, $y(0) = 5$ for $0 \leq t \leq 4$. After recording the instantaneous harvest $y(4) - 5$, restart with initial values $(x(4), 5)$ at $t = 4$ and run for $4 \leq t \leq 8$. Repeat this process several times and estimate the average instantaneous harvest.

6.4 Some Economic Aspects of Harvesting

The incorporation of economic considerations into resource-harvesting models leads to a rather new subject, bioeconomics. We shall only scratch the surface of this subject, considering some simple fisheries models and introducing the idea of an optimal control problem and illustrating the use of the maximum principle of Pontryagin (1962). For further exploration of bioeconomics, we refer the reader to the standard text reference by Clark (1990).

Let us consider a model for an open-access fishery, that is, a fishery in which the exploitation of resources is completely uncontrolled. Our model will be an idealization, but in practice, fisheries are usually relatively unregulated. However, the abuses of open access exploitation have led to the imposition of regulations in some instances.

We consider a fish population that is governed by a first-order differential equation,

$$y' = F(y),$$

and which is subjected to constant-effort harvesting, so that the model becomes

$$y' = F(y) - Ey. \quad (6.4)$$

The *yield*, or harvest in unit time, is

$$Y(E) = Ey.$$

Corresponding to an effort E there is an equilibrium population size $y_\infty = y_\infty(E)$ given by

$$F(y_\infty) = Ey_\infty$$

for $0 \leq E \leq E^* = F'(0)$. As we have seen in Section 1.5, we may plot $Y(E)$ as a function of E to obtain the *yield-effort curve* (Figure 6.18).

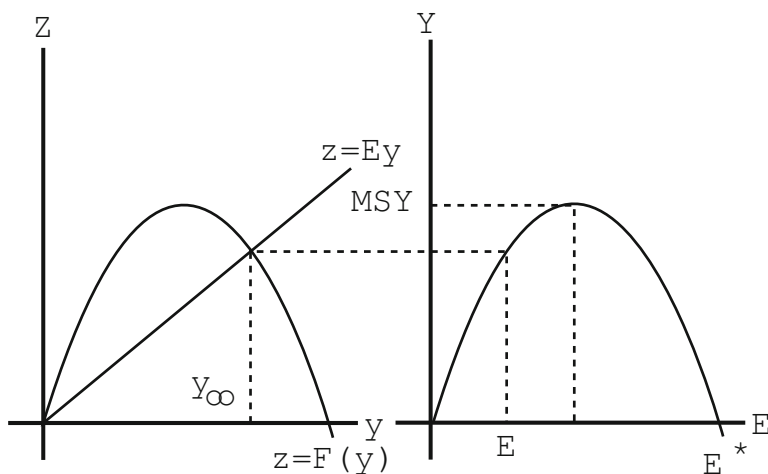


Fig. 6.18 The yield-effort curve.

In Figure 6.18, the yield-effort curve is shown for a compensation model, rising continuously to a maximum, called the *maximum sustainable yield* (MSY) and then decreasing continuously to zero for $E = E^*$. Recall from Section 1.5, however, that for a depensation model there may be a critical effort E^* at which the yield drops to zero discontinuously.

Example 1: For logistic growth, $F(y) = ry \left(1 - \frac{y}{K}\right)$ we have

$$\begin{aligned} F(y_\infty(E)) &= ry_\infty(E) \left[1 - \frac{y_\infty(E)}{K}\right] = Ey_\infty(E), \\ r \left[1 - \frac{y_\infty(E)}{K}\right] &= E, \\ y_\infty(E) &= K \left(1 - \frac{E}{r}\right). \end{aligned}$$

Thus

$$Y(E) = Ey_{\infty}(E) = KE \left(1 - \frac{E}{r}\right).$$

The maximum sustainable yield (MSY) is obtained by solving $\frac{dY}{dE} = 0$, giving $E_{MSY} = \frac{r}{2}$, $y(E_{MSY}) = \frac{K}{2}$, and $Y(E_{MSY}) = \frac{rK}{4}$.

Example 2: For the Ricker model $F(y) = rye^{-ay}$, we have

$$re^{-ay_{\infty}(E)} = E,$$

which leads to

$$y_{\infty}(E) = \frac{1}{a} \log \frac{r}{E} \quad \text{and} \quad Y(E) = \frac{E}{a} \log \frac{r}{E}.$$

Then

$$Y'(E) = \frac{E}{a} \log \frac{r}{E} - \frac{1}{a},$$

and maximum sustainable yield is obtained when $E \log \frac{r}{E} = 1$, which gives $Y(E_{MSY}) = \frac{1}{a}$.

To incorporate the economic aspects of a harvesting model we must attach a revenue from the sale of this fish or other item being harvested as well as a cost to carry out the harvesting. The simplest assumptions would be that the *cost* C of fishing is proportional to the effort,

$$C = cE,$$

and that the price does not depend on the supply of fish available to sell, so that the *revenue* R is a constant multiple of the quantity harvested,

$$R = pY(E).$$

This would give a *sustainable economic rent*, meaning a profit per unit time that can be maintained indefinitely, of

$$R - C = rY(E) - cE.$$

More generally, economic models may assume a functional relation, known as a *demand curve*, between the unit selling price p and the quantity Y available for sale. It would be reasonable to assume that p is a nonincreasing function of Y , or equivalently, that Y is a nonincreasing function of p .

A basic (but considerably oversimplified) principle is that in an open access fishery the effort tends to approach an equilibrium effort E_{∞} , called the *bionomic equilibrium effort*, at which the sustainable economic rent is zero. The argument is that at the bionomic equilibrium, $R = C$, and that $R > C$ if $E < E_{\infty}$, $R < C$ if $E > E_{\infty}$. If $E > E_{\infty}$, then $R < C$, and sustainable economic rent is negative. This means that some fisheries are losing money and therefore drop out of the market, thus decreas-

ing the total effort. If $E < E_\infty$, then $R > C$, and fishing is profitable, encouraging new fishers to join the market and increase the total effort. Thus there should be a tendency to approach the bionomic equilibrium, although we should recognize that the delay that represents the time needed to respond to the state of the market could produce oscillations around the equilibrium (Figure 6.19).

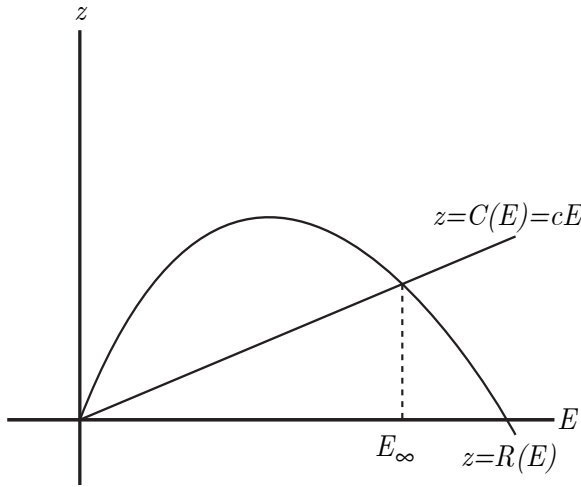


Fig. 6.19 Bionomic equilibrium.

There are many oversimplifications in this bionomic equilibrium principle. For example, the cost of fishing should also include the loss from not undertaking some alternative activity (opportunity cost), which in practice might be the most important component of fishing cost. If opportunity cost is included in the total cost, then the fishers who remain in an open access fishery would tend to be those with the poorest alternate economic opportunities.

Bionomic equilibrium is a situation in which there is *economic overfishing*, because a fishery that would produce positive economic rent if fishing effort were decreased is producing zero economic rent. Another possibility is *biological overfishing*, where $E_\infty > E_{MSY}$ and thus the sustainable yield and equilibrium population size are smaller than they would be if the effort were decreased.

Example 1: Consider the logistic model

$$y' = ry \left(1 - \frac{y}{K} \right) - Ey,$$

with $R = py_\infty(E)$ and $C = cE$. At equilibrium, $y_\infty = \frac{K}{r}(r - E)$, $R = \frac{pKE}{r}(r - E)$, and $C = cE$. Then the bionomic equilibrium E_∞ is determined from

$$\frac{pKE}{r}(r - E) = cE,$$

which gives

$$pK(r - E) = rc \quad \text{and} \quad r - E = \frac{rc}{pK},$$

so that

$$E_\infty = r - \frac{rc}{pK} = r \left(1 - \frac{c}{pK} \right).$$

The corresponding equilibrium population size is

$$y_\infty = K \left(1 - \frac{E_\infty}{r} \right) = K \left[1 - \left(1 - \frac{c}{pK} \right) \right] = \frac{c}{p}.$$

Thus if $\frac{c}{p} > K$, there is no exploitation at the bionomic equilibrium. If $\frac{c}{p} < \frac{K}{2}$, there is biological overfishing at bionomic equilibrium, but so long as $p > 0$, we do not have biological extinction at bionomic equilibrium. The bionomic equilibrium depends upon the cost–price ratio $\frac{c}{p}$ as well on the biological parameters r, K . In fact, under the assumptions of constant unit price p and constant cost c of unit effort, the equilibrium population size at bionomic equilibrium is equal to the cost–price ratio, and the harvesting effort at bionomic equilibrium may be found by simply calculating the effort that leads to this equilibrium population size.

Exercises

1. The model for fisheries

$$\frac{dx}{dt} = a - bx$$

was proposed by Schoener in 1973. Determine the equation of the yield–effort curve and explain its peculiar feature.

2. Find x_{MSY} (density of maximum sustainable yield) and $F(x_{MSY})$ (the maximum sustainable yield) for the Gompertz law of population growth:

$$\frac{dx}{dt} = rx \log \left(\frac{K}{x} \right).$$

Also sketch the yield–effort curve. What is E^* (effort at MSY)?

3. Find x_{\max} (density of maximum sustainable yield) and $F(x_{\max})$ (the maximum sustainable yield) for a population governed by the equation

$$\frac{dx}{dt} = rxe^{-x} - dx$$

with $r > d$.

4. What happens to the yield in a population governed by the equation

$$\frac{dx}{dt} = rxe^{-x} - dx$$

with $r < d$?

5. Show that if a population is governed by the equation

$$\frac{dx}{dt} = r\sqrt{x},$$

there is no maximum sustainable yield.

6.5 Optimization of Harvesting Returns

Consider an open-access fishery that is in bionomic equilibrium with biological overfishing. Let us suppose that we try to regulate this fishery by reducing the harvesting effort in order to increase the yield. The immediate effect of reducing effort is to reduce the yield, but the long-term effect will be to increase the yield by increasing the equilibrium population level. We would like to determine the optimal trade-off between current and future harvesting revenue. In order to do this we must take into account the time value of money because of interest. Since the quantity that we shall study is the present value of all future harvest rents, we shall do this by using a discount rate δ , and since we are studying a continuous model, we shall assume that this discount rate is compounded continuously.

When we speak of a discount rate δ , we mean that the present value of an amount that would have value 1 at a time t units in the future is $(1 - \delta)^t$. Normally the time unit is one year. If the discount is *compounded* K times per year, we use a discount rate of $\frac{\delta}{K}$ per discount period. Then t years would be tK discount periods, and the present value of an amount 1 at a time t units in the future would be $(1 - \frac{\delta}{K})^{tK}$. By continuous compounding, we mean that the present value would be the limit of this quantity as $K \rightarrow \infty$. Because

$$\lim_{K \rightarrow \infty} \left(1 - \frac{\delta}{K}\right)^{Kt} = e^{-\delta t},$$

the effect of a continuously compounded discount rate δ would be to introduce a factor $e^{-\delta t}$ for a rent at time t years in the future.

We now generalize the model (6.4) slightly by assuming that the harvest at time t is proportional to the effort E as before, but is not necessarily proportional to the population size $y(t)$. Thus we write the harvest at time t as

$$h(t) = EG\{y(t)\}, \quad (6.5)$$

where $G(y)$ is a nonnegative, nondecreasing function of y . We continue to assume a constant unit price p and a constant cost c of unit effort. Then the unit harvest cost when the population level is y is a function $c(y)$ given by

$$c(y) = \frac{c}{G(y)}. \quad (6.6)$$

An input of effort $E\Delta t$ from time t to time $t + \Delta t$ produces a harvest $EG\{y(t)\}\Delta t$, which produces a revenue of $pEG\{y(t)\}\Delta t$ at a cost of $cE\Delta t$. Thus the net revenue is

$$pEG\{y(t)\}\Delta t - cE\Delta t = [p - c\{y(t)\}]h(t)\Delta t, \quad (6.7)$$

making use of (6.5) and (6.6). The present value of all future harvesting effort is thus

$$P = \int_0^\infty e^{-\delta t} [p - c\{y(t)\}]h(t)dt. \quad (6.8)$$

Note that we must use a positive discount rate δ , since otherwise the infinite integral in (6.8) would diverge. If we choose to use a *finite time horizon* by considering only the harvesting effort until time T , then we would replace (6.8) by

$$P = \int_0^T e^{-\delta t} [p - c\{y(t)\}]h(t)dt. \quad (6.9)$$

With a finite time horizon we could choose to disregard the time value of money by taking $\delta = 0$ without introducing any problems with divergence of the integral.

The optimal control problem that we wish to solve is to choose the harvest rate function that will maximize the integral (6.8) subject to the *state equation*

$$y' = F(y) - h(t) = F\{y(t)\} - EG\{y(t)\} \quad (6.10)$$

and the constraints $y(t) \geq 0$ (because a realistic model requires a nonnegative population level) and $0 \leq h(t) \leq h_{\max}$ (because a fishery always has a maximum harvesting capability).

The solution of this optimal control problem, which we shall justify later, is that we should first choose the equilibrium population level y^* as the solution of the equation

$$F'(y) - \frac{c'(y)F(y)}{p - c(y)} = \delta,$$

or equivalently,

$$\frac{d}{dy}[F(y)\{p - c(y)\}] = \delta[p - c(y)]. \quad (6.11)$$

We then choose the effort E so that

$$h = EG(y^*) = F(y^*).$$

This still leaves open the question of the optimal way to reach this *optimal equilibrium population* level y^* . The answer to this question is that we should use the harvest rate $h^*(t)$ at time t that drives $y(t)$ to y^* as rapidly as possible,

$$h^*(t) = \begin{cases} h_{\max} & (y > y^*), \\ F(y^*) & (y = y^*), \\ 0 & (y < y^*). \end{cases}$$

This type of control, taking the extreme values 0 and h_{max} and switching when $y = y^*$, is called a *bang-bang control*. It may appear to be an extreme policy in practice, since it involves closing down a fishery completely, possibly for a long period. Such a policy is a consequence of the assumed linear dependence on E . In the next section we shall consider a nonlinear control problem for which the optimal control is less extreme.

At equilibrium of the fishery model (6.10), $F(y) = h$. Thus the sustainable economic rent at y is equal to

$$\rho(y) = [p - c(y)]F(y).$$

With this definition of $\rho(y)$, we may write the optimizing condition (6.11) as

$$\rho'(y) = \delta[p - c(y)]. \quad (6.12)$$

A marginal decrease of Δy in y gives an immediate rent of $[p - c(y)]\Delta y$ and decrease in sustained rent

$$\Delta p \approx \rho'(y)\Delta y$$

with a present value of

$$\int_0^\infty e^{-\delta t} \rho'(y)\Delta y dt = \frac{1}{\delta} \rho'(y)\Delta y.$$

Thus for the optimal solution, the marginal immediate gain is equal to the present value of the marginal future loss.

From (6.6) and the assumption that $G(y)$ is a nondecreasing function of y , it follows that $c(y)$ is nonincreasing, and $p - c(y)$ is nondecreasing. If $\rho'(y)$ is a decreasing function, the equation (6.12) has a unique solution (Figure 6.20).

The significance of $\delta = 0$ is that we should maximize the sustained economic rent, that is, we should maximize ρ by setting $\rho'(y) = 0$. The corresponding population level is y_0 . If we let $\delta \rightarrow \infty$, we set zero value on future revenues and set $p - c(y) = 0$ (no sustainable rent). The corresponding population level is y_∞ . Here y_0 is defined by $\rho'(y) = 0$ and y_∞ by $c(y_\infty) = p$; we have $y_\infty < y^* < y_0$.

The process described above is for a harvest $h(t) = EG\{y(t)\}$, which is not necessarily proportional to population size. In the special case $h(t) = Ey(t)$ that we considered earlier, (6.6) implies $c(y) = \frac{c}{y}$, and this is used in the determination of y^* . Once y^* has been determined, it is easy to calculate $F(y^*)$ and thus to determine the switching point in $h^*(t)$.

Example 1: Consider a logistic model with constant-effort harvest, so that

$$F(y) = ry \left(1 - \frac{y}{K}\right), \quad G(Y) = y, \quad c(y) = \frac{c}{y}.$$

The condition (6.12) becomes

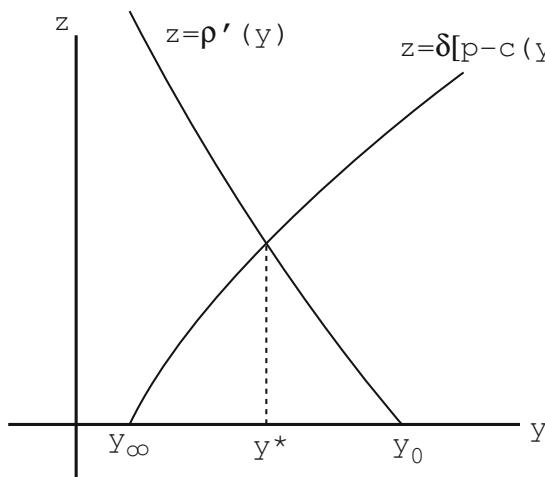


Fig. 6.20 Unique optimal equilibrium.

$$\begin{aligned}
 r \frac{d}{dy} \left[y \left(p - \frac{c}{y} \right) \left(1 - \frac{y}{K} \right) \right] &= \delta \left(p - \frac{c}{y} \right), \\
 r \frac{d}{dy} \left[y \left(p - \frac{c}{y} \right) - \frac{p}{K} y^2 + \frac{c}{K} y \right] &= \delta \left(p - \frac{c}{y} \right), \\
 2rpy^2 + y(\delta pK - rpK - rc) - \delta cK &= 0.
 \end{aligned}$$

This is a quadratic equation having one positive root and one negative root, and we take y^* to be the positive root,

$$y^* = \frac{K}{4} \left[\left(\frac{c}{pK} + 1 - \frac{\delta}{r} \right) + \left\{ \left(\frac{c}{pK} + 1 - \frac{\delta}{r} \right)^2 + \frac{8c\delta}{prK} \right\}^{\frac{1}{2}} \right].$$

For zero discount, $\delta = 0$, we obtain $y_0 = \frac{K}{2} \left(\frac{c}{pK} + 1 \right)$, and since $y_\infty = \frac{c}{p}$, we have

$$y_0 = \frac{1}{2}(y_\infty + K) = \frac{1}{2}y_\infty + y_{MSY} > y_{MSY}.$$

Thus we do not have biological overfishing if $\delta = 0$. However, increasing δ will decrease y^* from y_0 . An extreme case is $y_\infty = 0$, when fishing has zero cost, and then

$$y^* = \frac{K}{2} \left(1 - \frac{\delta}{r} \right) < y_{MSY}.$$

If the bionomic growth ratio $\frac{\delta}{r}$ exceeds 1, then $y^* = 0$, and the optimal harvest policy leads to the most rapid possible extinction of the resource. The interpretation is that if the discount rate δ is greater than the intrinsic growth rate r , the fisher should cash

in the resources as quickly as possible. This optimal fishing policy may of course not be considered ideal by those with concerns other than the economic rent of the fishery.

Exercise

1. For a population governed by the Gompertz law of population growth

$$\frac{dx}{dt} = rx \log\left(\frac{K}{x}\right)$$

find the bioeconomic equilibrium for a constant price p and a cost cE with c constant. Show that there is biological overfishing if $\frac{c}{p} < \frac{K}{e}$, but no biological extinction at bioeconomic equilibrium.

6.6 Justification of the Optimization Result

In the preceding section we stated a result that we used to provide the solution of an optimization problem. In this section we shall justify this result. In fact, we consider the more general

$$J[u] = \int_{t_0}^{t'} [f_0(t, y(t)) + g_0(t, y(t))u(t)] dt, \quad (6.13)$$

subject to a state equation

$$y' = f_1(t, y) + g_1(t, y)u, \quad (6.14)$$

with given values $y(t_0) = y_0$, $y(t_1) = y_1$, and with the *control* $u(t)$ chosen from an *admissible set* of functions that are piecewise continuous and satisfy

$$u_m \leq u(t) \leq u_M.$$

The problem considered in Section 6.6 was the special case

$$\begin{aligned} f_0(t, y) &= 0, & g_0(t, y) &= e^{-\delta t} [p - c(y)], \\ f_1(t, y) &= F(y), & g_1(t, y) &= -1, \\ u(t) &= h(t). \end{aligned}$$

However, in this problem we did not impose a *terminal condition* $y(t_1) = y_1$; the terminal condition is essentially replaced by the requirement that $y(t)$ approach an equilibrium as $t \rightarrow \infty$.

We begin the solution by solving the state equation (6.14) for u and substituting the result into (6.13), obtaining a form

$$\int_{t_0}^{t'} [G(t, y) + H(t, y)y'] dt \quad (6.15)$$

to be maximized under constraints

$$A(y, t) \leq y'(t) \leq B(y, t). \quad (6.16)$$

Theorem 6.3. *Let $y^*(t)$ be the unique solution of*

$$\frac{\partial G}{\partial y} = \frac{\partial H}{\partial t} \quad (6.17)$$

and suppose

$$\begin{aligned} \frac{\partial G}{\partial y} &\leq \frac{\partial H}{\partial t} & \text{if } y < y^*(t) \\ \frac{\partial G}{\partial y} &\geq \frac{\partial H}{\partial t} & \text{if } y > y^*(t). \end{aligned}$$

Then the optimal solution is the “closest possible trajectory to $y^*(t)$ ”, defined as follows:

- (i) If $y_0 > y^*(t_0)$, $y(t)$ uses the fastest possible descent $y' = A(y, t)$, $y(t_0) = y_0$ until $y^*(t)$ is reached.
- (ii) If $y_0 < y^*(t_0)$, $y(t)$ uses the fastest possible ascent $y' = B(y, t)$, $y(t_0) = y_0$ until $y^*(t)$ is reached.

Similarly, if $y_1 \neq y^*(t_1)$, use the same scheme backward from t_1 , until $y^*(t)$ is reached. Between these two points where $y^*(t)$ is reached, say t_a when $y(t)$ starting at t_0 reaches $y^*(t)$ and t_b when $y(t)$ ending at t_1 reaches $y^*(t)$, take $y(t) = y^*(t)$ [Figure 6.21].

Proof. We write the objective functional as a line integral

$$\int_{t_0}^{t_1} [G(t, y) + H(t, y)y'] dt = \int_C (Gdt + Hdy),$$

where C is the curve $y = y(t)$, $t_0 \leq t \leq t_1$. Suppose $y(t)$ is the claimed optimal trajectory and $y_1(t)$ is an arbitrary admissible curve from (t_0, y_0) to (t_1, y_1) . Because of the constraint (6.16) satisfied by $y_1(t)$, we must have $y_1(t) \leq y(t)$ at least for $t_0 \leq t \leq t_a$. Suppose $y_1(t) < y(t)$ on $t_0 < t < t_R$, $y_1(t_R) = y(t_R)$. Then

$$\int_{t_0}^{t_R} [G(t, y) + H(t, y)y'] dt - \int_{t_0}^{t_R} [G(t, y) + H(t, y)y'] dt = \int_{\Gamma} (Gdt + Hdy), \quad (6.18)$$

where Γ is the closed curve PRS as shown in Figure 6.22.

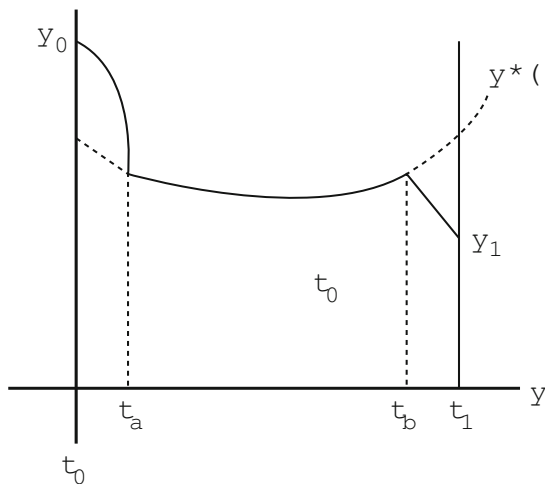


Fig. 6.21

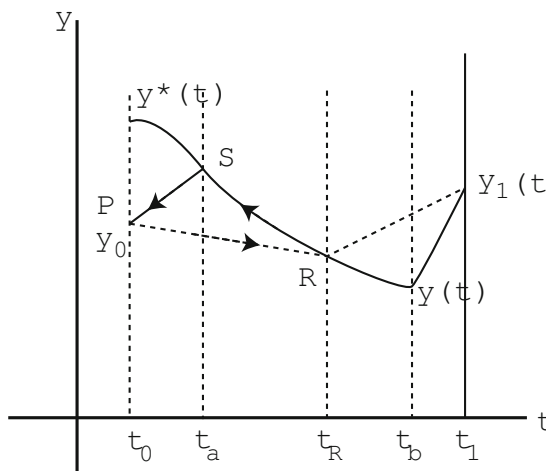


Fig. 6.22

By Green's theorem in the plane, this is equal to

$$\int_A \int \left(\frac{\partial G}{\partial y} - \frac{\partial H}{\partial t} \right) dy dt,$$

where A is the region in the interior of the curve Γ . Since this region lies below $y^*(t)$, $\frac{\partial G}{\partial y} \geq \frac{\partial H}{\partial t}$ in A , the integral (6.18) is nonnegative. This shows that $y(t)$ does better than $y_1(t)$ in optimizing on the interval $t_0 \leq t \leq t_R$, and a similar argument shows the optimality from t_R to t_1 . This completes the proof of the theorem. In fact, we

have established the stronger result that an optimal path must be optimal on every subinterval.

In the problem considered in Section 6.5, we used

$$u = h = F(y) - y',$$

$$J[h] = \int_0^\infty e^{-\delta t} [p - c\{y(t)\}][F\{y(t)\} - y'(t)] dt,$$

so that

$$G(t, y) = e^{-\delta t} [p - c(y)] F(y),$$

$$H(t, y) = -e^{-\delta t} [p - c(y)].$$

Then equation (6.17) becomes

$$e^{-\delta t} \frac{d}{dt} [\{p - c(y)\} F(y)] = \delta e^{-\delta t} [p - c(y)],$$

which is the same as equation (6.4) of Section 6.5, which we solved to determine the final value y^* .

If in the problem of Section 6.5 we had used a finite time horizon T instead of trying to optimize over $0 \leq t < \infty$ and had specified a terminal condition for $y(T)$, the theorem of this section would imply that it would be necessary to leave the path $y^*(t)$ before T and use a bang-bang control. Obviously, $h = 0$ would not be optimal, because $h = h_{\max}$ for t near T would provide a positive contribution to the present value. The “horizon effect” of rapid resource depletion for t near T is an artificial consequence of the finite time horizon. The practical effect of this can be eliminated by taking a sufficiently remote horizon (letting $T \rightarrow \infty$). In nonlinear models, such as will be considered in the next section, horizon effects typically tend to zero exponentially for finite t as $T \rightarrow \infty$ and are effectively eliminated by taking T large.

6.7 A Nonlinear Optimization Problem

In this section we consider a model

$$y' = F(y) - h(t), \quad y(0) = y_0, \quad y(T) = y_T, \quad (6.19)$$

on a finite interval for which the revenue is not a linear function of h . In most real market situations the price p is not a constant but depends on the supply h ; it is reasonable to assume that $p(h)$ is a nonincreasing function of h . The revenue $R(h) = hp(h)$ is then not proportional to h . We will assume that the function $R(h)$ has the properties

$$R(h) > 0, \quad R'(h) \geq 0, \quad R''(h) \leq 0.$$

We will neglect the cost of harvesting, but an approach similar to the one we employ could treat a more general situation that does include harvesting cost. The objective function now is

$$J[h] = \int_0^T e^{-\delta t} R(h) dt. \quad (6.20)$$

The nonlinearity in h implies that the direct line integral approach used in the preceding section cannot be used. In its stead, we use the maximum principle of Pontryagin (1962).

The Pontryagin maximum principle deals with the general control problem of maximizing an objective functional

$$J[u] = \int_0^T g[t, y(t), u(t)] dt, \quad (6.21)$$

where $u(t)$ is chosen from some class of admissible controls u on $0 \leq t \leq T$, with a state equation

$$y'(t) = f[t, y(t), u(t)], \quad y(0) = y_0. \quad (6.22)$$

In addition, the control u must also be such that the response $y(t)$ satisfies the terminal condition

$$y(T) = y_T. \quad (6.23)$$

We define the *Hamiltonian*

$$\mathcal{H}[t, y(t), u(t), \lambda(t)] = g[t, y(t), u(t)] + \lambda(t) f[t, y(t), u(t)], \quad (6.24)$$

where λ is an unknown function called the *adjoint variable*. The maximum principle says that if u is the optimal control and y the corresponding response, then there is an adjoint variable λ such that

$$\frac{d\lambda}{dt} = -\frac{\partial \mathcal{H}}{\partial y} = -\frac{\partial g}{\partial y} - \lambda \frac{\partial f}{\partial y} \quad (6.25)$$

and the optimal control maximizes the value of the Hamiltonian over all admissible controls for each t , that is,

$$\mathcal{H}[t, y(t), u(t), \lambda(t)] = \max_{u \in U} \mathcal{H}[t, y(t), u(t), \lambda(t)]. \quad (6.26)$$

If the optimal control is in the interior of the control interval, that is, if the control constraints are not *binding*, then this maximization statement is

$$\frac{\partial \mathcal{H}}{\partial u} = 0. \quad (6.27)$$

In general, we have three functions $y(t)$, $u(t)$, and $\lambda(t)$ to determine, and three equations, namely the state equation (6.22), the adjoint equation (6.25), and the maximum statement (6.26) or (6.27) can be used for their determination.

For the linear problem solved in Section 6.5 of maximizing

$$\int_0^\infty e^{-\delta t} [p - c(y)] h(t) dt$$

with the state equation

$$y' = F(y) - h(t),$$

we have

$$g(t, y, h) = e^{-\delta t} [p - c(y)] h, \quad f(t, y, h) = F(y) - h(t),$$

in (6.21) and (6.22). Then (6.24) gives the Hamiltonian

$$\mathcal{H}[t, y, h, \lambda] = e^{-\delta t} [p - c(y)] h(t) + \lambda(t) [F(y) - h(t)],$$

and the adjoint equation (6.25) is

$$\frac{d\lambda}{dt} = -c'(y) h(t) e^{-\delta t} + \lambda F'(y). \quad (6.28)$$

If the constraints are not binding, that is, if we can maximize by taking $\frac{\partial \mathcal{H}}{\partial u} = 0$ without being concerned that the resulting u will not be an admissible control, the maximum condition (6.27) gives

$$e^{-\delta t} [p - c(y)] - \delta(t) = 0,$$

or

$$\lambda(t) = e^{-\delta t} [p - c(y)]. \quad (6.29)$$

Substituting the expression (6.29) into the adjoint equation (6.28), we have

$$\begin{aligned} -\delta e^{-\delta t} [p - c(y)] &= c'(y) h e^{-\delta t} - \lambda F'(y) \\ &= c'(y) F(y) e^{-\delta t} - e^{-\delta t} [p - c(y)] F'(y), \end{aligned} \quad (6.30)$$

or

$$F'(y) - \frac{c'(y) F(y)}{p - c(y)} = \delta,$$

which is the solution obtained in Section 6.5. Note that we in effect set an equilibrium value terminal condition when we replaced h by $F(y)$ in (6.30).

In order to examine the situation in which the constraints are binding, we look at the linear problem of maximizing

$$J[u] = \int_{t_0}^{b_1} [f_0\{t, y(t)\} + g_0\{t, y(t)\} u(t)] dt \quad (6.31)$$

subject to a state equation

$$y' = F(y) - u(t), \quad (6.32)$$

with $y(t_0) = y_0$, $y(t_1) = y_1$, and the control constraints

$$u_m \leq u(t) \leq u_M. \quad (6.33)$$

This problem is the special case $f_1(t, y) = F(y)$, $y_1(t, y) = -1$ of problem (6.5), (6.6) in Section 6.6, which was analyzed there by the use of line integrals and Green's theorem. For this problem, the Hamiltonian is

$$\begin{aligned} \mathcal{H}[t, y, u, \lambda] &= f_0(t, y) + g_0(t, y)u + \lambda(t)[F(y) - u] \\ &= f_0(t, y) + \lambda(t)F(y) + u[g_0(t, y) - \lambda(t)]. \end{aligned}$$

We let

$$\sigma(t) = g_0\{t, y(t)\} - \lambda(t). \quad (6.34)$$

Since \mathcal{H} is linear in u , it is clear that in order to maximize \mathcal{H} we must take $u = u_M$ if $\sigma(t) > 0$ and $u = u_m$ if $\sigma(t) < 0$. This is a bang-bang control and σ is called the *switching function*. When the switching function vanishes, \mathcal{H} is independent of u . Thus binding constraints lead to bang-bang controls.

A singular case arises if $\sigma \equiv 0$ on an interval. To determine the corresponding *singular control*, we write $\sigma'(t) = 0$, or

$$\frac{\partial g_0}{\partial t} + \frac{\partial g_0}{\partial y} \frac{dy}{dt} - \frac{d\lambda}{dt} = 0.$$

From the state equation (6.32) and the adjoint equation (6.25) in the form

$$\frac{d\lambda}{dt} = -\frac{\partial}{\partial y}(f_0 + g_0u) - \lambda \frac{\partial}{\partial y}(F - u)$$

we have

$$\frac{\partial g_0}{\partial t} + \frac{\partial g_0}{\partial y}(F - u) + \frac{\partial f_0}{\partial y} + u \frac{\partial g_0}{\partial y} + \lambda \frac{\partial F}{\partial y} = 0.$$

From (6.34) and the switching condition $\sigma(t) = 0$, this is equivalent to

$$-\frac{\partial g_0}{\partial t} = F \frac{\partial g_0}{\partial y} + \frac{\partial f_0}{\partial y} + g_0 \frac{\partial F}{\partial y}. \quad (6.35)$$

In Section 6.6 we used the state equation to write the integral

$$\int_{t_0}^{t_1} [f_0\{t, y(t)\} + g_0\{t, y(t)\}u(t)] dt$$

in the form

$$\int_{t_0}^{t_1} [G(t, y) + H(t, y)y'] dt,$$

and here we have

$$\begin{aligned} f_0(t, y) + g_0(t, y)u &= f_0(t, y) + g_0(t, y)[F(y) - y'] \\ &= f_0(t, y) + g_0(t, y)F(y) - y'g_0(t, y). \end{aligned}$$

Thus

$$\begin{aligned} G(t, y) &= f_0(t, y) + g_0(t, y)F(y), \\ H(t, y) &= g_0(t, y) \end{aligned}$$

and we have

$$\begin{aligned} \frac{\partial G}{\partial t} &= \frac{\partial f_0}{\partial y} + g_0 \frac{\partial F}{\partial y} + F \frac{\partial g_0}{\partial y}, \\ \frac{\partial H}{\partial t} &= -\frac{\partial g_0}{\partial t}. \end{aligned}$$

Now (6.35) is equivalent to

$$\frac{\partial H}{\partial t} = -\frac{\partial G}{\partial y},$$

which was the equation in Section 6.6 that was solved to obtain the function $y^*(t)$. The content of the maximum principle is that for this problem the optimal control is a combination of bang-bang and singular controls.

The optimization problem (6.19) and (6.20) formulated at the beginning of this section is not linear in h , and the method used in Section 6.6 cannot be applied. Here we are forced to use the maximum principle. The Hamiltonian is

$$\mathcal{H}(t, y, u, \lambda) = e^{-\delta t} R(u) + \lambda(t)[F(y) - u].$$

If the control constraints are not binding, the maximum equation is

$$\frac{\partial \mathcal{H}}{\partial u} = e^{-\delta t} R'(u) - \lambda(t) = 0,$$

or

$$\lambda(t) = e^{-\delta t} R'\{u(t)\}. \quad (6.36)$$

The adjoint equation is

$$\frac{dy}{dt} = -\frac{\partial \mathcal{H}}{\partial y} = -\lambda(t)F'(y) = -e^{-\delta t} R'\{u(t)\}F'(y).$$

From (6.36) we obtain

$$\frac{d\lambda}{dt} = e^{-\delta t} \left[R''(u) \frac{du}{dt} - \delta R'(u) \right],$$

and combination with the adjoint equation gives

$$R''(u) \frac{du}{dt} - \delta R'(u) = -R'(u)F'(y).$$

We now have a two-dimensional autonomous system for the response y and the control u , namely

$$y' = F(y) - u, \quad (6.37)$$

$$u' = \frac{R'(u)}{R''(u)} [\delta - F'(y)]. \quad (6.38)$$

It is reasonable to assume that the revenue function $R(u)$ satisfies $R(u) > 0$, $R'(u) \geq 0$, and $R''(u) < 0$.

The y -isocline of the system (6.37) is the curve $u = F(y)$, and the u -isocline is the vertical line $F'(y) = \delta$. There is a unique equilibrium (y^*, u^*) given by

$$u^* = F(y^*), \quad F'(y^*) = \delta,$$

and we may verify that this equilibrium is a saddle point, and that the separatrices and phase portrait are as shown in [Figure 6.23](#)

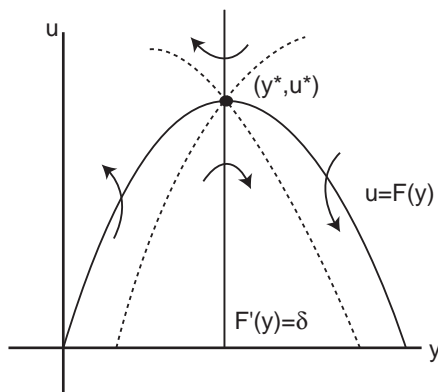


Fig. 6.23 Phase portrait with separatrices.

As in the linear problem of Section 6.5, there is an optimal equilibrium solution $y = y^*$, $u = u^*$. However, the optimal approach to equilibrium now is not a bang-bang control as in the linear case. The optimal trajectory must start on the line $y = y_0$ at time $t = 0$ and terminate on the line $y = y_T$ at time $t = T$. Since (y^*, u^*) is an equilibrium, the velocity along a trajectory near (y^*, u^*) is small. The closer an orbit is to the separatrices, the longer is the time taken to go from y_0 to y_T . Thus for small T , the trajectory will be far from the saddle point, and for large T , the trajectory will be close to the saddle point. For each point on $y = y_0$ there is a trajectory going

to the line $y = y_T$, and the optimal trajectory is the one taking the required time T (Figure 6.24).

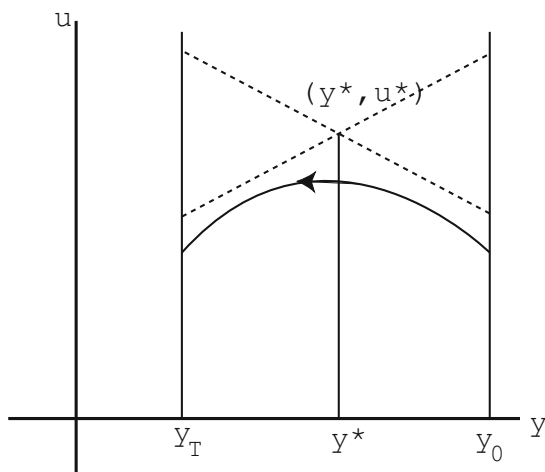


Fig. 6.24 Optimal trajectory.

As the time horizon T goes to infinity, the terminal condition drops out and the optimal trajectory is simply the part of the separatrix from $y = y_0$ to (y^*, u^*) . The optimal approach is more gradual than the bang-bang approach of the linear model. This reflects the dependence of prices on the harvest, which one would expect to cause more gradual changes in the optimal harvest rate.

6.8 Economic Interpretation of the Maximum Principle

We now give an interpretation of the maximum principle in economic terms. The computations that lead to this interpretation actually can be refined to give a proof of the maximum principle under the assumption that all functions involved are sufficiently smooth (twice differentiable). However, in practice the functions are not necessarily this smooth, as we have seen with bang-bang controls, and the general proof of the maximum principle is much more difficult. For a full proof, we refer the reader to the books of Pontryagin *et al.* (1962) and Lee & Markus (1968).

We consider the problem of maximizing

$$J[u] = \int_{t_0}^{t_1} g[s, y(s), u(s)] ds \quad (6.39)$$

subject to a state equation

$$y' = f(t, y, u) \quad (6.40)$$

with initial and terminal conditions

$$y(t_0) = y_0 \quad \text{and} \quad y(t_1) = y_1. \quad (6.41)$$

We view the terminal point (t_1, y_1) as fixed but treat the initial point (t_0, y_0) as variable, assuming that an optimal control $u^*(t)$ exists for all initial points (t_0, y_0) under consideration. For each initial point we define the function

$$w(t_0, y_0) = J[u^*] = \max J[u]. \quad (6.42)$$

This function $w(t_0, y_0)$ in economic problems represents the present value of the stock level y_0 at time t_0 under the assumption that an optimal exploitation policy will be used for $t_0 \leq t \leq t_1$. Thus $w(t_0, y_0)$ is the value of the capital asset y_0 at time t_0 with a given time horizon t_1 . Because the optimal policy must be optimal at each point of the interval $t_0 \leq t \leq t_1$, we must have

$$w(t_0, y_0) = \int_{t_0}^t g[s, y^*(s), u^*(s)] ds + u[t, y^*(t)] \quad (6.43)$$

for every t . Differentiation of (6.42) with respect to t gives

$$\begin{aligned} \frac{d}{dt} w[t, y^*(t)] &= \frac{\partial w}{\partial t} [t, y^*(t)] + \frac{\partial w}{\partial y} [t, y^*(t)] \frac{dy}{dt} \\ &= \frac{\partial w}{\partial t} [t, y^*(t)] + \frac{\partial w}{\partial y} [t, y^*(t)] f[t, y^*(t), u^*(t)] \\ &= -g[t, y^*(t), u^*(t)]. \end{aligned} \quad (6.44)$$

If we define

$$\lambda(t) = \frac{\partial w}{\partial y} [t, y^*(t)] \quad (6.45)$$

and the Hamiltonian

$$\mathcal{H}[t, y, u, \lambda] = g(t, y, u) + \lambda f(t, y, u), \quad (6.46)$$

the maximum principle implies that

$$\mathcal{H}[t, y^*(t), u^*(t), \lambda(t)] = \max_u \mathcal{H}[t, y, u, \lambda]. \quad (6.47)$$

The function $G(t, y, u)$ defined by

$$G(t, y, u) = \frac{\partial w}{\partial t} (t, y) + \frac{\partial w}{\partial y} (t, y) f(t, y, u) + g(t, y, u) \quad (6.48)$$

has the property that

$$G(t, y^*(t), u^*(t)) = 0 \quad (6.49)$$

because of (6.44). It is possible to show, by an argument that involves considering the effect of varying the control u in (6.43), that

$$G(t, y, u) \leq 0 \quad (6.50)$$

for all (t, y, u) . From (6.48) and (6.49) we see that $G(t, y, u)$ is maximized by the choice $y = y^*(t)$, $u = u^*(t)$. In particular, if we fix $y = y^*(t)$, then $G(t, y^*(t), u)$ is maximized by the choice $u = u^*(t)$, and if we fix $u = u^*(t)$ then $G(t, y, u^*(t))$ is maximized by the choice $y = y^*(t)$. As an alternative to this maximization argument, we could use the maximum principle to show that $G(t, y, u)$ is maximized by the choice $u = u^*(t)$, since

$$\begin{aligned} \mathcal{H}[t, y^*(t), u, \lambda] &= g(t, y^*(t), u) + \lambda f(t, y^*(t), u) \\ &= g(t, y^*(t), u) + \frac{\partial w}{\partial y}(t, y^*(t)) f(t, y^*(t), u) \\ &= G(t, y^*(t), u) - \frac{\partial w}{\partial t}(t, y^*(t)) \end{aligned}$$

by (6.45), (6.46), and (6.48); maximization of $G(t, y^*(t), u)$ is equivalent to maximization of $\mathcal{H}[t, y^*(t), u, \lambda]$.

Now we fix $u = u^*(t)$ and use the fact that $G(t, y^*(t), u)$ attains its maximum value when $y = y^*(t)$. If G is differentiable, we must have

$$\frac{\partial G}{\partial y}(t, y^*(t), u^*(t)) = 0. \quad (6.51)$$

Differentiation of (6.48) shows that (6.51) implies

$$\frac{\partial^2 w}{\partial y^2} f + \frac{\partial w}{\partial y} \frac{\partial f}{\partial y} + \frac{\partial^2 w}{\partial y \partial t} + \frac{\partial g}{\partial y} = 0, \quad (6.52)$$

with all functions being evaluated for $y = y^*(t)$, $u = u^*(t)$. Differentiation of (6.45) gives

$$\frac{d\lambda}{dt} = \frac{\partial^2 w}{\partial y^2} \frac{dy}{dt} f + \frac{\partial^2 w}{\partial t \partial y} = \frac{\partial^2 w}{\partial y^2} f + \frac{\partial^2 w}{\partial t \partial y}. \quad (6.53)$$

Combination of (6.52) and (6.53) gives

$$\frac{d\lambda}{dt} + \frac{\partial w}{\partial y} \frac{\partial f}{\partial y} + \frac{\partial g}{\partial y} = 0,$$

or

$$\frac{d\lambda}{dt} = -\lambda \frac{\partial f}{\partial y} - \frac{\partial g}{\partial y} = -\frac{\partial \mathcal{H}}{\partial y}, \quad (6.54)$$

which is the adjoint equation.

The above calculation shows that we may interpret the adjoint variable λ in the maximum principle as being given by the expression (6.45). Since $w(t, y)$ repre-

sents the value of the capital asset y at time t , the adjoint variable λ represents the *marginal value* of this capital asset at time t . Reduction of the capital by one unit would reduce the value by approximately λ ; for this reason $\lambda(t)$ is called the *shadow price* of the asset.

The Hamiltonian is the sum of two terms, each of which can be interpreted as a flow of values. The term $g(t, y, u)$ is the flow of revenues obtained from resource harvesting with the objective functional (6.39). The quantity $f(t, y, u)$ is the flow of investment in capital,

$$\frac{dy}{dt} = f(t, y, u).$$

The value of this investment flow is the flow multiplied by the shadow price $\lambda(t)$. Thus, the Hamiltonian represents the rate of increase of total assets, both accumulated dividends and capital assets. The maximum principle says that an optimal control must maximize the rate of increase of total assets, but we must know the shadow price in order to determine the optimal control. In effect, the maximum principle reduces the optimal control problem to the question of determining the shadow price. Unfortunately, the analytical determination of the shadow price in practice is a formidable problem, and it is often impossible to obtain an exact expression for the shadow price.

The adjoint equation (6.51) says that the rate of depreciation of capital $-\frac{d\lambda}{dt}$ consists of two parts, the marginal flow to accumulated dividends and the marginal flow to capital assets. Along the optimal path, this rate of depreciation must be equal to the sum of these marginal flows, $\frac{\partial g}{\partial y}$ and $\lambda \frac{\partial f}{\partial y}$, respectively. The economic literature usually speaks of a *current shadow price* $\tilde{\lambda}(t) = e^{\delta t} \lambda(t)$, rather than the discounted shadow price $\lambda(t)$ and of the *current-value* Hamiltonian

$$\tilde{\mathcal{H}} = e^{\delta t} \mathcal{H}.$$

In these terms the adjoint equation $\frac{d\lambda}{dt} = -\frac{\partial \mathcal{H}}{\partial y}$ becomes $\frac{d\tilde{\lambda}}{dt} = \delta \tilde{\lambda} - \frac{\partial \tilde{\mathcal{H}}}{\partial y}$.

6.9 Project: A Harvesting Model

Consider a predator–prey system described by the model

$$\begin{aligned} y' &= ry \left(1 - \frac{y}{K}\right) - \frac{ayz}{y+A} \\ z' &= sz \left(1 - \frac{z}{by}\right), \end{aligned}$$

where r, s, a, b, K are positive constants. Suppose the predator (z) species is harvested at a constant rate H .

1. Find the coexistence equilibrium of the unharvested system.

2. Show that under harvesting the equilibrium lies on a parabola opening downwards and passing through the points $(-A, 0)$, $(K, 0)$.
3. Show that for $H > 0$, the equilibrium (y, z) satisfies $z < by$.
4. Is it possible to drive the predator population to zero by harvesting? If so, what is the smallest harvest rate that achieves this?

6.10 Project: Harvesting of Two Species

Consider a predator–prey system given by the pair of equations

$$\begin{aligned}x' &= xf(x, y), \\ y' &= yg(x),\end{aligned}\tag{6.55}$$

with

$$f_x(x, y) < 0, \quad g_x(x, y) > 0, \quad f(K, 0) = 0, \quad g(J, 0) = 0.$$

Assume constant-yield harvesting of both species, to give a model

$$\begin{aligned}x' &= xf(x, y) - H_x, \\ y' &= yg(x) - H_y.\end{aligned}\tag{6.56}$$

Carry out as much as possible of the analogue to the theory of Section 6.2 for harvesting of only the predator species. Instead of a single critical harvest, you should expect to find a curve

$$h[H_x, H_y] = 0$$

of critical harvest rates.

Part III

Structured Population Models

Chapter 7

Models for Populations with Age Structure

7.1 Linear Discrete Models

In the preceding chapters we studied mainly models in which all members were alike, so that birth and death rates depended on total population size. However, we gave a few examples of populations with two classes of members and a birth rate that depended on the size of only one of the two classes, for discrete models in Section 2.6 and for continuous models in Section 3.3. These are examples of *structured* populations. In this chapter we shall study models for populations structured by age. In practice, animal populations are often measured by size with age structure used as an approximation to size structure. The study of age-structured models is considerably simpler than the study of general size-structured models, primarily because age increases linearly with the passage of time while the linkage of size with time may be less predictable. Age-structured models may be either discrete or continuous. We begin with linear models, for which total population size generally either increases or decreases exponentially over time.

First, we will consider a discrete model, for a population that is divided into a finite number of age classes labeled from zero to m , and will describe this population by giving the number of members in each class at a sequence of times. Thus, we let $\rho_{j,n}$ be the number of members in the j th class at the n th time ($j = 0, 1, \dots, m$; $n = 0, 1, 2, \dots$). We will assume that the length of time spent in each age class is the same and that this length of time is equal to the interval between measurements of population size. Thus, $\rho_{j,n+1}$, the number of members in the j th age class at the $(n+1)$ st time, is equal to $\rho_{j-1,n}$, the number of members in the $(j-1)$ st age class at the n th time minus the number of members of this age cohort who die before entering the next age class. We will assume that the probability of survival from one age class to the next depends only on age, that is, that there are constants p_0, p_1, \dots, p_{m-1} such that

$$\rho_{1,n+1} = p_0 \rho_{0,n}, \quad \rho_{2,n+1} = p_1 \rho_{1,n}, \quad \dots, \quad \rho_{m,n+1} = p_{m-1} \rho_{m-1,n},$$

for $n = 0, 1, 2, \dots$. Thus, p_j is the probability that a member of the j th age class survives to the $(j+1)$ st age class ($j = 0, 1, 2, \dots, m-1$); it is assumed that no member survives beyond the m th age class.

All recruitment of new members into the population is assumed to come from a birth process, with fecundity depending only on age. Thus, we assume that there are constants $\beta_0, \beta_1, \dots, \beta_m$ such that

$$\rho_{0,n+1} = \beta_0 \rho_{0,n} + \beta_1 \rho_{1,n} + \dots + \beta_m \rho_{m,n}.$$

The fecundities of the various age classes need not all be different from zero.

We now have a population model described by a system of $(m+1)$ linear difference equations

$$\begin{aligned} \rho_{0,n+1} &= \beta_0 \rho_{0,n} + \beta_1 \rho_{1,n} + \dots + \beta_m \rho_{m,n}, \\ \rho_{1,n+1} &= p_0 \rho_{0,n}, \\ \rho_{2,n+1} &= p_1 \rho_{1,n}, \\ &\vdots \\ \rho_{m,n+1} &= p_{m-1} \rho_{m-1,n}, \end{aligned}$$

to which we must add an initial condition and specify $\rho_{0,0}, \rho_{1,0}, \dots, \rho_{m,0}$. We assume that the survival probabilities p_0, p_1, \dots, p_{m-1} are positive and that the fecundity coefficients $\beta_0, \beta_1, \dots, \beta_m$ are nonnegative and not all zero. This population model is usually known as the *Leslie matrix model*, formulated in 1945, although similar forms appeared in earlier work of H. Bernardelli (1941) and E.G. Lewis (1942). The Leslie model may be written in vector-matrix form. We define the $(m+1)$ -dimensional column vector

$$\rho_n = \begin{bmatrix} \rho_{0,n} \\ \rho_{1,n} \\ \vdots \\ \rho_{m,n} \end{bmatrix}$$

and the Leslie matrix

$$A = \begin{bmatrix} \beta_0 & \beta_1 & \beta_2 & \dots & \beta_{m-1} & \beta_m \\ p_0 & 0 & 0 & \dots & 0 & 0 \\ 0 & p_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & p_{m-1} & 0 \end{bmatrix}.$$

Then the model consists of the vector difference equation

$$\rho_{n+1} = A \rho_n,$$

together with the specification of the initial vector ρ_0 . It is easy to solve this difference equation by induction and obtain the solution

$$\rho_n = A^n \rho_0.$$

This formal solution is of little use without an understanding of the nature of the matrix A^n for a given Leslie matrix A ; such an understanding requires knowledge of the eigenvalues and eigenvectors of the matrix A .

In the simplest case, in which the $(m+1) \times (m+1)$ matrix A has $(m+1)$ linearly independent eigenvectors $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_m$ corresponding to the eigenvalues $\lambda_0, \lambda_1, \dots, \lambda_m$, respectively, we may expand

$$\rho_n = \sum_{j=0}^m c_{j,n} \mathbf{v}_j,$$

because $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_m$ span the space \mathbb{R}^{m+1} of $(m+1)$ -dimensional vectors with real components. Then

$$\rho_{n+1} = \sum_{j=0}^m c_{j,n+1} \mathbf{v}_j = A \rho_n = \sum_{j=0}^m c_{j,n} A \mathbf{v}_j = \sum_{j=0}^m c_{j,n} \lambda_j \mathbf{v}_j,$$

and therefore

$$\sum_{j=0}^m (c_{j,n+1} - c_{j,n} \lambda_j) \mathbf{v}_j = \mathbf{0}.$$

Because the vectors $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_m$ are linearly independent, $c_{j,n+1} = \lambda_j c_{j,n}$ for $j = 0, 1, \dots, m$; $n = 0, 1, 2, \dots$, and this implies $c_{j,n} = \lambda_j^n c_{j,0}$ for $j = 0, 1, \dots, m$; $n = 0, 1, 2, \dots$. Now we may write

$$\rho_n = \sum_{j=0}^m c_{j,0} \lambda_j^n \mathbf{v}_j, \quad n = 0, 1, 2, \dots$$

We now suppose that the eigenvalues can be arranged in decreasing order, or at least that λ_0 is a real simple eigenvalue that is *dominant* in the sense that $|\lambda_j| < \lambda_0$ for $j = 1, 2, \dots, m$. Then we may write

$$\rho_n = c_{0,0} \lambda_0^n \mathbf{v}_0 + \mathbf{u}_n,$$

where $\lambda_0^{-n} \mathbf{u}_n \rightarrow \mathbf{0}$ as $n \rightarrow \infty$. In other words, ρ_n is λ_0^n multiplied by an eigenvector of A corresponding to λ_0 plus terms that are negligible compared to this term as $n \rightarrow \infty$. We define

$$P_n = \sum_{j=0}^m \rho_{j,n} \quad (\text{total population at time } n),$$

$$B_n = \sum_{j=0}^m \beta_j \rho_{j,n} = \rho_{0,n+1} \quad (\text{births into population at time } n).$$

Then we have

$$P_n = c_{0,0}\lambda_0^n \sum_{j=0}^m v_{0,j} + \sum_{j=0}^m u_{n,j},$$

and

$$\frac{\rho_n}{P_n} = \frac{c_{0,0}\mathbf{v}_0 + \lambda_0^{-n}\mathbf{u}_n}{c_{0,0}\sum_{j=0}^m v_{0,j} + \lambda_0^{-n}\sum_{j=0}^m u_{n,j}}.$$

Because $\lambda_0^{-n}\mathbf{u}_n \rightarrow 0$ as $n \rightarrow \infty$, we have

$$\frac{\rho_n}{P_n} \sim \frac{\mathbf{v}_0}{\sum_{j=0}^m v_{0,j}}.$$

Thus, the fraction of the population in each age class approaches a limit, and these limits are proportional to the components of the eigenvector \mathbf{v}_0 . Thus, there is a *stable age distribution*.

If the matrix A does not have $(m+1)$ linearly independent eigenvectors, the above reasoning must be modified, but we can still show that if there is a real simple dominant eigenvalue λ_0 with corresponding eigenvector \mathbf{v}_0 , then

$$\rho_n = c\lambda_0^n \mathbf{v}_0 + \mathbf{u}_n$$

with $\lim_{n \rightarrow \infty} \lambda_0^{-n}\mathbf{u}_n = \mathbf{0}$. This suffices to show that there is a stable age distribution.

Of course, the result makes no sense biologically unless every component of the eigenvector \mathbf{v}_0 is positive.

By examining the particular form of the Leslie matrix, we may show that the characteristic equation is

$$\lambda^{m+1} - \beta_0\pi_0\lambda^m - \beta_1\pi_1\lambda^{m-1} - \dots - \beta_m\pi_m = 0,$$

where $\pi_0 = 1, \pi_1 = p_0, \pi_2 = p_0p_1, \dots, \pi_m = p_0p_1 \dots p_{m-1}$. Thus, $\pi_j > 0$ is the probability of survival from birth to entry into the j th age class. Division by λ^{m+1} gives the equivalent characteristic equation

$$\sum_{j=0}^m \beta_j \pi_j \lambda^{-(j+1)} = 1.$$

Because $\sum_{j=0}^m \beta_j \pi_j \lambda^{-(j+1)}$ is unbounded as $\lambda \rightarrow 0^+$, tends to zero as $\lambda \rightarrow \infty$, and is monotone decreasing, there is a unique positive real root λ_0 . If \mathbf{v}_0 is the corresponding eigenvector, with components (v_0, v_1, \dots, v_m) , we may again use the form of the Leslie matrix to write the equations

$$\beta_0 v_0 + \beta_1 v_1 + \dots + \beta_m v_m = \lambda_0 v_0,$$

$$p_0 v_0 = \lambda_0 v_1,$$

$$p_1 v_1 = \lambda_0 v_2,$$

$$p_{m-1} v_{m-1} = \lambda_0 v_m.$$

If we take $v_0 = 1$, then from these equations we obtain

$$\begin{aligned} v_1 &= \frac{p_0}{\lambda_0} = \frac{\pi_1}{\lambda_0}, \\ v_2 &= \frac{p_1 v_1}{\lambda_0} = \frac{p_0 p_1}{\lambda_0^2} = \frac{\pi_2}{\lambda_0^2}, \\ &\vdots \\ v_m &= \frac{p_{m-1} v_{m-1}}{\lambda_0} = \frac{p_0 \cdots p_{m-1}}{\lambda_0^m} = \frac{\pi_m}{\lambda_0^m}. \end{aligned}$$

Thus, the components of \mathbf{v}_0 are proportional to π_j / λ_0^j ($j = 0, 1, \dots, m$), and these are the proportions in the stable age distribution, obviously positive. If $\lambda_0 > 1$, or equivalently if $\sum_{j=0}^m \beta_j \pi_j > 1$ (because then the value of λ that makes $\sum_{j=0}^m \pi_j \beta_j \lambda_j^{-(j+1)} = 1$ must be greater than one), the total population size grows like λ_0^n . If $\sum_{j=0}^m \beta_j \pi_j < 1$, so that $\lambda_0 < 1$, the total population size decreases like λ_0^n . If $\sum_{j=0}^m \beta_j \pi_j = 1$, so that $\lambda_0 = 1$, the total population size remains constant. In each case the proportion of the population in each age class tends to a limit.

While we have shown that for the Leslie matrix model there is a unique real simple eigenvalue λ_0 with a corresponding eigenvector whose components are all positive, we have *not* shown that this eigenvalue is dominant. In fact, additional conditions are needed to ensure the dominance of the eigenvalue λ_0 . The following algebraic result, known as the *Perron–Frobenius* theorem, is of some assistance.

Theorem 7.1. *Let A be a matrix all of whose elements are non-negative, and such that for some positive integer k every element of the matrix A^k is positive. Then A has a simple positive eigenvalue λ_0 with a corresponding eigenvector having all components positive, and $|\lambda_j| < \lambda_0$ for every other eigenvalue λ_j .*

The conclusions of the classical Perron–Frobenius theorem are exactly what we need, but the hypothesis is unsuitable for our purposes. Fortunately, it is possible to show that for a Leslie matrix model with two consecutive β_j different from zero, some power of the matrix has all elements positive, and thus the conclusion of the Perron–Frobenius theorem holds. Under these conditions we have as $n \rightarrow \infty$,

$$\begin{aligned} \rho_n &\sim c \lambda_0^n \mathbf{v}_0, \\ \rho_n &\sim c \sum_{j=0}^m \lambda_0^n \frac{\pi_j}{\lambda_0^j} = c \lambda_0^n \sum_{j=0}^m \frac{\pi_j}{\lambda_0^j}, \end{aligned}$$

so that $P_n \sim c \lambda_0^n$ if $\lambda_0 > 1$, but $P_n \sim c \pi_m \lambda_0^{n-m}$ if $\lambda_0 < 1$ and

$$B_n = c \sum_{j=0}^m \lambda_0^n \frac{\pi_j}{\lambda_0^j} \sim c \sum_{j=0}^m \beta_j \lambda_0^n \frac{\pi_j}{\lambda_0^j} = c \lambda_0^{n+1} \sum_{j=0}^m \beta_j \frac{\pi_j}{\lambda_0^{j+1}} = c_0 \lambda_0^{n+1},$$

because $\sum_{j=0}^m \beta_j \frac{\pi_j}{\lambda_0^{j+1}} = 1$.

There are age-structured populations for which the Perron–Frobenius result does not hold. The simplest example of a two-stage model with immature (nonreproducing) members and adults, which we have already encountered in our discussion of delayed-recruitment models, corresponds to a Leslie matrix of the form

$$A = \begin{bmatrix} 0 & \beta \\ \rho & 0 \end{bmatrix}$$

having two eigenvalues of equal absolute value and opposite sign.

If $\lambda = re^{i\theta}$ is a complex eigenvalue then $\lambda^n = r^n e^{in\theta} = r^n \cos n\theta + ir^n \sin n\theta$. Corresponding to the pair of complex conjugate eigenvalues $re^{\pm i\theta}$ there is a pair of real solutions of the difference equation containing terms $r^n \cos n\theta$ and $r^n \sin n\theta$ representing oscillations. If the absolute value r of this pair of eigenvalues is as large as λ_0 , then these terms cannot be neglected in the asymptotic expression for the solution vector ρ_n . An extreme case arises if the fertility is concentrated in a single age class, so that the birth rate might be periodic in time. A periodic age structure is known as a *Bernardelli population wave*. Such population waves have been observed in human populations. A simple example is given by the 3×3 Leslie matrix

$$A = \begin{bmatrix} 0 & 0 & \beta_2 \\ p_0 & 0 & 0 \\ 0 & p_1 & 0 \end{bmatrix}$$

with characteristic equation $\lambda^3 - \beta_2 p_0 p_1 = 0$. This equation has three roots, one real and positive and two nonreal, all having the same absolute value. The three cube roots of 1 are $-\frac{1}{2} + i\frac{\sqrt{3}}{2}$, $-\frac{1}{2} - i\frac{\sqrt{3}}{2}$, and 1. Letting $w = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$, we have $w^2 = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$, and $w^3 = 1$. Then the eigenvalues are $\sqrt[3]{\beta_2 p_0 p_1} w$, $\sqrt[3]{\beta_2 p_0 p_1} w^2$, and $\sqrt[3]{\beta_2 p_0 p_1}$.

Exercises

In Exercises 1 through 4, for the given Leslie matrix find the dominant eigenvalue (if there is one), the corresponding eigenvectors and the stable age distribution.

$$1. A = \begin{bmatrix} 0 & 1 & 1 \\ \frac{2}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{bmatrix}.$$

$$2. A = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & 0 \end{bmatrix}.$$

$$3. A = \begin{bmatrix} 1 & 2 & 1 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}.$$

$$4. A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}.$$

5. A population starts with 100 members of age zero. Assume that each member of age zero produces one offspring and that $2/3$ survive to age one. All members of age one produce 3 offspring and then die. Find the corresponding Leslie matrix, its dominant eigenvalue, and a corresponding eigenvector. Describe the stable age distribution and the asymptotic behavior of the population.
6. Consider the Leslie matrix

$$\begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{bmatrix}$$

[Bernardelli (1941)] describing an organism that matures in 2 years, reproduces, and then dies. Begin with an initial population with 100 members in each class and find the population densities for 5 time intervals.

7.2 Linear Continuous Models

It was a physician, Lt. Col. A.G. McKendrick who first introduced age structure into the dynamics of a one-sex population [McKendrick (1926)]. The McKendrick model assumes that the female population can be described by a function of two variables: age and time. Let $\rho(a, t)$ denote the *density* of individuals of age a at time t ; that is, the number of individuals with ages between a and $a + \Delta a$ at time t is approximately $\rho(a, t)\Delta a$. Then the total population at time t is approximately $\sum_a \rho(a, t)\Delta a$, whose “limit as $\Delta a \rightarrow 0$ ” is $\int_0^\infty \rho(a, t)da$, and we define the total population

$$P(t) = \int_0^\infty \rho(a, t)da.$$

In practice, it is reasonable to expect that $\rho(a, t) = 0$ for all sufficiently large a , so that this integral is not necessarily an infinite integral.

Age and time are related. For the *cohort* of members of the population born at the same time, say c , $da/dt = 1$, and $a = t - c$. More generally, if a represents some physiological measure such as size, then $da/dt = g(a)$. The cohort born at c has size $\rho(0, c)$. It proceeds along the *characteristic* $a = t - c$ and is altered only by deaths, since no new members can join the cohort. If $da/dt = g(a)$, the characteristics are the solution curves of

$$\frac{da}{dt} = g(a).$$

We will assume that members leave the population only through death, and that there is an age-dependent death rate $\mu(a)$. This means that over the time interval from t to $t + \Delta t$ a fraction $\mu(a)\Delta t$ of the members with ages between a and $a + \Delta a$ at time t die. At time t there are $\rho(a, t)\Delta a$ individuals with ages between a and $a + \Delta a$. Between the times t and $t + \Delta t$ the number of deaths from this age cohort

is $\rho(a, t)\Delta a\mu(a)\Delta t$, and the remainder survive, having ages between $a + \Delta t$ and $a + \Delta t + \Delta a$ at time $t + \Delta t$. Thus,

$$\rho(a + \Delta t, t + \Delta t)\Delta a \approx \rho(a, t)\Delta a - \rho(a, t)\mu(a)\Delta a\Delta t.$$

Division by $\Delta a\Delta t$ gives

$$\frac{\rho(a + \Delta t, t + \Delta t) - \rho(a, t)}{\Delta t} + \mu(a)\rho(a, t) \approx 0.$$

We then let $\Delta t \rightarrow 0$. If $\rho(a, t)$ is a differentiable function of a and t we have

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \frac{\rho(a + \Delta t, t + \Delta t) - \rho(a, t)}{\Delta t} &= \lim_{\Delta t \rightarrow 0} \frac{\rho(a + \Delta t, t + \Delta t) - \rho(a, t + \Delta t)}{\Delta t} \\ &\quad + \lim_{\Delta t \rightarrow 0} \frac{\rho(a, t + \Delta t) - \rho(a, t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} (\rho_a(a, t + \Delta t) + \rho_t(a, t)) \\ &= \rho_a(a, t) + \rho_t(a, t). \end{aligned}$$

Thus we obtain the *McKendrick equation* (1926)

$$\rho_a(a, t) + \rho_t(a, t) + \mu(a)\rho(a, t) = 0. \quad (7.1)$$

This equation is also known as the *von Foerster equation* (1959), because the same equation arises in cellular biology. The function $\mu(a) \geq 0$ is called the *mortality function* or *death modulus*. If $y(\alpha)$ is the number of individuals starting at age a who survive to age α , then

$$y(\alpha + \Delta\alpha) - y(\alpha) \approx -\mu(\alpha)y(\alpha)\Delta\alpha.$$

If we divide by $\Delta\alpha$ and let $\Delta\alpha \rightarrow 0$, we obtain $y'(\alpha) = -\mu(\alpha)y(\alpha)$, and this implies $y(a_2) = y(a_1)e^{-\int_{a_1}^{a_2} \mu(\alpha)d\alpha}$. Thus the probability that an individual of age a_1 will survive to age a_2 is $e^{-\int_{a_1}^{a_2} \mu(\alpha)d\alpha}$. In particular,

$$\pi(a) = e^{-\int_0^a \mu(\alpha)d\alpha}$$

is the probability of survival from birth to age a .

Next, we assume that the birth process is governed by a function $\beta(a)$ called the *birth modulus*; that is, that $\beta(a)\Delta t$ is the number of offspring produced by members with ages between a and $a + \Delta a$ in the time interval from t to $t + \Delta t$. Thus, the total number of births between time t and time $t + \Delta t$ is $\Delta t \sum \beta(a)\rho(a, t)\Delta a$, which “tends as $\Delta a \rightarrow 0$ ” to $\Delta t \int_0^\infty \beta(a)\rho(a, t)da$. Since this quantity must also be $\rho(0, t)\Delta t$, we obtain the *renewal condition*

$$B(t) = \rho(0, t) = \int_0^\infty \beta(a)\rho(a, t)da.$$

Observe that if we know $B(t)$, then we can calculate $\rho(a, t)$ as the number of births at time $t - a$ multiplied by the survival fraction to age a plus the number of survivors from the initial population born before time $t - a$ and surviving to time t . The upper triangle of the first quadrant of the (a, t) plane represents members of the population who were already alive at time $t = 0$.

In order to complete the model, we must specify an initial age distribution (at time zero)

$$\rho(a, 0) = \varphi(a).$$

Then the full model consists of a partial differential equation and two auxiliary conditions, one of which is an integral condition. The full model is

$$\begin{aligned} \rho_a(a, t) + \rho_t(a, t) + \mu(a)\rho(a, t) &= 0, \\ \rho(0, t) &= \int_0^\infty \beta(a)\rho(a, t)da, \\ \rho(a, 0) &= \phi(a). \end{aligned} \tag{7.2}$$

This is analogous to the discrete Leslie model: The function $\beta(a)$ corresponds to the sequence $\beta_0, \beta_1, \dots, \beta_m$; the function $\mu(a)$ corresponds to the values p_0, p_1, \dots, p_{m-1} (more precisely, $e^{-\int_0^a \mu(\alpha)d\alpha}$ corresponds to the sequence of survival probabilities π_1, \dots, π_m). The solution of the model (7.2) is carried out using the *method of characteristics*.

7.3 The Method of Characteristics

In order to transform the problem (7.2) into a more manageable form, we will integrate along characteristics of the partial differential equation. The characteristics are the lines $t = a + c$. Their importance lies in the fact that the value of the function ρ at a point (a, t) is determined by the values of ρ on the characteristic through (a, t) because a member of the population of age a at time t must have been of age $a - \alpha$ at time $t - \alpha$ for every $\alpha \geq 0$ such that $\alpha \leq a$ and $\alpha \leq t$. In other words, the points on a given characteristic $t = a + c$ all correspond to the same age cohort.

If $t \geq a$, then $\rho(a, t)$ is just the number of survivors to age a of the $\rho(0, t - a)$ members born at time $(t - a)$. Since the fraction surviving to age a is $\pi(a)$, we have $\rho(a, t) = \rho(0, t - a)\pi(a)$ if $t \geq a$. If $t < a$, then $\rho(a, t)$ is just the number of survivors to age a of the $\phi(a - t)$ members who were of age $(a - t)$ at time zero. Since the fraction surviving from age $(a - t)$ to age a is $\pi(a)/\pi(a - t)$, we have $\rho(a, t) = \phi(a - t)\pi(a)/\pi(a - t)$ if $t < a$. Thus

$$\rho(a, t) = \begin{cases} \rho(0, t - a)\pi(a) & \text{for } t \geq a, \\ \phi(a - t)\pi(a)/\pi(a - t) & \text{for } t < a. \end{cases} \tag{7.3}$$

In terms of μ , we may write this as

$$\rho(a, t) = \begin{cases} \rho(0, t-a)e^{-\int_0^a \mu(\alpha) d\alpha} & \text{for } t \geq a, \\ \phi(a-t)e^{-\int_{a-t}^a \mu(\alpha) d\alpha} & \text{for } t < a. \end{cases}$$

We have defined the function $B(t)$ representing the number of births in unit time t , by

$$B(t) = \rho(0, t).$$

Then the original problem is

$$\begin{aligned} \rho_a(a, t) + \rho_t(a, t) + \mu(a)\rho(a, t) &= 0, \\ B(t) &= \int_0^\infty \beta(a)\rho(a, t)da, \\ \rho(a, 0) &= \phi(a). \end{aligned}$$

and we have obtained the representation

$$\rho(a, t) = \begin{cases} B(t-a)e^{-\int_0^a \mu(\alpha) d\alpha} & \text{for } t \geq a, \\ \phi(a-t)e^{-\int_{a-t}^a \mu(\alpha) d\alpha} & \text{for } t < a. \end{cases} \quad i$$

We let $\psi(t)$ be the rate of births from members who were present in the population at time zero. From (7.1) and (7.3) we have

$$\begin{aligned} \rho(0, t) &= \int_0^\infty \beta(a)\rho(a, t)da \\ &= \int_0^t \beta(a)\rho(a, t)da + \int_t^\infty \beta(a)\rho(a, t)da \\ &= \int_0^t \beta(a)\rho(0, t-a)e^{-\int_0^a \mu(\alpha) d\alpha}da + \int_t^\infty \beta(a)\phi(a-t)e^{-\int_{a-t}^a \mu(\alpha) d\alpha}da \\ &= \int_0^t \beta(a)e^{-\int_0^a \mu(\alpha) d\alpha}B(t-a)da + \int_t^\infty \beta(a)\phi(a-t)e^{-\int_{a-t}^a \mu(\alpha) d\alpha}da. \end{aligned}$$

Thus, we may evaluate ψ in terms of the birth and death moduli and the initial age distribution

$$\begin{aligned} \psi(t) &= \int_t^\infty \beta(a)\phi(a-t)e^{-\int_{a-t}^a \mu(\alpha) d\alpha}da \\ &= \int_0^\infty \beta(t+s)\phi(s)e^{-\int_s^{s+t} \mu(\alpha) d\alpha}ds. \end{aligned} \quad (7.4)$$

We also write $\pi(a) = e^{-\int_0^a \mu(\alpha) d\alpha}$, and we see that $B(t)$ is a solution of the *renewal equation*

$$B(t) = \psi(t) + \int_0^t \beta(a)\pi(a)B(t-a)da,$$

a linear Volterra integral equation of convolution type with kernel $\beta(a)\pi(a)$. Conversely, if $B(t)$ is a solution of the renewal equation, then we have a solution of the original problem given by

$$\rho(a, t) = \begin{cases} B(t-a)e^{-\int_0^a \mu(\alpha) d\alpha} = B(t-a)\pi(a) & \text{for } t \geq a, \\ \phi(a-t)e^{-\int_{a-t}^a \mu(\alpha) d\alpha} & \text{for } t < a. \end{cases}$$

The problem that now faces us is to describe the behavior of solutions of the renewal equation under the assumptions that $\phi(t) \rightarrow 0$ as $t \rightarrow \infty$, and, in fact, $\int_0^\infty \phi(t) dt < \infty$ and $R = \int_0^\infty \beta(a)\pi(a) da < \infty$. The number R is the expected number of offspring for each individual over a lifetime, being the sum over all ages a of probability of survival to age a multiplied by the number of offspring at age a .

One approach to this problem [Feller (1941)] involves taking Laplace transforms and using the convolution property to obtain

$$\hat{B}(p) = \hat{\Phi}(p) + \hat{F}(p)\hat{B}(p),$$

or

$$\hat{B}(p) = \frac{\hat{\Phi}(p)}{1 - \hat{F}(p)},$$

where \hat{B} is the Laplace transform of B , $\hat{\Phi}$ is the Laplace transform of ϕ , and \hat{F} is the Laplace transform of $\beta\pi$, that is,

$$\hat{F}(p) = \int_0^\infty \beta(a)\pi(a)e^{-pa} da.$$

Then

$$\hat{F}(0) = \int_0^\infty \beta(a)\pi(a) da = R, \quad \lim_{p \rightarrow \infty} \hat{F}(p) = 0,$$

and $\hat{F}(p)$ is a monotone decreasing function of p . Thus there is a unique real solution p_0 of $\hat{F}(p) = 1$, which is positive if $R > 1$, zero if $R = 1$, and negative if $R < 1$. For every complex $p = \alpha + i\gamma$ we have

$$\begin{aligned} \hat{F}(p) &= \hat{F}(\alpha + i\gamma) = \int_0^\infty \beta(a)\pi(a)e^{-(\alpha+i\gamma)a} da \\ &= \int_0^\infty \beta(a)\pi(a)e^{-\alpha a}(\cos \gamma a - i \sin \gamma a) da, \\ \Re \hat{F}(\alpha + i\gamma) &= \int_0^\infty \beta(a)\pi(a)e^{-\alpha a} \cos \gamma a da. \end{aligned}$$

If $\alpha \geq p_0$ and $\beta \neq 0$, we obtain

$$|\Re \hat{F}(\alpha + i\gamma)| < \int_0^\infty \beta(a)\pi(a)e^{-\alpha a} da \leq \int_0^\infty \beta(a)\pi(a)e^{-p_0 a} da = 1,$$

and thus p_0 is a *dominant root* of the characteristic equation $\hat{F}(p) = 1$. From this it is possible to conclude that $\hat{B}(p)$ is equal to the sum of two terms, one of which has the form $B/(p - p_0)$ for some constant B and the other of which is analytic for $\Re p \geq p_0$. From this it is possible to deduce that

$$B(t) = Be^{p_0 t} + E(t),$$

where $|E(t)| \leq ce^{(p_0 - \varepsilon)t}$ as $t \rightarrow \infty$. From the representation of $\rho(a, t)$ in terms of $B(t)$ we obtain

$$\rho(a, t) = Be^{p_0(t-a)}\pi(a) + E(t-a),$$

and from $P(t) = \int_0^\infty \rho(a, t)da$ we obtain

$$P(t) = Pe^{p_0 t} + E(t)$$

for some constant P . In each of these expressions, $E(t)$ represents a function that grows no faster than $e^{(p_0 - \varepsilon)t}$ as $t \rightarrow \infty$.

A *stable age distribution* (or *persistent age distribution*) is defined to be a solution $\rho(a, t)$ of the form $\rho(a, t) = A(a)T(t)$. We may shift a constant factor between $A(a)$ and $T(t)$ and thus assume $\int_0^\infty A(a)da = 1$; then we have

$$P(t) = \int_0^\infty \rho(a, t)da = T(t) \int_0^\infty A(a)da = T(t).$$

Thus $\rho(a, t) = P(t)A(a)$, and in a stable age distribution the proportion $\rho(a, t)/P(t)$ of age a is $A(a)$, independent of time.

Substitution of the form $\rho(a, t) = A(a)P(t)$ into the McKendrick equation (7.1) gives

$$A'(a)P(t) + A(a)P'(t) + \mu(a)A(a)P(t) = 0,$$

or

$$-\mu(a) - \frac{A'(a)}{A(a)} = \frac{P'(t)}{P(t)}.$$

Here primes denote derivatives with respect to the appropriate variable. Because the left side of this equation is a function of a only, while the right side is a function of t only, each side must be equal to a constant p_0 (as yet undetermined). Thus we have two separate ordinary differential equations:

$$P'(t) - p_0 P(t) = 0,$$

with solution

$$P(t) = P(0)e^{p_0 t},$$

and

$$A'(a) + (\mu(a) + p_0)A(a) = 0,$$

with solution

$$A(a) = A(0)e^{-p_0 a}e^{-\int_0^a \mu(\alpha)d\alpha} = A(0)e^{-p_0 a}\pi(a).$$

To satisfy the renewal condition $\rho(0, t) = \int_0^\infty \beta(a)\rho(a, t)da$ we must have

$$P(t)A(0) = \int_0^\infty \beta(a)A(a)P(t)da = P(t) \int_0^\infty \beta(a)A(0)e^{-p_0 a}\pi(a)da$$

or

$$\int_0^\infty \beta(a)\pi(a)e^{-p_0a}da = 1,$$

which is known as the *Lotka-Sharpe equation* [Lotka and Sharpe (1911)] for p_0 . As we have already remarked, this has a unique real root, which is positive if $R = \int_0^\infty \beta(a)\pi(a)da > 1$, zero if $R = 1$, and negative if $R < 1$. In a stable age distribution,

$$\rho(a, t) = ce^{p_0t}e^{-p_0a}\pi(a) = ce^{p_0(t-a)}\pi(a),$$

and the content of our analysis of the renewal equation earlier is that asymptotically (as $t \rightarrow \infty$) every age distribution tends to a stable age distribution. The Lotka-Sharpe equation is the analogue of the equation $\sum_{j=0}^m \beta_j \pi_j \lambda_0^{-(j+1)} = 1$ obtained in the discrete case, and the result that every age distribution tends to a stable age distribution is the analogue of the result that in the discrete case $\rho_n \sim c\lambda_0^n \mathbf{v}_0$.

If $R = 1$ then the total population size $P(t)$ is constant,

$$\rho(a, t) = B\pi(a),$$

and the birth rate is also a constant,

$$B(t) = \int_0^\infty \beta(a)P(0)A(0)\pi(a)da = P(0)A(0) \int_0^\infty \beta(a)\pi(a)da = P(0)A(0).$$

In this case we have an *equilibrium age distribution*. It is easy to see that, conversely, if the total population size is constant then the birth rate is also constant and $\rho(a, t)$ is independent of t .

Example 1. In the “genesis” model we assume $\phi(a) = \delta(a)$, the Dirac delta function with $\delta(a) = 0$ for $a \neq 0$, $\int_0^\infty \delta(a)da = 1$. Thus the initial population is all at age zero. Let us assume also that the birth modulus $\beta(a)$ and the death modulus $\mu(a)$ are both constants, β and μ , respectively. Then the renewal equation takes the form

$$B(t) = \phi(t) + \int_0^t \beta e^{-\int_0^a \mu d\alpha} B(t-a)da,$$

with

$$\psi(t) = \int_0^\infty \beta \phi(s) e^{-\int_s^{s+t} \mu d\alpha} ds = \int_0^\infty \delta(s) e^{-\mu t} ds = \beta e^{-\mu t}.$$

Also

$$\beta e^{-\int_0^a \mu d\alpha} = \beta e^{-\mu a}.$$

Thus, $B(t)$ satisfies

$$\begin{aligned}
B(t) &= \beta e^{-\mu t} + \beta \int_0^t e^{-\mu a} B(t-a) da \\
&= \beta e^{-\mu t} + \beta \int_0^t e^{-\mu(t-s)} B(s) ds \\
&= \beta e^{-\mu t} + \beta e^{-\mu t} \int_0^t e^{\mu s} B(s) ds.
\end{aligned}$$

Differentiation gives

$$\begin{aligned}
B'(t) &= -\mu \beta e^{-\mu t} + \beta e^{-\mu t} e^{\mu t} B(t) - \beta \mu e^{-\mu t} \int_0^t e^{\mu s} B(s) ds \\
&= -\mu \left(\beta e^{-\mu t} + \beta e^{-\mu t} \int_0^t e^{\mu s} B(s) ds \right) + \beta B(t) \\
&= -\mu B(t) + \beta B(t) = (\beta - \mu) B(t).
\end{aligned}$$

From the renewal equation with $t = 0$ we see that $B(0) = \beta$. Now $B'(t) = (\beta - \mu)B(t)$, $B(0) = \beta$ implies $B(t) = \beta e^{(\beta - \mu)t}$, and this gives the age distribution function

$$\rho(a, t) = \begin{cases} \beta e^{(\beta - \mu)(t-a) - \mu a} & \text{for } t \geq a, \\ \delta(a-t) e^{-\mu t} & \text{for } t < a. \end{cases}$$

The total population size is

$$\begin{aligned}
P(t) &= \int_0^\infty \rho(a, t) da = \int_0^t \rho(a, t) da + \int_t^\infty \rho(a, t) da \\
&= \int_t^0 \beta e^{(\beta - \mu)t} e^{-\beta a} da + \int_t^\infty \delta(a-t) e^{-\mu t} da \\
&= \beta e^{(\beta - \mu)t} \int_0^t e^{-\beta a} da + e^{-\mu t} \\
&= e^{(\beta - \mu)t} (1 - e^{-\beta t}) + e^{-\mu t} = e^{(\beta - \mu)t}.
\end{aligned}$$

Exercises

1. Consider a model with β and μ constant and the initial age distribution $\phi(a)$ arbitrary.
 - (i) Show that $\psi(t) = \beta e^{-\mu t} \int_0^\infty \phi(s) ds$.
 - (ii) Obtain an integral equation for $B(t)$.
 - (iii) Show by differentiation of the integral equation obtained in part (i) that

$$B'(t) = (\beta - \mu)B(t), \quad B(0) = \beta \int_0^\infty \phi(s) ds.$$

- (iv) Solve for $B(t)$.

(v) What can you say about the total population size as $t \rightarrow \infty$?

2.* Consider a model with $\phi(a)$ arbitrary, μ constant, and $\beta(a) = \beta e^{-\mu a}$.

(i) Show that $\psi(t) = \beta \int_0^\infty e^{-\mu s} \phi(s) ds$.

(ii) Show that $B(t)$ satisfies the initial value problem

$$B'(t) = (\beta - 2\mu)B(t), \quad B(0) = \beta \int_0^\infty e^{-\mu s \phi(s)} ds.$$

(iii) Solve the initial value problem obtained in part (ii) to find that

$$B(t) = \beta e^{-(\beta-2\mu)t} \int_0^\infty e^{-\mu s} \phi(s) ds.$$

3.* Consider a model with μ constant and an arbitrary initial age distribution $\phi(a)$ for which $\beta(a) = 0$ if $a \leq T$ and $\beta(a)$ is a constant β if $a > T$. Show that the integral equation satisfied by $B(t)$ for $t > T$ is

$$B(t) = \beta e^{-\mu t} \int_0^\infty \phi(s) ds + \beta e^{-\mu t} \int_0^{t-T} e^{\mu s} B(s) ds$$

and deduce that $B(t)$ satisfies the differential-difference equation

$$B'(t) = \beta e^{-\mu T} B(t-T) - \mu B(t).$$

4. Suppose that a constant b can be found such that

$$\rho(0, t) = b \int_0^\infty \rho(a, t) da \equiv bP(t),$$

and that nobody lives forever (that is, there is a maximal life span). Obtain an integral equation for $P(t)$.

7.4 Nonlinear Continuous Models

In the previous section, the birth and death rates were linear, and this implies that total population size grows exponentially, dies out exponentially, or remains constant. In studying models without age structure, we considered situations in which populations have a carrying capacity that is approached as $t \rightarrow \infty$. In order to allow this possibility for age-structured models we must assume some nonlinearity. We now consider the possibility of birth and death moduli of the form $\beta(a, P(t))$ and $\mu(a, P(t))$, depending on total population size. A variant, which can be developed by analogous methods, would be to allow the birth modulus (and possibly also the death modulus) to depend on $\rho(a, t)$, the number of members in the same age cohort. We will consider only the continuous case, because the methods of linear al-

gebra used to treat the linear discrete case have no direct adaptation to the nonlinear discrete model.

If the birth and death moduli are allowed to depend on total population size the description of the model must include the definition of $P(t)$. Thus, the full model is now

$$\begin{aligned}\rho_a(a, t) + \rho_t(a, t) + \mu(a, P(t))\rho(a, t) &= 0 \\ B(t) = \rho(0, t) &= \int_0^\infty \beta(a, P(t))\rho(a, t)da \quad (7.5) \\ P(t) &= \int_0^\infty \rho(a, t)da \\ \rho(a, 0) &= \phi(a).\end{aligned}$$

We can transform the problem just as in the linear case by integrating along characteristics. If we define

$$\mu^*(\alpha) = \mu(\alpha, P(t - a + \alpha)),$$

the same calculations give

$$\rho(a, t) = \begin{cases} B(t - a)e^{-\int_0^a \mu^*(\alpha)d\alpha} & \text{for } t \geq a, \\ \phi(a - t)e^{-\int_{a-t}^a \mu^*(\alpha)d\alpha} & \text{for } t < a. \end{cases} \quad (7.6)$$

When we substitute these expressions into (7.5) and (7.6), we obtain a pair of functional equations for $B(t)$ and $P(t)$, whose solution gives an explicit solution for $\rho(a, t)$, namely

$$\begin{aligned}B(t) &= b(t) + \int_0^t \beta(a, P(t))e^{-\int_0^a \mu^*(\alpha)d\alpha}B(t - a)da, \\ P(t) &= p(t) + \int_0^t e^{-\int_0^a \mu^*(\alpha)d\alpha}B(t - a)da,\end{aligned}$$

where

$$\begin{aligned}\mu^*(\alpha) &= \mu(\alpha, P(t - a + \alpha)), \\ b(t) &= \int_t^\infty \beta(a, P(t))\phi(a - t)e^{-\int_{a-t}^a \mu^*(\alpha)d\alpha}da, \\ p(t) &= \int_t^\infty \phi(a - t)e^{-\int_{a-t}^a \mu^*(\alpha)d\alpha}da.\end{aligned}$$

It is reasonable to assume that $\int_0^\infty \phi(a)da < \infty$ and that the functions $\beta(a, P), \mu(a, P)$ are continuous and nonnegative; under these hypotheses it is easy to verify that $b(t)$ and $p(t)$ are continuous and nonnegative and tend to zero as $t \rightarrow \infty$, and that $b(0) > 0, p(0) > 0$. Without additional assumptions it is not necessarily true that the pair of functional equations has a solution for $0 \leq t < \infty$, but it is possible to

prove that if $\sup_{a \geq 0, P \geq 0} \beta(a, P) < \infty$, then there is a unique continuous nonnegative solution on $0 \leq t < \infty$. This model is due to M.L. Gurtin and R.C. MacCamy (1974).

A solution $\rho(a, t)$ that is independent of t is called an *equilibrium* age distribution. If $\rho(a, t) = \rho(a)$ is an equilibrium age distribution then both $P(t) = \int_0^\infty \rho(a) da$ and $B(t) = \int_0^\infty \beta(a, P(t)) \rho(a) da$ are constant. Conversely, if $P(t)$ and $B(t)$ are constant then $\rho(a, t)$ is independent of t and thus is an equilibrium age distribution.

If $\rho(a)$ is an equilibrium age distribution, the McKendrick equation becomes an ordinary differential equation $\rho'(a) + \mu(a, P)\rho(a) = 0$, with initial condition $\rho(0) = B$, whose solution is

$$\rho(a) = B e^{-\int_0^a \mu(\alpha, P) d\alpha}.$$

If we define

$$\pi(a, P) = e^{-\int_0^a \mu(\alpha, P) d\alpha} \quad (7.7)$$

the probability of survival from birth to age a when the population size is the constant P , then

$$\rho(a) = B \pi(a, P).$$

From $P = \int_0^\infty \rho(a) da$ we obtain

$$P = B \int_0^\infty \pi(a) da,$$

and from $B = \int_0^\infty \beta(a, P) \rho(a) da$, we obtain

$$P = B \int_0^\infty \beta(a, P) \pi(a, P) da.$$

Thus, for an equilibrium age distribution with birth rate B and population size P , P must satisfy the equation

$$R(P) = \int_0^\infty \beta(a, P) \pi(a, P) da = 1,$$

and then B is given by

$$B = \frac{P}{\int_0^\infty \pi(a, P) da}.$$

Then $1 / \int_0^\infty \pi(a, P) da$ is the average life expectancy and the equilibrium age distribution is $\rho(a) = B \pi(a, P)$. The quantity $R(P)$, called the *reproductive number*, is the expected number of offspring that an individual has over its lifetime when the total population size is P .

Example 1. Suppose that β is independent of age and is a function of P only. Then

$$B(t) = \int_0^\infty \beta(P(t)) \rho(a, t) da = \beta(P(t)) \int_0^\infty \rho(a, t) da = P(t) \beta(P(t)), \quad (7.8)$$

and the problem is reduced to a single functional equation for $P(t)$ together with this explicit formula for $B(t)$. If we define $g(P) = P\beta(P)$, the equation for $P(t)$ is

$$P(t) = p(t) + \int_0^t e^{-\int_0^a \mu^*(\alpha) d\alpha} g(P(t-a)) da.$$

Example 2. Suppose that β is independent of age and in addition that μ is independent of population size and is a function of age only. Then, instead of $\mu^*(\alpha) = \mu(\alpha, P(t-a+\alpha))$, we have $\mu(\alpha)$, and the equation for $P(t)$ is a Volterra integral equation called the *nonlinear renewal equation*,

$$P(t) = p(t) + \int_0^t e^{-\int_0^a \mu(\alpha) d\alpha} g(P(t-a)) da = p(t) + \int_0^t \pi(\alpha) g(P(t-a)) da. \quad (7.9)$$

In order to analyze (7.9), we first apply a theorem of Levin and Shea (1972) stating that the asymptotic behavior of (7.9) is the same as that of the *limit equation*

$$P(t) = \int_0^\infty \pi(\alpha) g(P(t-a)) da = \int_{-\infty}^t \pi(t-u) g(P(u)) du. \quad (7.10)$$

In other words, the particular choice

$$p(t) = \int_t^\infty g(P(t-a)) \pi(a) da$$

of initial function is general in the sense that the behavior of (7.9) is the same for all initial functions that tend to zero.

It is possible to prove [Londen (1974)] that every bounded solution of (7.10) approaches a limit P_∞ as $t \rightarrow \infty$ with

$$P_\infty = g(P_\infty) \int_0^\infty \pi(a) da.$$

or

$$\beta(P_\infty) \int_0^\infty \pi(a) da = 1.$$

If we differentiate the integral equation after writing

$$\int_0^t \pi(a) g(P(t-a)) da = \int_0^t \pi(t-\tau) g(P(\tau)) d\tau$$

and using the relationship

$$\begin{aligned} \frac{d}{dt} \int_0^t \pi(t-\tau) g(P(\tau)) d\tau &= \pi(0) g(P(t)) + \int_0^t \pi'(t-\tau) g(P(\tau)) d\tau \\ &= \pi(0) g(P(t)) + \int_0^t \pi'(a) g(P(t-a)) da, \end{aligned}$$

we obtain the integrodifferential equation

$$P'(t) = p'(t) + \pi(0)g(P(t)) + \int_0^t g(P(t-a))\pi'(a)da,$$

whose linearization about the equilibrium P_∞ is

$$u'(t) = \pi(0)g'(P_\infty)u(t) + g'(P_\infty) \int_0^t u(t-a)\pi'(a)da.$$

We recall that for the linear integrodifferential equation $u'(t) = \alpha u(t) + \beta \int_0^t u(t-a)k(a)da$ with $\int_0^\infty k(a)da = 1$ if $\alpha + \beta \geq 0$ then solutions do *not* tend to zero (Section 3.4). Here $\alpha = \pi(0)g'(P_\infty)$,

$$k(a) = \frac{\pi'(a)}{\int_0^\infty \pi'(a)da} = \frac{\pi'(a)}{\pi(0)},$$

so that $\beta = g'(P_\infty) \int_0^\infty \pi'(a)da = -g(P_\infty)\pi(0)$ and $\alpha + \beta = 0$. Thus, the “equilibrium” P_∞ of the nonlinear renewal equation cannot be asymptotically stable. However, it can be shown [Brauer (1976, 1987b)] that if $g'(P_\infty) \int_0^\infty \pi(a)da < 1$, the equilibrium P_∞ is stable in a weaker sense, namely that small disturbances do not alter the solution very much, while if $g'(P_\infty) \int_0^\infty \pi(a)da > 1$, solutions tend away from P_∞ . Thus, the condition $g'(P_\infty) \int_0^\infty \pi(a)da < 1$ is necessary for the limit P_∞ to be meaningful biologically. Since

$$g'(P_\infty) = \beta(P_\infty) + P_\infty\beta'(P_\infty)$$

and $\beta(P_\infty) \int_0^\infty \pi(a)da = 1$, this condition is equivalent to $\beta'(P_\infty) < 0$.

For age-structured populations with μ a function of age only and β a function of population size only, we have shown that every solution tends to an equilibrium age distribution, and we have

$$\lim_{t \rightarrow \infty} P(t) = P_\infty, \quad \lim_{t \rightarrow \infty} B(t) = P_\infty\beta(P_\infty),$$

while

$$\rho(a, t) = B(t-a)\pi(a) \sim P_\infty\beta(P_\infty)\pi(a).$$

Such models are more realistic than models described by ordinary differential equations in which age dependence is ignored completely. However, the assumption that mortality is independent of population size is quite unrealistic, especially in populations whose members compete for resources.

In the special case of a constant death rate

$$\pi(a) = e^{-\mu a},$$

differentiation of the equation (7.10) gives

$$P'(t) = g(P(t)) - \mu P(t),$$

an ordinary differential model with a density-dependent birth rate.

The nonlinear renewal equation (7.9) as a model for a population with a birth rate depending only on total population size and a death rate depending only on age can be constructed directly, without starting from an age-structured model. If the number of births in unit time when total population size is P is $B = g(P) = P\beta(P)$, and the fraction of the population surviving to age a is $\pi(a)$, then the number of members born at time $(t - a)$ and having age between a and $a + \Delta a$ at time t is approximately $\pi(a)g(P(ta))\Delta a$. Thus, the total population size $P(t)$ at time t is the sum of two terms. The first term $p(t)$ represents the number of members born before the time $t = 0$ and still surviving at time t , and the second term $\int_0^t \pi(a)g(P(t - a))da$ represents the number of members born between time zero and time t and surviving to time t . This explains the model (7.9).

We may proceed in the other direction to derive the age distribution from the nonlinear renewal equation (7.9) under the assumptions that the birth rate depends on total population size and the death rate depends on age. From the relations (7.1), (7.2), and (7.3) we have, for $t \geq a$,

$$\rho(a, t) = B(t - a)\pi(a) = \mathbb{I}(t - a)\beta(P(t - a))\pi(a) = g(P(t - a))\pi(a).$$

Example 3. Suppose that μ is independent of age and is a function of P only. Then the McKendrick equation is

$$\rho_a(a, t) + \rho_t(a, t) + \mu(P(t))\rho(a, t) = 0.$$

We integrate with respect to a from zero to ∞ and assume $\lim_{a \rightarrow \infty} \rho(a, t) = 0$. Then

$$\begin{aligned} \int_0^\infty \rho_a(a, t) da &= \rho(a, t) \Big|_{a=0}^{a=\infty} = -\rho(0, t) = -B(t), \\ \int_0^\infty \rho_t(a, t) da &= \frac{d}{dt} \int_0^\infty \rho(a, t) da = P'(t), \end{aligned}$$

and

$$\int_0^\infty \mu(P(t))\rho(a, t) da = \mu(P(t)) \int_0^\infty \rho(a, t) da = P(t)\mu(P(t)).$$

Thus, the problem can be reformulated as a functional equation for $P(t)$ together with an ordinary differential equation $P' + \mu(P)P = B$. We now assume in addition that β is independent of population size and is a function of age only. Then the system takes the form

$$\begin{aligned} P'(t) + P(t)\mu(P(t)) &= B(t) \\ B(t) &= b(t) + \int_0^t \beta(a)e^{-\int_0^a \mu^*(\alpha) d\alpha} B(t - a) da, \end{aligned}$$

with

$$\mu^*(\alpha) = \mu(P(t - a + \alpha)).$$

It is convenient to make a change of variables and write

$$\int_0^a \mu^*(\alpha) d\alpha = \int_0^a \mu(P(t-a+\alpha)) d\alpha = \int_{t-a}^t \mu(P(u)) du,$$

so that

$$B(t) = b(t) + \int_0^t \beta(a) e^{-\int_{t-a}^t \mu(P(u)) du} B(t-a) da.$$

Now we let

$$B^*(t) = B(t) e^{\int_0^t \mu(P(u)) du}, \quad P^*(t) = P(t) e^{\int_0^t \mu(P(u)) du}$$

and substitute, obtaining the system

$$\begin{aligned} B^*(t) &= b^*(t) + \int_0^t \beta(a) B^*(t-a) da, \\ P^{*'}(t) &= B^*(t), \end{aligned}$$

with

$$b^*(t) = \int_t^\infty \beta(a) \phi(a-t) da.$$

If we search for stable age distributions by substituting $\rho(a, t) = A(a)P(t)$ into the McKendrick equation, we obtain

$$-\frac{A'(a)}{A(a)} = \frac{P'(t)}{P(t)} + \mu(P(t)),$$

and see as before that each side of this equation must be a constant p_0 . Thus, $A'(a) + p_0 A(a) = 0$, which gives $A(a) = A(0)e^{-p_0 a}$, and $P'(t) + P(t)\mu(P(t)) = p_0 P(t)$. The renewal condition

$$\rho(0, t) = \int_0^\infty \beta(a) \rho(a, t) da$$

gives

$$P(t)A(0) = \int_0^\infty \beta(a) P(t) A(0) e^{-p_0 a} da,$$

which implies

$$\int_0^\infty \beta(a) e^{-p_0 a} da = 1.$$

In the linear theory we showed that every solution approaches a stable age distribution. A similar but weaker result is true in the nonlinear case. By the use of the Laplace transform and the equation for $B^*(t)$, we can show that $B^*(t) \sim ce^{p_0 t}$ as $t \rightarrow \infty$ and because $P^{*'}(t) = B^*(t)$ we have $P^*(t) \sim ce^{p_0 t}/p_0$ (unless $p_0 = 0$). We now obtain $\rho(a, t) \sim A(a)Q(t)$, $B(t) \sim p_0 Q(t)$ as $t \rightarrow \infty$, where

$$Q(t) = \frac{c}{p_0} e^{p_0 t - \int_0^t \mu(P(u)) du}.$$

This differs from the result in the linear case because $Q(t)$ does not in general satisfy the differential equation $y' + y\mu(y) = p_0y$ and thus $A(a)Q(t)$ is not a solution of the model. It is, however, true that $Q(t)/P(t) \rightarrow 1$ as $t \rightarrow \infty$.

Exercises

- 1.* Assume that μ is independent of age and is a function of P only. Assume also that $\lim_{a \rightarrow \infty} \rho(a, t) = 0$. Integrate (7.5) with respect to a from zero to ∞ and deduce that

$$P'(t) = B(t) - P(t)\mu(P(t)).$$

- 2.* Assume that μ is a function of P only and that $\beta(a, P) = \beta_0 e^{-\alpha a}$, with β_0 and α positive constants. Multiply (7.5) by $e^{-\alpha a}$ and integrate with respect to a from zero to ∞ and deduce that

$$B'(t) = [\beta_0 - \alpha - \mu(P(t))]B(t).$$

- 3.* Exercises 1 and 2 show that if μ is a function of P only and $\beta(a, P) = \beta_0 e^{-\alpha a}$, then the model (7.5) is equivalent to the system of ordinary differential equations

$$\begin{aligned} P' &= B - P\mu(P), \\ B' &= [\beta_0 - \alpha - \mu(P)]B. \end{aligned}$$

- (i) Show that $P = 0, B = 0$ is an equilibrium that is asymptotically stable if and only if $\beta_0 - \alpha < \mu(0)$.
- (ii) Show that there is an equilibrium (P_∞, B_∞) with $\mu(P_\infty) = \beta_0 - \alpha$ and $B_\infty = P_\infty \mu(P_\infty) = (\beta_0 - \alpha)P_\infty$.
- (iii) Show that the equilibrium (P_∞, B_∞) is asymptotically stable if and only if $\alpha - \beta_0 < P_\infty \mu'(P_\infty)$, $\mu'(P_\infty) > 0$.

7.5 Models with Discrete Age Groups

Under certain conditions, the age-structured partial differential equation model (7.5) can be reduced to a system of ordinary differential equations [Hethcote (1997), Li and Hallam (1988), Tudor (1985)].

Partition the age interval into a finite number n of subintervals $[a_0, a_1), [a_1, a_2), \dots, [a_{n-1}, a_n)$, where $a_0 = 0$ and $a_n \leq \infty$. Denote the number of individuals with ages in interval $[a_{i-1}, a_i]$ by $H_i(t)$, so that $H_i(t) = \int_{a_{i-1}}^{a_i} \rho(t, a) da$, $i = 1, \dots, n$. Then integrating the partial differential equation in (7.5) from a_0 to a_1 , we have

$$\frac{dH_1(t)}{dt} + \rho(t, a_1) - \rho(t, a_0) + \int_{a_0}^{a_1} \mu(a, P) \rho(t, a) da = 0. \quad (7.11)$$

Assume that individuals with ages in each interval have the same vital rates such that $\beta(a, P) = \beta_i$, $\mu(a, P) = \mu_i$, for a in $[a_{i-1}, a_i]$, $i = 1, \dots, n$. Here β_i and μ_i are age-independent, but may be density-dependent. Then, in the age interval $[0, a_1]$, we have

$$\rho(t, 0) = \sum_1^n \beta_i H_i(t), \quad \int_{a_0}^{a_1} \mu \rho(t, a) da = \mu_1 H_1(t),$$

which leads to

$$\frac{dH_1}{dt} = \sum_1^n \beta_i H_i - (m_1 + \mu_1) H_1. \quad (7.12)$$

Here m_1 is the progression rate from groups 1 to 2, defined by $m_1 = \rho(t, a_1)/H_1(t)$, and we assume it is time-independent.

Integrating the partial differential equation in (7.5) from a_{i-1} to a_i for $2 \leq i \leq \infty$, we have

$$\frac{dH_i}{dt} = m_{i-1} H_{i-1} - (m_i + \mu_i) H_i, \quad i = 2, \dots, n, \quad (7.13)$$

where m_i is the age progression rate from groups i to $i+1$, and we let $m_n = 0$. Then the system in (7.5) is reduced to a system of n ordinary differential equations.

We provide a simple example to demonstrate how the discrete age group model described by equations (7.12) and (7.13) can be applied to populations and infectious diseases.

Example 1. A two-age-group population model. There are many means by which individuals of a species might compete for resources and by which intraspecific competition might express itself. Organisms that do not undergo such radical changes during their life cycles (e.g., birds, mammals, most reptiles, fishes, and hemimetabolous insects such as aphids, true bugs, and grasshoppers) can experience considerable competition between juveniles and adults for common resources. Intraspecific competition can also occur to organisms with simple life cycles. The well-studied flour beetles of genus *Tribolium*, whose adult and larval stages utilize food resources in common, provide a case in point [Costantino and Desharnais (1991)].

Let $J(t)$ and $A(t)$ denote the densities of juveniles and adults at time t , respectively. Using the model in (7.12) and (7.13), with $n = 2$, we have the two-age-group model

$$\begin{aligned} J'(t) &= \beta(J, A)A - (m(J, A) + \mu_1(J, A))J, \\ A'(t) &= m(J, A)J - \mu_2(J, A)A, \end{aligned} \quad (7.14)$$

where $'$ denotes d/dt , β is the birth rate of adults, μ_i , $i = 1, 2$, are death rates for juveniles and adults, respectively, and m is the age progression rate.

Models similar to (7.14) have been studied intensively. Readers are referred to [Cushing and Li (1991), Gurney and Nisbet (1998), Kostova and Li (1996), Kostova, Li and Friedman (1999), Nisbet and Onyiah (1994), Thieme (2003), Tschumy (1982)].

7.6 Project: Ordinary Differential Equations with Age Structure

Begin with the age-structured population model (7.2). We divide the population into n age groups defined by the age intervals, $[a_{i-1}, a_i]$ for $i = 1, \dots, n$, where $0 = a_0 < a_1 < \dots < a_{n-1} < a_n = \infty$. For each age interval, assume that the birth and death moduli are constants, i.e., $\beta(a) = \beta_i$ and $\mu(a) = \mu_i$. Assume also that the population has reached an equilibrium age distribution with exponential growth, so that $\rho(a, t) = e^{qt} A(a)$. Let

$$\rho_i(t) = \int_{a_{i-1}}^{a_i} \rho(a, t) da = e^{qt} \int_{a_{i-1}}^{a_i} A(a) da = e^{qt} P_i,$$

where $\rho_i(t)$ is the number of individuals in the age bracket $[a_{i-1}, a_i]$ and P_i is the size of the i th age group at time 0.

1. Substitute $\rho(a, t)$ into the McKendrick equation (7.1) to obtain an ordinary differential equation for $A(a)$,

$$A'(a) = -[\mu(a) + q]A(a).$$

2. Find the solution on the interval $[a_{i-1}, a_i]$ given by

$$A(a) = A(a_{i-1}) \exp[-(\mu_i + q)(a - a_{i-1})].$$

3. Define the constants c_i by

$$c_i = \frac{A(a_i)}{P_i},$$

and obtain the system of ordinary differential equations

$$\begin{aligned} \frac{dN_1}{dt} &= \sum_{j=1}^n f_j P_j - [c_1 + \mu_1] \rho_1, \\ \frac{dN_i}{dt} &= c_{i-1} \rho_{i-1} - (c_i + \mu_i) \rho_i, \quad 2 \leq i \leq n-1, \\ \frac{dN_n}{dt} &= c_{n-1} \rho_{n-1} - \mu_n \rho_n. \end{aligned} \tag{7.15}$$

7.7 Project: Nonlinear Age Structured Population Growth

The *limit equation* of the nonlinear renewal equation is

$$P(t) = \int_0^\infty \pi(a) g(P(t-a)) da.$$

Suppose we accept as fact that if $p(t) \rightarrow 0$ as $t \rightarrow \infty$, then the asymptotic behavior of the nonlinear renewal equation is the same as that of the limit equation.

1. Show that if the solution $P(t)$ of the limit equation has a limit P_∞ as $t \rightarrow \infty$, then

$$P_\infty = g(P_\infty) \int_0^\infty \pi(a) da.$$

2. Find the integral equation that is the linearization of the limit equation about its equilibrium.
3. Find the characteristic equation of the linearization (i.e., the condition on the parameter λ that the linearization has a solution that is a constant multiple of $e^{\lambda t}$).
4. Show that all roots of this characteristic equation have negative real part if

$$g'(P_\infty) \int_0^\infty \pi(a) da < 1.$$

7.8 Project: A Size Structured Population Model

Suppose that a population is structured by size (or some other physiological quantity) m and that the size of an individual grows at a rate $g(m)$ depending only on the size. Suppose that there is a size-dependent mortality rate $\theta(m)$ and a size-dependent fecundity rate $\phi(m)$. Formulate a model analogous to the McKendrick age - structured population model, that is, assume that individuals grow larger or die and that those who survive may reproduce.

Chapter 8

Models for Populations with Spatial Structure

8.1 Introduction

Populations may be structured by spatial location. There are two common different ways to include spatial location in a population. One way is by means of *metapopulations*, that is, populations of populations, with links between them such as a collection of towns and cities connected by a transportation network. The air transport subnetwork includes connecting links between distant communities, and we may study the dynamics of populations of different cities as a function of the flow of people between them and their own local dynamics in this framework. A metapopulation may be divided into *patches*, with each patch corresponding to a separate location. The corresponding models may be systems of ordinary differential equations, with the population size of each species in each patch as a variable. Thus metapopulation models are often systems of ordinary differential equations of high dimension. Some basic references are Hanski (1999), Hanski and Gilpin (1997), Levin, Powell and Steele (1993), Neuhauser (2001).

A second way of including spatial dependence in a population is to allow motion of individuals in the population, describing the population size as a function of time and spatial location under some specific assumptions on the nature of the motion of individuals. Since population size is now a function of more than one variable (a time variable and a number of spatial variables, the number depending on the dimension of the space in which individuals move), this approach will lead to *partial differential equation* models. These equations are often of *reaction-diffusion* type, with a reaction term corresponding to the (local) population dynamics and a diffusion term describing the motion of individuals in space.

8.2 Some Simple Examples of Metapopulation Models

Insect populations have been observed that become extinct in some locations but persist globally because of recolonization from other locations [Andrewartha and Birch (1954)]. An early metapopulation model [Levins (1969)] considered a population divided among distinct patches that are either occupied or vacant. It is expressed in terms of the fraction $p(t)$ of patches occupied at time t . The assumptions are that there is a colonization or recolonization rate c and an extinction rate e associated with each patch, and that the rate at which vacant patches are occupied is proportional to the fraction p of occupied patches and the fraction $(1 - p)$ of vacant patches. Thus p satisfies the differential equation

$$p' = cp(1 - p) - ep. \quad (8.1)$$

It is easy to verify that if $c < e$, the only equilibrium of equation (8.1) is $p = 0$, and that this equilibrium is asymptotically stable. If $c > e$, the equilibrium $p = 0$ is unstable, but there is an asymptotically stable equilibrium

$$p = 1 - \frac{e}{c} > 0.$$

The model (8.1) is an occupancy model that assumes that the colonization rate in a patch depends only on the fraction of occupied patches, not on any properties of individual patches. In addition, it ignores the local population dynamics of each patch. A more detailed description would describe the population dynamics of each patch and would include travel of individuals between patches. We will restrict ourselves to metapopulations containing only a single species, but it would be natural to extend to competing species or predator–prey metapopulations.

As a very simple example, consider two populations in separate patches governed by logistic equations with the same intrinsic growth rate but different carrying capacities,

$$\begin{aligned} y_1' &= ry_1 \left(1 - \frac{y_1}{K_1} \right) \\ y_2' &= ry_2 \left(1 - \frac{y_2}{K_2} \right). \end{aligned} \quad (8.2)$$

If there is no interaction between the two populations, the two equations of (8.2) can be solved separately, and it is easy to show that there is an asymptotically stable equilibrium $y_1 = K_1$, $y_2 = K_2$.

Now let us add the assumption that there is a proportional travel rate a from the first patch to the second patch to give a simple metapopulation model

$$\begin{aligned} y_1' &= ry_1 \left(1 - \frac{y_1}{K_1}\right) - ay_1, \\ y_2' &= ry_2 \left(1 - \frac{y_2}{K_2}\right) + ay_1. \end{aligned} \quad (8.3)$$

In order to describe the behavior of the model (8.3), we need to find equilibria of the system and analyze their stability. Equilibria are solutions of the pair of equations

$$\begin{aligned} ry_1 \left(1 - \frac{y_1}{K_1}\right) - ay_1 &= 0, \\ ry_2 \left(1 - \frac{y_2}{K_2}\right) + ay_1 &= 0. \end{aligned}$$

One possibility from the first of these equations is that $y_1 = 0$, and then the second equation implies either $y_2 = 0$ or $y_2 = K_2$. A second possibility is

$$y_1 = K_1 \left(1 - \frac{a}{r}\right), \quad (8.4)$$

provided $r > a$. This implies

$$ry_2 \left(1 - \frac{a}{r}\right) + ay_1 = 0,$$

which can be solved for y_2 . Note that this solution satisfies $y_2 > K_2$.

To determine the stability of an equilibrium, we form the matrix of the linearization of (8.3) at an equilibrium (y_1, y_2) , namely

$$\begin{bmatrix} r \left(1 - \frac{2y_1}{K_1}\right) - a & 0 \\ a & r \left(1 - \frac{2y_2}{K_2}\right) \end{bmatrix}.$$

The eigenvalues of this matrix are the diagonal elements. At the equilibrium $(0, 0)$ these are $r - a, r$, and therefore this equilibrium is unstable. At the equilibrium $(0, K_2)$ these are $r - a, -r$. Thus this equilibrium is asymptotically stable if and only if $r < a$. At the equilibrium with both population sizes positive, the eigenvalues are both negative if and only if the element in the first row, first column is negative. The equilibrium is asymptotically stable because the equilibrium, which exists if $r > a$, satisfies

$$y_1 > \frac{K_1}{2} \left(1 - \frac{a}{r}\right).$$

Thus the model (8.3) has one asymptotically stable equilibrium, namely $(0, K_2)$ if $r < a$, and an equilibrium (y_1, y_2) with $y_1 > 0, y_2 > 0$ if $r > a$.

The analysis of a model like (8.3) but with travel between patches in both directions such as

$$\begin{aligned}y_1' &= ry_1 \left(1 - \frac{y_1}{K_1}\right) - a_1 y_1 + a_2 y_2, \\y_2' &= ry_2 \left(1 - \frac{y_2}{K_2}\right) + a_1 y_1 - a_2 y_2,\end{aligned}\tag{8.5}$$

may be carried out in a similar manner, but the algebra is more complicated.

Exercises

1. Consider a metapopulation model in which the patches are islands and there is also a mainland with a large population that has a negligible risk of extinction,

$$p' = m(1 - p) + cp(1 - p) - ep.$$

Find the biologically meaningful equilibria and analyze their stability.

2. For the model (8.3) find the value of y_2 for the equilibrium when $r > a$.
3. Carry out the analysis of the model (8.5) with travel between patches, finding equilibria and determining their stability for all nonnegative values of the parameters r, a_1, a_2 .
4. [Britton (2002)] Consider a generalization of the model (8.1) in which there is a fraction p_0 of uninhabitable patches as well as a fraction p_1 of occupied patches and a fraction p_2 of habitable but uninhabited patches, with a rate λ of creation of habitable but uninhabited patches, a colonization rate c , a rate d of population extinction, and a rate e of patch destruction [Keymer et al. (2000)]. The corresponding model is

$$\begin{aligned}p_0' &= e(p_1 + p_2) - \lambda p_0, \\p_1' &= cp_1 p_2 - (d + e)p_1, \\p_2' &= \lambda p_0 - cp_1 p_2 + dp_1 - ep_2.\end{aligned}$$

- (i) Reduce the model to a two - dimensional system by letting $p_0 = 1 - p_1 - p_2$.
- (ii) Show that the reduced system has two equilibria with $p_2 > 0$, one of which has $p_1 = 0$ and one of which has $p_1 \neq 0$.
- (iii) Find the threshold quantity depending on the model parameters that distinguishes which of these two equilibria is asymptotically stable, and interpret this quantity ecologically. [Hint: Begin by finding the condition for the asymptotic stability of the equilibrium with $p_1 = 0$.]

8.3 A General Metapopulation Model

A metapopulation model may be regarded abstractly as a graph. We think of each population patch as a vertex of the graph, and each travel route between two patches as an edge of the graph. Since travel may be in either direction, the graph is *bi-directional*. The population of a patch is said to have direct access to another patch if there is a route linking the two patches. If there is a sequence of more than one route to another patch, the population is said to have indirect access to the second patch. For simplicity, we will always assume that a metapopulation is fully connected, that is, that there is a route linking every pair of patches.

Consider a metapopulation with n patches. It is convenient to define an ordering P_1, P_2, \dots, P_n of the patches, and then we define the *connection matrix* to be the matrix C whose (j, i) element is 1 if P_i has direct access to P_j and zero if P_i does not have direct access to P_j . Our description of a metapopulation model will include a movement matrix that combines the connection matrix with a description of the intensity of the connection.

Let N_i be a function of t denoting the number of individuals in patch i at time t . Then we may write a metapopulation model as the system of n differential equations

$$N'_i = f_i(N_i) + \sum_{j=1}^n m_{ij}N_j - \sum_{j=1}^n m_{ji}N_i. \quad (8.6)$$

Here, f_i describes the population dynamics within the patch P_i , the term $\sum_{j=1}^n m_{ij}N_j$ describes the inflow of individuals into the patch P_i from other patches, and the term $-\sum_{j=1}^n m_{ji}N_i$ describes the outflow of individuals from the patch P_i to other patches. It is assumed here that $m_{ii} = 0$ for $i = 1, 2, \dots, n$, since there is no travel assumed within a patch. If we write

$$m_{ii} = - \sum_{j=1, j \neq i}^n m_{ji},$$

we may rewrite (8.6) as

$$N'_i = f_i(N_i) + \sum_{j=1}^n m_{ij}N_j.$$

In vector form, this becomes

$$N' = f(N) + MN. \quad (8.7)$$

Here N is the (column) vector with components (N_1, N_2, \dots, N_n) and M is the matrix

$$\begin{bmatrix} -\sum_{j=1, j \neq 1}^n m_{j1} & m_{12} & \cdots & m_{1n} \\ m_{21} & -\sum_{j=1, j \neq 2}^n m_{j2} & \cdots & m_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ m_{n1} & m_{n2} & \cdots & -\sum_{j=1, j \neq n}^n m_{jn} \end{bmatrix}.$$

The sum of the elements of each column of M is zero, and it is possible to show, using some advanced results from matrix theory, that M has zero as an eigenvalue and all eigenvalues have nonpositive real part. If there are no internal dynamics in each patch, or if the population in each patch approaches an equilibrium, then the migration component, the solution of the linear system

$$N' = MN,$$

approaches a limiting value. Over a short time interval, it may be reasonable to neglect the internal population dynamics in each patch and consider only the migration component of the dynamics.

Example 1. Consider a metapopulation with two patches and no internal patch dynamics. The corresponding movement matrix is

$$\begin{bmatrix} -m_{21} & m_{12} \\ m_{21} & -m_{12} \end{bmatrix},$$

and the governing model is the two-dimensional linear system

$$\begin{aligned} N_1' &= -m_{21}N_1 + m_{12}N_2, \\ N_2' &= m_{21}N_1 - m_{12}N_2. \end{aligned}$$

Since $(N_1 + N_2)' = 0$, the total population size is a constant K ; we may let $N_2 = K - N_1$ to reduce the system to a single first-order linear differential equation

$$N_1' = -m_{21}N_1 + m_{12}(K - N_1) = m_{12}K - (m_{12} + m_{21})N_1.$$

The solution of this equation is

$$N_1(t) = N_1(0) + \frac{Km_{12}}{m_{12} + m_{21}}(1 - e^{-(m_{12} + m_{21})t}),$$

with limit as $t \rightarrow \infty$

$$N_1(\infty) = N_1(0) + \frac{Km_{12}}{m_{12} + m_{21}}.$$

Then

$$\begin{aligned} N_2(t) &= K - N_1(t) = K - N_1(0) - \frac{Km_{12}}{m_{12} + m_{21}}(1 - e^{-(m_{12} + m_{21})t}) \\ &= N_2(0) - \frac{Km_{12}}{m_{12} + m_{21}}(1 - e^{-(m_{12} + m_{21})t}), \end{aligned}$$

with limit as $t \rightarrow \infty$

$$N_2(\infty) = N_2(0) - \frac{Km_{12}}{m_{12} + m_{21}}.$$

Exercises

1. For a two-patch model with no internal patch dynamics in which the only travel is from patch 1 to patch 2 find the limiting population sizes for the two patches.
2. Consider an n patch model with no internal patch dynamics in which the only travel is from patch i to patch $i + 1$ ($i = 1, 2, \dots, n - 1$). Find the limiting population size of each patch.

8.4 A Metapopulation Model with Residence and Travel

In the previous section we formulated a metapopulation model for individuals who travel between patches, but do not necessarily have a home patch. Travel by humans follows a different pattern. Normally, each person has a home patch and may visit other patches but tends to return home. A description of such populations must distinguish between residents of a patch and residents of a different patch who are visiting.

Let $N_i(t)$ be the number of residents of the patch P_i who are present in their home patch at time t , and let $N_{ij}(t)$ be the number of residents of the patch P_i who are visiting the patch P_j at time t . Residents of P_i travel to other patches at a per capita rate σ_i , and the probability that a trip is to P_j is v_{ij} , with $v_{ii} = 0$ and $\sum_{j=1}^n v_{ij} = 1$. Persons who travel from P_i to P_j return to P_i at a rate ρ_{ij} , with $\rho_{ii} = 0$. Then the travel pattern of individuals satisfies the equations

$$N'_{ii} = \sum_{j=1}^n \rho_{ij} N_{ij} - \sigma_i N_{ii}, \quad (8.8)$$

$$N'_{ij} = \sigma_i v_{ij} N_{ii} - \rho_{ij} N_{ij}.$$

At equilibrium, since $v_{ii} = 0$,

$$N_{ij} = \frac{\sigma_i v_{ij}}{\rho_{ij}} N_{ii}, \quad (8.9)$$

$$\sum_{j \neq i} N_{ij} = \sum_{j \neq i} \frac{\sigma_i v_{ij}}{\rho_{ij}} N_{ii} = N_{ii} \sigma_i \sum_{j=1}^n \frac{v_{ij}}{\rho_{ij}}. \quad (8.10)$$

Let N_i be the total number of residents of P_i , a constant if we ignore internal patch dynamics and consider only the mobility pattern. Then $N_i = N_{ii} + \sum_{j \neq i} N_{ij}$, and combining this with (8.9), we obtain

$$N_i = \sigma_i N_{ii} \sum_{j=1}^n \frac{v_{ij}}{\rho_{ij}} + N_{ii} = N_{ii} \Gamma_i,$$

where

$$\Gamma_i = 1 + \sigma_i \sum_{j=1}^n \frac{v_{ij}}{\rho_{ij}}.$$

Thus

$$N_{ii} = \frac{N_i}{\Gamma_i}, \quad N_{ij} = N_i \frac{\sigma_i v_{ij}}{\rho_{ij} \Gamma_i} \quad (j \neq i). \quad (8.11)$$

The expressions (8.11) give the fraction of residents of a patch who are present in each patch at equilibrium.

There are some special cases of the general expressions (8.11) that represent different patterns. For example, the assumption of instantaneous return home ($\rho_{ij} \rightarrow 0$ for each i, j) gives the equilibrium population sizes $N_{ii} = N_i, N_{ij} = 0 (j \neq i)$, and in the limit all members of the population are in their home patch. The same is true if $\sigma \rightarrow \infty$, corresponding to the assumption of infinite waiting time between trips, or equivalently, no travel.

If everyone returns home at the same rate, $\rho_{ij} = \rho$ for all i, j , we obtain

$$N_{ii} = N_i \left(\frac{\rho}{\rho + \sigma_i} \right), \quad N_{ij} = N_i \left(\frac{\sigma_i v_{ij}}{\rho + \sigma_i} \right) \quad (j \neq i).$$

If visitors tend to stay at their destination permanently, corresponding to $\rho_{ij} \rightarrow 0$ for all i, j , we have a migration model. In this case, the model (8.8) can be solved explicitly to give

$$N_{ii}(t) = N_i e^{-\sigma_i t},$$

so that $N_{ii}(t) \rightarrow 0$. The equilibrium value of N_{ij} is $v_{ij} N_i$ because $v_{ii} = 0$ and $\sum_{j=1}^n v_{ij} = 1$. At equilibrium, the original inhabitants of a patch are distributed among the other patches, and the number in each patch is determined by the distribution of visitors v_{ij} .

Example 1. Consider a residency metapopulation with two patches and no internal patch dynamics. Then, since $v_{12} = v_{21} = 1$, the corresponding model is

$$\begin{aligned} N'_{11} &= \rho_{12} N_{12} - \sigma_1 N_{11}, \\ N'_{12} &= \sigma_1 N_{11} - \rho_{12} N_{12}, \\ N'_{21} &= \sigma_2 N_{22} - \rho_{21} N_{21}, \\ N'_{22} &= \rho_{21} N_{21} - \sigma_2 N_{22}. \end{aligned} \quad (8.12)$$

Since

$$(N_{11} + N_{12})' = (N_{21} + N_{22})' = 0,$$

the numbers of residents of the two patches are constants K_1, K_2 respectively. Thus

$$N_{12} = K_1 - N_{11}, \quad N_{21} = K_2 - N_{22},$$

and we may reduce (8.12) to a two-dimensional linear system

$$\begin{aligned} N'_{11} &= \rho_{12}K_1 - (\rho_{12} + \sigma_1)N_{11}, \\ N'_{22} &= \rho_{21}K_2 - (\rho_{21} + \sigma_2)N_{22}. \end{aligned} \quad (8.13)$$

The equilibrium of this system is

$$\begin{aligned} N_{11} &= \frac{\rho_{12}}{\rho_{12} + \sigma_1}K_1, & N_{12} &= \frac{\sigma_1}{\rho_{12} + \sigma_1}K_1, \\ N_{21} &= \frac{\sigma_2}{\rho_{21} + \sigma_2}K_2, & N_{22} &= \frac{\rho_{21}}{\rho_{21} + \sigma_2}K_2. \end{aligned}$$

Alternative formulations of mobility metapopulation models in the context of disease transmission may be found in [Castillo-Chavez, Song, and Zhang (2003), Khan, Arino, et al. (2009), Rios-Soto, Song, and Castillo-Chavez (2011)].

Exercises

1. Consider two populations with the same birth and death rates in two patches, with proportional travel rates in both directions, given by a system

$$\begin{aligned} y'_1 &= \Lambda - \mu y_1 - \sigma_1 y_1 + \sigma_2 y_2, \\ y'_2 &= \Lambda - \mu y_2 - \sigma_2 y_2 + \sigma_1 y_1. \end{aligned}$$

- (i) Find all biologically meaningful equilibria and analyze their stability.
 - (ii) Suppose the travel rate σ_2 is kept fixed but $\sigma_1 \rightarrow \infty$. What is the equilibrium state to which the system tends?
2. Suppose that the environment in patch 2 in the previous exercise is poor, so that there is an increased death rate, giving a model

$$\begin{aligned} y'_1 &= \Lambda - \mu y_1 - \sigma_1 y_1 + \sigma_2 y_2, \\ y'_2 &= \Lambda - (\mu + d)y_2 - \sigma_2 y_2 + \sigma_1 y_1. \end{aligned}$$

- (i) Find all biologically meaningful equilibria and analyze their stability.
- (ii) Suppose the additional death rate d tends to $+\infty$. What is the equilibrium state to which the system tends?

8.5 The Diffusion Equation

Partial differential equations arise in situations in which an unknown function depends on more than one variable. Many applications of mathematics in the physical and biological sciences involve partial differential equations, and there is a great variety of techniques for obtaining information about the behavior of solutions.

An example of a partial differential equation arising in a biological context comes from a population that may depend continuously on both time and location in space. Let us begin by considering the motion of particles. Here, by a particle we might mean an individual cell, a member of a population, or any object of a set in whose spatial distribution as a function of time we are interested.

Our approach is to take a small region of space and to form a balance equation that says that the rate of change of the number of particles in the region is equal to the rate at which particles flow out of the region minus the rate at which particles flow into the region plus the rate of creation of particles in the region.

We shall confine ourselves mainly to the case in which the dependence is with respect to a single space coordinate. Let us think of a tube whose cross section has constant area A , and let x denote the distance along the tube from some arbitrary starting point $x = 0$. We assume that the tube is a bounded region described by the inequalities $0 \leq x \leq L$. In a later section we will describe the analogous problem for a population that may move in a two-dimensional planar region.

Let $u(x, t)$ be the concentration of particles (number per unit volume) at location x at time t , meaning that in the portion of the tube between x and $x + h$, with volume Ah , the number of particles is approximately $Ahu(x, t)$. By “approximately” we mean that if h is small, the error in this approximation $Ahu(x, t)$ is smaller than a constant multiple of h^2 .

We let $J(x, t)$ be the flux of particles at location x at time t , by which we mean the time rate of the number of particles crossing a unit area in the positive direction. For every x_0 the net rate of flow into the region between x_0 and $x_0 + h$ is $AJ(x_0, t) - AJ(x_0 + h, t)$. We let $Q(x_0, t, u)$ be the net growth rate per unit length at location x_0 at time t , representing births and deaths.

We have a balance relation on the interval $x_0 \leq x \leq x_0 + h$, expressing the fact that the rate of change of population size at time t in this interval is equal to the growth rate of population in this interval plus the net flux, so that

$$hAu_t(x_0, t) = hAQ(x_0, t, u) - A[J(x_0 + h, t) - J(x_0, t)]. \quad (8.14)$$

If we divide (8.14) by Ah and then take the limit as $h \rightarrow 0$, we obtain the conservation law

$$u_t(x_0, t) = Q(x_0, t, u) - \frac{\partial J}{\partial x}(x_0, t),$$

valid for all t at every point x_0 . Since x_0 is arbitrary, this relation holds at every point, and we have the equation

$$u_t(x, t) = Q(x, t, u) - \frac{\partial J}{\partial x}(x, t) \quad (8.15)$$

We now specialize, temporarily, to the case in which there is no net growth rate in the interior of the tube, so that $Q(x, t) = 0$ for all $t \geq 0$, $0 \leq L$. In order to obtain a model which describes the population density $u(x, t)$, we must make some assumption that relates the rate of change of flux density $\frac{\partial J}{\partial x}$ and the population density $u(x, t)$. If the motion is random, then *Fick's law* says that the flux due to

random motion is approximately proportional to the rate of change of particle concentration, that is, that J is proportional to u_x . If population density decreases as x increases ($u_x < 0$), we would expect $J > 0$, so that J and u_x have opposite sign, and thus

$$J = -Du_x,$$

with D a constant called the diffusivity or diffusion coefficient. More generally, D could be a function of the location x , but we shall confine our attention to constant diffusivity. Equation (8.15) then becomes a second order-partial differential equation

$$u_t(x, t) = Du_{xx}(x, t). \quad (8.16)$$

The equation (8.16) is called the *heat* or *diffusion* equation.

An alternative derivation of the diffusion equation using random walks may be found in [Levin (1986)], an excellent introduction that highlights connections to biology.

In addition to the partial differential equation (8.16) we need to impose additional conditions to give a problem with a unique solution. It is possible to establish the following result.

Theorem: The diffusion equation (8.16) has a unique solution for $0 \leq x \leq L, 0 \leq t < \infty$, with a specified initial condition $u(x, 0) = f(x)$ for $0 \leq x \leq L$ and with specified boundary conditions for $x = 0$ and $x = L$. These boundary conditions may specify the value of u or the value of the partial derivative u_x for $x = 0$ and $x = L$.

This theorem tells us that a suitable problem consists of the diffusion equation together with an initial condition (a condition for $t = 0$) and boundary conditions (conditions for $x = 0, x = L$). Such a problem is called an *initial boundary value problem*.

Problems in an infinite tube defined by $-\infty < x < \infty$ for which no boundary conditions are required, or a semi-infinite tube $0 \leq x < \infty$ for which a boundary condition is required only at $x = 0$, will also be considered. In each case there is a unique bounded solution for a specified initial condition $u(x, 0) = f(x)$ for $-\infty \leq x < \infty$ or $u(x, 0) = f(x)$ for $0 \leq x < \infty$.

A boundary condition specifying that the solution must vanish at a boundary (called an absorbing boundary) may be taken to say that an individual leaving the region must die immediately or can never return. This is an idealization, but we may think of a large region with $u = 0$ far enough away. A boundary condition specifying that the partial derivative of the solution must vanish at the boundary may be taken to say that the population is confined to the region and there is no flow across the boundary.

There are several types of initial condition that may arise. One possibility is that particles are absent initially, $u(x, 0) = 0$, and enter through the boundary. A second possibility is that particles are inserted at a single point x_0 , $u(x, 0) = u_0 \delta(x - x_0)$. Here, $\delta(x)$ denotes the delta “function”, which is zero except for $x = 0$ and infinite at $x = 0$ in such a way that

$$\int_{-\infty}^{\infty} \delta(x) dx = 1, \quad (8.17)$$

and if f is any continuous function, then

$$\int_{-\infty}^{\infty} \delta(x - \xi) f(\xi) dx = f(x). \quad (8.18)$$

Obviously, this “function” is not a proper function, but if it is used in calculations using only the properties (8.17) and (8.18), it can actually lead to correct and useful results. A third kind of initial condition would be to specify a constant initial concentration, $u(x, 0) = u_0$ for $0 \leq x \leq L$.

Exercises

1. Suppose that the diffusivity is a function of population density, $D = D(u)$. Show that the partial differential equation (8.16) is replaced by

$$u_t = D(u)u_{xx} + D'(u)(u_x)^2.$$

8.6 Solution by Separation of Variables

In this section we will obtain solutions of some initial boundary value problems for the diffusion equation (8.16) that will be needed in applications in later sections. We use the method of separation of variables, which is used for solving many second order partial differential equations. It is **not** the same as the method of separation of variables for solving first order ordinary differential equations.

Example 1. Find a solution of the diffusion equation (8.16) for $0 \leq x \leq L$, $0 \leq t < \infty$, subject to the boundary conditions

$$u(0, t) = u(L, t) = 0 \quad (8.19)$$

and the initial condition

$$u(x, 0) = f(x), \quad 0 \leq x \leq L \quad (8.20)$$

with f a given smooth function.

Solution: We look for a solution of the form

$$u(x, t) = X(x)T(t), \quad (8.21)$$

where $X(x)$ is a smooth function defined for $0 \leq x \leq L$ and $T(t)$ is a smooth function defined for $t \geq 0$. From (8.21) we have

$$u_t(x, t) = X(x)T'(t), \quad u_x(x, t) = X'(x)T(t), \quad u_{xx}(x, t) = X''(x)T(t),$$

with the primes denoting ordinary derivatives. We substitute these expressions into (8.16), and see that (8.21) satisfies (8.16) if and only if

$$X(x)T'(t) = DX''(x)T(t), \quad 0 \leq x \leq L, \quad t \geq 0. \quad (8.22)$$

Since we are looking for nonzero solutions, neither $X(x)$ nor $T(t)$ can be identically zero. We may divide (8.22) by $DX(t)T(t)$ to obtain the separated form

$$\frac{T'(t)}{DT(t)} = \frac{X''(x)}{X(x)}, \quad (8.23)$$

for those values of t for which $T(t) \neq 0$ and those values of x for which $X(x) \neq 0$. The left side of (8.23) depends only on t and the right side of (8.23) depends only on x . Since these expressions must be equal for all values of x and t , each side of (8.23) must be equal to a constant, which we call k for the moment. Thus (8.23) implies the two relations

$$\frac{X''(x)}{X(x)} = \frac{T'(t)}{DT(t)} = k$$

or equivalently, the pair of ordinary differential equations

$$\begin{aligned} X''(x) - kX(x) &= 0, \\ T'(t) - kDT(t) &= 0. \end{aligned}$$

Because every solution of the second of these has the form $T(t) = ce^{kDt}$ and because we want solutions to be bounded as functions of t we require the constant k to be negative or zero. Therefore we replace k by $-\lambda^2$, where $\lambda \geq 0$, to give the pair of ordinary differential equations

$$\begin{aligned} X''(x) + \lambda^2 X(x) &= 0 \\ T'(t) + D\lambda^2 T(t) &= 0. \end{aligned} \quad (8.24)$$

The next step is to use the boundary conditions (8.19) to determine all possible values of λ that are compatible with the problem (8.16), (8.19). Substitution of the assumed form (8.21) into the boundary conditions (8.19) gives

$$\begin{aligned} u(0, t) &= X(0)T(t) = 0, \\ u(L, t) &= X(L)T(t) = 0. \end{aligned}$$

Since $T(t)$ is not identically zero, we must have

$$X(0) = 0, \quad X(L) = 0. \quad (8.25)$$

Every solution of the first differential equation in (8.24) has the form

$$X(x) = A + Bx \quad (\lambda = 0)$$

or

$$X(x) = A \cos \lambda x + B \sin \lambda x \quad (\lambda > 0).$$

If $\lambda = 0$, then (8.25) implies $A = 0$, $A + BL = 0$, so that $A = B = 0$ and $X(x)$ is identically zero. Thus we must have $\lambda > 0$, and every solution of the first differential equation in (8.24) has the form

$$X(x) = A \cos \lambda x + B \sin \lambda x.$$

Now (8.25) implies $A = 0$ and $B \sin \lambda L = 0$. We must have $B \neq 0$, since otherwise $X(x)$ would be identically zero. In order to have $B \sin \lambda L = 0$ we must have $\sin \lambda L = 0$ or $\lambda L = n\pi$ with n a positive integer. We obtain a sequence of admissible values, called eigenvalues, $\lambda_n = \frac{n\pi}{L}$ ($n = 1, 2, \dots$) and corresponding solutions

$$X_n(x) = B_n \sin \frac{n\pi x}{L}.$$

For each eigenvalue λ_n there is a corresponding solution $T_n(t)$ of $T'(t) + \lambda_n^2 DT(t) = 0$, namely $T_n(t) = e^{-D(\frac{n\pi}{L})^2 t}$, leading to a solution

$$u_n(x, t) = X_n(x)T_n(t) = B_n e^{-D(\frac{n\pi}{L})^2 t} \sin \frac{n\pi x}{L}$$

of the diffusion equation (8.16) and the boundary conditions (8.19).

It remains to try to satisfy the initial condition (8.20). In order to do this, we take a sum of solutions $u_n(x, t)$ for each positive integer n , also a solution of (8.16), (8.19) because (8.16) is a linear partial differential equation, and try to choose the coefficients B_n to satisfy (8.20). We have

$$\begin{aligned} u(x, t) &= \sum_{n=1}^{\infty} B_n e^{-D(\frac{n\pi}{L})^2 t} \sin \frac{n\pi x}{L}, \\ u(x, 0) &= \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{L} \quad (0 \leq x \leq L). \end{aligned}$$

Thus our problem is to choose B_n ($n = 1, 2, \dots$) so that

$$f(x) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{L} \quad (0 \leq x \leq L). \quad (8.26)$$

If this is possible, we may multiply both sides of (8.26) by $\sin \frac{k\pi x}{L}$ and integrate from 0 to L to obtain

$$\int_0^L f(x) \sin \frac{k\pi x}{L} dx = \sum_{n=1}^{\infty} B_n \int_0^L \sin \frac{k\pi x}{L} \sin \frac{n\pi x}{L} dx. \quad (8.27)$$

Here, we are assuming also that we may interchange the order of the summation and integration on the right side of (8.27). Since

$$\int_0^L \sin \frac{k\pi x}{L} \sin \frac{n\pi x}{L} dx = \begin{cases} 0 & (k \neq n), \\ \frac{L}{2} & (k = n), \end{cases}$$

(8.27) reduces to

$$\int_0^L f(x) \sin \frac{k\pi x}{L} dx = \frac{L}{2} B_k,$$

or

$$B_k = \frac{2}{L} \int_0^L f(x) \sin \frac{k\pi x}{L} dx. \quad (8.28)$$

Then

$$f(x) = \frac{2}{L} \sum_{n=1}^{\infty} \left(\int_0^L f(\xi) \sin \frac{\eta\pi\xi}{L} d\xi \right) \sin \frac{\eta\pi x}{L} \quad (8.29)$$

and

$$u(x, t) = \frac{2}{L} \sum_{n=1}^{\infty} e^{-D(\frac{n\pi}{L})^2 t} \left(\int_0^L f(\xi) \sin \frac{n\pi\xi}{L} d\xi \right) \sin \frac{\eta\pi x}{L} \quad (8.30)$$

is the solution of (8.16), (8.19), (8.20).

The series (8.29) is called the Fourier sine series expansion of $f(x)$ on the interval $0 \leq x \leq L$ and the expressions in (8.28) are called the Fourier coefficients. There is a remarkable theorem stating that if $f(x)$ is a reasonably well behaved function (specifically, if $f(x)$ and $f'(x)$ are piecewise continuous) on $0 \leq x \leq L$, then the Fourier series of $f(x)$ converges to $f(x)$ at every point of continuity of $f(x)$. At a point of discontinuity of $f(x)$ the Fourier series of $f(x)$ converges to the average of the two one-sided limits of $f(x)$ at this point of discontinuity. This means that the series (8.30) is a genuine solution of the initial boundary value problem (8.16), (8.19), (8.20). We may write the solution (8.30) in the form

$$u(x, t) = \int_0^L K(x, \xi, t) f(\xi) d\xi$$

with kernel

$$K(x, \xi, t) = \frac{2}{L} \sum_{n=1}^{\infty} e^{-(\frac{n\pi}{L})^2 Dt} \sin \frac{n\pi x}{L} \sin \frac{n\pi \xi}{L}.$$

Example 2. Find a solution of the diffusion equation (8.16) for $0 \leq x \leq L$, $0 \leq t < \infty$ subject to the boundary conditions

$$u_x(0, t) = u_x(L, t) = 0 \quad (8.31)$$

and the initial condition (8.20).

Solution: The separation of variables proceeds exactly as in Example 1 up to the point of examining the boundary conditions. These become

$$u_x(0, t) = X'(0)T(t), \quad u_x(L, t) = X'(L)T(t) = 0,$$

from which we obtain

$$X'(0) = 0, \quad X'(L) = 0. \quad (8.32)$$

Every solution of the differential equation $X''(x) + \lambda^2 X(x) = 0$ has the form

$$X(x) = A + Bx \quad (\lambda = 0)$$

or

$$X(x) = A \cos \lambda x + B \sin \lambda x \quad (\lambda > 0).$$

If $\lambda = 0$, (8.32) implies $B = 0$, giving the solution $X_0(x) = A_0$. If $\lambda > 0$, $X'(0) = 0$ implies $B = 0$, and $X'(L) = 0$ then implies $\sin \lambda L = 0$, or $\lambda L = n\pi$ with n a positive integer. We obtain a sequence of eigenvalues $\lambda_n = (\frac{n\pi}{L})^2$ ($n = 1, 2, \dots$) and corresponding solutions

$$X_n(x) = A_n \cos \frac{n\pi x}{L}.$$

For each eigenvalue λ_n there is a corresponding solution $T_n(t) = e^{-(\frac{n\pi}{L})^2 Dt}$ of $T'(t) + \lambda_n^2 DT(t) = 0$, leading to a solution

$$u_n(x, t) = X_n(x)T_n(t) = B_n e^{-D(\frac{n\pi}{L})^2 t} \cos \frac{n\pi x}{L}$$

of the diffusion equation (8.16) and the boundary conditions (8.31).

In order to satisfy the initial condition (8.20) as well, we take a sum of solutions $u_n(x, t)$ for each nonnegative integer n . We have

$$\begin{aligned} u(x, t) &= A_0 + \sum_{n=1}^{\infty} A_n e^{-D(\frac{n\pi}{L})^2 t} \cos \frac{n\pi x}{L}, \\ u(x, 0) &= A_0 + \sum_{n=1}^{\infty} A_n \cos \frac{n\pi x}{L} = f(x) \quad (0 \leq x \leq L). \end{aligned}$$

Integration and use of the relation

$$\int_0^L \cos \frac{n\pi x}{L} dx = 0 \quad (n = 1, 2, \dots)$$

gives $\int_0^L f(x) dx = LA_0$. Multiplication by $\cos \frac{k\pi x}{L}$ and integration gives, with the aid of

$$\int_0^L \cos \frac{k\pi x}{L} \cos \frac{n\pi x}{L} dx = \begin{cases} 0 & (k \neq n), \\ \frac{L}{2} & (k = n), \end{cases}$$

the result

$$A_n = \frac{2}{L} \int_0^L f(x) \cos \frac{n\pi x}{L} dx.$$

Then we have

$$f(x) = \frac{1}{L} \int_0^L f(x) dx + \frac{2}{L} \sum_{n=1}^{\infty} \left(\int_0^L f(\xi) \cos n\pi\xi d\xi \right) \cos \frac{n\pi x}{L} \quad (8.33)$$

and

$$u(x, t) = \frac{1}{L} \int_0^L f(s) dx + \frac{2}{L} \sum_{n=1}^{\infty} e^{-D(\frac{n\pi}{L})^2 t} \left(\int_0^L f(\xi) \cos n\pi\xi d\xi \right) \cos \frac{n\pi x}{L}. \quad (8.34)$$

The Fourier cosine series (8.33) converges to $f(x)$ if $f(x)$ is a piecewise smooth function and then (8.34) gives a genuine solution of the initial boundary value problem (8.16), (8.31), (8.20). Just as in Example 1, we may write this solution in an integral form, but now the kernel has the form

$$K(x, \xi, t) = \frac{1}{L} + \frac{2}{L} \sum_{n=1}^{\infty} e^{-(\frac{n\pi}{L})^2 Dt} \cos \frac{n\pi\xi}{L} \cos \frac{n\pi x}{L}.$$

Example 3. Find a solution of the diffusion equation (8.16) for $0 \leq x \leq L$, $0 \leq t < \infty$ subject to the boundary conditions

$$u(0, t) = u_x(L, t) = 0 \quad (8.35)$$

and the initial condition $u(x, 0) = u_0 \delta(x - x_0)$.

Solution: The separation of variables proceeds exactly as in Example 1 up to the point of examining the boundary conditions. These become

$$u(0, t) = X(0)T(t) = 0, \quad u_x(L, t) = X'(L)T(t) = 0,$$

from which we obtain

$$X(0) = 0, \quad X'(L) = 0. \quad (8.36)$$

Every solution of the differential equation $X''(x) + \lambda^2 X(x) = 0$ has the form

$$X(x) = A + Bx \quad (\lambda = 0)$$

or

$$X(x) = A \cos \lambda x + B \sin \lambda x \quad (\lambda > 0).$$

If $\lambda = 0$, (8.36) implies $A = 0$, and there is no non-trivial solution. If $\lambda > 0$, $X(0) = 0$ implies $A = 0$, and $X'(L) = 0$ then implies $\cos \lambda L = 0$, or

$$\sqrt{\lambda} L = \frac{(2n+1)\pi}{2}.$$

We obtain a sequence of eigenvalues

$$\lambda_n = \left[\frac{(2n+1)\pi}{2L} \right]^2$$

and corresponding solutions

$$X_n(x) = B_n \sin(\lambda_n x).$$

For each eigenvalue λ_n there is a corresponding solution $T_n(t) = e^{-\lambda_n D t}$ of $T'(t) + \lambda_n^2 D T(t) = 0$, leading to a solution

$$u_n(x, t) = X_n(x) T_n(t) = A_n e^{-D \frac{(2n+1)\pi}{2L} t} \sin \left(\frac{(2n+1)\pi}{2L} x \right)$$

of the diffusion equation (8.16) and the boundary conditions (8.35).

In order to satisfy the initial condition as well, we take a sum of solutions $u_n(x, t)$ for each nonnegative integer n . We have

$$\begin{aligned} u(x, t) &= \sum_{n=1}^{\infty} B_n e^{-D(\lambda_n)^2 t} \sin(\lambda_n x), \\ u(x, 0) &= \sum_{n=1}^{\infty} B_n \sin(\lambda_n x) = f(x) \quad (0 \leq x \leq L). \end{aligned}$$

Multiplication by $\sin(\lambda_k x)$ and integration gives, with the aid of

$$\int_0^L \sin(\lambda_k x) \sin(\lambda_n x) dx = \begin{cases} 0 & (k \neq n) \\ \frac{L}{2} & (k = n), \end{cases}$$

the result

$$B_n = 2L u_0 \int_0^L \delta(x - x_0) \sin(\lambda_n x) dx = \frac{2}{L} u_0 \sin(\lambda_n x_0).$$

Then we have

$$f(x) = \frac{2}{L} u_0 \sum_{n=1}^{\infty} \sin(\lambda_n x_0) \sin(\lambda_n x) \quad (8.37)$$

and

$$\begin{aligned} u(x, t) &= \frac{2}{L} u_0 \sum_{n=1}^{\infty} e^{-D \lambda_n^2 t} \sin(\lambda_n x) \sin(\lambda_n x_0) \\ &= \frac{2}{L} u_0 \sum_{n=1}^{\infty} e^{-D \left[\frac{(2n+1)\pi}{2L} \right]^2 t} \sin \left(\left[\frac{(2n+1)\pi}{2L} \right] x \right) \sin \left(\left[\frac{(2n+1)\pi}{2L} \right]^2 x_0 \right). \end{aligned} \quad (8.38)$$

The Fourier sine series (8.37) converges to $f(x)$ if $f(x)$ is a piecewise smooth function and then (8.38) gives a genuine solution of the initial boundary value problem.

Example 4. Find the solution of the diffusion equation (8.16), the initial condition (8.20), and the inhomogeneous boundary conditions

$$u(0, t) = u_1, \quad u(L, t) = u_2. \quad (8.39)$$

Solution: The solution (8.30) obtained in Example 1 is a transient, that is, a function that approaches zero as $t \rightarrow \infty$. Here, we would not expect the solution to be a pure transient. Accordingly, we try to find a solution of the form

$$u(x, t) = v(x) + w(x, t)$$

where $w(x, t)$ is a transient solution obtained as in Example 1 and $v(x)$ is a steady-state solution independent of time. Both the steady state solution $v(x)$ and the transient solution must satisfy the diffusion equation. Thus $v''(x) = 0$ for $0 < x < L$, so that $v(x) = c_1 + c_2x$. We require $v(x)$ to satisfy the boundary conditions (8.39), giving $v(0) = c_1 = u_1$ and $v(L) = c_1 + c_2L = u_2$, or $c_1 = u_1$, $c_2 = \frac{u_2 - u_1}{L}$. The steady state solution is

$$v(x) = u_1 + \frac{u_2 - u_1}{L}x.$$

The transient solution $w(x, t)$ satisfies the diffusion equation (8.16) and the homogeneous boundary conditions (8.19). The initial condition (8.20) becomes

$$u(x, 0) = v(x) + w(x, 0) = f(x)$$

or

$$w(x, 0) = f(x) - v(x) = f(x) - u_1 - \frac{u_2 - u_1}{L}x.$$

We now see that the transient solution $w(x, t)$ may be obtained just as in Example 1 using the Fourier sine series expansion of $f(x) - u_1 - \frac{u_2 - u_1}{L}x$. If we carry out the expansion, we obtain the solution

$$u(x, t) = u_1 + \frac{u_2 - u_1}{L}x + \frac{2}{L} \sum_{n=1}^{\infty} \left(\int_0^L \left[f(s) - \left\{ u_1 + \frac{u_2 - u_1}{L}s \right\} \right] \sin \frac{n\pi s}{L} ds \right) \sin \frac{n\pi x}{L}.$$

Example 5. Find the solution of the diffusion equation (8.16), the initial condition (8.20) and the periodic boundary conditions

$$u(L, t) = u(-L, t), \quad u_x(L, t) = u_x(-L, t), \quad (8.40)$$

in the region $-L \leq x \leq L$, $0 \leq t < \infty$.

Solution: The separation of variables proceeds as in Example 1, leading to boundary conditions

$$\begin{aligned} u(L, t) &= X(L)T(t) = u(-L, t) = X(-L)T(t), \\ u_x(L, t) &= X'(L)T(t) = u_x(-L, t) = X'(-L)T(t), \end{aligned}$$

giving

$$X(L) = X(-L), \quad X'(L) = X'(-L).$$

If $\lambda = 0$, we have $X(x) = A + Bx$, $X'(x) = B$. The periodicity conditions give $A + BL = A - BL$ or $B = 0$, so that $X_0(x) = A$. If $\lambda > 0$, the solution has the form $X(x) = A \cos \lambda x + B \sin \lambda x$, $X'(x) = -A\lambda \sin \lambda x + B\lambda \cos \lambda x$. The periodicity conditions give

$$\begin{aligned} A \cos L + B \sin \lambda L &= A \cos \lambda L - B \sin \lambda L \\ A\lambda \sin \lambda L + B\lambda \cos \lambda L &= -A\lambda \sin \lambda L + B\lambda \cos \lambda L. \end{aligned}$$

from which we have

$$B \sin \lambda L = 0, \quad A\lambda \sin \lambda L = 0$$

Since A and B cannot both be zero, it follows that $\sin \lambda L = 0$. We now have a sequence of eigenvalues $\lambda_n = (\frac{n\pi}{L})^2$ ($n = 1, 2, \dots$) and corresponding solutions

$$X_n(x) = A_n \cos \frac{n\pi x}{L} + B_n \sin \frac{n\pi x}{L}.$$

For each eigenvalue λ_n with $n > 0$ there is a corresponding solution $T_n(t)$ of $T'(t) + \lambda_n^2 DT(t) = 0$, namely $T_n(t) = e^{-(\frac{n\pi}{L})^2 Dt}$, leading to a solution of the partial differential equation (8.16) and the boundary conditions (8.40)

$$u_n(x, t) = X_n(x)T_n(t) = e^{-D(\frac{n\pi}{L})^2 t} [A_n \cos \frac{n\pi x}{L} + B_n \sin \frac{n\pi x}{L}]$$

For $\lambda = 0$ there is a constant solution $u_0(x, t) = A_0$.

To satisfy the initial condition (8.20) we need the *full Fourier series expansion* of $f(x)$ on the interval $-L \leq x \leq L$

$$f(x) = A_0 + \sum_{n=1}^{\infty} \left[A_n \cos \frac{n\pi x}{L} + B_n \sin \frac{n\pi x}{L} \right]. \quad (8.41)$$

In order to calculate the coefficients in this series, we proceed much as in Examples 1 and 2, multiplying (8.41) by $\cos \frac{k\pi x}{L}$ and by $\sin \frac{k\pi x}{L}$ and integrating from $-L$ to L , using the relations

$$\begin{aligned} \int_{-L}^L \sin \frac{k\pi x}{L} \sin \frac{n\pi x}{L} dx &= \begin{cases} 0 & (k \neq n), \\ L & (k = n), \end{cases} \\ \int_{-L}^L \cos \frac{k\pi x}{L} \cos \frac{n\pi x}{L} dx &= \begin{cases} 0 & (k \neq n), \\ L & (k = n), \end{cases} \\ \int_{-L}^L \sin \frac{n\pi x}{L} \cos \frac{n\pi x}{L} dx &= 0, \\ \int_{-L}^L 1 dx &= 2L. \end{aligned}$$

We obtain

$$\begin{aligned}
A_0 &= \frac{2}{L} \int_{-L}^L f(x) dx, \\
A_n &= \frac{1}{L} \int_{-L}^L f(x) \cos \frac{n\pi x}{L} dx, \\
B_n &= \frac{1}{L} \int_{-L}^L f(x) \sin \frac{n\pi x}{L} dx.
\end{aligned}$$

We now have the full Fourier series expansion

$$\begin{aligned}
f(x) &= \frac{2}{L} \int_{-L}^L f(\xi) d\xi \\
&+ \frac{1}{L} \sum_{n=1}^{\infty} \left[\cos \frac{n\pi x}{L} \int_{-L}^L f(\xi) \cos \frac{n\pi \xi}{L} d\xi + \sin \frac{n\pi x}{L} \int_{-L}^L f(\xi) \sin \frac{n\pi \xi}{L} d\xi \right].
\end{aligned}$$

We may use the trigonometric identity

$$\cos \frac{n\pi y}{L} \cos \frac{n\pi \xi}{L} + \sin \frac{n\pi y}{L} \sin \frac{n\pi \xi}{L} = \cos \frac{n\pi}{L} (x - \xi)$$

to write this as

$$f(x) = \frac{2}{L} \int_{-L}^L f(\xi) d\xi + \frac{1}{L} \sum_{n=1}^{\infty} \int_{-L}^L f(\xi) \cos \frac{n\pi}{L} (x - \xi) d\xi.$$

This gives the solution

$$u(x, t) = \frac{2}{L} \int_{-L}^L f(\xi) d\xi + \frac{1}{L} \sum_{n=1}^{\infty} e^{-D(\frac{n\pi}{L})^2 t} \int_{-L}^L f(\xi) \cos \frac{n\pi(x-\xi)}{L} d\xi. \quad (8.42)$$

This solution may be written in the form

$$u(x, t) = \int_{-L}^L K(x - \xi, t) f(\xi) d\xi$$

with kernel

$$K(x, t) = \frac{2}{L} + \frac{1}{L} \sum_{n=1}^{\infty} e^{-D(\frac{n\pi}{L})^2 t} \cos \frac{n\pi x}{L}.$$

Exercises

In Exercises 1–4, find separable solutions of the diffusion equation for $0 \leq x \leq L$, $0 \leq t < \infty$, satisfying the given boundary conditions:

1. $u(0, t) = 0, \quad u_x(L, t) = 0.$
2. $u_x(0, t) = 0, \quad u(L, t) = 0.$

3. $u_x(0,t) - \beta u(0,t) = 0, \quad u(L,t) = 0, [\beta > 0].$
4. $u_x(0,t) - \beta u(0,t) = 0, \quad u_x(L,t) + \gamma u(L,t) = 0,$
 $:(\beta > 0, \gamma > 0).$

In Exercises 5–10 find the solution of the diffusion equation satisfying the given boundary and initial conditions

5. $u(0,t) = 0, \quad u_x(L,t) = 0, \quad u(x,0) = x.$
6. $u_x(0,t) = 0, \quad u(L,t) = 0, \quad u(x,0) = x.$
7. $u(0,t) = 0, \quad u_x(L,t) = 0, \quad u(x,0) = f(x).$
8. $u_x(0,t) = 0, \quad u(L,t) = 0, \quad u(x,0) = f(x).$
9. $u(0,t) = u_1, \quad u_x(L,t) = 0, \quad u(x,0) = f(x).$
10. $u_x(0,t) = 0, \quad u(L,t) = u_2, \quad u(x,0) = f(x).$
11. For the diffusion equation on $0 \leq x \leq L, 0 \leq t < \infty$, with the boundary conditions $u(0,t) = 0, u_x(L,t) + \alpha u(L,t) = 0$, and $\alpha > 0$, show that the separated solutions are $u(x,t) = B_n e^{-D\lambda_n t} \sin \lambda_n x$, with eigenvalues λ_n determined by $\alpha \tan \lambda L = -\lambda$.
12. Find the solution of the diffusion equation on $0 \leq x \leq L, 0 \leq t < \infty$, with boundary conditions $u(0,t) = 0, u(L,t) = 0$, and initial condition $u(x,0) = u_0 \delta(x - \frac{L}{2})$, corresponding to a point release at the midpoint of the tube.
13. Show that $u(x,t) = u_0 [1 + \cos \frac{\pi x}{L} e^{-D(\frac{\pi^2}{L^2})t}]$ is a solution of the diffusion equation on $0 \leq x \leq L, 0 \leq t < \infty$ with boundary conditions $u_x(0,t) = 0, u_x(L,t) = 0$. What is the initial value $u(x,0)$?
14. Find the separated solutions of the diffusion equation for $0 \leq x \leq L, 0 \leq t < \infty$, that satisfy the boundary conditions $u_x(0,t) = \alpha u(0,t), u_x(L,t) + \alpha u(L,t) = 0$, where $\alpha > 0$.

8.7 Solutions in Unbounded Regions

In the previous section we applied the method of separation of variables to problems in finite x -regions $0 \leq x \leq L$ or $-L \leq x \leq L$. The boundary conditions gave a sequence of values for the separation constant, and this led to a sequence of solutions of the diffusion equation together with the boundary conditions. We then formed a sum of these solutions and were able to use the Fourier series of the initial function to determine coefficients to make this sum satisfy the initial condition as well.

For the diffusion equation in the region $-\infty < x < \infty, 0 \leq t < \infty$, there are no longer any boundary conditions. This means that there are no restrictions on the separation constant. The analogue of the sum used for a finite x -interval will be an integral over the parameter λ . In order to satisfy an initial condition $u(x,0) = f(x)$ for $-\infty < x < \infty$, we will need an integral analogue of the Fourier series expansion. There is such an analogue, called the Fourier integral, which we will use without proof. The solutions that we obtain in this section will be integral representations, and we may use the rules for differentiation of integrals to verify that they are in fact

solutions. The reader who is willing to accept these integral representations without motivation (although, of course, not without verification) may omit the derivation using the Fourier integral.

The Fourier integral representation that we need is

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi) \cos \lambda(x - \xi) d\xi d\lambda. \quad (8.43)$$

Another way to describe this would be to define

$$A(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\xi) \cos \lambda \xi d\xi, \quad (8.44)$$

$$B(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\xi) \sin \lambda \xi d\xi,$$

and then write

$$f(x) = \int_{-\infty}^{\infty} [A(\lambda) \cos \lambda x + B(\lambda) \sin \lambda x] d\lambda. \quad (8.45)$$

To see that (8.44) and (8.45) are equivalent to (8.43) we calculate

$$\begin{aligned} & \int_{-\infty}^{\infty} [A(\lambda) \cos \lambda x + B(\lambda) \sin \lambda x] d\lambda \\ &= \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \{f(\xi) \cos \lambda \xi \cos \lambda x + f(\xi) \sin \lambda \xi \sin \lambda x\} d\xi \right] d\lambda \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi) \cos \lambda(x - \xi) d\xi d\lambda. \end{aligned}$$

It is possible to prove that the expansion (8.43) is valid if $f(x)$ is piecewise differentiable on $-\infty < x < \infty$ and the infinite integral $\int_{-\infty}^{\infty} |f(x)| dx$ converges.

We now consider the problem on $-\infty < x < \infty$, $0 \leq t < \infty$, consisting of the diffusion equation (8.16) and the initial condition (8.20) on $-\infty < x < \infty$. The usual separation of variables procedure leads to solutions $u(x, t) = X(x)T(t) = [A(\lambda) \cos \lambda x + B(\lambda) \sin \lambda x] e^{-D\lambda^2 t}$ for every λ , $-\infty < \lambda < \infty$. In order to satisfy the initial condition, we form the “sum”

$$u(x, t) = \int_{-\infty}^{\infty} e^{-D\lambda^2 t} [A(\lambda) \cos \lambda x + B(\lambda) \sin \lambda x] d\lambda$$

with

$$u(x, 0) = f(x) = \int_{-\infty}^{\infty} [A(\lambda) \cos \lambda x + B(\lambda) \sin \lambda x] d\lambda.$$

Then $A(\lambda)$ and $B(\lambda)$ are given by (8.44), and $f(x)$ is given by (8.45). Thus we have

$$\begin{aligned}
 u(x, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-D\lambda^2 t} \left[\int_{-\infty}^{\infty} f(\xi) \cos \lambda(x - \xi) d\xi \right] d\lambda \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\xi) \left[\int_{-\infty}^{\infty} e^{-D\lambda^2 t} \cos \lambda(x - \xi) d\lambda \right] d\xi.
 \end{aligned} \tag{8.46}$$

In order to evaluate the inner integral in (8.46) we use the formula

$$\int_{-\infty}^{\infty} e^{-ax^2} \cos bx dx = \sqrt{\frac{\pi}{a}} e^{-\frac{b^2}{4a}}, \tag{8.47}$$

valid for $a > 0$, with $a = Dt$, $b = x - \xi$, to give

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\xi) \sqrt{\frac{\pi}{Dt}} e^{-\frac{(x-\xi)^2}{4Dt}} d\xi = \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} e^{-\frac{(x-\xi)^2}{4Dt}} f(\xi) d\xi. \tag{8.48}$$

We may write (8.48) in the form

$$u(x, t) = \int_{-\infty}^{\infty} K(x - \xi, t) f(\xi) d\xi \tag{8.49}$$

with

$$K(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}. \tag{8.50}$$

The function $K(x, t)$ has the property that

$$\int_{-\infty}^{\infty} K(x, t) dx = 1,$$

since

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} \quad (a > 0).$$

If $x \neq 0$, we have

$$K(x, 0) = \lim_{t \rightarrow 0} K(x, t) = \lim_{t \rightarrow 0} \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} = \lim_{s \rightarrow \infty} \frac{\sqrt{s}}{\sqrt{4\pi D}} e^{-\frac{x^2}{4D}s} = 0.$$

However, if $x = 0$, $\lim_{t \rightarrow 0} K(x, 0)$ is infinite. This indicates that $K(x, 0) = \delta(x)$.

To show that $K(x, t)$ is a solution of the diffusion equation, we use the rules for derivatives of integrals and the chain rule to calculate

$$\begin{aligned}
K_t(x, t) &= \frac{1}{\sqrt{4\pi D}} \left(-\frac{1}{2} t^{-3/2} e^{-\frac{x^2}{4Dt}} \right) + \frac{1}{\sqrt{4\pi D}} \left(t^{-1/2} e^{-\frac{x^2}{4Dt}} \frac{x^2}{4Dt^2} \right) \\
&= \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \left(\frac{x^2}{4Dt^2} - \frac{1}{2t} \right), \\
K_x(x, t) &= -\frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \frac{1}{2Dt} + \frac{1}{\sqrt{4\pi Dt}} \frac{x}{2Dt} e^{-\frac{x^2}{4Dt}} \frac{x}{2Dt} \\
&= \frac{1}{4\pi Dt} e^{-\frac{x^2}{4Dt}} \left(\frac{x^2}{4D^2t^2} - \frac{1}{2Dt} \right), \\
K_{xx}(x, t) &= -\frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \frac{1}{2Dt} + \frac{1}{\sqrt{4\pi Dt}} \frac{x}{2Dt} e^{-\frac{x^2}{4Dt}} \frac{x}{2Dt} \\
&= \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \left(\frac{x^2}{4D^2t^2} - \frac{1}{2Dt} \right) = \frac{1}{D} K_t(x, t).
\end{aligned}$$

We have now established that the function $K(x, t)$ is the solution of the diffusion equation on $-\infty < x < \infty$, $0 \leq t < \infty$ that satisfies the initial condition $u(x, 0) = \delta(x)$. It is called the fundamental solution of the diffusion equation for this region because it has the property that for an arbitrary initial function $f(x)$, the solution of the diffusion equation that satisfies the initial condition $u(x, 0) = f(x)$ is given by (8.49). It is easy to see, using the rules for differentiation of integrals and the fact that $K(x, t)$ satisfies the diffusion equation, that the function $u(x, t)$ defined by (8.49) is a solution of the diffusion equation. In addition,

$$u(x, 0) = \int_{-\infty}^{\infty} K(x - \xi, 0) f(\xi) d\xi = f(x),$$

using the property (8.18) of the delta function. Thus knowledge of the fundamental solution for $-\infty < x < \infty$, $0 \leq t < \infty$ enables us to solve the initial value problem in this region for any initial function.

Example 1. Solve the initial value problem

$$u_t = Du_{xx} + au, \quad u(x, 0) = f(x),$$

for $-\infty < x < \infty$, $0 \leq t < \infty$.

Solution: We make the change of dependent variable $u(x, t) = e^{at} v(x, t)$. Since

$$\begin{aligned}
u_t(x, t) &= e^{at} v_t(x, t) + ae^{at} v(x, t), \\
u_{xx}(x, t) &= e^{at} v_{xx}(x, t),
\end{aligned}$$

the partial differential equation reduces to

$$e^{at} v_t(x, t) + ae^{at} v(x, t) = De^{at} v_{xx}(x, t) + ae^{at} v(x, t),$$

or

$$v_t(x, t) = Dv_{xx}(x, t). \quad (8.51)$$

The solution of the partial differential equation (8.51) together with the initial condition $v(x, 0) = u(x, 0) = f(x)$ is, by (8.48),

$$v(x, t) = \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} e^{-\frac{(x-\xi)^2}{4Dt}} f(\xi) d\xi.$$

Thus the solution of the problem is

$$u(x, t) = \frac{e^{at}}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} e^{-\frac{(x-\xi)^2}{4Dt}} f(\xi) d\xi.$$

In particular, the solution with initial condition $u(x, 0) = u_0\delta(x)$, corresponding to an initial release of u_0 members at $x = 0$, is

$$u(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{\left(at - \frac{x^2}{4Dt}\right)}.$$

Another type of unbounded region is $0 \leq x < \infty$, $0 \leq t < \infty$, corresponding to a semi-infinite tube beginning at $x = 0$. For the diffusion equation in this region it is appropriate to assign a boundary condition such as $u(0, t) = 0$ or $u_x(0, t) = 0$ for $0 \leq t < \infty$ at $x = 0$, and an initial condition $u(x, 0) = f(x)$ for $0 \leq x < \infty$. In order to solve the initial boundary value problem consisting of the diffusion equation (8.16) and the boundary condition

$$u(0, t) = 0, \quad 0 \leq t < \infty, \quad (8.52)$$

with the initial condition

$$u(x, 0) = f(x), \quad 0 \leq x < \infty, \quad (8.53)$$

we will need to make use of the *Fourier sine integral* representation. This representation is

$$f(x) = \frac{2}{\pi} \int_0^{\infty} \sin \lambda x \left(\int_0^{\infty} f(\xi) \sin \lambda \xi d\xi \right) d\lambda,$$

or

$$f(x) = \int_0^{\infty} B(\lambda) \sin \lambda x d\lambda,$$

where

$$B(\lambda) = \frac{2}{\pi} \int_0^{\infty} f(\xi) \sin \lambda \xi d\xi, \quad (8.54)$$

and it is valid for every piecewise differentiable function $f(x)$ on $0 \leq x < \infty$ for which the infinite integral $\int_0^{\infty} |f(\xi)| d\xi$ converges.

The process of separation of variables for the diffusion equation (8.16) together with the boundary condition (8.52) leads to solutions

$$u(x, t) = X(x)T(t) = B(\lambda) \sin \lambda x e^{-D\lambda^2 t}$$

for every λ , $0 < \lambda < \infty$. In order to satisfy the initial condition (8.53) we form the “sum”

$$u(x, t) = \int_{-\infty}^{\infty} e^{-D\lambda^2 t} B(\lambda) \sin \lambda x d\lambda$$

with

$$u(x, 0) = f(x) = \int_{-\infty}^{\infty} B(\lambda) \sin \lambda x dx.$$

Then $B(\lambda)$ is given by (8.54) and $f(x)$ is given by (8.54). Thus we have

$$\begin{aligned} u(x, t) &= \frac{2}{\pi} \int_0^{\infty} e^{-D\lambda^2 t} \left(\int_0^{\infty} f(\xi) \sin \lambda \xi d\xi \right) \sin \lambda x d\lambda \\ &= \frac{2}{\pi} \int_0^{\infty} f(\xi) \left(\int_0^{\infty} e^{-D\lambda^2 t} \sin \lambda x \sin \lambda \xi d\lambda \right) d\xi. \end{aligned} \quad (8.55)$$

In order to evaluate the inner integral in (8.55), we use the trigonometric identity

$$\sin \lambda x \sin \lambda \xi = \frac{1}{2} [\cos \lambda (x - \xi) - \cos \lambda (x + \xi)]$$

and the integration formula

$$\int_0^{\infty} e^{-ax^2} \cos bx dx = \frac{1}{2} \sqrt{\frac{\pi}{a}} e^{-\frac{b^2}{4a}},$$

analogous to (8.47), to obtain

$$\begin{aligned} &\int_0^{\infty} e^{-D\lambda^2 t} \sin \lambda x \sin \lambda \xi d\lambda \\ &= \frac{1}{2} \int_0^{\infty} e^{-D\lambda^2 t} \cos \lambda (x - \xi) d\lambda - \frac{1}{2} \int_0^{\infty} e^{-D\lambda^2 t} \cos \lambda (x + \xi) d\lambda \\ &= \frac{1}{4} \sqrt{\frac{\pi}{Dt}} \left(e^{-\frac{(x-\xi)^2}{4Dt}} - e^{-\frac{(x+\xi)^2}{4Dt}} \right). \end{aligned}$$

Then we have

$$\begin{aligned} u(x, t) &= \frac{2}{\pi} \frac{1}{4} \sqrt{\frac{\pi}{Dt}} \int_0^{\infty} f(\xi) \left(e^{-\frac{(x-\xi)^2}{4Dt}} - e^{-\frac{(x+\xi)^2}{4Dt}} \right) d\xi \\ &= \frac{1}{2\sqrt{\pi Dt}} \int_0^{\infty} f(\xi) \left(e^{-\frac{(x-\xi)^2}{4Dt}} - e^{-\frac{(x+\xi)^2}{4Dt}} \right) d\xi. \end{aligned} \quad (8.56)$$

Example 2. Find the solution of the diffusion equation (8.16), the boundary condition (8.52), and the initial condition (8.53) with a constant initial distribution $f(x) = u_0$ ($0 \leq x < \infty$).

Solution: The representation (8.56) gives the solution

$$u(x, t) = \frac{u_0}{2\sqrt{\pi Dt}} \int_0^\infty \left(e^{-\frac{(x-\xi)^2}{4Dt}} - e^{-\frac{(x+\xi)^2}{4Dt}} \right) d\xi$$

The changes of variable $x - \xi = -z\sqrt{4Dt}$ in the first integral and $x + \xi = z\sqrt{4Dt}$ in the second integral reduces this to

$$\begin{aligned} u(x, t) &= \frac{u_0}{2\sqrt{\pi Dt}} \sqrt{4Dt} \left[\int_{-\frac{x}{\sqrt{4Dt}}}^\infty e^{-z^2} dz - \int_{\frac{x}{\sqrt{4Dt}}}^\infty e^{-z^2} dz \right] \\ &= \frac{u_0}{\sqrt{\pi Dt}} \left[\int_{-\frac{x}{\sqrt{4Dt}}}^\infty e^{-z^2} dz - \int_{\frac{x}{\sqrt{4Dt}}}^\infty e^{-z^2} dz \right] \\ &= \frac{u_0}{\sqrt{\pi}} \int_{-\frac{x}{\sqrt{4Dt}}}^{\frac{x}{\sqrt{4Dt}}} e^{-z^2} dz = \frac{2u_0}{\sqrt{\pi}} \int_0^{\frac{x}{\sqrt{4Dt}}} e^{-z^2} dz. \end{aligned} \quad (8.57)$$

The integral in (8.57) cannot be evaluated analytically but has been tabulated. The *error function* $\text{erf}(x)$ is defined by

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz,$$

and we may write the solution (8.57) as

$$u(x, t) = u_0 \text{erf} \left(\frac{x}{\sqrt{4Dt}} \right).$$

It should be noted that the initial function in this example does not satisfy the requirement of the Fourier sine integral representation that $\int_0^\infty |f(x)| dx$ converge. Nevertheless, it may be verified that (8.57) is actually a solution of the given problem.

Example 3. Find the solution of the diffusion equation (8.16), the boundary condition

$$u(0, t) = u_0 \quad (0 \leq t < \infty),$$

and the initial condition

$$u(x, 0) = 0 \quad (0 \leq x < \infty).$$

Solution: Let $u_0(x) = u_0$ be the steady-state solution, satisfying $u_0(0, t) = u_0$, $u_0(x, 0) = u_0$. Then the desired solution $u(x, t)$ may be written

$$u(x, t) = u_0(x) + v(x, t),$$

where $v(x, t)$ satisfies the boundary condition $v(0, t) = 0$ and the initial condition $v(x, 0) = -u_0$. Then $v(x, t)$ is just the solution obtained in Example 2 with u_0 replaced by $-u_0$, namely

$$v(x, t) = -u_0 \operatorname{erf} \frac{x}{\sqrt{4Dt}},$$

and the desired solution is

$$u(x, t) = u_0 \left[1 - \operatorname{erf} \left(\frac{x}{\sqrt{4Dt}} \right) \right].$$

Exercises

1. Verify that $u(x, t) = u_0[1 - \operatorname{erf}(\frac{x}{\sqrt{4Dt}})]$ satisfies the diffusion equation.
2. Find the solution of the diffusion equation on $-\infty < x < \infty$, $0 \leq t < \infty$ such that

$$u(x, 0) = \begin{cases} 0 & (x < a), \\ u_0 & (a \leq x \leq b), \\ 0 & (x > b). \end{cases}$$

3. Find the solution of the diffusion equation on $-\infty < x < \infty$ such that

$$u(x, 0) = \begin{cases} u_0 & (x < 0), \\ 0 & (x > 0). \end{cases}$$

4. Find the solution of the diffusion equation on $0 \leq x < \infty$, $0 \leq t < \infty$ such that $u(0, t) = 0$ and

$$u(x, 0) = \begin{cases} 1 & (0 \leq x \leq L), \\ 0 & (x > L). \end{cases}$$

8.8 Linear Reaction–Diffusion Equations

If there is a creation term $Q(x, t, u) \neq 0$ in the partial differential equation (8.15) and the flux is diffusive, then (8.15) becomes

$$u_t(x, t) = Du_{xx}(x, t) + Q(x, t, u). \quad (8.58)$$

The equation (8.58) is called a *reaction–diffusion* equation; the term $Q(x, t, u)$ represents reactions such as the birth and death of population members, while the term $Du_{xx}(x, t)$ represents diffusion of the population in space. In addition to the equation (8.58), the model may require initial conditions at $t = 0$ and/or boundary conditions at $x = 0$ and $x = L$. For example, if we do not allow individuals to cross the boundaries at $x = 0$ and $x = L$, we would impose the boundary conditions

$$u_x(t, 0) = 0, \quad u_x(L, 0) = 0. \quad (8.59)$$

In general, problems consisting of a partial differential equation (8.58) together with the boundary condition (8.59) cannot be solved explicitly, but it is sometimes possible to state results about the qualitative nature of the solutions.

In this section we shall describe some properties of solutions of reaction–diffusion equations when the creation term $Q(x, t, u)$ is a linear function of the population density u . One of the first applications of equations of reaction–diffusion type to a population model was the work of Skellam (1951) on the dispersal of a population that grows exponentially, so that $f(u) = au$. We consider the release of a population with u_0 members at time $t = 0$ from the location $x = 0$, governed by the equation

$$u_t(x, t) = Du_{xx}(x, t) + au(x, t) \quad (8.60)$$

with initial condition $u(x, 0) = \delta(x)$. It is possible to solve this problem analytically, as we have seen in Example 1 of the preceding section obtaining

$$u(x, t) = \frac{u_0}{2\sqrt{\pi Dt}} e^{\left(at - \frac{x^2}{4Dt}\right)}. \quad (8.61)$$

We may deduce from the solution (8.60) that the population spreads with an asymptotic speed $2\sqrt{aD}$. To see this, we write the level curves of the solution given by (8.61) in the form

$$\frac{u_0}{2\sqrt{\pi Dt}} e^{\left(at - \frac{x^2}{4Dt}\right)} = c.$$

Taking natural logarithms of both sides of this equation, we have

$$\frac{x^2}{t^2} = 4aD - \frac{4D}{t} \log \left(\frac{c}{u_0} \sqrt{4\pi Dt} \right),$$

Since

$$\lim_{t \rightarrow \infty} \frac{4D}{t} \log \left(\frac{c}{u_0} \sqrt{4\pi Dt} \right) = \lim_{t \rightarrow \infty} \frac{4D}{t} \left[\log \frac{c}{u_0} + \frac{1}{2} \log(4\pi Dt) \right] = 0$$

we have $\lim_{t \rightarrow \infty} \frac{x^2}{t^2} = 4aD$. Thus as $t \rightarrow \infty$, $x \approx 2\sqrt{aDt}$. This shows that the population spreads out from the origin with asymptotic velocity $2\sqrt{aD}$.

The same partial differential equation was studied on the interval $0 \leq x \leq L$ by Kierstad & Slobodkin (1953) with absorbing boundary conditions

$$u(0, t) = u(L, t) = 0. \quad (8.62)$$

The problem (8.60), (8.62) has been solved by separation of variables in Example 1, Section 8.6, where we obtained the separable solutions

$$e^{\left(a - \frac{n^2\pi^2 D}{L^2}\right)t} \sin(n\pi x/L).$$

The most general solution is a superposition of such solutions, a sum over n of terms of this form. The population grows only if, at least for some n , there is a positive exponent. In other words, growth of the population requires $a > \pi^2 D/L^2$, or $L > \pi \sqrt{D/a}$. Thus a positive population size can be maintained in the region only if the region is large enough. If the diffusion is rapid enough that individuals leave the region before reproducing, the population will tend to die out. This suggests a possible explanation for the phenomenon of populations that are distributed in discrete patches, namely that survival is possible only in sufficiently large regions. There are other possible explanations that have been advanced, such as the idea of diffusive instability with diffusion tending to destabilize a population which would tend to an asymptotically stable equilibrium in the absence of diffusion.

Exercises

1. Solve the initial value problem

$$u_t = Du_{xx} + g(t), \quad u(x, 0) = f(x),$$

where $g(t)$ is a given function of t , for $-\infty < x < \infty$, $0 \leq t < \infty$. [Hint: Begin by making a change of variable $u(x, t) = v(x, t) + h(t)$ with the function $h(t)$ chosen to reduce the problem as in Example 1, Section 8.6.]

8.9 Nonlinear Reaction–Diffusion Equations

In this section we shall consider reaction–diffusion equations containing a nonlinear growth rate $g(u)$ with $g(0) = g(K) = 0$ and $g(u) > 0$ for $0 < u < K$. Thus, we shall consider the equation

$$u_t(x, t) = Du_{xx}(x, t) + g(u). \quad (8.63)$$

We begin by looking for solutions $u(t)$ that are independent of x . Then $u_{xx} = 0$ and the equation (8.63) reduces to the ordinary differential equation

$$u' = g(u). \quad (8.64)$$

The total population size is $\frac{1}{L} \int_0^L u(x, t) dx$, and if u is independent of x , this reduces to $u(t)$. Thus we have an ordinary differential equation model for the total population size of the type studied in Chapter 1. As we know, every bounded solution of (8.64) approaches the equilibrium K as $t \rightarrow \infty$. In order to show that this equilibrium solution is asymptotically stable in the sense that solutions of (8.63) whose initial total population size is sufficiently close to K approach it as $t \rightarrow \infty$, we linearize about the equilibrium K , obtaining a linearized equation of the form (8.60) with $a = g'(K) < 0$. From the fundamental solution (8.61) we see that every solution of

this linearization approaches zero as $t \rightarrow \infty$, and this proves the asymptotic stability of the equilibrium K . However, we must remember that this asymptotic stability refers only to solutions starting close enough to the equilibrium and does not rule out the possibility of other kinds of solution.

We next consider the possibility of patterns in space, that is, solutions that are independent of t . Then $u_t = 0$, and the equation (8.63) reduces to the second-order ordinary differential equation

$$Du'' + g(u) = 0.$$

The equivalent first-order system is $u' = v$, $v' = -g(u)/D$ with equilibria at $(0,0)$ and $(K,0)$. The matrix of the linearization at an equilibrium x_∞ is

$$\begin{bmatrix} 0 & 1 \\ -g'(u_\infty) & 0 \end{bmatrix}. \quad (8.65)$$

Since $g'(0) > 0$, the equilibrium $(0,0)$ is a center, and the equilibrium $(K,0)$ is a saddle point for (8.65). By an analysis of the trajectories of (8.65) it is possible to show that all nonconstant solutions are unstable and thus that no pattern is biologically realistic. For reaction–diffusion systems governing interacting populations in more than one space dimension, the variety of possible behaviors is great, and stable patterns can occur.

Nonlinear reaction–diffusion equations may have traveling wave solutions. A wave is a function of the form $u(x,t) = U(x-ct)$ for some constant c . The reason for the description of such a function as a wave is that if $x-ct$ is constant, then u is constant, and this implies a motion in the direction of increasing x with velocity c . If $u(x,t)$ is a wave, then $u_x(x,t) = U'(x-ct)$, $u_{xx}(x,t) = U''(x-ct)$, and $u_t(x,t) = (-c)U'(x-ct)$. Thus for such a wave to be a solution of (8.63), we must have

$$-cU'(x-ct) = g[U(x-ct)] + DU''(x-ct).$$

If we let $z = x-ct$, then the function $U(z)$ must satisfy the following second order ordinary differential equation in z :

$$DU'' + cU' + g(U) = 0.$$

The equivalent system is

$$U' = V, \quad V' = -\frac{g(U)}{D} - c\frac{V}{D}, \quad (8.66)$$

whose equilibria are solutions of the pair of equations $V = 0$, $g(U) = 0$. Thus the equilibria of (8.66) are $(0,0)$ and $(K,0)$. The matrix of the linearization at an equilibrium $(U_\infty, 0)$ is

$$\begin{bmatrix} 0 & 1 \\ -\frac{g'(U_\infty)}{D} & -\frac{c}{D} \end{bmatrix}.$$

Since $g'(K) < 0$, it is easy to see that $(K, 0)$ is a saddle point. The equilibrium at $(0, 0)$ is asymptotically stable, a node if $c^2 > 4Dg'(0)$ and a spiral point if $c^2 < 4Dg'(0)$. By studying the phase portrait of the system it is possible to show that there is an orbit starting at the saddle point as $z \rightarrow -\infty$ that tends to the equilibrium at $(0, 0)$ as $z \rightarrow \infty$ if the equilibrium at $(0, 0)$ is a node, that is, if $c^2 > 4Dg'(0)$. This orbit corresponds to a wave traveling to the right. Every other orbit of the system (8.66) either has U unbounded as $z \rightarrow -\infty$ or as $z \rightarrow \infty$ or includes negative values of U . Such orbits are not reasonable biologically; only the connecting orbit is acceptable for representing a wave. In fact, for the connecting orbit, U remains between 0 and K . For the system (8.66) the independent variable z represents distance along the wave, and the independent variables U and V represent population density and rate of change of population density in space respectively (Figure 8.1).

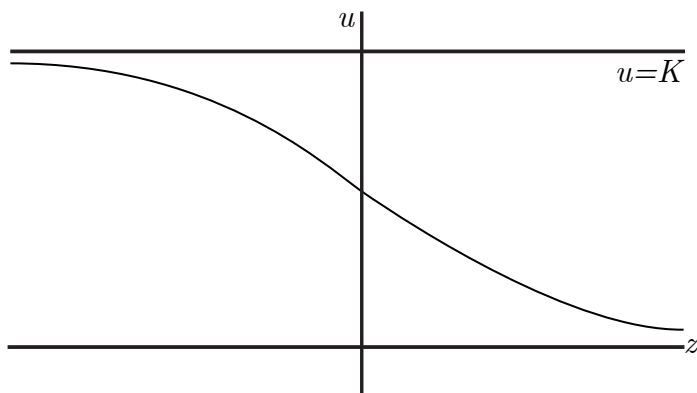


Fig. 8.1 A connecting orbit.

Example 1. Consider a logistic population growth model with diffusion,

$$u_t(x, t) = Du_{xx}(x, t) + ru(x, t) \left(1 - \frac{u(x, t)}{K} \right). \quad (8.67)$$

The same partial differential equation occurs as the Fisher equation for the spatial distribution of an advantageous gene. This example fits the general framework above exactly, and we may deduce that there are traveling wave solutions.

Because of the requirement that the equilibrium $(0, 0)$ must be a node in order to have a connecting orbit, the minimum velocity for a traveling-wave solution is $c^* = 2\sqrt{g'(0)D}$. It is possible to show that in some cases every solution behaves as $t \rightarrow \infty$ like a wave of the form $U(x - x_0 - c^*t)$ for some x_0 .

8.9.1 Two-Species Interactions

Let us consider a system of two interacting populations as discussed in Chapter 3, either a predator–prey system or two species in competition, with both populations diffusing in a one-dimensional space, possibly with different diffusivities. Such a system is governed by a system of two reaction–diffusion equations, of the form

$$\begin{aligned}u_t(x, t) &= D_1 u_{xx}(x, t) + F(u, v), \\v_t(x, t) &= D_2 v_{xx}(x, t) + G(u, v).\end{aligned}\tag{8.68}$$

Suppose that there is a spatially uniform steady state, that is, an equilibrium (u_∞, v_∞) of the system of ordinary differential equations

$$\begin{aligned}u' &= F(u, v), \\v' &= G(u, v).\end{aligned}\tag{8.69}$$

If we linearize about this equilibrium, we obtain a linear reaction–diffusion system of the form

$$\begin{aligned}u_t(x, t) &= D_1 u_{xx}(x, t) + au + bv, \\v_t(x, t) &= D_2 v_{xx}(x, t) + cu + dv.\end{aligned}\tag{8.70}$$

We look for separated solutions of (8.70) in the form

$$u(x, t) = U(x)T(t), \quad v(x, t) = V(x)T(t).$$

We obtain

$$UT' = D_1 U''T + aUT + bVT, \quad VT' = d_2 V''T + cUT + dVT,$$

or

$$\frac{T'}{T} = D_1 \frac{U''}{U} + a + b \frac{V}{U} = d_2 \frac{V''}{V} + c \frac{U}{V} + d = \lambda,$$

with separation constant λ . Now we have $T' = \lambda T$, and we look for solutions of the pair of equations

$$\begin{aligned}D_1 U'' + (a - \lambda)U + bV &= 0, \\D_2 V'' + cU + (d - \lambda)V &= 0,\end{aligned}\tag{8.71}$$

of the form

$$U = c_1 \cos px + c_2 \sin px, \quad V = c_3 \cos px + c_4 \sin px.$$

When we substitute this form into the system (8.71), we obtain the equations

$$\begin{aligned}
D_1[-c_1 p^2 \cos px - c_2 p^2 \sin px] + (a - \lambda)c_1 \cos px + (a - \lambda)c_2 \sin px \\
+ bc_3 \cos px + bc_4 \sin px = 0, \quad (8.72) \\
D_2[-C_3 p^2 \cos px - c_4 p^2 \sin px] + cc_1 \cos px + cc_2 \sin px + (d - \lambda)c_3 \cos px \\
+ (d - \lambda)c_4 \sin px = 0.
\end{aligned}$$

Since (8.72) must hold for all x , the coefficients of $\cos px$ and $\sin px$ must vanish. This leads to the four equations

$$\begin{aligned}
(a - \lambda - p^2 D_1)c_1 + bc_3 &= 0, \\
(a - \lambda - p^2 D_1)c_2 + bc_4 &= 0, \\
cc_1 + (d - \lambda - p^2 D_2)c_3 &= 0, \\
(d - \lambda - p^2 D_2)c_4 + cc_2 &= 0.
\end{aligned}$$

These four equations are satisfied with c_1, c_2, c_3, c_4 not all zero if the two equations

$$\begin{aligned}
(a - \lambda - p^2 D_1)c_1 + bc_3 &= 0, \quad (8.73) \\
cc_1 + (d - \lambda - p^2 D_2)c_3 &= 0,
\end{aligned}$$

have a solution with c_1 and c_3 not both zero. The condition for this is that the determinant of the matrix

$$\begin{bmatrix} a - p^2 D_1 - \lambda & b \\ c & d - p^2 D_2 - \lambda \end{bmatrix}$$

be zero. This is equivalent to the statement that λ is an eigenvalue of the matrix

$$\begin{bmatrix} a - p^2 D_1 & b \\ c & d - p^2 D_2 \end{bmatrix}. \quad (8.74)$$

One question that arises for the system (8.68) is *diffusive instability*, meaning that the equilibrium (u_∞, v_∞) is asymptotically stable for the system of ordinary differential equations (8.69) but unstable for the system with diffusion (8.68). The asymptotic stability of this equilibrium for (8.69) is determined by the matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

and this equilibrium is asymptotically stable for (8.69) if and only if

$$a + d < 0, \quad ad - bc > 0, \quad (8.75)$$

and is asymptotically stable for (8.68) if and only if the eigenvalues of the matrix (8.74) have negative real part for some value of p , or

$$a + d - p^2(D_1 + D_2) < 0, \quad (a - p^2 D_1)(d - p^2 D_2) - bc > 0. \quad (8.76)$$

If $a + d < 0$, then $a + d - p^2(D_1 + D_2) < 0$. Thus diffusive instability would require that $ad - bc > 0$ but $(a - p^2D_1)(d - p^2D_2) - bc < 0$, or

$$\begin{aligned} H(p^2) &= (a - p^2D_1)(d - p^2D_2) - bc \\ &= D_1D_2p^4 - (D_1d + D_2a)p^2 + (ad - bc) < 0. \end{aligned}$$

Since $ad - bc > 0$, to have $H(p^2) < 0$ it is necessary that

$$D_1d + D_2a > 0.$$

If $D_1 = D_2$, this would contradict $a + d < 0$; thus diffusive instability requires $D_1 \neq D_2$. The minimum value of $H(p^2)$ as a function of p is attained for

$$p^2 = \frac{D_1d + D_2a}{2D_1D_2}$$

and is

$$ad - bc - \frac{(D_1d + D_2a)^2}{4D_1D_2}.$$

If p can be chosen to make this quantity negative, diffusive instability is possible. Thus diffusive instability is possible if and only if

$$ad - bc < \frac{(D_1d + D_2a)^2}{4D_1D_2}.$$

Because diffusive instability requires $D_1d + D_2a > 0$, we may write the necessary and sufficient conditions for diffusive instability as

$$a + d < 0, \quad ad - bc > 0, \quad aD_2 + dD_1 > \sqrt{D_1D_2(ad - bc)}. \quad (8.77)$$

If $ad > 0$, then the condition $a + d < 0$ implies $a < 0, d < 0$. But then $D_1d + D_2a < 0$, which would contradict $H(p^2) < 0$. Thus diffusive instability requires $ad < 0$, and then $ad - bc > 0$ implies $bc < 0$.

We may assume, with no loss of generality, that $a > 0$, and then $d < 0$. Since $bc < 0$, the matrix A has one of the two forms

$$(i) \begin{bmatrix} + & - \\ + & - \end{bmatrix}, \quad (ii) \begin{bmatrix} + & + \\ - & - \end{bmatrix}.$$

The species u is an *activator*, since it promotes an increase in its population size, while species v is an *inhibitor*, since it inhibits its own rate of growth. In case (i), an increase in u leads to an increase in both species, while an increase in v leads to a decrease in both species (activator–inhibitor). In case (ii), an increase in either species leads to an increase in u and a decrease in v (positive feedback). If we define

$$\tau_1 = \frac{1}{a}, \quad \tau_2 = \frac{1}{|d|} = -\frac{1}{d},$$

then (8.77) implies

$$\sqrt{D_1 \tau_1} < \sqrt{D_2 \tau_2}.$$

The quantities $\sqrt{D_1 \tau_1}$, $\sqrt{D_2 \tau_2}$ have the dimensions of distance, and are called the ranges of activation and inhibition respectively. Thus diffusive instability requires that the range of inhibition be greater than the range of activation.

If the two species are in a predator–prey relationship, then it is possible to have $a > 0, b < 0, c > 0, d < 0$, and thus diffusive instability is possible. If the interaction between the two species is competitive, then, as we have seen in Section 5.3, $a < 0, b < 0, c < 0, d < 0$, and thus diffusive instability is not possible. In fact, if $a < 0, b < 0, c < 0, d < 0$, it would be possible to have $ad - bc < 0$ (instability for (8.64)) but $H(p^2) > 0$ for all p (asymptotic stability for (8.68)). In general, diffusion tends to have a stabilizing effect. Diffusive instability requires very specific conditions on the coefficients. In particular, diffusive instability requires that the diffusion coefficients be unequal.

Exercises

1. [Allen (2007)] Consider a single-species population model on a line segment $0 \leq x \leq L$ that would have an asymptotically stable equilibrium K if there were no diffusion, whose population size $y(x, t)$ is described by the reaction–diffusion equation

$$y_t = f(y) + Dy_{xx}$$

with boundary conditions

$$\frac{\partial y}{\partial x}(0, t) = \frac{\partial y}{\partial x}(L, t) = 0.$$

Since the equilibrium K without diffusion is assumed asymptotically stable, $f'(K) < 0$.

- (i) Show that the linearization of the model at the equilibrium $y = K$ is

$$u_t = f'(K)u + Du_{xx}.$$

- (ii) Look for solutions of the linearization with the form

$$u(x, t) = e^{\sigma t} \cos(kx),$$

a reasonable form, since the linearization without diffusion has exponential solutions and the boundary conditions imply that these are cosines. Show that for every such solution the real part of σ is negative. [This shows that diffusive instability is not possible for a single-species model.]

8.10 Diffusion in Two Dimensions

In Section 8.4 we described the motion of a population that depends on both time and location in space with the restriction that the space be one-dimensional. It is natural to try to extend this description to a two-dimensional space.

Our approach is to take a small plane rectangle and to form a balance equation that says that the rate of change of the number of particles in the rectangle is equal to the rate at which particles flow out of the rectangle minus the rate at which particles flow into the rectangle plus the rate of creation of particles in the rectangle.

We think of a rectangle defined by inequalities $x_0 \leq x \leq x_0 + h$, $y_0 \leq y \leq y_0 + k$. Let $u(x, y, t)$ be the concentration of population members (number per unit area) at location (x_0, y_0) at time t , meaning that in the rectangle $x_0 \leq x \leq x_0 + h$, $y_0 \leq y \leq y_0 + k$, the number of members is approximately $hku(x, y, t)$. By “approximately” we mean that if h and k are small, the error in this approximation $hku(x, y, t)$ is smaller than a constant multiple of $h^2 + k^2$. The flux $J(x, y, t)$ in two dimensions must be described as a vector, which we may do by giving its horizontal (in the x -direction) component $J_1(x, y, t)$ and its vertical (in the y -direction) component $J_2(x, y, t)$. The net horizontal flow out of the rectangle $x_0 \leq x \leq x_0 + h$, $y_0 \leq y \leq y_0 + k$ is $k[J_1(x_0, y_0, t) - J_1(x_0 + h, y_0, t)]$, and the net vertical flow out of this rectangle is $h[J_2(x_0, y_0, t) - J_2(x_0, y_0 + k, t)]$. If there are no births or deaths in the rectangle, we have a balance equation for the rectangle expressing the fact that the rate of change of population size at time t is equal to the net flux, so that

$$hku_t(x_0, y_0, t) = -k[J_1(x_0 + h, y_0, t) - J_1(x_0, y_0, t)] - h[J_2(x_0, y_0 + k, t) - J_2(x_0, y_0, t)]. \quad (8.78)$$

If we divide (8.78) by hk and then take the limit as $h \rightarrow 0, k \rightarrow 0$, we obtain the conservation law

$$u_t(x_0, y_0, t) = - \left[\frac{\partial J_1}{\partial x}(x_0, y_0, t) + \frac{\partial J_2}{\partial y}(x_0, y_0, t) \right],$$

valid for all t at every point (x_0, y_0) . Since x_0 and y_0 are arbitrary, this relation holds at every point, and we have the equation

$$u_t(x, y, t) = - \left[\frac{\partial J_1}{\partial x}(x, y, t) + \frac{\partial J_2}{\partial y}(x, y, t) \right]. \quad (8.79)$$

In vector language, the quantity $\frac{\partial J_1}{\partial x}(x, y, t) + \frac{\partial J_2}{\partial y}(x, y, t)$ is called the *divergence* of the vector $\mathbf{J}(x, y, t)$ and is denoted by $\text{div } \mathbf{J}$.

Just as in one dimension we must make some assumption that relates the rate of change of flux density $\frac{\partial J_1}{\partial x} + \frac{\partial J_2}{\partial y}$ and the population density $u(x, y, t)$. If the motion is random, then *Fick's law* says that the flux due to random motion is approximately proportional to the rate of change of particle concentration, that is, that J_1 is proportional to u_x and J_2 is proportional to u_y with the same constant of proportionality.

This leads to

$$J_1 = -Du_x, \quad J_2 = -Du_y$$

with D a constant called the diffusivity or diffusion coefficient. (We restrict our attention to constant diffusivity.) Then equation (8.79) then becomes a second order partial differential equation

$$u_t(x, y, t) = D \left[\frac{\partial u_x(x, y, t)}{\partial x} + \frac{\partial u_y(x, y, t)}{\partial y} \right] = D[u_{xx}(x, y, t) + u_{yy}(x, y, t)]. \quad (8.80)$$

Equation (8.80) is called the *heat* or *diffusion* equation in two space dimensions. The quantity $u_{xx}(x, y, t) + u_{yy}(x, y, t)$ is called the *Laplacian* of the function $u(x, y, t)$ and is often written $\nabla^2 u$. In addition to the partial differential equation (8.80) we need to impose additional conditions to give a problem with a unique solution. It is possible to establish a result analogous to the existence and uniqueness theorem stated in Section 4.4 for the diffusion equation in one space dimension, namely that the equation (8.80) has a unique solution in every region in the plane with a piecewise smooth boundary with a specified boundary condition giving the value of the solution or its normal derivative at every point of the boundary and with a specified initial condition.

If the given region is a rectangle, it is possible to use the method of separation of variables to find solutions analytically. The process is more complicated than in one space dimension because one must first separate the time variable from the two space variables and then separate the two space variables. For nonrectangular regions it may be possible to use a coordinate system in which the boundary takes a convenient form. For example, a region bounded by a circle with center at the origin may be described in polar coordinates (r, θ) with boundary $r = r_0$. The diffusion equation may be transformed into the appropriate coordinate system, and it may be possible to carry out separation of variables in this coordinate system. In general, separation of variables in more than two variables is very complicated technically, and we shall not explore this subject here.

Another complication that arises in two dimensions for the diffusion equation is that there are many possible steady-state solutions. In one space dimension, a steady-state solution is a solution of the ordinary differential equation $u'' = 0$ and must be a linear function. In two dimensions, a steady-state solution is a solution of the *Laplace equation*

$$u_{xx}(x, y, t) + u_{yy}(x, y, t) = \nabla^2 u = 0.$$

The Laplace equation has many solutions. In fact, there is a unique solution of the Laplace equation for every given boundary condition. The Laplace equation may be solved by separation of variables in some coordinate systems.

If there are creation terms as well as flux, we obtain a reaction–diffusion equation in two space dimensions, of the form

$$u_t(x, y, t) = D[u_{xx}(x, y, t) + u_{yy}(x, y, t)] + g(u(x, y, t)) = \nabla^2 u + g(u) = 0.$$

Reaction–diffusion equations in two space dimensions may exhibit all the possible behaviors that we found in one space dimension. More complicated behavior is possible, including stable patterns, periodic wave trains, and spiral waves. These behaviors have been exhibited experimentally in some situations. There are many open problems in the mathematical study of the asymptotic behavior of solutions of reaction–diffusion equations, especially in more than one space dimension and for systems of more than one equation.

More general models may be appropriate in some situations. Diffusion may vary in different directions, and for models of diffusion in aquatic systems it may be necessary to incorporate an advection term in the partial differential equations to describe the water current velocity.

Reaction diffusion models arise in a variety of other biological situations, including the spread of a favorable allele in genetics, spatial pattern formation in developmental biology, nerve action potentials, and autocatalytic chemical reactions. The study of reaction–diffusion models in biology is a new and rapidly developing field; we have barely scratched the surface here.

8.11 Project: Cats and Birds in Space

This project is based on the MTBI report Gambino et al. (2007).

Consider the mean field approximation to a group of interacting populations on a bounded lattice, which can be written as

$$\begin{aligned}\dot{p}_0 &= \mu_1 p_1 + \mu_2 p_2 + \mu_3 p_3 + nu_2(1 - \eta_2)p_2 p_1 + nu_3(1 - \eta_{31})p_3 p_1 \\ &\quad + nu_3(1 - \eta_{32})p_3 p_2 - (\phi_1 p_1 + \phi_2 p_2 + \phi_3 p_3)p_0, \\ \dot{p}_1 &= \phi_1 p_1 p_0 - (\mu_1 + v_2 p_2 + v_3 p_3)p_1, \\ \dot{p}_2 &= \phi_2 p_2 p_0 + v_2 \eta_2 p_1 p_2 - (\mu_2 + v_3 p_3)p_1, \\ \dot{p}_3 &= \phi_3 p_2 p_0 - \mu_3 p_3 + v_3 \eta_{31} p_3 p_1,\end{aligned}$$

where p_0 represents the proportion of empty sites on the lattice and p_i represents the proportion of sites occupied by site i ($i \in \{1, 2, 3\}$).

The system of equations describes a predator–prey relationship between three populations. p_1 represents the prey population, p_2 represents the mesopredator population that prey solely on p_1 , and p_3 represents the superpredator population that preys on both p_1 and p_2 . A description of the parameters is given below:

1. Interpret all the parameters (which we assume to be positive) and discuss the relationships that exist between the interacting populations.
2. Show that the dynamics of the above system can be studied with just the equations of motion for p_1 , p_2 , and p_3 . (Hint: Add the derivatives.)
3. Using the reduced model from part b., derive the species-free and single species equilibria and interpret the conditions for the existence of each equilibrium.

Parameters	Description
ϕ_i	Intrinsic birth rate of species i
μ_i	Death rate of species i
v_i	Foraging rate for species i ($i \in \{2, 3\}$)
η_2	Probability that successful mesopredator predation leads to birth
η_{3i}	Probability that successful superpredator predation on species i leads to birth ($i \in \{1, 2\}$)

4. Derive and interpret necessary conditions for stability of each equilibrium found above. (Hint: Compute the Jacobian and evaluate it at each equilibrium.)

8.12 Project: The Cable Equation

This project is based on the MTBI report Foster et al. (2010).

Neurons are excitable cells that communicate with other neurons by receiving and sending electrical signals. High levels of voltage are essential to propagating a signal and are most often transmitted along the dendritic branches of a neuron. A neural network is often described as a collection of long cylindrical cables that propagate electrical signals, and cable theory has been applied to measure voltages in the cell membrane, where the cable equation is

$$\lambda^2 \frac{\partial^2 V}{\partial x^2} = \tau \frac{\partial V}{\partial t} + V.$$

To derive the cable equation, assume that the current travels through a long and cylindrical cable that is uniform in diameter. The voltage (V) is the difference between the interior and exterior voltage, as a deviation from resting value and a resting membrane battery ($V = V_i - V_e - E_r$). While the interior voltage is dependent on space, x , and time, t , neither the exterior voltage nor the resting membrane battery is. Intracellular current, i_i , travels from left to right as x increases, and intracellular resistance per unit length, r_i , is constant (see [Figure 8.2](#)).

1. Use Ohm's law ($I = \frac{V}{R}$) to describe the relationship between the change in voltage across some length of the cable (Δx) and i_i . (Hint: the change in voltage can be written as Δx and the resistance is the product of the per unit length resistance and the unit of length, Δx . Take the limit as Δx gets small.)
2. Use Kirchhoff's current law ($i_{net} = i_{in} - i_{out}$) to describe the relationship between the interior current i_i and the axial or membrane current i_m (Hint: if the net current is conserved, then the change in current, Δi_i should be balanced by

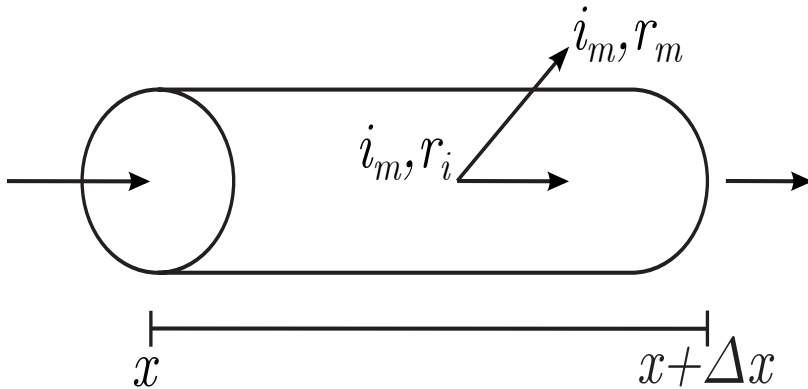


Fig. 8.2 Schematic for the dendrite as a cable. Current comes in from the left and travels a distance $x + \Delta x$, losing current through the membrane as it travels, but adhering to Kirchhoff's current law.

the membrane current across the entire length increment, $i_m \Delta x$.)

3. Express the per unit length membrane current density i_m as a sum of the parallel capacitive and conductive currents. (Hint: use the current-voltage formulation ($I = C \frac{dV}{dt}$) for the capacitance and Ohm's law for the conductance.)
4. Combine the first three results to derive the cable equation. What do the dimensionless quantities λ and τ equal in this derivation? (Hint: take the derivative of **a**) with respect to x .)
5. How could you modify the cable equation to represent the change in voltage along a dendritic branch?

Consider a passive cable for a very long dendrite with constant current I injected at position $x = 0$. Assume that the cable is a uniform semi-infinite circular cylinder ($0 \leq x \leq \infty$). The cable equation is given by:

$$\lambda^2 \frac{\partial^2 V}{\partial x^2} = \tau \frac{\partial V}{\partial t} + V,$$

where λ is a resistance constant and τ a time constant. Write down the boundary conditions. Assume that the potential (voltage) V and V_x are bounded as $x \rightarrow \infty$. Find the steady-state solution $V(x)$.

Consider a passive cable of finite length ($0 \leq x \leq L$). Assume when the nerve is at rest that $V(x) = 0$. At $t = 0$, a steady current of I_{app} is injected into the right end of the cable, and the left end of the cable is sealed to current flow (i.e., $I_x(0) = 0$). Find the steady-state value of the potential V at the sealed end assuming unit resistance (i.e., $I(t, x) = V(t, x)$).

8.13 Project: Some Equations of Diffusion Type

Problem 1 Show that the nonautonomous and inhomogeneous diffusion-type equation

$$\frac{\partial u}{\partial t} = a(t) \frac{\partial^2 u}{\partial x^2} - (g(t) - c(t)x) \frac{\partial u}{\partial x} + (-b(t)x^2 + f(t)x + d(t))u, \quad (8.81)$$

where a, b, c, d, f, g are suitable functions of time t only, can be reduced to the standard autonomous form

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial \xi^2} \quad (8.82)$$

with the help of the substitution [Suazo et al. (2011a)]

$$u(x, t) = \frac{1}{\sqrt{\mu(t)}} e^{\alpha(t)x^2 + \delta(t)x + \kappa(t)} v(\xi, \tau), \quad (8.83)$$

$$\xi = \beta(t)x + \varepsilon(t), \quad \tau = \gamma(t).$$

Here, $\mu, \alpha, \beta, \gamma, \delta, \varepsilon, \kappa$ are functions of t that satisfy

$$\frac{\mu'}{2\mu} + 2a\alpha + d = 0, \quad (8.84)$$

and

$$\begin{aligned} \frac{d\alpha}{dt} + b - 2c\alpha - 4a\alpha^2 &= 0, \\ \frac{d\beta}{dt} - (c + 4a\alpha)\beta &= 0, \\ \frac{d\gamma}{dt} - a\beta^2 &= 0, \\ \frac{d\delta}{dt} - (c + 4a\alpha)\delta &= f - 2\alpha, \\ \frac{d\varepsilon}{dt} + (g - 2a\delta)\beta &= 0, \\ \frac{d\kappa}{dt} + g\delta - a\delta^2 &= 0. \end{aligned} \quad (8.85)$$

(The first equation of (8.85) is called the *Riccati nonlinear differential equation* [Watson (1944)] and we shall refer to the system (8.85) as a *Riccati-type system*.)

Problem 2 Show that the substitution (8.83) reduces the nonlinear Riccati equation to the second-order linear equation

$$\mu'' - \tau(t)\mu' - 4\sigma(t)\mu = 0, \quad (8.86)$$

where

$$\tau(t) = \frac{a'}{a} + 2c - 4d, \quad \sigma(t) = ab + cd - d^2 + \frac{d}{2} \left(\frac{a'}{a} - \frac{d'}{d} \right), \quad (8.87)$$

which we refer to as a *characteristic equation* [Suazo et al. (2011a)].

Problem 3 Show that the diffusion-type equation (8.81) has a particular solution of the form

$$u = \frac{1}{\sqrt{\mu(t)}} e^{\alpha(t)x^2 + \beta(t)xy + \gamma(t)y^2 + \delta(t)x + \varepsilon(t)y + \kappa(t)}, \quad (8.88)$$

provided that the time-dependent functions $\mu, \alpha, \beta, \gamma, \delta, \varepsilon, \kappa$ satisfy the Riccati-type system (10.29), (8.85) [Suazo et al. (2011a)].

Problem 4 Use the *superposition principle* in order to solve (formally) the Cauchy initial value problem for the diffusion-type equation (8.81) subject to initial data $u(x, 0) = \varphi(x)$ on the entire real line $-\infty < x < \infty$ in an integral form

$$u(x, t) = \int_{-\infty}^{\infty} K_0(x, y, t) \varphi(y) dy \quad (8.89)$$

with the fundamental solution (heat kernel) [Suazo et al. (2011a)]

$$K_0(x, y, t) = \frac{1}{\sqrt{2\pi\mu_0(t)}} e^{\alpha_0(t)x^2 + \beta_0(t)xy + \gamma_0(t)y^2 + \delta_0(t)x + \varepsilon_0(t)y + \kappa_0(t)}, \quad (8.90)$$

where the particular solution of the Riccati-type system (8.83), (8.85) is given by

$$\begin{aligned}
\alpha_0(t) &= -\frac{1}{4a(t)} \frac{\mu'_0(t)}{\mu_0(t)} - \frac{d(t)}{2a(t)}, \\
\beta_0(t) &= \frac{h(t)}{\mu_0(t)}, \quad h(t) = \exp\left(\int_0^t (c(s) - 2d(s)) ds\right), \\
\gamma_0(t) &= \frac{d(0)}{2a(0)} - \frac{a(t)h^2(t)}{\mu_0(t)\mu'_0(t)} - 4 \int_0^t \frac{a(s)\sigma(s)h(s)}{(\mu'_0(s))^2} ds \\
&= \frac{d(0)}{2a(0)} - \frac{1}{2\mu_1(0)} \frac{\mu_1(t)}{\mu_0(t)} \tag{8.91} \\
\delta_0(t) &= \frac{h(t)}{\mu_0(t)} \int_0^t \left[\left(f(s) + \frac{d(s)}{a(s)} g(s) \right) \mu_0(s) + \frac{g(s)}{2a(s)} \mu'_0(s) \right] \frac{ds}{h(s)}, \\
\varepsilon_0(t) &= -\frac{2a(t)h(t)}{\mu'_0(t)} \delta_0(t) - 8 \int_0^t \frac{a(s)\sigma(s)h(s)}{(\mu'_0(s))^2} (\mu_0(s)\delta_0(s)) ds \\
&\quad + 2 \int_0^t \frac{a(s)h(s)}{\mu'_0(s)} \left[f(s) + \frac{d(s)}{a(s)} g(s) \right] ds, \\
\kappa_0(t) &= -\frac{a(t)\mu_0(t)}{\mu'_0(t)} \delta_0^2(t) - 4 \int_0^t \frac{a(s)\sigma(s)}{(\mu'_0(s))^2} (\mu_0(s)\delta_0(s))^2 ds \\
&\quad + 2 \int_0^t \frac{a(s)}{\mu'_0(s)} (\mu_0(s)\delta_0(s)) \left[f(s) + \frac{d(s)}{a(s)} g(s) \right] ds,
\end{aligned}$$

with $\delta(0) = g(0)/(2a(0))$, $\varepsilon(0) = -\delta(0)$, $\kappa(0) = 0$. Here, μ_0 and μ_1 are the so-called *standard solutions* of the characteristic equation (8.86) subject to the initial data

$$\mu_0(0) = 0, \quad \mu'_0(0) = 2a(0) \neq 0, \quad \mu_1(0) \neq 0, \quad \mu'_1(0) = 0. \tag{8.92}$$

(See Problem 7 below for the corresponding asymptotics.)

Problem 5 Show that the Riccati-type system (10.29), (8.85) has the (general) solution [Suazo et al. (2011b)]

$$\begin{aligned}
mu(t) &= -2\mu(0)\mu_0(t)(\alpha(0) + \gamma_0(t)), \\
\alpha(t) &= \alpha_0(t) - \frac{\beta_0^2(t)}{4(\alpha(0) + \gamma_0(t))}, \\
\beta(t) &= -\frac{\beta(0)\beta_0(t)}{2(\alpha(0) + \gamma_0(t))}, \\
\gamma(t) &= \gamma(0) - \frac{\beta^2(0)}{4(\alpha(0) + \gamma_0(t))},
\end{aligned} \tag{8.93}$$

and

$$\begin{aligned}
\delta(t) &= \delta_0(t) - \frac{\beta_0(t)(\delta(0) + \varepsilon_0(t))}{2(\alpha(0) + \gamma_0(t))}, \\
\varepsilon(t) &= \varepsilon(0) - \frac{\beta(0)(\delta(0) + \varepsilon_0(t))}{2(\alpha(0) + \gamma_0(t))}, \\
\kappa(t) &= \kappa(0) + \kappa_0(t) - \frac{(\delta(0) + \varepsilon_0(t))^2}{4(\alpha(0) + \gamma_0(t))},
\end{aligned} \tag{8.94}$$

in terms of the coefficients of the fundamental solution (8.91) subject to arbitrary initial data $\mu(0), \alpha(0), \beta(0), \gamma(0), \delta(0), \varepsilon(0), \kappa(0)$.

Hint: Use (8.88)–(8.90) and the elementary integral

$$\int_{-\infty}^{\infty} e^{-ay^2+2by} dy = \sqrt{\frac{\pi}{a}} e^{b^2/a}, \quad a > 0. \tag{8.95}$$

Problem 6 Use the transformation (8.83) with the standard heat kernel

$$K_0(\xi, \eta, \tau) = \frac{1}{\sqrt{4\pi(\tau - \tau_0)}} \exp \left[-\frac{(\xi - \eta)^2}{4(\tau - \tau_0)} \right] \tag{8.96}$$

for the diffusion equation (8.82), and (8.93), (8.94) in order to derive the fundamental solution (8.90) for the diffusion-type equation (8.81) [Suazo et al. (2011b)].

Problem 7 Show that the solution (8.93), (8.94) implies [Suazo et al. (2011a)]

$$\begin{aligned}
\mu_0 &= \frac{2\mu}{\mu(0)\beta^2(0)} (\gamma - \gamma(0)), \\
\alpha_0 &= \alpha_0(t) - \frac{\beta^2}{4(\gamma - \gamma(0))}, \\
\beta_0 &= \frac{\beta(0)\beta}{2(\gamma - \gamma(0))}, \\
\gamma_0 &= -\alpha(0) - \frac{\beta^2(0)}{4(\gamma - \gamma(0))},
\end{aligned}$$

and

$$\begin{aligned}
\delta_0 &= \delta - \frac{\beta(\varepsilon - \varepsilon(0))}{2(\gamma - \gamma(0))}, \\
\varepsilon_0 &= -\delta(0) + \frac{\beta(0)(\varepsilon - \varepsilon(0))}{2(\gamma - \gamma(0))}, \\
\kappa_0 &= \kappa - \kappa(0) - \frac{(\varepsilon - \varepsilon(0))^2}{4(\gamma - \gamma(0))},
\end{aligned}$$

and use this representation in order to derive the asymptotics

$$\begin{aligned}
\alpha_0(t) &= -\frac{1}{4a(0)t} - \frac{c(0)}{4a(0)} + \frac{a'(0)}{8a^2(0)} + \mathcal{O}(t), \\
\beta_0(t) &= \frac{1}{2a(0)t} - \frac{a'(0)}{4a^2(0)} + \mathcal{O}(t), \\
\gamma_0(t) &= -\frac{1}{4a(0)t} + \frac{c(0)}{4a(0)} + \frac{a'(0)}{8a^2(0)} + \mathcal{O}(t), \\
\delta_0(t) &= \frac{g(0)}{2a(0)} + \mathcal{O}(t), \quad \varepsilon_0(t) = -\frac{g(0)}{2a(0)} + \mathcal{O}(t), \\
\kappa_0(t) &= \mathcal{O}(t),
\end{aligned} \tag{8.97}$$

as $t \rightarrow 0^+$. Use these formulas in order to establish the asymptotics of the fundamental solution (8.90),

$$\begin{aligned}
K_0(x, y, t) &\sim \frac{1}{\sqrt{4\pi a(0)t}} \exp \left[-\frac{(x-y)^2}{4a(0)t} \right] \\
&\times \exp \left[\frac{a'(0)}{8a^2(0)} (x-y)^2 - \frac{c(0)}{4a(0)} (x^2 - y^2) \right] \exp \left[\frac{g(0)}{2a(0)} (x-y) \right].
\end{aligned} \tag{8.98}$$

(Here, $f \sim g$ as $t \rightarrow 0^+$ if $\lim_{t \rightarrow 0^+} (f/g) = 1$.)

Problem 8 Check by a direct substitution that (8.93), (8.94) satisfy the Riccati-type system (10.29), (8.85). Use the asymptotics (8.97) in order to verify the continuity with respect to initial data [Suazo et al. (2011a)]

$$\lim_{t \rightarrow 0^+} \mu(t) = \mu(0), \quad \lim_{t \rightarrow 0^+} \alpha(t) = \alpha(0), \quad \text{etc.} \tag{8.99}$$

This transformation property allows us to find a solution of the initial value problem in terms of the fundamental solution and may be referred to as a *nonlinear superposition principle* for the Riccati-type system).

Problem 9 The nonlinear equation

$$\frac{\partial v}{\partial t} + a(t) \left(v \frac{\partial v}{\partial x} - \frac{\partial^2 v}{\partial x^2} \right) - c(t) \left(x \frac{\partial v}{\partial x} + v \right) + g(t) \frac{\partial v}{\partial x} = 2(2b(t)x - f(t)), \tag{8.100}$$

when $a = 1$ and $b = c = f = g = 0$, is known as *Burgers' equation*. (We shall refer to (8.100) as a nonautonomous Burgers-type equation.) Verify that [Suazo et al. (2011a)]

$$v_t + a(vv_x - v_{xx}) + (g - cx)v_x - cv + 2(f - 2bx) = -2 \left(\frac{u_t - Qu}{u} \right)_x$$

if

$$v = -2 \frac{u_x}{u} \quad (\text{The Cole-Hopf transformation}) \tag{8.101}$$

and

$$Qu = au_{xx} - (g - cx)u_x + (d + fx - bx^2)u$$

(a, b, c, d, f, g are functions of t only). The substitution (8.101) turns the nonlinear Burgers-type equation (8.100) into the diffusion-type equation (8.81). Show that the solution of the Cauchy initial value problem can be represented as

$$v(x, t) = -2 \frac{\partial}{\partial x} \ln \left[\int_{-\infty}^{\infty} K_0(x, y, t) \exp \left(-\frac{1}{2} \int_{-\infty}^y v(z, 0) dz \right) dy \right],$$

where the heat kernel is given by (8.90), for suitable initial data $v(z, 0)$ on \mathbb{R} .

Now we consider from a united viewpoint several elementary diffusion-type equations (8.81) that are important in applications.

Problem 10 Show that for the standard *diffusion equation*

$$\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2}, \quad a = \text{constant} > 0, \quad (8.102)$$

the heat kernel is given by

$$K(x, y, t) = \frac{1}{\sqrt{4\pi at}} \exp \left(-\frac{(x-y)^2}{4at} \right), \quad t > 0. \quad (8.103)$$

(See [Cannon (1984)] and references therein for a detailed investigation of the classical one-dimensional heat equation.)

Problem 11 In the mathematical description of the nerve cell, a dendritic branch is typically modeled using the cylindrical *cable equation* [Jack, Noble, & Tsien (1983)]

$$\tau \frac{\partial u}{\partial t} = \lambda^2 \frac{\partial^2 u}{\partial x^2}, \quad \tau = \text{constant} > 0. \quad (8.104)$$

Show that the fundamental solution on \mathbb{R} is given by

$$K_0(x, y, t) = \frac{\sqrt{\tau} e^{t/\tau}}{\sqrt{4\pi \lambda^2 t}} \exp \left[-\frac{\tau (x-y)^2}{4\lambda^2 t} \right], \quad t > 0. \quad (8.105)$$

Problem 12 Show that the fundamental solution of the *Fokker–Planck equation*

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + x \frac{\partial u}{\partial x} + u \quad (8.106)$$

on \mathbb{R} is given by [Suazo et al. (2011a)]

$$K_0(x, y, t) = \frac{1}{\sqrt{2\pi(1-e^{-2t})}} \exp \left[-\frac{(x-e^{-t}y)^2}{2(1-e^{-2t})} \right], \quad t > 0. \quad (8.107)$$

Here,

$$\lim_{t \rightarrow 0^+} K_0(x, y, t) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}, \quad y = \text{constant.} \quad (8.108)$$

Problem 13 Equation

$$\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + (g - kx) \frac{\partial u}{\partial x}, \quad a, k > 0, \quad g \geq 0, \quad (8.109)$$

corresponds to the heat equation with linear drift when $g = 0$ [Miller (1977)]. In stochastic differential equations this equation corresponds to the Kolmogorov forward equation for the regular Ornstein–Uhlenbeck process [Craddock (2009)]. Show that the fundamental solution is given by

$$K_0(x, y, t) = \frac{\sqrt{k}e^{kt/2}}{\sqrt{4\pi a \sinh(kt)}} \exp \left[-\frac{(k(xe^{-kt/2} - ye^{kt/2}) + 2g \sinh(kt/2))^2}{4ak \sinh(kt)} \right], \quad t > 0.$$

See [Craddock (2009), Suazo, Suslov, Vega-Guzmán (2010)] for more details.

Part IV

Disease Transmission Models

Chapter 9

Epidemic Models

9.1 Introduction to Epidemic Models

Communicable diseases such as measles, influenza, and tuberculosis are a fact of life. We will be concerned with both epidemics, which are sudden outbreaks of a disease, and endemic situations, in which a disease is always present. The AIDS epidemic, the recent SARS epidemic, recurring influenza pandemics, and outbursts of diseases such as the Ebola virus are events of concern and interest to many people. The prevalence and effects of many diseases in less-developed countries are probably not as well known but may be of even more importance. Every year millions, of people die of measles, respiratory infections, diarrhea, and other diseases that are easily treated and not considered dangerous in the Western world. Diseases such as malaria, typhus, cholera, schistosomiasis, and sleeping sickness are endemic in many parts of the world. The effects of high disease mortality on mean life span and of disease debilitation and mortality on the economy in afflicted countries are considerable.

We give a brief introduction to the modeling of epidemics; more thorough descriptions may be found in such references as [Anderson & May (1991), Diekmann & Heesterbeek (2000)]. This chapter will describe models for epidemics, and the next chapter will deal with models for endemic situations, but we begin with some general ideas about disease transmission.

The idea of invisible living creatures as agents of disease goes back at least to the writings of Aristotle (384 BC–322 BC). It developed as a theory in the sixteenth century. The existence of microorganisms was demonstrated by van Leeuwenhoek (1632–1723) with the aid of the first microscopes. The first expression of the germ theory of disease by Jacob Henle (1809–1885) came in 1840 and was developed by Robert Koch (1843–1910), Joseph Lister (1827–1912), and Louis Pasteur (1822–1875) in the latter part of the nineteenth century and the early part of the twentieth century.

The mechanism of transmission of infections is now known for most diseases. Generally, diseases transmitted by viral agents, such as influenza, measles, rubella

(German measles), and chicken pox, confer immunity against reinfection, while diseases transmitted by bacteria, such as tuberculosis, meningitis, and gonorrhea, confer no immunity against reinfection. Other human diseases, such as malaria, are transmitted not directly from human to human but by vectors, agents (usually insects) that are infected by humans and who then transmit the disease to humans. There are also diseases such as West Nile virus, that are transmitted back and forth between animals and vectors. Heterosexual transmission of HIV/AIDS is also a vector process in which transmission goes back and forth between males and females.

We will focus on the transmission dynamics of an infection from individual to individual in a population, but many of the same ideas arise in transmission of a genetic characteristic, such as gender, race, genetic diseases, a cultural “characteristic,” such as language or religion, an addictive activity, such as drug use, and the gain or loss of information communicated through gossip, rumors, and so on.

Similarly, many of the ideas arise also with different characterizations of what is meant by an individual, including the types of cells in the study of disease dynamics of the immune system. In the study of *Chagas* disease, a “house” (infested houses may correspond to “infected” individuals) may be chosen as an epidemiological unit; in tuberculosis, a household or community or a group of strongly linked individuals (“cluster”) may be the chosen unit.

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

The historian W.H. McNeill argues, especially in his book *Plagues and Peoples* (1976), that the spread of communicable diseases frequently has been an important influence in history. For example, there was a sharp population increase throughout the world in the eighteenth century; the population of China increased from 150 million in 1760 to 313 million in 1794, and the population of Europe increased from 118 million in 1700 to 187 million in 1800. There were many factors involved in this increase, including changes in marriage age and technological improvements leading to increased food supplies, but these factors are not sufficient to explain the increase. Demographic studies indicate that a satisfactory explanation requires recognition of a decrease in the mortality caused by periodic epidemic infections. This decrease came about partly through improvements in medicine, but a more important influence was probably the fact that more people developed immunities against infection as increased travel intensified the circulation and cocirculation of diseases.

There are many biblical references to diseases as historical influences. The Book of Exodus describes the plagues that were brought down upon Egypt in the time of Moses. Another example is the decision of Sennacherib, the king of Assyria, to abandon his attempt to capture Jerusalem about 700 BC because of the illness of his soldiers (Isaiah 37, 36-38), and there are several other biblical descriptions of epidemic outbreaks.

The fall of empires has been attributed directly or indirectly to epidemic diseases. In the second century AD, the so-called Antonine plagues (possibly measles and) invaded the Roman Empire, causing drastic population reductions and economic hardships leading to disintegration of the empire because of disorganization, which facilitated invasions of barbarians. The Han Empire in China collapsed in the third century AD after a very similar sequence of events. The defeat of a population of millions of Aztecs by Cortez and his 600 followers can be explained, in part, by a smallpox epidemic that devastated the Aztecs but had almost no effect on the invading Spaniards, thanks to their built-in immunities. The Aztecs were not only weakened by disease but also confounded by what they interpreted as a divine force favoring the invaders. Smallpox then spread southward to the Incas in Peru and was an important factor in the success of Pizarro's invasion a few years later. Smallpox was followed by other diseases such as measles and diphtheria imported from Europe to North America. In some regions, the indigenous populations were reduced to one-tenth of their previous levels by these diseases: Between 1519 and 1530 the Indian population of Mexico was reduced from 30 million to 3 million.

The Black Death (probably bubonic plague) spread from Asia throughout Europe in several waves during the fourteenth century, beginning in 1346, and is estimated to have caused the death of as much as one-third of the population of Europe between 1346 and 1350. The disease recurred regularly in various parts of Europe for more than 300 years, notably as the Great Plague of London of 1665–1666. It then gradually withdrew from Europe. Since the plague struck some regions harshly while avoiding others, it had a profound effect on political and economic developments in medieval times. In the last bubonic plague epidemic in France (1720–1722), half the population of Marseilles, 60 percent of the population in nearby Toulon, 44 per cent of the population of Arles, and 30 percent of the population of Aix and Avignon died, but the epidemic did not spread beyond Provence. Expansions and interpretations of these historical remarks may be found in McNeill (1976), which was our primary source on the history of the spread and effects of diseases.

The above examples depict the *sudden* dramatic impact that diseases have had on the demography of human populations via disease-induced mortality. In considering the combined role of diseases, war, and natural disasters on mortality rates, one may conclude that historically humans who are more likely to survive and reproduce have either a good immune system, a propensity to avoid war and disasters, or, nowadays, excellent medical care and/or health insurance.

Descriptions of epidemics in ancient and medieval times frequently used the term “plague” because of a general belief that epidemics represented divine retribution for sinful living. More recently, some have described AIDS as punishment for sinful activities. Such views have often hampered or delayed attempts to control this modern epidemic.

There are many questions of interest to public health physicians confronted with a possible epidemic. For example, how severe will an epidemic be? This question may be interpreted in a variety of ways. For example, how many individuals will be affected and require treatment? What is the maximum number of people needing

care at any particular time? How long will the epidemic last? How much good would quarantine of victims do in reducing the severity of the epidemic?

Scientific experiments usually are designed to obtain information and to test hypotheses. Experiments in epidemiology with controls are often difficult or impossible to design, and even if it is possible to arrange an experiment, there are serious ethical questions involved in withholding treatment from a control group. Sometimes data may be obtained after the fact from reports of epidemics or of endemic disease levels, but the data may be incomplete or inaccurate. In addition, data may contain enough irregularities to raise serious questions of interpretation, such as whether there is evidence of chaotic behavior [Ellner, Gallant, and Theiler (1995)]. Hence, parameter estimation and model fitting are very difficult. These issues raise the question whether mathematical modeling in epidemiology is of value.

Mathematical modeling in epidemiology provides understanding of the underlying mechanisms that influence the spread of disease, and in the process, it suggests control strategies. In fact, models often identify behaviors that are unclear in experimental data—often because data are nonreproducible and the number of data points is limited and subject to errors in measurement. For example, one of the fundamental results in mathematical epidemiology is that *most* mathematical epidemic models, including those that include a high degree of heterogeneity, usually exhibit “threshold” behavior, which in epidemiological terms can be stated as follows: *If the average number of secondary infections caused by an average infective is less than one, a disease will die out, while if it exceeds one there will be an epidemic.* This broad principle, consistent with observations and quantified via epidemiological models, has been used routinely to estimate the effectiveness of vaccination policies and the likelihood that a disease may be eliminated or eradicated. Hence, even if it is not possible to verify hypotheses accurately, agreement with hypotheses of a qualitative nature is often valuable. Expressions for the basic reproductive number for HIV in various populations is being used to test the possible effectiveness of vaccines that may provide temporary protection by reducing either HIV-infectiousness or susceptibility to HIV. Models are used to estimate how widespread a vaccination plan must be to prevent or reduce the spread of HIV.

In the mathematical modeling of disease transmission, as in most other areas of mathematical modeling, there is always a trade-off between simple models, which omit most details and are designed only to highlight general qualitative behavior, and detailed models, usually designed for specific situations including short-term quantitative predictions. Detailed models are generally difficult or impossible to solve analytically and hence their usefulness for theoretical purposes is limited, although their strategic value may be high. For public health professionals, who are faced with the need to make recommendations for strategies to deal with a specific situation, simple models are inadequate and numerical simulation of detailed models is necessary. In this chapter, we concentrate on simple models in order to establish broad principles. Furthermore, these simple models have additional value since they are the building blocks of models that include detailed structure. A specific goal is to compare the dynamics of simple and slightly more detailed models primarily

to see whether slightly different assumptions can lead to significant differences in qualitative behavior.

Many of the early developments in the mathematical modeling of communicable diseases are due to public health physicians. The first known result in mathematical epidemiology is a defense of the practice of inoculation against smallpox in 1760 by Daniel Bernouilli, a member of a famous family of mathematicians (eight spread over three generations) who had been trained as a physician. The first contributions to modern mathematical epidemiology are due to P.D. En'ko between 1873 and 1894 [Dietz (1988)], and the foundations of the entire approach to epidemiology based on compartmental models were laid by public health physicians such as Sir R.A. Ross, W.H. Hamer, A.G. McKendrick, and W.O. Kermack between 1900 and 1935, along with important contributions from a statistical perspective by J. Brownlee. A particularly instructive example is the work of Ross on malaria. Dr. Ross was awarded the second Nobel Prize in Medicine for his demonstration of the dynamics of the transmission of malaria between mosquitoes and humans. Although his work received immediate acceptance in the medical community, his conclusion that malaria could be controlled by controlling mosquitoes was dismissed on the grounds that it would be impossible to rid a region of mosquitoes completely and that in any case, mosquitoes would soon invade the region. After Ross formulated a mathematical model that predicted that malaria outbreaks could be avoided if the mosquito population could be reduced below a critical threshold level, field trials supported his conclusions and led to sometimes brilliant successes in malaria control. Unfortunately, the Garki project provides a dramatic counterexample. This project worked to eradicate malaria from a region temporarily. However, people who have recovered from an attack of malaria have a temporary immunity against reinfection. Thus elimination of malaria from a region leaves the inhabitants of this region without immunity when the campaign ends, and the result can be a serious outbreak of malaria.

We formulate our descriptions as *compartmental models*, with the population under study being divided into compartments and with assumptions about the nature and time rate of transfer from one compartment to another. Diseases that confer immunity have a different compartmental structure from diseases without immunity and from *diseases transmitted by vectors*. The rates of transfer between compartments are expressed mathematically as derivatives with respect to time of the sizes of the compartments, and as a result our models are formulated initially as *differential equations*. Models in which the rates of transfer depend on the sizes of compartments over the past as well as at the instant of transfer lead to more general types of functional equations, such as differential–difference equations and integral equations.

In this chapter we describe models for epidemics, acting on a sufficiently rapid time scale that demographic effects, such as births, natural deaths, immigration into and emigration out of a population may be ignored. In the next chapter we will describe models in which demographic effects are included.

9.2 The Simple Kermack–McKendrick Epidemic Model

Throughout history, epidemics have had major effects on the course of events. For example, the Black Death, now identified as probably having been the bubonic plague which had actually invaded Europe as early as the sixth century, spread from Asia throughout Europe in several waves during the fourteenth century, beginning in 1346, and is estimated to have caused the death of as much as one third of the population of Europe between 1346 and 1350. The disease recurred regularly in various parts of Europe for more than 300 years, notably as the Great Plague of London of 1665–1666. It then gradually withdrew from Europe.

More than 15% of the population of London died in the Great Plague (1665–1666). It appeared quite suddenly, grew in intensity, and then disappeared, leaving part of the population untouched. One of the early triumphs of mathematical epidemiology was the formulation of a simple model by Kermack and McKendrick (1927) whose predictions are very similar to this behavior, observed in countless epidemics. The Kermack–McKendrick model is a compartmental model based on relatively simple assumptions on the rates of flow between different classes of members of the population.

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

We will use the terminology SIR to describe a disease that confers immunity against reinfection, to indicate that the passage of individuals is from the susceptible class S to the infective class I to the removed class R . Epidemics are usually diseases of this type. We would use the terminology SIS to describe a disease with no immunity against re-infection, to indicate that the passage of individuals is from the susceptible class to the infective class and then back to the susceptible class. Usually, diseases caused by a virus are of SIR type, while diseases caused by bacteria are of SIS type.

In addition to the basic distinction between diseases for which recovery confers immunity against reinfection and diseases for which recovered members are susceptible to reinfection, and the intermediate possibility of temporary immunity signified by a model of SIRS type, more complicated compartmental structure is possible. For example, there are SEIR and SEIS models, with an exposed period between being infected and becoming infective.

When there are only a few infected members, the start of a disease outbreak depends on random contacts between small numbers of individuals. In the next section we will use this to describe an approach to the study of the beginning of a disease outbreak by means of branching processes, but we begin with a description of deterministic compartmental models.

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

We are assuming that the epidemic process is *deterministic*, that is, that the behavior of a population is determined completely by its history and by the rules that describe the model. In formulating models in terms of the derivatives of the sizes of each compartment we are also assuming that the number of members in a compartment is a differentiable function of time. This assumption is plausible once a disease outbreak has become established but is not valid at the beginning of a disease outbreak when there are only a few infectives. In the next section we will describe a different approach for the initial stage of a disease outbreak.

The basic compartmental models to describe the transmission of communicable diseases are contained in a sequence of three papers by W.O. Kermack and A.G. McKendrick in 1927, 1932, and 1933. The first of these papers described epidemic models. What is often called the Kermack–McKendrick epidemic model is actually a special case of the general model introduced in this paper. The general model included dependence on age of infection, that is, the time since becoming infected, and can be used to provide a unified approach to compartmental epidemic models.

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

$$\begin{aligned} S' &= -\beta SI, \\ I' &= \beta SI - \alpha I, \\ R' &= \alpha I. \end{aligned} \tag{9.1}$$

A flow chart is shown in Figure 9.1. It is based on the following assumptions:

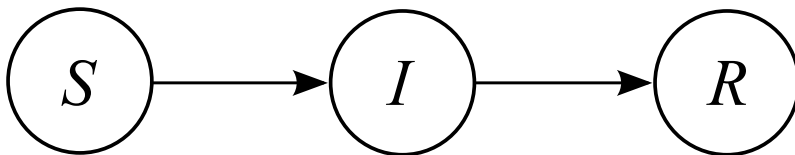


Fig. 9.1 Flow chart for the SIR model.

- (i) An average member of the population makes contact sufficient to transmit infection with βN others per unit time, where N represents total population size (mass action incidence).

- (ii) Infectives leave the infective class at rate αI per unit time.
- (iii) There is no entry into or departure from the population, except possibly through death from the disease.
- (iv) There are no disease deaths, and the total population size is a constant N .

According to (i), since the probability that a random contact by an infective is with a susceptible, who can then transmit infection, is S/N , the number of new infections in unit time per infective is $(\beta N)(S/N)$, giving a rate of new infections $(\beta N)(S/N)I = \beta SI$. Alternatively, we may argue that for a contact by a susceptible the probability that this contact is with an infective is I/N and thus the rate of new infections per susceptible is $(\beta N)(I/N)$, giving a rate of new infections $(\beta N)(I/N)S = \beta SI$. Note that both approaches give the same rate of new infections; in models with more complicated compartmental structure one may be more appropriate than the other.

We need not give an algebraic expression for N , since it cancels out of the final model, but we should note that for an SIR disease model, $N = S + I + R$. Later, we will allow the possibility that some infectives recover while others die of the disease. The hypothesis (iii) really says that the time scale of the disease is much faster than the time scale of births and deaths, so that demographic effects on the population may be ignored. An alternative view is that we are interested only in studying the dynamics of a single epidemic outbreak.

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

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

and the solution of this elementary differential equation is

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

Thus, the fraction of infectives remaining infective s time units after having become infective is $e^{-\alpha s}$, so that the length of the infective period is distributed exponentially with mean $\int_0^\infty e^{-\alpha s} ds = 1/\alpha$, and this is what (ii) really assumes. If we assume, instead of (ii), that the fraction of infectives remaining infective a time τ after having become infective is $P(\tau)$, the second equation of (9.1) would be replaced by the integral equation

$$I(t) = I_0(t) + \int_0^\infty \beta S(t-\tau)I(t-\tau)P(\tau)d\tau,$$

where $I_0(t)$ represents the members of the population who were infective at time $t = 0$ and are still infective at time t .

The assumptions of a rate of contacts proportional to population size N with constant of proportionality β and of an exponentially distributed recovery rate are

unrealistically simple. More general models can be constructed and analyzed, but our goal here is to show what may be deduced from extremely simple models. It will turn out that many more realistic models exhibit very similar qualitative behaviors.

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

$$\begin{aligned} S' &= -\beta SI, \\ I' &= (\beta S - \alpha)I, \end{aligned} \tag{9.2}$$

together with initial conditions

$$S(0) = S_0, \quad I(0) = I_0, \quad S_0 + I_0 = N.$$

We think of introducing a small number of infectives into a population of susceptibles and ask whether there will be an epidemic. We remark that the model makes sense only so long as $S(t)$ and $I(t)$ remain nonnegative. Thus if either $S(t)$ or $I(t)$ reaches zero, we consider the system to have terminated. We observe that $S' < 0$ for all t and $I' > 0$ if and only if $S > \alpha/\beta$. Thus I increases so long as $S > \alpha/\beta$, but since S decreases for all t , I ultimately decreases and approaches zero. If $S_0 < \alpha/\beta$, I decreases to zero (no epidemic), while if $S_0 > \alpha/\beta$, I first increases to a maximum attained when $S = \alpha/\beta$ and then decreases to zero (epidemic).

The quantity $\beta S_0/\alpha$ is a threshold quantity, called the *basic reproduction number* [Heesterbeek (1996)] and denoted by \mathcal{R}_0 , which determines whether there is an epidemic. If $\mathcal{R}_0 < 1$, the infection dies out, while if $\mathcal{R}_0 > 1$, there is an epidemic. The definition of the basic reproduction number \mathcal{R}_0 is that it is the number of secondary infections caused by a single infective introduced into a wholly susceptible population of size $N \approx S_0$ over the course of the infection of this single infective. In this situation, an infective makes βN contacts in unit time, all of which are with susceptibles and thus produce new infections, and the mean infective period is $1/\alpha$; thus the basic reproduction number is actually $\beta N/\alpha$ rather than $\beta S_0/\alpha$. Another way to view this apparent discrepancy is to consider two ways in which an epidemic may begin. One way is an epidemic started by a member of the population being studied, for example by returning from travel with an infection acquired away from home. In this case we would have $I_0 > 0, S_0 + I_0 = N$. A second way is for an epidemic to be started by a visitor from outside the population. In this case, we would have $S_0 = N$.

Since (9.2) is a two-dimensional autonomous system of differential equations, the natural approach would be to find equilibria and linearize about each equilibrium to determine its stability. However, since every point with $I = 0$ is an equilibrium, the system (9.2) has a line of equilibria, and this approach is not applicable (the linearization matrix at each equilibrium has a zero eigenvalue).

Fortunately, there is an alternative approach that enables us to analyze the system (9.2). The sum of the two equations of (9.2) is

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

Thus $S + I$ is a nonnegative smooth decreasing function and therefore tends to a limit as $t \rightarrow \infty$. Also, it is not difficult to prove that the derivative of a nonnegative smooth decreasing function must tend to zero, and this shows that

$$I_\infty = \lim_{t \rightarrow \infty} I(t) = 0.$$

Thus $S + I$ has limit S_∞ .

Integration of the sum of the two equations of (9.2) from 0 to ∞ gives

$$\alpha \int_0^\infty (S(t) + I(t)) dt = S_0 + I_0 - S_\infty = N - S_\infty.$$

Division of the first equation of (9.2) by S and integration from 0 to ∞ gives

$$\log \frac{S_0}{S_\infty} = \beta \int_0^\infty I(t) dt = \frac{\beta}{\alpha} [N - S_\infty] = \mathcal{R}_0 \left[1 - \frac{S_\infty}{N} \right]. \quad (9.3)$$

Equation (9.3) is called the *final size relation*. It gives a relationship between the basic reproduction number and the size of the epidemic. Note that the final size of the epidemic, the number of members of the population who are infected over the course of the epidemic, is $N - S_\infty$. This is often described in terms of the *attack rate* $(1 - S_\infty/N)$. [Technically, the attack rate should be called an attack ratio, since it is dimensionless and is not a rate.]

The final size relation (9.3) can be generalized to epidemic models with more complicated compartmental structure than the simple SIR model (9.2), including models with exposed periods, treatment models, and models including quarantine of suspected individuals and isolation of diagnosed infectives. The original Kermack–McKendrick model (1927) included dependence on the time since becoming infected (age of infection), and this includes such models.

Integration of the first equation of (9.2) from 0 to t gives

$$\log \frac{S_0}{S(t)} = \beta \int_0^t I(t) dt = \frac{\beta}{\alpha} [N - S(t) - I(t)],$$

and this leads to the form

$$I(t) + S(t) - \frac{\alpha}{\beta} \log S(t) = N - \frac{\alpha}{\beta} \log S_0. \quad (9.4)$$

This implicit relation between S and I describes the orbits of solutions of (9.2) in the (S, I) plane.

In addition, since the right side of (9.3) is finite, the left side is also finite, and this shows that $S_\infty > 0$. The final size relation (9.3) is valid for a large variety of epidemic models, as we shall see in later sections.

It is not difficult to prove that there is a unique solution of the final size relation (9.3). To see this, we define the function

$$g(x) = \log \frac{S_0}{x} - \mathcal{R}_0 \left[1 - \frac{x}{N} \right].$$

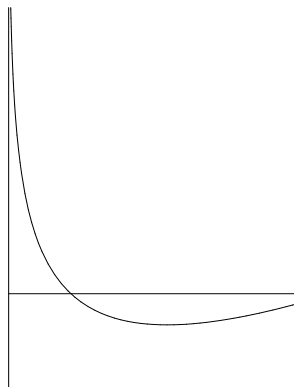


Fig. 9.2 The function $g(x)$.

Then, as shown in [Figure 9.2](#),

$$g(0+) > 0, \quad g(N) < 0,$$

and $g'(x) < 0$ if and only if

$$0 < x < \frac{N}{\mathcal{R}_0}.$$

If $\mathcal{R}_0 \leq 1$, $g(x)$ is monotone decreasing from a positive value at $x = 0+$ to a negative value at $x = N$. Thus there is a unique zero S_∞ of $g(x)$ with $S_\infty < N$.

If $\mathcal{R}_0 > 1$, $g(x)$ is monotone decreasing from a positive value at $x = 0+$ to a minimum at $x = N/\mathcal{R}_0$ and then increases to a negative value at $x = N$. Thus there is a unique zero S_∞ of $g(x)$ with

$$S_\infty < \frac{N}{\mathcal{R}_0}.$$

In fact,

$$g\left(\frac{S_0}{\mathcal{R}_0}\right) = \log \mathcal{R}_0 - \mathcal{R}_0 + \frac{S_0}{N} \leq \log \mathcal{R}_0 - \mathcal{R}_0 + 1.$$

Since $\log \mathcal{R}_0 < \mathcal{R}_0 - 1$ for $\mathcal{R}_0 > 0$, we actually have

$$g\left(\frac{S_0}{\mathcal{R}_0}\right) < 0,$$

and

$$S_\infty < \frac{S_0}{\mathcal{R}_0}.$$

It is generally difficult to estimate the contact rate β , which depends on the particular disease being studied but may also depend on social and behavioral factors. The quantities S_0 and S_∞ may be estimated by serological studies (measurements of immune responses in blood samples) before and after an epidemic, and from these data the basic reproduction number \mathcal{R}_0 may be estimated using (9.3). This estimate, however, is a retrospective one, which can be derived only after the epidemic has run its course.

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

$$I_{\max} = S_0 + I_0 - \frac{\alpha}{\beta} \log S_0 - \frac{\alpha}{\beta} + \frac{\alpha}{\beta} \log \frac{\alpha}{\beta}, \quad (9.5)$$

obtained by substituting $S = \alpha/\beta$, $I = I_{\max}$ into (9.4).

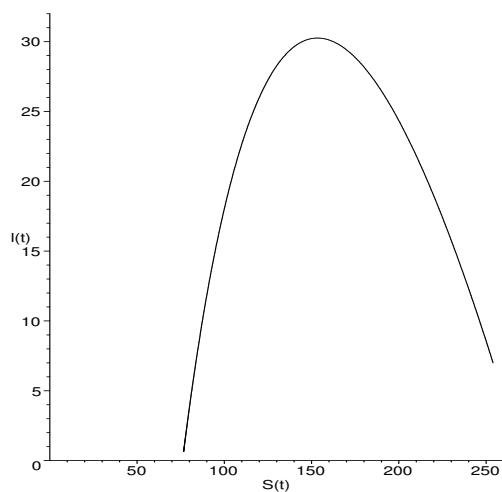
Example 1. A study of Yale University freshmen [Evans (1982), reported by Hethcote(1989)] described an influenza epidemic with $S_0 = 0.911$, $S_\infty = 0.513$. Here we are measuring the number of susceptibles as a fraction of the total population size, or using the population size K as the unit of size. Substitution into the final size relation gives the estimate $\beta/\alpha = 1.18$ and $R_0 = 1.18$. Since we know that τ is approximately 3 days for influenza, we see that β is approximately 0.48 contacts per day per member of the population.

Example 2. (The Great Plague in Eyam) The village of Eyam near Sheffield, England, suffered an outbreak of bubonic plague in 1665–1666 the source of which is generally believed to be the Great Plague of London. The Eyam plague was survived by only 83 of an initial population of 350 persons. Since detailed records were preserved and the community was persuaded to quarantine itself to try to prevent the spread of disease to other communities, the disease in Eyam has been used as a case study for modeling [Raggett (1982)]. Detailed examination of the data indicates that there were actually two outbreaks, of which the first was relatively mild. Thus we shall try to fit the model (9.2) over the period from mid-May to mid-October 1666, measuring time in months with an initial population of 7 infectives and 254 susceptibles, and a final population of 83. Raggett (1982) gives values of susceptibles and infectives in Eyam on various dates, beginning with $S(0) = 254$, $I(0) = 7$, shown in Table 9.1.

The final size relation with $S_0 = 254$, $I_0 = 7$, $S_\infty = 83$ gives $\beta/\alpha = 6.54 \times 10^{-3}$, $\alpha/\beta = 153$. The infective period was 11 days, or 0.3667 month, so that $\alpha = 2.73$. Then $\beta = 0.0178$. The relation (9.5) gives an estimate of 30.4 for the maximum number of infectives. We use the values obtained here for the parameters β and τ in the model (9.2) for simulations of both the phase plane, here the (S, I) -plane, and for graphs of S and I as functions of t (Figures 9.3, 9.4, 9.5). Figure 9.6 plots these data points together with the phase portrait given in Figure 9.3 for the model (9.2).

The actual data for the Eyam epidemic are remarkably close to the predictions of this very simple model. However, the model is really too good to be true. Our model assumes that infection is transmitted directly between people. While this is

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

Table 9.1 Eyam Plague data.**Fig. 9.3** The S - I plane.

possible, bubonic plague is transmitted mainly by rat fleas. When an infected rat is bitten by a flea, the flea becomes extremely hungry and bites the host rat repeatedly, spreading the infection in the rat. When the host rat dies, its fleas move on to other rats, spreading the disease further. As the number of available rats decreases, the fleas move to human hosts, and this is how plague starts in a human population (although the second phase of the epidemic may have been the pneumonic form of bubonic plague, which can be spread from person to person). One of the main reasons for the spread of plague from Asia into Europe was the passage of many trading ships; in medieval times ships were invariably infested with rats. An accurate model of plague transmission would have to include flea and rat populations, as well as movement in space. Such a model would be extremely complicated, and its predictions might well not be any closer to observations than our simple unrealistic model. Very recent study of the data from Eyam suggests that the rat population may not have been large enough to support the epidemic and human to human transmission

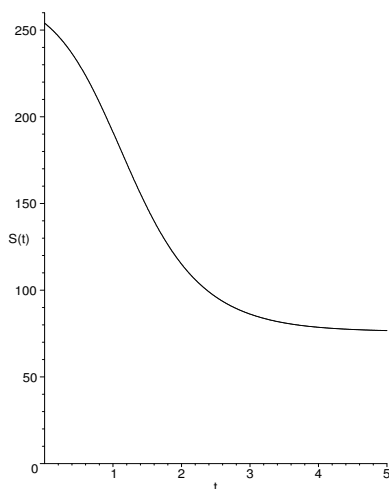


Fig. 9.4 S as a function of t .

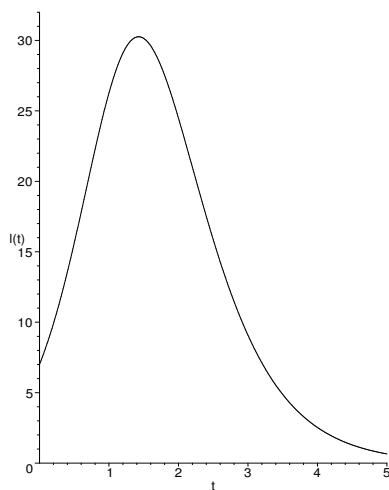


Fig. 9.5 I as a function of t .

was also a factor. Raggett (1982) also used a stochastic model to fit the data, but the fit was rather poorer than the fit for the simple deterministic model(9.2).

In the village of Eyam the rector persuaded the entire community to quarantine itself to prevent the spread of disease to other communities. One effect of this policy was to increase the infection rate in the village by keeping fleas, rats, and people in close contact with one another, and the mortality rate from bubonic plague was much higher in Eyam than in London. Further, the quarantine could do nothing to prevent the travel of rats and thus did little to prevent the spread of disease to other communities. One message this suggests to mathematical modelers is that control

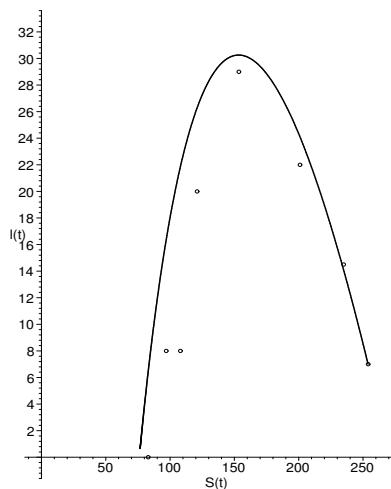


Fig. 9.6 The S - I plane, model and data.

strategies based on false models may be harmful, and it is essential to distinguish between assumptions that simplify but do not alter the predicted effects substantially, and wrong assumptions that make an important difference.

In order to prevent the occurrence of an epidemic if infectives are introduced into a population, it is necessary to reduce the basic reproductive number R_0 below one. This may sometimes be achieved by immunization, which has the effect of transferring members of the population from the susceptible class to the removed class and thus of reducing $S(0)$. If a fraction p of the population is successfully immunized, the effect is to replace $S(0)$ by $S(0)(1 - p)$, and thus to reduce the basic reproductive number to $\beta S(0)(1 - p)/\alpha$. The requirement $\beta S(0)(1 - p)/\alpha < 1$ gives $1 - p < \alpha/\beta S(0)$, or

$$p > 1 - \frac{\alpha}{\beta S(0)} = 1 - \frac{1}{R_0}.$$

A large basic reproductive number means that the fraction that must be immunized to prevent the spread of infection is large. This relation is connected to the idea of herd immunity, which we shall introduce in the next chapter.

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

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

and the initial growth rate is

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

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

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

However, because of incomplete data and underreporting of cases, this estimate may not be very accurate. This inaccuracy is even more pronounced for an outbreak of a previously unknown disease, where early cases are likely to be misdiagnosed. Because of the final size relation, estimation of β or \mathcal{R}_0 is an important question that has been studied by a variety of approaches.

There are serious shortcomings in the simple Kermack–McKendrick model as a description of the beginning of a disease outbreak, and a very different kind of model is required.

Exercises

1. The same survey of Yale students described in Example 1 reported that 91.1 percent were susceptible to influenza at the beginning of the year and 51.4 percent were susceptible at the end of the year. Estimate the basic reproductive number β/α and decide whether there was an epidemic.
2. What fraction of Yale students in Exercise 1 would have had to be immunized to prevent an epidemic?
3. What was the maximum number of Yale students in Exercises 1 and 2 suffering from influenza at any time?
4. An influenza epidemic was reported at an English boarding school in 1978 that spread to 512 of the 763 students. Estimate the basic reproductive number β/α .
5. What fraction of the boarding school students in Exercise 4 would have had to be immunized to prevent an epidemic?
6. What was the maximum number of boarding school students in Exercises 4 and 5 suffering from influenza at any time?
7. A disease is introduced by two visitors into a town with 1200 inhabitants. An average infective is in contact with 0.4 inhabitants per day. The average duration of the infective period is 6 days, and recovered infectives are immune against reinfection. How many inhabitants would have to be immunized to avoid an epidemic?
8. Consider a disease with $\beta = 1/3000$, $1/\alpha = 6$ days in a population of 1200 members. Suppose the disease conferred immunity on recovered infectives. How many members would have to be immunized to avoid an epidemic?
9. A disease begins to spread in a population of 800. The infective period has an average duration of 14 days and the average infective is in contact with 0.1 persons per day. What is the basic reproductive number? To what level must the average rate of contact be reduced so that the disease will die out?
10. European fox rabies is estimated to have a transmission coefficient β of $80 \text{ km}^2 \text{ years/fox}$ and an average infective period of 5 days. There is a critical carrying capacity K_c measured in foxes per km^2 , such that in regions with fox density less than K_c , rabies tends to die out, while in regions with fox density greater

than K_c , rabies tends to persist. Estimate K_c . [Remark: It has been suggested in Great Britain that hunting to reduce the density of foxes below the critical carrying capacity would be a way to control the spread of rabies.]

11. A large English estate has a population of foxes with a density of 1.3 foxes/km². A large fox hunt is planned to reduce the fox population enough to prevent an outbreak of rabies. Assuming that the contact number β/α is 1 km²/fox, find what fraction of the fox population must be caught.
12. Following a complaint from the SPCA, organizers decide to replace the fox hunt of Exercise 1 by a mass inoculation of foxes for rabies. What fraction of the fox population must be inoculated to prevent a rabies outbreak?
13. What actually occurs on the estate of these exercises is that 10 percent of the foxes are killed and 15 percent are inoculated. Is there danger of a rabies outbreak.
14. Here is another approach to the analysis of the SIR model (9.2).

(i) Divide the two equations of the model to give

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

(ii) Integrate to find the orbits in the (S, I) -plane,

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

with c an arbitrary constant of integration.

(iii) Define the function

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

and show that each orbit is given implicitly by the equation $V(S, I) = c$ for some choice of the constant c .

(iv) Show that no orbit reaches the I -axis and deduce that $S_\infty = \lim_{t \rightarrow \infty} S(t) > 0$, which implies that part of the population escapes infection.

9.3 A Branching-Process Disease-Outbreak Model

The Kermack–McKendrick compartmental epidemic model assumes that the sizes of the compartments are large enough that the mixing of members is homogeneous, or at least that there is homogeneous mixing in each subgroup if the population is stratified by activity levels. However, at the beginning of a disease outbreak, there is a very small number of infective individuals, and the transmission of infection is a stochastic event depending on the pattern of contacts between members of the population; a description should take this pattern into account.

Our approach will be to give a stochastic-branching process description of the beginning of a disease outbreak to be applied as long as the number of infectives remains small, distinguishing a (minor) disease outbreak confined to this stage from a (major) epidemic, which occurs if the number of infectives begins to grow at an exponential rate. Once an epidemic has started, we may switch to a deterministic compartmental model, arguing that in a major epidemic, contacts would tend to be more homogeneously distributed. Implicitly, we are thinking of an infinite population, and by a major epidemic we mean a situation in which a nonzero fraction of the population is infected, and by a minor outbreak we mean a situation in which the infected population may grow but remains a negligible fraction of the population.

There is an important difference between the behavior of branching process models and the behavior of models of Kermack–McKendrick type, namely, as we shall see in this section that for a stochastic disease outbreak model if $\mathcal{R}_0 < 1$, the probability that the infection will die out is 1, but if $\mathcal{R}_0 > 1$, there is a positive probability that the infection will increase initially but will produce only a minor outbreak and will die out before triggering a major epidemic.

We describe the network of contacts between individuals by a graph with members of the population represented by vertices and with contacts between individuals represented by edges. The study of graphs originated with the abstract theory of Erdős and Rényi of the 1950s and 1960s [Erdős and Rényi (1959, 1960, 1961)]. It has become important in many areas of application, including social contacts and computer networks, as well as the spread of communicable diseases. We will think of networks as bidirectional, with disease transmission possible in either direction along an edge.

An edge is a contact between vertices that can transmit infection. The number of edges of a graph at a vertex is called the *degree* of the vertex. The degree distribution of a graph is $\{p_k\}$, where p_k is the fraction of vertices having degree k . The degree distribution is fundamental in the description of the spread of disease.

We think of a small number of infectives in a population of susceptibles large enough that in the initial stage, we may neglect the decrease in the size of the susceptible population. Our development begins along the lines of that of [Diekmann and Heesterbeek (2000)] and then develops along the lines of [Callaway, Newman, Strogatz, and Watts (2000), Newman (2002), Newman, Strogatz, and Watts (2002)]. We assume that the infectives make contacts independently of one another and let p_k denote the probability that the number of contacts by a randomly chosen individual is exactly k , with $\sum_{k=0}^{\infty} p_k = 1$. In other words, $\{p_k\}$ is the degree distribution of the vertices of the graph corresponding to the population network. For the moment, we assume that every contact leads to an infection, but we will relax this assumption later.

It is convenient to define the *probability generating function*

$$G_0(z) = \sum_{k=0}^{\infty} p_k z^k.$$

Since $\sum_{k=0}^{\infty} p_k = 1$, this power series converges for $0 \leq z \leq 1$, and may be differentiated term by term. Thus

$$p_k = \frac{G_0^{(k)}(0)}{k!}, \quad k = 0, 1, 2, \dots$$

It is easy to verify that the generating function has the properties

$$G_0(0) = p_0, \quad G_0(1) = 1, \quad G_0'(z) > 0, \quad G_0''(z) > 0.$$

The mean degree, which we denote by $\langle k \rangle$ or z_1 , is

$$\langle k \rangle = \sum_{k=1}^{\infty} k p_k = G_0'(1).$$

More generally, we define the moments

$$\langle k^j \rangle = \sum_{k=1}^{\infty} k^j p_k, \quad j = 1, 2, \dots, \infty.$$

When a disease is introduced into a network, we think of it as starting at a vertex (patient zero) that transmits infection to every individual to whom this individual is connected, that is, along every edge of the graph from the vertex corresponding to this individual. We may think of this individual as being inside the population, as when a member of a population returns from travel after being infected, or as being outside the population, as when someone visits a population and brings an infection. For transmission of disease after this initial contact we need to use the *excess degree* of a vertex. If we follow an edge to a vertex, the excess degree of this vertex is one less than the degree. We use the excess degree because infection cannot be transmitted back along the edge whence it came. The probability of reaching a vertex of degree k , or excess degree $(k-1)$, by following a random edge is proportional to k , and thus the probability that a vertex at the end of a random edge has excess degree $(k-1)$ is a constant multiple of $k p_k$ with the constant chosen to make the sum over k of the probabilities equal to 1. Then the probability that a vertex has excess degree $(k-1)$ is

$$q_{k-1} = \frac{k p_k}{\langle k \rangle}.$$

This leads to a generating function $G_1(z)$ for the excess degree,

$$G_1(z) = \sum_{k=1}^{\infty} q_{k-1} z^{k-1} = \sum_{k=1}^{\infty} \frac{k p_k}{\langle k \rangle} z^{k-1} = \frac{1}{\langle k \rangle} G_0'(z),$$

and the mean excess degree, which we denote by $\langle k_e \rangle$, is

$$\begin{aligned}
\langle k_e \rangle &= \frac{1}{\langle k \rangle} \sum_{k=1}^{\infty} k(k-1)p_k \\
&= \frac{1}{\langle k \rangle} \sum_{k=1}^{\infty} k^2 p_k - \frac{1}{\langle k \rangle} \sum_{k=1}^{\infty} k p_k \\
&= \frac{\langle k^2 \rangle}{\langle k \rangle} - 1 = G'_1(1).
\end{aligned}$$

We let $\mathcal{R}_0 = G'_1(1)$, the mean excess degree. This is the mean number of secondary cases infected by patient zero and is the basic reproduction number as usually defined; the threshold for an epidemic is determined by \mathcal{R}_0 . The quantity $\langle k_e \rangle = G'_1(1)$ is sometimes written in the form

$$\langle k_e \rangle = G'_1(1) = \frac{z_2}{z_1},$$

where $z_2 = \sum_{k=1}^{\infty} k(k-1)p_k = \langle k^2 \rangle - \langle k \rangle$ is the mean number of second neighbors of a random vertex.

Our next goal is to calculate the probability that the infection will die out and will not develop into a major epidemic, proceeding in two steps. First we find the probability that a secondary infected vertex (a vertex that has been infected by another vertex in the population) will not spark a major epidemic.

Suppose that the secondary infected vertex has excess degree j . We let z_n denote the probability that this infection dies out within the next n generations. For the infection to die out in n generations, each of the j secondary infections coming from the initial secondary infected vertex must die out in $(n-1)$ generations. The probability of this is z_{n-1} for each secondary infection, and the probability that all secondary infections will die out in $(n-1)$ generations is z_{n-1}^j . Now z_n is the sum over j of these probabilities, weighted by the probability q_j of j secondary infections. Thus

$$z_n = \sum_{j=0}^{\infty} q_j z_{n-1}^j = G_1(z_{n-1}).$$

Since $G_1(z)$ is an increasing function, the sequence z_n is an increasing sequence and has a limit z_{∞} , which is the probability that this infection will die out eventually. Then z_{∞} is the limit as $n \rightarrow \infty$ of the solution of the difference equation

$$z_n = G_1(z_{n-1}), \quad z_0 = 0.$$

Thus z_{∞} must be an equilibrium of this difference equation, that is, a solution of $z = G_1(z)$. Let w be the smallest positive solution of $z = G_1(z)$. Then, because $G_1(z)$ is an increasing function of z , $z \leq G_1(z) \leq G_1(w) = w$ for $0 \leq z \leq w$. Since $z_0 = 0 < w$ and $z_{n-1} \leq w$ implies

$$z_n = G_1(z_{n-1}) \leq G_1(w) = w,$$

it follows by induction that

$$z_n \leq w, \quad n = 0, 1, \dots, \infty.$$

From this we deduce that

$$z_\infty = w.$$

The equation $G_1(z) = z$ has a root $z = 1$, since $G_1(1) = 1$. Because the function $G_1(z) - z$ has a positive second derivative, its derivative $G'_1(z) - 1$ is increasing and can have at most one zero. This implies that the equation $G_1(z) = z$ has at most two roots in $0 \leq z \leq 1$. If $\mathcal{R}_0 < 1$, the function $G_1(z) - z$ has a negative first derivative

$$G'_1(z) - 1 \leq G'_1(1) - 1 = \mathcal{R}_0 - 1 < 0,$$

and the equation $G_1(z) = z$ has only one root, namely $z = 1$. On the other hand, if $\mathcal{R}_0 > 1$, the function $G_1(z) - z$ is positive for $z = 0$ and negative near $z = 1$ since it is zero at $z = 1$, and its derivative is positive for $z < 1$ and z near 1. Thus in this case the equation $G_1(z) = z$ has a second root $z_\infty < 1$.

This root z_∞ is the probability that an infection transmitted along one of the edges at the initial secondary vertex will die out, and this probability is independent of the excess degree of the initial secondary vertex. It is also the probability that an infection originating outside the population, such as an infection brought from outside into the population under study, will die out.

Next, we calculate the probability that an infection originating at a primary infected vertex, such as an infection introduced by a visitor from outside the population under study, will die out. The probability that the disease outbreak will die out eventually is the sum over k of the probabilities that the initial infection in a vertex of degree k will die out, weighted by the degree distribution $\{p_k\}$ of the original infection, and this is

$$\sum_{k=0}^{\infty} p_k z_\infty^k = G_0(z_\infty).$$

To summarize this analysis, we see that if $\mathcal{R}_0 < 1$, the probability that the infection will die out is 1. On the other hand, if $\mathcal{R}_0 > 1$, there is a solution $z_\infty < 1$ of

$$G_1(z) = z,$$

and there is a probability $1 - G_0(z_\infty) > 0$ that the infection will persist, and will lead to an epidemic. However, there is a positive probability $G_0(z_\infty)$ that the infection will increase initially but will produce only a minor outbreak and will die out before triggering a major epidemic. This distinction between a minor outbreak and a major epidemic, and the result that if $\mathcal{R}_0 > 1$ there may be only a minor outbreak and not a major epidemic, are aspects of stochastic models not reflected in deterministic models.

If contacts between members of the population are random, corresponding to the assumption of mass action in the transmission of disease, then the probabilities p_k are given by the *Poisson distribution*

$$p_k = \frac{e^{-c} c^k}{k!}$$

for some constant c [Brauer, van den Driessche, and Wu (2008), pp. 142–143]. The generating function for the Poisson distribution is $e^{c(z-1)}$. Then $G_1(z) = G_0(z)$, and $\mathcal{R}_0 = c$, so that

$$G_1(z) = G_0(z) = e^{\mathcal{R}_0(z-1)}.$$

The commonly observed situation that most infectives do not pass on infection but there are a few “superspreading events” [Riley et al. (2003)] corresponds to a probability distribution that is quite different from a Poisson distribution, and could give a quite different probability that an epidemic will occur. For example, if $\mathcal{R}_0 = 2.5$, the assumption of a Poisson distribution gives $z_\infty = 0.107$ and $G_0(z_\infty) = 0.107$, so that the probability of an epidemic is 0.893. The assumption that nine out of ten infectives do not transmit infection while the tenth transmits 25 infections gives

$$G_0(z) = (z^{25} + 9)/10, \quad G_1(z) = z^{24}, \quad z_\infty = 0, \quad G_0(z_\infty) = 0.9,$$

from which we see that the probability of an epidemic is 0.1. Another example, possibly more realistic, is to assume that a fraction $(1 - p)$ of the population follows a Poisson distribution with constant r , while the remaining fraction p consists of superspreaders each of whom makes L contacts. This would give a generating function

$$\begin{aligned} G_0(z) &= (1 - p)e^{r(z-1)} + pz^L, \\ G_1(z) &= \frac{r(1 - p)e^{r(z-1)} + pLz^{L-1}}{r(1 - p) + pL}, \end{aligned}$$

and

$$\mathcal{R}_0 = \frac{r^2(1 - p) + pL(L - 1)}{r(1 - p) + pL}.$$

For example, if $r = 2.2$, $L = 10$, $p = 0.01$, numerical simulation gives

$$\mathcal{R}_0 = 2.5, \quad z_\infty = 0.146,$$

so that the probability of an epidemic is 0.849.

These examples demonstrate that the probability of a major epidemic depends strongly on the nature of the contact network. Simulations suggest that for a given value of the basic reproduction number, the Poisson distribution is the one with the maximum probability of a major epidemic.

It has been observed that in many situations there is a small number of long-range connections in the graph, allowing rapid spread of infection. There is a high degree of clustering (some vertices with many edges), and there are short path lengths. Such a situation may arise if a disease is spread to a distant location by an air traveler. This type of network is called a *small-world* network. Long range connections in a network can increase the likelihood of an epidemic dramatically.

These examples indicate that the probability of an epidemic depends strongly on the contact network at the beginning of a disease outbreak. We will not explore network models further here, but we point out that this is an actively developing field of science. Some basic references are [Newman (2003, Strogatz (2001))].

9.3.1 Transmissibility

Contacts do not necessarily transmit infection. For each contact between individuals of whom one has been infected and the other is susceptible, there is a probability that infection will actually be transmitted. This probability depends on such factors as the closeness of the contact, the infectivity of the member who has been infected, and the susceptibility of the susceptible member. We assume that there is a mean probability T , called the *transmissibility*, of transmission of infection. The transmissibility depends on the rate of contacts, the probability that a contact will transmit infection, the duration time of the infection, and the susceptibility. The development in Section 9.2 assumed that all contacts transmit infection, that is, that $T = 1$.

In this section, we will continue to assume that there is a network describing the contacts between members of the population whose degree distribution is given by the generating function $G_0(z)$, but we will assume in addition that there is a mean transmissibility T .

When disease begins in a network, it spreads to some of the vertices of the network. Edges that are infected during a disease outbreak are called *occupied*, and the size of the disease outbreak is the cluster of vertices connected to the initial vertex by a continuous chain of occupied edges.

The probability that exactly m infections are transmitted by an infective vertex of degree k is

$$\binom{k}{m} T^m (1 - T)^{k-m}.$$

We define $I_0(z, T)$ to be the generating function for the distribution of the number of occupied edges attached to a randomly chosen vertex, which is the same as the distribution of the infections transmitted by a randomly chosen individual for any (fixed) transmissibility T . Then

$$\begin{aligned} I_0(z, T) &= \sum_{m=0}^{\infty} \left[\sum_{k=m}^{\infty} p_k \binom{k}{m} T^m (1 - T)^{k-m} \right] z^m \\ &= \sum_{k=0}^{\infty} p_k \left[\sum_{m=0}^k \binom{k}{m} (zT)^m (1 - T)^{k-m} \right] \\ &= \sum_{k=0}^{\infty} p_k [zT + (1 - T)]^k = G_0(1 + (z - 1)T). \end{aligned} \tag{9.6}$$

In this calculation we have used the binomial theorem to see that

$$\sum_{m=0}^k \binom{k}{m} (zT)^m (1-T)^{(k-m)} = [zT + (1-T)]^k.$$

Note that

$$\Gamma_0(0, T) = G_0(1 - T), \quad \Gamma_0(1, T) = G_0(1) = 1, \quad \Gamma'_0(z, T) = TG'_0(1 + (z - 1)T).$$

For secondary infections we need the generating function $\Gamma_1(z, T)$ for the distribution of occupied edges leaving a vertex reached by following a randomly chosen edge. This is obtained from the excess degree distribution in the same way,

$$\Gamma_1(z, T) = G_1(1 + (z - 1)T)$$

and

$$\Gamma_1(0, T) = G_1(1 - T), \quad \Gamma_1(1, T) = G_1(1) = 1, \quad \Gamma'_1(z, T) = TG'_1(1 + (z - 1)T).$$

The basic reproduction number is now

$$\mathcal{R}_0 = \Gamma'_1(1, T) = TG'_1(1).$$

The calculation of the probability that the infection will die out and will not develop into a major epidemic follows the same lines as the argument for $T = 1$. The result is that if $\mathcal{R}_0 = TG'_1(1) < 1$, the probability that the infection will die out is 1. If $\mathcal{R}_0 > 1$, there is a solution $z_\infty(T) < 1$ of

$$\Gamma_1(z, T) = z,$$

and a probability $1 - \Gamma_0(z_\infty(T), T) > 0$ that the infection will persist, and will lead to an epidemic. However, there is a positive probability $\Gamma_1(z_\infty(T), T)$ that the infection will increase initially but will produce only a minor outbreak and will die out before triggering a major epidemic.

Another interpretation of the basic reproduction number is that there is a *critical transmissibility* T_c defined by

$$T_c G'_1(1) = 1.$$

In other words, the critical transmissibility is the transmissibility that makes the basic reproduction number equal to 1. If the mean transmissibility can be decreased below the critical transmissibility, then an epidemic can be prevented.

The measures used to try to control an epidemic may include contact interventions, that is, measures affecting the network such as avoidance of public gatherings and rearrangement of the patterns of interaction between caregivers and patients in a hospital, and transmission interventions such as careful hand washing or face masks to decrease the probability that a contact will lead to disease transmission.

Exercises

In each exercise, assume that the transmissibility is 1.

1. Show that it is not possible for a major epidemic to develop unless at least one member of the contact network has degree at least 3.
2. What is the probability of a major epidemic if every member of the contact network has degree 3.
3. Consider a truncated Poisson distribution, with

$$p_k = \begin{cases} \frac{e^{-c} c^k}{k!}, & k \leq 10, \\ 0, & k > 10. \end{cases}$$

Estimate (numerically) the probability of a major epidemic if $c = 1.5$.

4. Show that the probability generating function for an exponential distribution, given by

$$p_k = (1 - e^{-1/r})e^{-k/r},$$

is

$$G_0(z) = \frac{1 - e^{-1/r}}{1 - ze^{-1/r}}.$$

5. A power law distribution is given by

$$p_k = Ck^{-\alpha}.$$

For what values of α is it possible to normalize this (i.e., choose C to make $\sum p_k = 1$)?

9.4 Network and Compartmental Epidemic Models

Compartmental models for epidemics are not suitable for describing the beginning of a disease outbreak because they assume that all members of a population are equally likely to make contact with a very small number of infectives. Thus, as we have seen in the preceding section, stochastic branching process models are better descriptions of the beginning of an epidemic. They allow the possibility that even if a disease outbreak has a reproduction number greater than 1, it may be only a minor outbreak and may not develop into a major epidemic. One possible approach to a more realistic description of an epidemic would be to use a branching process model initially and then make a transition to a compartmental model when the epidemic has become established and there are enough infectives that mass action mixing in the population is a reasonable approximation. Another approach would be to continue to use a network model throughout the course of the epidemic. In this section we shall indicate how a compartmental approach and a network approach are related.

The development is taken from Volz (2008), Miller (2010), and Miller and Volz (2011).

We assume that there is a known static *configuration model* (CM) network in which the probability that a node u has degree k_u is $P(k_u)$. We let $G_0(z)$ denote the probability generating function of the degree distribution,

$$G_0(z) = \sum_{k=0}^{\infty} p_k z^k,$$

with mean degree $\langle \rangle = G'_0(1)$.

The per-edge from an infected node is assumed to be β , and it is assumed that infected nodes recover at a rate α . We use an edge-based compartmental model because the probability that a random neighbor is infected is not necessarily the same as the probability that a random individual is infected. We let $S(t)$ denote the fraction of nodes that are susceptible at time t , $I(t)$ the fraction of nodes that are infective at time t , and $R(t)$ the fraction of nodes that are recovered at time t . It is easy to write an equation for R' , the rate at which infectives recover. If we know $S(t)$, we can find $I(t)$, because a decrease in S gives a corresponding increase in I . Since

$$S(t) + I(t) + R(t) = 1,$$

we need only find the probability that a randomly selected node is susceptible.

We assume that the hazard of infection for a susceptible node u is proportional to the degree k_u of the node. Each contact is represented by an edge of the network joining u to a neighboring node. We let φ_I denote the probability that this neighbor is infective. Then the per-edge hazard of infection is

$$\lambda_E = \beta \varphi_I.$$

Assuming that edges are independent, u 's hazard of infection at time t is

$$\lambda_u(t) = k_u \lambda_E(t) = k_u \beta \varphi_I(t).$$

Consider a randomly selected node u and let $\theta(t)$ be the probability that a random neighbor has not transmitted infection to u . Then the probability that u is susceptible is θ^{k_u} . Averaging over all nodes, we see that the probability that a random node u is susceptible is

$$S(t) = \sum_{k=0}^{\infty} P(k) [\theta(t)]^k = G_0(\theta(t)). \quad (9.7)$$

We break θ into three parts,

$$\theta = \varphi_S + \varphi_I + \varphi_R,$$

with φ_S the probability that a random neighbor v of u is susceptible, φ_I the probability that a random neighbor v of u is infective but has not transmitted infection to u ,

and φ_R the probability that a random neighbor v has recovered without transmitting infection to u . Then the probability that v has transmitted infection to u is $1 - \theta$.

Since infected neighbors recover at rate α , the flux from φ_I to φ_R is $\alpha\varphi_I$. Thus

$$\varphi'_R = \alpha\varphi_I.$$

It is easy to see from this that

$$R' = \alpha I. \quad (9.8)$$

Since edges from infected neighbors transmit infection at rate β , the flux from φ_I to $(1 - \theta)$ is $\beta\varphi_I$. Thus

$$\theta' = -\beta\varphi_I. \quad (9.9)$$

To obtain φ'_I we need the flux into and out of the φ_I compartment. The incoming flux from φ_S results from infection of the neighbor. The outgoing flux to φ_R corresponds to recovery of the neighbor without having transmitted infection, and the outgoing flux to $(1 - \theta)$ corresponds to transmission without recovery. The total outgoing flux is $(\alpha + \beta)\varphi_I$.

To determine the flux from φ_S to φ_I , we need the rate at which a neighbor changes from susceptible to infective. Consider a random neighbor v of u ; the probability that v has degree k is $kp(k)/\langle k \rangle$. Since there are $(k - 1)$ neighbors of v that could have infected v , the probability that v is susceptible is θ^{k-1} . Averaging over all k , we see that the probability that a random neighbor v of u is susceptible is

$$\varphi_S = \sum_{k=0}^{\infty} \frac{kp(k)}{\langle k \rangle} \theta^{k-1} = \frac{G'_0(\theta)}{G'_0(1)}. \quad (9.10)$$

To calculate φ_R , we note that the flux from φ_I to φ_R and the flux from φ_I to $(1 - \theta)$ are proportional with proportionality constant α/β . Since both φ_R and $(1 - \theta)$ start at zero,

$$\varphi_R = \frac{\alpha}{\beta}(1 - \theta). \quad (9.11)$$

Now, using (9.9), (9.10), (9.11), and

$$\varphi_I = \theta - \varphi_S - \varphi_R,$$

we obtain

$$\theta' = -\beta\varphi_I = -\beta\theta + \beta\varphi_S + \beta\varphi_R = -\beta\theta + \beta \frac{G'_0(\theta)}{G'_0(1)} + \alpha(1 - \theta). \quad (9.12)$$

We now have a dynamic model consisting of equations (9.12), (9.7), (9.8), and $S + I + R = 1$. We wish to show a relationship between this set of equations and the simple Kermack–McKendrick compartmental model (9.2). In order to accomplish this, we need only show under what conditions we would have $S' = -\beta SI$.

Differentiating (9.7) and using (9.9), we obtain

$$S' = G'_0(\theta)\theta' = -G'_0(\theta)\beta\varphi_I.$$

Consider a large population with N members, each making $C \leq N - 1$ contacts, so that

$$S = \theta^C, \quad G'_0(\theta) = \frac{CS}{\theta}S'(\theta),$$

and

$$S' = -\beta CS \frac{\varphi_I}{\theta}.$$

We now let $C \rightarrow \infty$ (which implies $N \rightarrow \infty$) in such a way that

$$\hat{\beta} = \beta C$$

remains constant. Then

$$S' = -\hat{\beta} \frac{\varphi_I}{\theta}.$$

We will now show that

$$\frac{\varphi_I}{\theta} \approx 1,$$

and this will yield the desired approximation

$$S' = -\hat{\beta} SI. \tag{9.13}$$

The probability that an edge to a randomly chosen node has not transmitted infection is θ (assuming that the given target node cannot transmit infection), and the probability that in addition it is connected to an infected node is φ_I . Because $\hat{\beta} = \beta C$ is constant and therefore bounded as C grows, only a fraction no greater than a constant multiple of I/C of edges to the target node may have transmitted infection from a node that is still infected. For large values of C , φ_I is approximately I . Similarly, θ is approximately 1 as $C \rightarrow \infty$. Thus $\varphi_I/\theta \approx I$ as $C \rightarrow \infty$. This gives the desired approximate equation for S . The result remains valid if all degrees are close to the average degree as the average degree grows.

The edge-based compartmental modeling approach that we have used can be generalized in several ways. For example, heterogeneity of mixing can be included. In general, one would expect that early infections would be in individuals having more contacts, and thus that an epidemic would develop more rapidly than a mass action compartmental model would predict. When contact duration is significant, as would be the case in sexually transmitted diseases, an individual with a contact would play no further role in disease transmission until a new contact is made, and this can be incorporated in a network model.

The network approach to disease modeling is a rapidly developing field of study, and there will undoubtedly be fundamental developments in our understanding of the modeling of disease transmission. Some useful references are [Bansal, Read, Pourbohloul, and Meyers (2010), Meyers (2007), Meyers et al. (2006), Meyers et al. (2005), Newman(2001), Newman (2002), Newman, Strogatz, and Watts (2001), Strogatz (2001)].

In the remainder of this chapter, we assume that we are in an epidemic situation following a disease outbreak that has been modeled initially by a branching process. Thus we return to the study of compartmental models.

9.5 More Complicated Epidemic Models

We have established that the simple Kermack–McKendrick epidemic model (9.2) has the following basic properties:

1. There is a basic reproduction number \mathcal{R}_0 such that if $\mathcal{R}_0 < 1$, the disease dies out while if $\mathcal{R}_0 > 1$, there is an epidemic.
2. The number of infectives always approaches zero and the number of susceptibles always approaches a positive limit as $t \rightarrow \infty$.
3. There is a relationship between the reproduction number and the final size of the epidemic, which is an equality if there are no disease deaths.

In fact, these properties hold for epidemic models with more complicated compartmental structure. We will describe some common epidemic models as examples.

9.5.1 Exposed Periods

In many infectious diseases there is an exposed period after the transmission of infection from susceptibles to potentially infective members but before these potential infectives develop symptoms and can transmit infection. To incorporate an exposed period with mean exposed period $1/\kappa$, we add an exposed class E and use compartments S, E, I, R and total population size $N = S + E + I + R$ to give a generalization of the epidemic model (9.2)

$$\begin{aligned} S' &= -\beta SI, \\ E' &= \beta SI - \kappa E, \\ I' &= \kappa E - \alpha I. \end{aligned} \tag{9.14}$$

A flow chart is shown in [Figure 9.7](#).

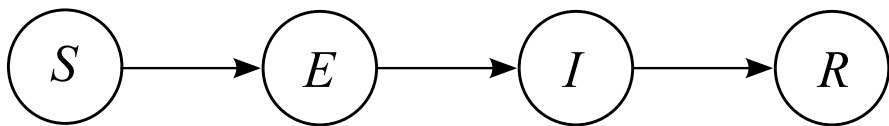


Fig. 9.7 Flow chart for the SEIR model.

The analysis of this model is the same as the analysis of (9.2), but with I replaced by $E + I$. That is, instead of using the number of infectives as one of the variables, we use the total number of infected members, whether or not they are capable of transmitting infection.

In some diseases there is some infectivity during the exposed period. This may be modeled by assuming infectivity reduced by a factor ε during the exposed period. A calculation of the rate of new infections per susceptible leads to a model

$$\begin{aligned} S' &= -\beta S(I + \varepsilon E), \\ E' &= \beta S(I + \varepsilon E) - \kappa E, \\ I' &= \kappa E - \alpha I. \end{aligned} \tag{9.15}$$

We take initial conditions

$$S(0) = S_0, \quad E(0) = E_0, \quad I(0) = I_0.$$

For this model,

$$\mathcal{R}_0 = \frac{\beta N}{\alpha} + \varepsilon \frac{\beta N}{\kappa}.$$

Integration of the sum of the equations of (9.14) from 0 to ∞ gives

$$N - S_\infty = \alpha \int_0^\infty I(s) ds.$$

Integration of the third equation of (9.15) gives

$$\kappa \int_0^\infty E(s) ds = \alpha \int_0^\infty I(s) ds - I_0,$$

and division of the first equation of (9.15) by S followed by integration from 0 to ∞ gives

$$\begin{aligned} \log \frac{S_0}{S_\infty} &= \int_0^\infty \beta [I(s) + \varepsilon E(s)] ds \\ &= \beta \int_0^\infty [I(s) + \varepsilon E(s)] ds \\ &= \beta \left[\varepsilon + \frac{\kappa}{\alpha} \right] \int_0^\infty E(s) ds - \frac{\varepsilon \beta I_0}{\kappa} \\ &= \mathcal{R}_0 \left[1 - \frac{S_\infty}{N} \right] - \frac{\varepsilon \beta I_0}{\kappa}. \end{aligned}$$

In this final size relation there is an initial term $\beta I_0/\alpha$, caused by the assumption that there are individuals infected originally who are beyond the exposed stage in which they would have had some infectivity. In order to obtain a final size relation without such an initial term it is necessary to assume $I(0) = 0$, that initial infectives

are in the first stage in which they can transmit infection. If $I(0) = 0$, the final size relation has the form (9.3).

9.5.2 Treatment Models

One form of treatment that is possible for some diseases is vaccination to protect against infection before the beginning of an epidemic. For example, this approach is commonly used for protection against annual influenza outbreaks. A simple way to model this would be to reduce the total population size by the fraction of the population protected against infection.

In reality, such inoculations are only partly effective, decreasing the rate of infection and also decreasing infectivity if a vaccinated person does become infected. This may be modeled by dividing the population into two groups with different model parameters, which would require some assumptions about the mixing between the two groups. This is not difficult, but we will not explore this direction here.

If there is a treatment for infection once a person has been infected, this may be modeled by supposing that a fraction γ per unit time of infectives is selected for treatment, and that treatment reduces infectivity by a fraction δ . Suppose that the rate of removal from the treated class is η . This leads to the Sitr model, where T is the treatment class, given by

$$\begin{aligned} S' &= -\beta S[I + \delta T], \\ I' &= \beta S[I + \delta T] - (\alpha + \gamma)I, \\ T' &= \gamma I - \eta T. \end{aligned} \tag{9.16}$$

A flow chart is shown in [Figure 9.8](#).

It is not difficult to prove, much as was done for the model (9.2), that

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

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

$$\mathcal{R}_0 = \frac{\beta N}{\alpha + \gamma} + \frac{\gamma}{\alpha + \gamma} \frac{\delta\beta N}{\eta}. \tag{9.17}$$

It is also possible to establish the final size relation (9.3) by means very similar to those used for the simple model (9.2). We integrate the first equation of (9.16) to obtain

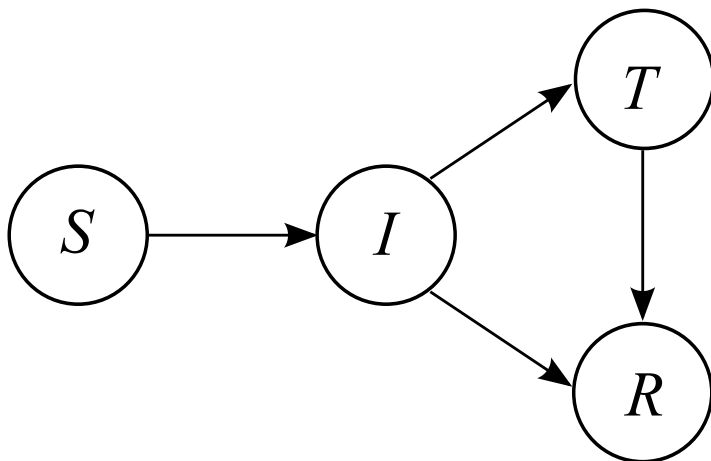


Fig. 9.8 Flow chart for the SITR model.

$$\log \frac{S_0}{S_\infty} = \int_0^\infty \beta[I(t) + \delta T(t)]dt = \beta \int_0^\infty [I(t) + \delta T(t)]dt.$$

Integration of the third equation of (9.16) gives

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

Integration of the sum of the first two equations of (9.16) gives

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

Combination of these three equations and (9.17) gives (9.3).

9.5.3 An Influenza Model

In some diseases, such as influenza, at the end of a stage individuals may proceed to one of two stages. There is a latent period after which a fraction p of latent individuals L proceeds to an infective stage I , while the remaining fraction $(1 - p)$ proceeds to an asymptomatic stage A , with infectivity reduced by a factor δ and a different period $1/\eta$. The influenza model of [Arino et al. (2006, 2007)] is

$$\begin{aligned} S' &= -\beta S[I + \delta A], \\ L' &= \beta S[I + \delta A] - \kappa L, \\ I' &= p\kappa L - \alpha I, \\ A' &= (1 - p)\kappa L - \eta A, \end{aligned} \tag{9.18}$$

and

$$\mathcal{R}_0 = \beta N \left[\frac{p}{\alpha} + \frac{\delta(1-p)}{\eta} \right].$$

A flow chart is shown in Figure 9.9.

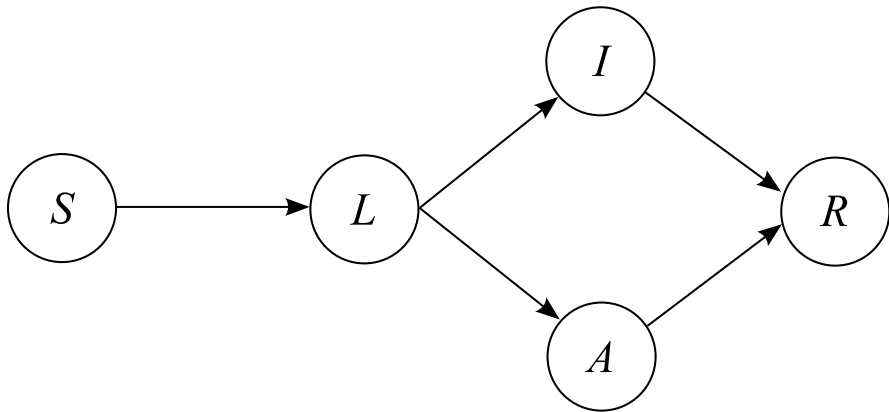


Fig. 9.9 Influenza model flowchart.

The same approach used in earlier examples leads to the same final size relation (9.3).

The model (9.18) is an example of a differential infectivity model. In such models, also used in the study of HIV/AIDS [Hyman, Li and Stanley (1999)], individuals enter a specific group when they become infected and stay in that group over the course of the infection. Different groups may have different parameter values. For example, for influenza infective and asymptomatic members may have different infectivities and different periods of stay in the respective stages.

9.5.4 A Quarantine-Isolation Model

For an outbreak of a new disease, where no vaccine is available, isolation of diagnosed infectives and quarantine of people who are suspected of having been infected (usually by tracing of contacts of diagnosed infectives) are the only control measures available. We formulate a model to describe the course of an epidemic, originally introduced for modeling the SARS epidemic of 2002-2003 [Gumel et al. (2004)], when control measures are begun under the following assumptions:

1. Exposed members may be infective with infectivity reduced by a factor ϵ_E , $0 \leq \epsilon_E < 1$.
2. Exposed members who are not isolated become infective at rate κ_E .

3. We introduce a class Q of quarantined members and a class J of isolated (hospitalized) members, and exposed members are quarantined at a proportional rate γ_Q in unit time (in practice, a quarantine will also be applied to many susceptibles, but we ignore this in the model). Quarantine is not perfect, but it reduces the contact rate by a factor ε_Q . The effect of this assumption is that some susceptibles make fewer contacts than the model assumes.
4. Infectives are diagnosed at a proportional rate γ_I per unit time and isolated. Isolation is imperfect, and there may be transmission of disease by isolated members, with an infectivity factor of ε_J .
5. Quarantined members are monitored, and when they develop symptoms at rate κ_Q they are isolated immediately.
6. Infectives leave the infective class at rate α_I and isolated members leave the isolated class at rate α_J .

These assumptions lead to the SEQIJR model [Gumel et al. (2004)]:

$$\begin{aligned}
 S' &= -\beta S[\varepsilon_E E + \varepsilon_E \varepsilon_Q Q + I + \varepsilon_J J], \\
 E' &= \beta S[\varepsilon_E E + \varepsilon_E \varepsilon_Q Q + I + \varepsilon_J J] - (\kappa_E + \gamma_Q)E, \\
 Q' &= \gamma_Q E - \kappa_Q Q, \\
 I' &= \kappa_E E - (\alpha_I + \gamma_I)I, \\
 J' &= \kappa_Q Q + \gamma_I I - \alpha_J J.
 \end{aligned} \tag{9.19}$$

The model before control measures are begun is the special case

$$\gamma_Q = \gamma_I = \kappa_Q = \alpha_J = 0, \quad Q = J = 0$$

of (9.19). It is the same as (9.15).

A flow chart is shown in [Figure 9.10](#).

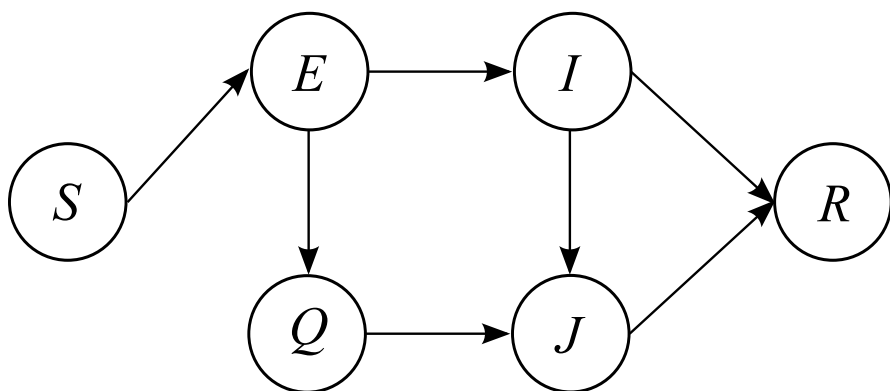


Fig. 9.10 Flow chart for the SEQIJR model.

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

$$\gamma_Q = \gamma_I = \kappa_Q = \alpha_J = 0.$$

We have already calculated \mathcal{R}_0 for (9.15), and we may calculate \mathcal{R}_c in the same way but using the full model with quarantined and isolated classes. We obtain

$$\mathcal{R}_c = \frac{\varepsilon_E \beta N}{D_1} + \frac{\beta N \kappa_E}{D_1 D_2} + \frac{\varepsilon_Q \varepsilon_E \beta N \gamma_Q}{D_1 \kappa_Q} + \frac{\varepsilon_I \beta N \kappa_E \gamma_I}{\alpha_J D_1 D_2} + \frac{\varepsilon_J \beta N \gamma_Q}{\alpha_J D_1},$$

where $D_1 = \gamma_Q + \kappa_E$, $D_2 = \gamma_I + \alpha_J$.

Each term of \mathcal{R}_c has an epidemiological interpretation. The mean duration in E is $1/D_1$ with contact rate $\varepsilon_E \beta$, giving a contribution to \mathcal{R}_c of $\varepsilon_E \beta N/D_1$. A fraction κ_E/D_1 goes from E to I , with contact rate β and mean duration $1/D_2$, giving a contribution of $\beta N \kappa_E/D_1 D_2$. A fraction γ_Q/D_1 goes from E to Q , with contact rate $\varepsilon_E \varepsilon_Q \beta$ and mean duration $1/\kappa_Q$, giving a contribution of $\varepsilon_E \varepsilon_Q \beta N \gamma_Q/D_1 \kappa_Q$. A fraction $\kappa_E \gamma_I/D_1 D_2$ goes from E to I to J , with a contact rate of $\varepsilon_J \beta$ and a mean duration of $1/\alpha_J$, giving a contribution of $\varepsilon_J \beta N \kappa_E \gamma_I/\alpha_J D_1 D_2$. Finally, a fraction γ_Q/D_1 goes from E to Q to J with a contact rate of $\varepsilon_J \beta$ and a mean duration of $1/\alpha_J$, giving a contribution of $\varepsilon_J \beta N \gamma_Q/D_1 \alpha_J$. The sum of these individual contributions gives \mathcal{R}_c .

In the model (9.19) the parameters γ_Q and γ_I are *control* parameters, which may be chosen in the attempt to manage the epidemic. The parameters ε_Q and ε_I depend on the strictness of the quarantine and isolation processes and are thus also control measures in a sense. The other parameters of the model are specific to the disease being studied. While they are not variable, their measurements are subject to experimental error.

The linearization of (9.19) at the disease-free equilibrium $(N, 0, 0, 0, 0)$ has matrix

$$\begin{bmatrix} \varepsilon_E \beta N - (\kappa_E + \gamma_Q) & \varepsilon_E \varepsilon_Q \beta & \beta N & \varepsilon_J \beta N \\ \gamma_Q & -\kappa_Q & 0 & 0 \\ \kappa_E & 0 & -(\alpha_I + \gamma_I) & 0 \\ 0 & \kappa_Q & \gamma_I & -\alpha_J \end{bmatrix}.$$

The corresponding characteristic equation is a fourth-degree polynomial equation whose leading coefficient is 1 and whose constant term is a positive constant multiple of $1 - \mathcal{R}_c$, thus positive if $\mathcal{R}_c < 1$ and negative if $\mathcal{R}_c > 1$. If $\mathcal{R}_c > 1$, there is a positive eigenvalue, corresponding to an initial exponential growth rate of solutions of (9.19). If $\mathcal{R}_c < 1$, it is possible to show that all eigenvalues of the coefficient matrix have negative real part, and thus solutions of (9.19) die out exponentially [van den Driessche and Watmough (2002)].

In order to show that analogues of the relation (9.3) and $S_\infty > 0$ derived for the model (9.2) are valid for the management model (9.19), we begin by integrating the equations for $S + E, Q, I, J$, of (9.19) with respect to t from $t = 0$ to $t = \infty$, using the initial conditions

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

We continue by integrating the equation for S , and then an argument similar to the one used for (9.2) but technically more complicated may be used to show that $S_\infty > 0$ for the treatment model (9.19) and also to establish the final size relation

$$\log \frac{S_0}{S_\infty} = \mathcal{R}_c \left[1 - \frac{S_\infty}{N} \right].$$

Thus the asymptotic behavior of the management model (9.19) is the same as that of the simpler model (9.2).

In the various compartmental models that we have studied, there are significant common features. This suggests that compartmental models can be put into a more general framework. In fact, this general framework is the age of infection epidemic model originally introduced by Kermack and McKendrick in 1927. However, we will not explore this generalization here.

Exercises

1. Compare the qualitative behaviors of the models

$$S' = -\beta SI, \quad I' = \beta SI - \alpha I,$$

and

$$S' = -\beta SI, \quad E' = \beta SI - \kappa E, \quad I' = \kappa E - \alpha I,$$

with

$$\beta = 1/3000, \quad \alpha = 1/6, \quad \kappa = 1/2, \quad S(0) = 999, \quad I(0) = 1.$$

These models represent an SIR epidemic model and an SEIR epidemic model respectively with a mean infective period of 6 days and a mean exposed period of 2 days. Do numerical simulations to decide whether the exposed period affects the behavior of the model noticeably.

2. Consider three basic epidemic models—the simple SIR model,

$$\begin{aligned} S' &= -\beta SI, \\ I' &= \beta SI - \alpha I, \end{aligned}$$

the SEIR model with some infectivity in the exposed period,

$$\begin{aligned}
S' &= -\beta S(I + \varepsilon E), \\
E' &= \beta S(I + \varepsilon E) - \kappa E, \\
I' &= \kappa E - \alpha I,
\end{aligned}$$

and the SIR model with treatment,

$$\begin{aligned}
S' &= -\beta S(I + \delta T), \\
I' &= \beta S(I + \delta T) - (\alpha + \varphi)I, \\
T' &= \varphi I - \eta T.
\end{aligned}$$

Use the parameter values

$$\beta = \frac{1}{3000}, \quad \alpha = \frac{1}{4}, \quad \varepsilon = \frac{1}{2}, \quad \kappa = \frac{1}{2}, \quad \delta = \frac{1}{2}, \quad \eta = \frac{1}{4}, \quad \varphi = 1,$$

and the initial values

$$S(0) = 995, \quad E(0) = 0, \quad I(0) = 5, \quad T(0) = 0.$$

For each model,

- (i) Calculate the reproduction number and the epidemic size.
- (ii) Do some numerical simulations to obtain the epidemic size by determining the change in S , the maximum number of infectives by measuring I , and the duration of the epidemic.

If you feel really ambitious, formulate and analyze an SEIR model with infectivity in the exposed period and treatment.

3. Consider an SIR model in which a fraction θ of infectives is isolated in a perfectly quarantined class Q with standard incidence (meaning that individuals make a contacts in unit time of which a fraction $I/(N - Q)$ are infective), given by the system

$$\begin{aligned}
S' &= -aS \frac{I}{N - Q}, \\
I' &= aS \frac{I}{N - Q} - (\theta + \alpha)I, \\
Q' &= \theta I - \gamma Q, \\
R' &= \alpha I + \gamma Q.
\end{aligned}$$

- (i) Find the equilibria.
- (ii) Find the basic reproduction number \mathcal{R}_0 .
- (iii) For influenza-like parameters, take $\alpha = 0.5$, $\theta = 1, 2, 4$, $\gamma = 0.4$, and $\mathcal{R}_0 = 2.5$, sketch the phase plane of the system and observe what is happening.

4. Isolation/quarantine is a complicated process because we don't live in a perfect world. In hospitals, patients may inadvertently or deliberately break from isolation and in the process have casual contacts with others including medical personnel and visitors. Taking this into account, we are led to the model

$$\begin{aligned} S' &= -aS \frac{[I + \rho \tau Q]}{N - \sigma Q}, \\ I' &= aS \frac{[I + \rho \tau Q]}{N - \sigma Q} - (\theta + \alpha)I, \\ Q' &= \theta I - \gamma Q, \\ R' &= \alpha I + \gamma Q. \end{aligned}$$

- (i) Determine all the parameters in the system and define each parameter.
 - (ii) Show that the population is constant.
 - (iii) Find all equilibria.
 - (iv) Find the reproductive number \mathcal{R}_0 .
 - (v) Describe the asymptotic behavior of the model, including its dependence on the basic reproduction number.
5. Formulate a model analogous to (9.16) for which treatment is not started immediately, but begins at time $\tau > 0$. Can you say anything about the dependence of the reproduction number on τ ?

9.6 An SIR Model with a General Infectious Period Distribution

In the simple model (9.2) studied in Section 9.2 we have assumed that the infective period is exponentially distributed. Now let us consider an SIR epidemic model in a population of constant size N with mass action incidence in which $P(\tau)$ is the fraction of individuals who are still infective a time τ after having become infected. The model is

$$\begin{aligned} S' &= -\beta S(t)I(t), \\ I(t) &= I_0(t) + \int_0^t [-S'(t - \tau)]P(\tau) d\tau. \end{aligned} \tag{9.20}$$

Here, $I_0(t)$ is the number of individuals who were infective initially at $t = 0$ who are still infective at time t . Then

$$I_0(t) \leq (N - S_0)P(t),$$

because if all initial infectives were newly infected we would have equality in this relation, and if some initial infectives had been infected before the starting time $t = 0$, they would recover earlier.

We assume that $P(\tau)$ is a nonnegative, nonincreasing function with $P(0) = 1$. We assume also that the mean infective period $\int_0^\infty P(\tau) d\tau$ is finite. Since a single infective causes βN new infections in unit time and $\int_0^\infty P(\tau) d\tau$ is the mean infective period, it is easy to calculate

$$\mathcal{R}_0 = \beta N \int_0^\infty P(s) ds.$$

Since S is a nonnegative decreasing function, it follows as for (9.2) that $S(t)$ decreases to a limit S_∞ as $t \rightarrow \infty$, but we must proceed differently to show that $I(t) \rightarrow 0$. This will follow if we can prove that $\int_0^t I(s) ds$ is bounded as $t \rightarrow \infty$. We have

$$\begin{aligned} \int_0^t I(s) ds &= \int_0^t I_0(\tau) ds + \int_0^t \int_0^s [-S'(s-\tau)] P(\tau) d\tau ds \\ &\leq (N - S_0) \int_0^t P(\tau) d\tau + \int_0^t \int_\tau^t [-S'(s-\tau)] ds P(\tau) d\tau \\ &\leq (N - S_0) \int_0^t P(\tau) ds + \int_0^t [S_0 - S(t-\tau)] P(\tau) d\tau \\ &\leq N \int_0^t P(\tau) d\tau. \end{aligned}$$

Since $\int_0^\infty P(\tau) d\tau$ is assumed to be finite, it follows that $\int_0^t I(s) ds$ is bounded, and thence that $I(t) \rightarrow 0$.

Now integration of the first equation in (9.20) from 0 to ∞ gives

$$\log \frac{S_0}{S_\infty} = \beta \int_0^\infty I(\tau) d\tau < \infty,$$

and this shows that $S_\infty > 0$.

If all initially infected individuals are newly infected, so that $I_0(t) = (N - S_0)P(t)$, integration of the second equation of (9.20) gives

$$\begin{aligned} \int_0^\infty I(s) ds &= \int_0^\infty I_0(s) ds + \int_0^\infty \int_0^s [-S'(s-\tau)] P(\tau) d\tau ds \\ &= (N - S_0) \int_0^\infty P(\tau) d\tau + \int_0^\infty \int_\tau^\infty [-S'(t-\tau)] ds P(\tau) d\tau \\ &= (N - S_0) \int_0^\infty P(\tau) d\tau + \int_0^\infty [S_0 - S_\infty] P(\tau) d\tau \\ &= (N - S_\infty) \int_0^\infty P(\tau) d\tau \\ &= \mathcal{R}_0 \left[1 - \frac{S_\infty}{N} \right], \end{aligned}$$

and this is the final size relation, identical to (9.3). If there are individuals who were infected before time $t = 0$, a positive term

$$(N - S_0) \int_0^\infty P(t) dt - \int_0^\infty I_0(t) dt$$

must be subtracted from the right side of this equation.

The generalization to arbitrary infectious periods in this section is a component of the age of infection epidemic model of Kermack and McKendrick (1927), which also incorporates general compartmental structures. The examples of this section and the previous section are all special cases of the age of infection model.

Exercises

1. Formulate a description of the model (9.20) with an infective period of fixed length σ and calculate its basic reproduction number.

9.7 The Age of Infection Epidemic Model

The general epidemic model described by Kermack and McKendrick (1927) included a dependence of infectivity on the time since becoming infected (age of infection). We let $S(t)$ denote the number of susceptibles at time t and let $\varphi(t)$ be the total infectivity at time t , defined as the sum of products of the number of infected members with each infection age and the mean infectivity for that infection age. We assume that on average, members of the population make a constant number a of contacts in unit time. We let $B(\tau)$ be the fraction of infected members remaining infected at infection age τ and let $\pi(\tau)$ with $0 \leq \pi(\tau) \leq 1$ be the mean infectivity at infection age τ . Then we let

$$A(\tau) = \pi(\tau)B(\tau),$$

the mean infectivity of members of the population with infection age τ . We assume that there are no disease deaths, so that the total population size is a constant N .

The age of infection epidemic model is

$$\begin{aligned} S' &= -\beta S \varphi, \\ \varphi(t) &= \varphi_0(t) + \int_0^t \beta S(t-\tau) \varphi(t-\tau) A(\tau) d\tau \\ &= \varphi_0(t) + \int_0^t [-S'(t-\tau)] A(\tau) d\tau. \end{aligned} \tag{9.21}$$

The basic reproduction number is

$$\mathcal{R}_0 = \beta N \int_0^\infty A(\tau) d\tau.$$

We write

$$-\frac{S'(t)}{S(t)} = \beta \varphi_0(t) + \beta \int_0^t [-S'(t-\tau)]A(\tau)d\tau.$$

Integration with respect to t from 0 to ∞ gives

$$\begin{aligned} \log \frac{S_0}{S_\infty} &= \beta \int_0^\infty \varphi_0(t)dt + \beta \int_0^\infty \int_0^t [-S'(t-\tau)]A(\tau)d\tau dt \\ &= \beta \int_0^\infty \varphi_0(t)dt + \beta \int_0^\infty A(\tau) \int_\tau^\infty [-S'(t-\tau)]dt d\tau \\ &= \beta \int_0^\infty \varphi_0(t)dt + [S_0 - S_\infty] \int_0^\infty A(\tau)d\tau \\ &= \beta [N - S_\infty] \int_0^\infty A(\tau)d\tau + \beta \int_0^\infty [\varphi_0(t) - (N - S_0)A(t)]dt \\ &= \mathcal{R}_0 \left[1 - \frac{S_\infty}{N} \right] - \beta \int_0^\infty [(N - S_0)A(t) - \varphi_0(t)]dt. \end{aligned} \quad (9.22)$$

Here, $\varphi_0(t)$ is the total infectivity of the initial infectives when they reach age of infection t . If all initial infectives have infection age zero at $t = 0$, then $\varphi_0(t) = [N - S_0]A(t)$, and

$$\int_0^\infty [\varphi_0(t) - (N - S_0)A(t)]dt = 0.$$

Then (9.22) takes the form

$$\log \frac{S_0}{S_\infty} = \mathcal{R}_0 \left(1 - \frac{S_\infty}{N} \right), \quad (9.23)$$

and this is the general final size relation. If there are initial infectives with infection age greater than zero, let $u(\tau)$ be the fraction of these individuals with infection age τ , $\int_0^\infty u(\tau)d\tau = 1$. At time t these individuals have infection age $t + \tau$ and mean infectivity $A(t + \tau)$. Thus

$$\varphi_0(t) = (N - S_\infty) \int_0^\infty u(\tau)A(t + \tau)d\tau,$$

and

$$\begin{aligned} \int_0^\infty \varphi_0(t)dt &= (N - S_\infty) \int_0^\infty \int_0^\infty u(\tau)A(t + \tau)d\tau dt \\ &= (N - S_\infty) \int_0^\infty u(\tau) \left[\int_\tau^\infty A(v)dv \right] d\tau \\ &= (N - S_\infty) \int_0^\infty A(v) \left[\int_0^v u(\tau)d\tau \right] dv \\ &\leq (N - S_\infty) \int_0^\infty A(v)dv, \end{aligned}$$

since $\int_0^v u(\tau)d\tau \leq 1$.

Thus, the initial term satisfies

$$\int_0^\infty [(N - S_0)A(t) - \varphi_0(t)]dt \geq 0.$$

The final size relation is sometimes presented in the form

$$\log \frac{S_0}{S_\infty} = \mathcal{R}_0 \left(1 - \frac{S_\infty}{S_0} \right); \quad (9.24)$$

see, for example [Arino et al. (2007), Heffernan, Smith?, and Wahl (2005)]. Without the initial term this form would represent the final size relation for an epidemic started by someone outside the population under study, so that $S_0 = N, I_0 = 0$.

Example 1. The SEIR model (9.15) can be viewed as an age of infection model with $\varphi = \varepsilon E + I$. To use the age of infection interpretation, we need to determine the kernel $A(\tau)$ in order to calculate its integral. We let $u(\tau)$ be the fraction of infected members with infection age τ who are not yet infective and $v(\tau)$ the fraction of infected members who are infective. Then the rate at which members become infective at infection age τ is $\kappa u(\tau)$, and we have

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

The solution of this system is

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

Thus we have

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

and it is easy to calculate

$$\int_0^\infty A(\tau) d\tau = \frac{1}{\alpha} + \frac{\varepsilon}{\kappa}.$$

This gives the same value for \mathcal{R}_0 as was calculated directly.

The age of infection model also includes the possibility of disease stages with distributions that are not exponential [Feng (2007), Feng, Xu, and Zhao (2007)].

Example 2. Consider an SEIR model in which the exposed stage has an exponential distribution but the infective stage has a period distribution given by a function P ,

$$\begin{aligned} S' &= -\beta SI, \\ E' &= \beta SI - \kappa E, \\ I(t) &= \int_0^t \kappa E(s) P(t-s) ds, \end{aligned} \quad (9.25)$$

with initial conditions

$$S(0) = S_0, \quad E(0) = E_0, \quad I(0) = 0.$$

If we define $u(\tau), v(\tau)$ as in Example 1, we again obtain $u(\tau) = e^{-\kappa\tau}$, and v satisfies

$$A(\tau) = v(\tau) = \int_0^\tau \kappa u(s)P(\tau-s)ds = \int_0^\tau \kappa e^{-\kappa s}P(\tau-s)ds.$$

Now,

$$\begin{aligned} \int_0^\infty A(\tau)d\tau &= \int_0^\infty \int_0^\tau \kappa e^{-\kappa s}P(\tau-s)dsd\tau \\ &= \int_0^\infty \left[\int_s^\infty P(\tau-s)d\tau \right] \kappa e^{-\kappa s}ds \\ &= \int_0^\infty \left[\int_0^\infty P(u)du \right] \kappa e^{-\kappa s}ds \\ &= \int_0^\infty P(u)du. \end{aligned}$$

For period distributions that are not exponential, it is possible to calculate

$$\int_0^\infty A(\tau)d\tau$$

without having to calculate the function $A(\tau)$ explicitly.

Example 3. Consider an SEIR model in which the exposed period has a distribution given by a function Q and the infective period has a distribution given by a function P . Then

$$\begin{aligned} S' &= -\beta SI, \\ E(t) &= E_0Q(t) + \int_0^t [-S'(s)]Q(t-s)ds. \end{aligned}$$

In order to obtain an equation for I , we differentiate the equation for E , obtaining

$$E'(t) = E_0Q'(t) - S'(t) + E_0 \int_0^t [-S'(s)]Q'(t-s)ds.$$

Thus the input to I at time t is

$$E_0Q'(t) + E_0 \int_0^t [-S'(s)]Q'(t-s)ds,$$

and

$$I(t) = E_0 \int_0^t Q'(u)P(t-u)du + E_0 \int_0^t [-S'(s)]Q'(u-s)dsP(t-u)du.$$

The first term in this expression may be written as $I_0(t)$, and the second term may be simplified, using interchange of the order of integration in the iterated integral, to yield

$$\int_0^t \int_0^u [-S'(s)] Q'(u-s) ds P(t-u) du = \int_0^t \int_s^t Q'(u-s) du P(t-u) [-S'(s)] ds.$$

If we define

$$A(t-s) = \int_s^t Q'(u-s) P(t-u) du = \int_0^{t-s} Q'(t-s-v) P(v) dv,$$

we obtain

$$I(t) = I_0(t) + \int_0^t [-S'(s)] A(t-s) ds.$$

Then the model is

$$\begin{aligned} S' &= -\beta SI, \\ E(t) &= E_0 Q(t) + \int_0^t [-S'(s)] Q(t-s) ds, \\ I(t) &= I_0(t) + \int_0^t [-S'(s)] A(t-s) ds, \end{aligned} \tag{9.26}$$

which is in age of infection form with $\phi = I$, and we have an explicit expression for $A(\tau)$.

Exercises

1. Interpret the models (9.16), (9.18), and (9.19) introduced earlier as age of infection models and use this interpretation to calculate their reproduction numbers.
2. Calculate the basic reproduction number for the model (9.26) but with infectivity in the exposed class having a reduction factor ε .

9.8 Models with Disease Deaths

The assumption in the model (9.2) of a rate of contacts per infective that is proportional to population size N , called *mass action incidence* or bilinear incidence, was used in all the early epidemic models. However, it is quite unrealistic, except possibly in the early stages of an epidemic in a population of moderate size. It is more realistic to assume a contact rate that is a nonincreasing function of total population size. For example, a situation in which the number of contacts per infective in unit time is constant, called *standard incidence*, is a more accurate description for sexu-

ally transmitted diseases. If there are no disease deaths, so that the total population size remains constant, such a distinction is unnecessary.

We generalize the model (9.2) by dropping assumption (iv) and replacing assumption (i) by the assumption that an average member of the population makes $C(N)$ contacts in unit time with $C'(N) \geq 0$ [Castillo-Chavez, Cooke, Huang, and Levin (1989a), Dietz (1982)], and we define

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

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

$$\beta(N) + N\beta'(N) \geq 0. \quad (9.27)$$

Some epidemic models [Dietz (1982)] have used a Michaelis–Menten type of interaction of the form

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

Another form based on a mechanistic derivation for pair formation [Heesterbeek and Metz (1993)] leads to an expression of the form

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

Data for diseases transmitted by contact in cities of moderate size [Mena-Lorca and Hethcote (1992)] suggests that data fit the assumption of the form

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

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

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

$$\begin{aligned} S' &= -\beta(N)SI, \\ I' &= \beta(N)SI - \alpha I, \\ N' &= -(1 - f)\alpha I. \end{aligned} \quad (9.28)$$

Since N is now a decreasing function, we define $N(0) = N_0 = S_0 + I_0$. We also have the equation $R' = -f\alpha I$, but we need not include it in the model, since R is determined when S , I , and N are known. We should note that if $f = 1$, the total population size remains equal to the constant N , and the model (9.28) reduces to the simpler model (9.2) with β replaced by the constant $\beta(N_0)$.

We wish to show that the model (9.28) has the same qualitative behavior as the model (9.2), namely that there is a basic reproduction number that distinguishes between disappearance of the disease and an epidemic outbreak, and that some members of the population are left untouched when the epidemic passes. These two properties are the central features of all epidemic models.

For the model (9.28) the basic reproduction number is given by

$$\mathcal{R}_0 = \frac{N_0\beta(N_0)}{\alpha}$$

because a single infective introduced into a wholly susceptible population makes $C(N_0) = N_0\beta(N_0)$ contacts in unit time, all of which are with susceptibles and thus produce new infections, and the mean infective period is $1/\alpha$.

We assume that $\beta(0)$ is finite, thus ruling out standard incidence (standard incidence does not appear to be realistic if the total population N approaches zero, and it would be more natural to assume that $C(N)$ grows linearly with N for small N). If we let $t \rightarrow \infty$ in the sum of the first two equations of (9.28), we obtain

$$\alpha \int_0^\infty I(s)ds = S_0 + I_0 - S_\infty = N - S_\infty.$$

The first equation of (9.28) may be written as

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

Since

$$\beta(N) \geq \beta(N_0),$$

integration from 0 to ∞ gives

$$\log \frac{S_0}{S_\infty} = \int_0^\infty \beta(N(t))I(t)dt \geq \beta(N_0) \int_0^\infty I(t)dt = \frac{\beta(N_0)(N_0 - S_\infty)}{\alpha N_0}.$$

We now obtain a final size inequality

$$\log \frac{S_0}{S_\infty} = \int_0^\infty \beta(N(t))I(t)dt \geq \beta(N_0) \int_0^\infty I(t)dt = \mathcal{R}_0 \left[1 - \frac{S_\infty}{N_0} \right].$$

If the disease death rate is small, the final size inequality is an approximate equality.

It is not difficult to show that $N(t) \geq fN_0$, and then a similar calculation using the inequality $\beta(N) \leq \beta(fN_0) < \infty$ shows that

$$\log \frac{S_0}{S_\infty} \leq \beta(fN_0) \int_0^\infty I(t)dt,$$

from which we may deduce that $S_\infty > 0$.

Exercises

1. For the model (9.28) show that the final total population size is given by

$$N_\infty = fN_0 + (1 - f)S_\infty.$$

9.9 A Vaccination Model

To cope with annual seasonal influenza epidemics there is a program of vaccination before the “flu” season begins. Each year, a vaccine is produced aimed at protecting against the three influenza strains considered most dangerous for the coming season. We formulate a model to add vaccination to the simple SIR model (9.2) under the assumption that vaccination reduces susceptibility (the probability of infection if a contact with an infected member of the population is made).

We consider a population of total size N and assume that a fraction γ of this population is vaccinated prior to a disease outbreak. Thus we have a subpopulation of size $N_U = (1 - \gamma)N$ of unvaccinated members and a subpopulation of size $N_V = \gamma N$ of vaccinated members. We assume that vaccinated members have susceptibility to infection reduced by a factor σ , $0 \leq \sigma \leq 1$, with $\sigma = 0$ describing a perfectly effective vaccine and $\sigma = 1$ describing a vaccine that has no effect. We assume also that vaccinated individuals who are infected have infectivity reduced by a factor δ and may also have a recovery rate α_V that is different from the recovery rate of infected unvaccinated individuals α_U .

We let S_U, S_V, I_U, I_V denote the number of unvaccinated susceptibles, the number of vaccinated susceptibles, the number of unvaccinated infectives, and the number of vaccinated infectives respectively.

The resulting model is

$$\begin{aligned} S'_U &= -\beta S_U(I_U + \delta I_V), \\ S'_V &= -\sigma\beta S_V(I_U + \delta I_V), \\ I'_U &= \beta S_U(I_U + \delta I_V) - \alpha_U I_U, \\ I'_V &= \sigma\beta S_V(I_U + \delta I_V) - \alpha_V I_V. \end{aligned} \tag{9.29}$$

The initial conditions prescribe $S_U(0), S_V(0), I_U(0), I_V(0)$, with

$$S_U(0) + I_U(0) = N_U, \quad S_V(0) + I_V(0) = N_V.$$

Since the infection now is beginning in a population that is not fully susceptible, we speak of the control reproduction number \mathcal{R}_c rather than the basic reproduction number. However, as we will soon see, calculation of the control reproduction number will require a more general definition and a considerable amount of technical computation. The computation method is applicable to both basic and control reproduction numbers. We will use the term *reproduction number* to denote either a basic reproduction number or a control reproduction number. We are able to obtain final size relations without knowledge of the reproduction number, but these final size relations do contain information about the reproduction number, and more.

Since S_U and S_V are decreasing nonnegative functions they have limits $S_U(\infty)$ and $S_V(\infty)$ respectively as $t \rightarrow \infty$. The sum of the equations for S_U and I_U in (9.29) is

$$(S_U + I_U)' = -\alpha_U I_U,$$

from which we conclude, just as in the analysis of (9.2), that $I_U(t) \rightarrow 0$ as $t \rightarrow \infty$, and that

$$\alpha \int_0^\infty I_U(t) dt = N_U - S_U(\infty). \quad (9.30)$$

Similarly, using the sum of the equations for S_V and I_V , we see that $I_V(t) \rightarrow 0$ as $t \rightarrow \infty$, and that

$$\alpha \int_0^\infty I_V(t) dt = N_V - S_V(\infty). \quad (9.31)$$

Integration of the equation for S_U in (9.29) and use of (??) gives

$$\begin{aligned} \log \frac{S_U(0)}{S_U(\infty)} &= \beta \left[\int_0^\infty I_U(t) dt + \delta \int_0^\infty I_V(t) dt \right] \\ &= \frac{\beta N_U}{\alpha} \left[1 - \frac{S_U(\infty)}{N_U} \right] + \frac{\delta \beta N_V}{\alpha} \left[1 - \frac{S_V(\infty)}{N_V} \right]. \end{aligned} \quad (9.32)$$

A similar calculation using the equation for S_V gives

$$\log \frac{S_V(0)}{S_V(\infty)} = \frac{\sigma \beta N_U}{\alpha} \left[1 - \frac{S_U(\infty)}{N_U} \right] + \frac{\delta \sigma \beta N_V}{\alpha} \left[1 - \frac{S_V(\infty)}{N_V} \right]. \quad (9.33)$$

This pair of equations (9.32), (9.33) are the final size relations. They make it possible to calculate $S_U(\infty), S_V(\infty)$ if the parameters of the model are known.

It is convenient to define the matrix

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} = \begin{bmatrix} \frac{\beta N_U}{\alpha_U} & \frac{\delta \beta N_V}{\alpha_V} \\ \frac{\sigma \beta N_U}{\alpha_U} & \frac{\delta \sigma \beta N_V}{\alpha_V} \end{bmatrix}.$$

Then the final size relations (9.32), (9.33) may be written

$$\begin{aligned}\log \frac{S_U(0)}{S_U(\infty)} &= K_{11} \left[1 - \frac{S_U(\infty)}{N_U} \right] + K_{12} \left[1 - \frac{S_V(\infty)}{N_V} \right], \\ \log \frac{S_V(0)}{S_V(\infty)} &= K_{21} \left[1 - \frac{S_U(\infty)}{N_U} \right] + K_{22} \left[1 - \frac{S_V(\infty)}{N_V} \right].\end{aligned}\tag{9.34}$$

The matrix K is closely related to the reproduction number. In the next section we describe a general method for calculating reproduction numbers that will involve this matrix.

Exercises

1. Suppose we want to model the spread of influenza in a city using an SLIAR model (susceptible-latent-infectious-asymptomatic-recovered, respectively). Then our system of equation would be

$$\begin{aligned}S' &= -\beta(I + \delta A)S, \\ L' &= \beta(I + \delta A)S - \kappa L, \\ I' &= p\kappa L - \gamma I, \\ A' &= (1 - p)\kappa L - \eta A, \\ R' &= \eta A + \gamma I,\end{aligned}$$

where β is the transmission coefficient, δ is the reduced transmissibility factor from asymptomatic contacts, κ is the rate of disease progression from the latent class, p is the proportion of individuals that are clinically diagnosed, η is the recovery rate from the asymptomatic class, γ is the recovery rate from the infectious (clinically diagnosed) class, and N is the total population size.

- (i) Add a vaccination class to the model. Assume that the vaccine imparts partial protection until it becomes fully effective. Is the population of the new system constant? Are there any endemic equilibria?
- (ii) Vary the vaccination rate from 0.2 to 0.8 and determine how the number of infected individuals changes compared with the model without vaccination. Does vaccination prevent the outbreak?

9.10 The Next Generation Matrix

Up to this point, we have calculated reproduction numbers by following the secondary cases caused by a single infective introduced into a population. However, if there are subpopulations with different susceptibilities to infection, as in the vaccination model introduced in Section 9.9, it is necessary to follow the secondary infections in the subpopulations separately, and this approach will not yield the re-

production number. It is necessary to give a more general approach to the meaning of the reproduction number, and this is done through the *next generation matrix* [Diekmann and Heesterbeek (2000), Diekmann, Heesterbeek, and Metz (1990), van den Driessche and Watmough (2002)]. The underlying idea is that we must calculate the matrix whose (i, j) entry is the number of secondary infections caused in compartment i by an infected individual in compartment j . The procedure applies both to epidemic models, as studied in this chapter, and to models with demographics for endemic diseases, to be studied in the next chapter.

In a compartmental disease transmission model we sort individuals into compartments based on a single, discrete state variable. A compartment is called a *disease compartment* if the individuals therein are infected. Note that this use of the term disease is broader than the clinical definition and includes stages of infection such as exposed stages in which infected individuals are not necessarily infective. Suppose there are n disease compartments and m nondisease compartments, and let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ be the subpopulations in each of these compartments. Further, we denote by \mathcal{F}_i the rate at which secondary infections increase the i -th disease compartment and by \mathcal{V}_i the rate at which disease progression, death, and recovery decrease the i -th compartment. The compartmental model can then be written in the form

$$\begin{aligned} x'_i &= \mathcal{F}_i(x, y) - \mathcal{V}_i(x, y), \quad i = 1, \dots, n, \\ y'_j &= g_j(x, y), \quad j = 1, \dots, m. \end{aligned} \tag{9.35}$$

Note that the decomposition of the dynamics into \mathcal{F} and \mathcal{V} and the designation of compartments as infected or uninfected may not be unique; different decompositions correspond to different epidemiological interpretations of the model. The definitions of \mathcal{F} and \mathcal{V} used here differ slightly from those in [van den Driessche and Watmough (2002)].

The derivation of the basic reproduction number is based on the linearization of the ODE model about a disease-free equilibrium. For an epidemic model with a line of equilibria, it is customary to use the equilibrium with all members of the population susceptible. We assume:

- $\mathcal{F}_i(0, y) = 0$ and $\mathcal{V}_i(0, y) = 0$ for all $y \geq 0$ and $i = 1, \dots, n$.
- The disease-free system $y' = g(0, y)$ has a unique equilibrium that is asymptotically stable, that is, all solutions with initial conditions of the form $(0, y)$ approach a point $(0, y_o)$ as $t \rightarrow \infty$. We refer to this point as the disease-free equilibrium.

The first assumption says that all new infections are secondary infections arising from infected hosts; there is no immigration of individuals into the disease compartments. It ensures that the disease-free set, which consists of all points of the form $(0, y)$, is invariant. That is, any solution with no infected individuals at some point in time will be free of infection for all time. The second assumption ensures that the disease-free equilibrium is also an equilibrium of the full system. The uniqueness of the disease-free equilibrium in the second assumption is required for models

with demographics, to be studied in the next chapter. Although it is not satisfied in epidemic models, the specification of a specific disease-free equilibrium with all members of the population susceptible is sufficient to validate the results.

Next, we assume:

- $\mathcal{F}_i(x, y) \geq 0$ for all nonnegative x and y and $i = 1, \dots, n$.
- $\mathcal{V}_i(x, y) \leq 0$ whenever $x_i = 0$, $i = 1, \dots, n$.
- $\sum_{i=1}^n \mathcal{V}_i(x, y) \geq 0$ for all nonnegative x and y .

The reasons for these assumptions are that the function \mathcal{F} represents new infections and cannot be negative, each component \mathcal{V}_i represents a net outflow from compartment i and must be negative (inflow only) whenever the compartment is empty, and the sum $\sum_{i=1}^n \mathcal{V}_i(x, y)$ represents the total outflow from all infected compartments. Terms in the model leading to increases in $\sum_{i=1}^n x_i$ are assumed to represent secondary infections and therefore belong in \mathcal{F} .

Suppose that a single infected person is introduced into a population originally free of disease. The initial ability of the disease to spread through the population is determined by an examination of the linearization of (9.35) about the disease-free equilibrium $(0, y_0)$. It is easy to see that the assumption $\mathcal{F}_i(0, y) = 0$, $\mathcal{V}_i(0, y) = 0$ implies

$$\frac{\partial \mathcal{F}_i}{\partial y_j}(0, y_0) = \frac{\partial \mathcal{V}_i}{\partial y_j}(0, y_0) = 0$$

for every pair (i, j) . This implies that the linearized equations for the disease compartments x are decoupled from the remaining equations and can be written as

$$x' = (F - V)x, \quad (9.36)$$

where F and V are the $n \times n$ matrices with entries

$$F = \frac{\partial \mathcal{F}_i}{\partial x_j}(0, y_0) \quad \text{and} \quad V = \frac{\partial \mathcal{V}_i}{\partial x_j}(0, y_0).$$

Because of the assumption that the disease-free system $y' = g(0, y)$ has a unique asymptotically stable equilibrium, the linear stability of the system (9.35) is completely determined by the linear stability of the matrix $(F - V)$ in (9.36).

The number of secondary infections produced by a single infected individual can be expressed as the product of the expected duration of the infectious period and the rate at which secondary infections occur. For the general model with n disease compartments, these are computed for each compartment for a hypothetical index case. The expected time the index case spends in each compartment is given by the integral $\int_0^\infty \phi(t, x_0) dt$, where $\phi(t, x_0)$ is the solution of (9.36) with $F = 0$ (no secondary infections) and nonnegative initial conditions x_0 representing an infected index case:

$$x' = -Vx, \quad x(0) = x_0. \quad (9.37)$$

In effect, this solution shows the path of the index case through the disease compartments from the initial exposure through to death or recovery with the $i - th$

component of $\varphi(t, x_0)$ interpreted as the probability that the index case (introduced at time $t = 0$) is in disease state i at time t . The solution of (9.37) is $\phi(t, x_0) = e^{-Vt}x_0$, where the exponential of a matrix is defined by the Taylor series

$$e^A = I + A + \frac{A^2}{2} + \frac{A^3}{3!} + \cdots + \frac{A^k}{k!} + \cdots.$$

This series converges for all t (see, for example, [Hirsch and Smale (1974)]). Thus $\int_0^\infty \varphi(t, x_0) dt = V^{-1}x_0$, and the (i, j) entry of the matrix V^{-1} can be interpreted as the expected time an individual initially introduced into disease compartment j spends in disease compartment i .

The (i, j) entry of the matrix F is the rate at which secondary infections are produced in compartment i by an index case in compartment j . Hence, the expected number of secondary infections produced by the index case is given by

$$\int_0^\infty F e^{-Vt} x_0 dt = F V^{-1} x_0.$$

Following Diekmann and Heesterbeek (2000), the matrix $K = FV^{-1}$ is referred to as the *next generation matrix* for the system at the disease-free equilibrium. The (i, j) entry of K is the expected number of secondary infections in compartment i produced by individuals initially in compartment j , assuming, of course, that the environment experienced by the individual remains homogeneous for the duration of its infection.

Shortly, we will describe some results from matrix theory that imply that the matrix $K_L = FV^{-1}$, called the *next generation matrix with small domain*, is nonnegative and therefore has a nonnegative eigenvalue, $\mathcal{R}_0 = \rho(FV^{-1})$, such that there are no other eigenvalues of K with modulus greater than \mathcal{R}_0 and there is a nonnegative eigenvector ω associated with \mathcal{R}_0 [Berman and Plemmons (1970), Theorem 1.3.2]. This eigenvector is in a sense the distribution of infected individuals that produces the greatest number \mathcal{R}_0 of secondary infections per generation. Thus, \mathcal{R}_0 and the associated eigenvector ω suitably define a “typical” infective, and the basic reproduction number can be rigorously defined as the spectral radius of the matrix K_L . The spectral radius of a matrix K_L , denoted by $\rho(K_L)$, is the maximum of the moduli of the eigenvalues of K_L . If K_L is irreducible, then \mathcal{R}_0 is a simple eigenvalue of K_L and is strictly larger in modulus than all other eigenvalues of K_L . However, if K_L is reducible, which is often the case for diseases with multiple strains, then K_L may have several positive real eigenvalues corresponding to reproduction numbers for each competing strain of the disease.

We have interpreted the reproduction number for a disease as the number of secondary infections produced by an infected individual in a population of susceptible individuals. If the reproduction number $\mathcal{R}_0 = \rho(FV^{-1})$ is consistent with the differential equation model, then it should follow that the disease-free equilibrium is asymptotically stable if $\mathcal{R}_0 < 1$ and unstable if $\mathcal{R}_0 > 1$.

This is shown through a sequence of lemmas.

The spectral bound (or abscissa) of a matrix A is the maximum real part of all eigenvalues of A . If each entry of a matrix T is nonnegative, we write $T \geq 0$ and refer to T as a nonnegative matrix. A matrix of the form $A = sI - B$, with $B \geq 0$, is said to have the Z sign pattern. These are matrices whose off-diagonal entries are negative or zero. If in addition, $s \geq \rho(B)$, then A is called an M-matrix. Note that in this section, I denotes an identity matrix, not a population of infectious individuals. The following lemma is a standard result from [Berman and Plemmons (1970)].

Lemma 9.1. *If A has the Z sign pattern, then $A^{-1} \geq 0$ if and only if A is a nonsingular M-matrix.*

The assumptions we have made imply that each entry of F is nonnegative and that the off-diagonal entries of V are negative or zero. Thus V has the Z sign pattern. Also, the column sums of V are positive or zero, which, together with the Z sign pattern, implies that V is a (possibly singular) M-matrix [Berman and Plemmons (1970), condition M_{35} of Theorem 6.2.3]. In what follows, it is assumed that V is nonsingular. In this case, $V^{-1} \geq 0$, by Lemma 9.1. Hence, $K_L = FV^{-1}$ is also nonnegative.

Lemma 9.2. *If F is nonnegative and V is a nonsingular M-matrix, then $\mathcal{R}_0 = \rho(FV^{-1}) < 1$ if and only if all eigenvalues of $(F - V)$ have negative real parts.*

Proof. Suppose $F \geq 0$ and V is a nonsingular M-matrix. By Lemma 9.1, $V^{-1} \geq 0$. Thus, $(I - FV^{-1})$ has the Z sign pattern, and by Lemma 1, $(I - FV^{-1})^{-1} \geq 0$ if and only if $\rho(FV^{-1}) < 1$. From the equalities $(V - F)^{-1} = V^{-1}(I - FV^{-1})^{-1}$ and $V(V - F)^{-1} = I + F(V - F)^{-1}$, it follows that $(V - F)^{-1} \geq 0$ if and only if $(I - FV^{-1})^{-1} \geq 0$. Finally, $(V - F)$ has the Z sign pattern, so by Lemma 9.1, $(V - F)^{-1} \geq 0$ if and only if $(V - F)$ is a nonsingular M-matrix. Since the eigenvalues of a nonsingular M-matrix all have positive real parts, this completes the proof. \square

Theorem 9.1. *Consider the disease transmission model given by (9.35). The disease-free equilibrium of (9.35) is locally asymptotically stable if $\mathcal{R}_0 < 1$, but unstable if $\mathcal{R}_0 > 1$.*

Proof. Let F and V be as defined as above, and let J_{21} and J_{22} be the matrices of partial derivatives of g with respect to x and y evaluated at the disease-free equilibrium. The Jacobian matrix for the linearization of the system about the disease-free equilibrium has the block structure

$$J = \begin{bmatrix} F - V & 0 \\ J_{21} & J_{22} \end{bmatrix}.$$

The disease-free equilibrium is locally asymptotically stable if the eigenvalues of the Jacobian matrix all have negative real parts. Since the eigenvalues of J are those of $(F - V)$ and J_{22} , and the latter all have negative real parts by assumption, the disease-free equilibrium is locally asymptotically stable if all eigenvalues of $(F - V)$ have negative real parts. By the assumptions on \mathcal{F} and \mathcal{V} , F is nonnegative and V is a nonsingular M-matrix. Hence, by Lemma 2 all eigenvalues of $(F - V)$ have negative

real parts if and only if $\rho(FV^{-1}) < 1$. It follows that the disease-free equilibrium is locally asymptotically stable if $\mathcal{R}_0 = \rho(FV^{-1}) < 1$.

Instability for $\mathcal{R}_0 > 1$ can be established by a continuity argument. If $\mathcal{R}_0 \leq 1$, then for any $\varepsilon > 0$, $((1 + \varepsilon)I - FV^{-1})$ is a nonsingular M-matrix and by Lemma 9.1, $((1 + \varepsilon)I - FV^{-1})^{-1} \geq 0$. By Lemma 9.2, all eigenvalues of $((1 + \varepsilon)V - F)$ have positive real parts. Since $\varepsilon > 0$ is arbitrary, and eigenvalues are continuous functions of the entries of the matrix, it follows that all eigenvalues of $(V - F)$ have nonnegative real parts. To reverse the argument, suppose all the eigenvalues of $(V - F)$ have nonnegative real parts. For any positive ε , $(V + \varepsilon I - F)$ is a nonsingular M-matrix, and by Lemma 9.2, $\rho(F(V + \varepsilon I)^{-1}) < 1$. Again, since $\varepsilon > 0$ is arbitrary, it follows that $\rho(FV^{-1}) \leq 1$. Thus, $(F - V)$ has at least one eigenvalue with positive real part if and only if $\rho(FV^{-1}) > 1$, and the disease-free equilibrium is unstable whenever $\mathcal{R}_0 > 1$.

These results validate the extension of the definition of the reproduction number to more general situations. In the vaccination model (9.29) of the previous section we calculated a pair of final size relations that contained the elements of a matrix K . This matrix is precisely the next generation matrix with large domain $K_L = FV^{-1}$ that has been introduced in this section.

Example 1. Consider the *SEIR* model with infectivity in the exposed stage,

$$\begin{aligned} S' &= -\beta S(I + \varepsilon E), \\ E' &= \beta S(I + \varepsilon E) - \kappa E, \\ I' &= \kappa E - \alpha I, \\ R' &= \alpha I. \end{aligned} \tag{9.38}$$

Here the disease states are E and I ,

$$\mathcal{F} = \begin{bmatrix} \varepsilon E \beta N + I \beta N \\ 0 \end{bmatrix},$$

and

$$F = \begin{bmatrix} \varepsilon \beta N & \beta N \\ 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} \kappa & 0 \\ -\kappa & \alpha \end{bmatrix}, \quad V^{-1} = \begin{bmatrix} \frac{1}{\kappa} & 0 \\ \frac{1}{\alpha} & \frac{1}{\alpha} \end{bmatrix}.$$

Then we may calculate

$$K_L = FV^{-1} = \begin{bmatrix} \frac{\varepsilon \beta N}{\kappa} + \frac{\beta N}{\alpha} & \frac{\beta N}{\alpha} \\ 0 & 0 \end{bmatrix}.$$

Since FV^{-1} has rank 1, it has only one nonzero eigenvalue, and since the trace of the matrix is equal to the sum of the eigenvalues, it is easy to see that

$$\mathcal{R}_0 = \frac{\varepsilon\beta N}{\kappa} + \frac{\beta N}{\alpha},$$

the element in the first row and first column of FV^{-1} . If all new infections are in a single compartment, as is the case here, the basic reproduction number is the trace of the matrix FV^{-1} .

In general, it is possible to reduce the size of the next generation matrix to the number of *states at infection* [Diekmann and Heesterbeek (2000)]. The states at infection are those disease states in which there can be new infections. Suppose that there are n disease states and k states at infection with $k < n$. Then we may define an auxiliary $n \times k$ matrix E in which each column corresponds to a state at infection and has 1 in the corresponding row and 0 elsewhere. Then the next generation matrix is the $k \times k$ matrix

$$K = E^T K_L E.$$

It is easy to show, using the fact that $EE^T K_L = K_L$, that the $n \times n$ matrix K_L and the $k \times k$ matrix K have the same nonzero eigenvalues and therefore the same spectral radius. Construction of the next generation matrix that has lower dimension than the next generation matrix with large domain may simplify the calculation of the basic reproduction number.

In Example 1 above, the only disease state at infection is E , the matrix A is

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

and the next generation matrix K is the 1×1 matrix

$$K = \left[\frac{\varepsilon\beta N}{\kappa} + \frac{\beta N}{\alpha} \right].$$

Example 2. Consider the vaccination model (9.29). The disease states are I_U and I_V . Then

$$\mathcal{F} = \begin{bmatrix} \beta N_U(I_U + \delta I_V) \\ \sigma\beta N_V(I_U + \delta I_V) \end{bmatrix},$$

and

$$F = \begin{bmatrix} \beta N_U & \delta\beta N_U \\ \sigma\beta N_V & \sigma\delta\beta N_V \end{bmatrix} \quad V = \begin{bmatrix} \alpha_U & 0 \\ 0 & \alpha_V \end{bmatrix}.$$

It is easy to see that the next generation matrix with large domain is the matrix K calculated in Section 3.3. Since each disease state is a disease state at infection, the next generation matrix is K , the same as the next generation matrix with large domain. As in Example 1, the determinant of K is zero and K has rank 1. Thus the control reproduction number is the trace of K ,

$$\mathcal{R}_c = \frac{\beta N_U}{\alpha_U} + \frac{\delta \sigma N_V}{\alpha_V}.$$

The next two examples include demographics (births and deaths).

Example 3. (A multi-strain model of gonorrhea) The following example comes from Lima and Torres (1997). The system of equations is

$$\begin{aligned} \frac{ds}{dt} &= \rho - \mu S(t) - \frac{c\lambda_1 S(t)I_1(t)}{N(t)} - \frac{c\lambda_2 S(t)I_2(t)}{N(t)} + (1-p)\gamma_1 I_1(t) + \gamma_2 I_2(t), \\ \frac{dI_1}{dt} &= \frac{c\lambda_1 S(t)I_1(t)}{N(t)} - (\mu + \gamma_1)I_1(t), \\ \frac{dI_2}{dt} &= \frac{c\lambda_2 S(t)I_2(t)}{N(t)} - (\mu + \gamma_2)I_2(t) + p\gamma_1 I_1(t), \end{aligned}$$

where S is the susceptible class, I_1 is the class infected with strain 1, and I_2 is the class of individuals infected with a mutated strain. The “birth” rate of the population is ρ , μ is the natural mortality rate, c is the probability of successful contact, λ_i is the of strain i , γ_i is the recovery rate of strain i , and p is the proportion of the original infected population that become infected by the mutated strain.

The disease-free equilibrium for this model is $[S = N, I_1 = 0, I_2 = 0]$. Next we will reorder our variables: $\left[\frac{dI_1}{dt} \frac{dI_2}{dt} \right]^T$ and note that we need only the infected classes to calculate \mathcal{R}_0 . Then the new infection terms are $\frac{c\lambda_1 S(t)I_1(t)}{N(t)}$ in the $\frac{dI_1}{dt}$ equation and $\frac{c\lambda_2 S(t)I_2(t)}{N(t)}$ in the $\frac{dI_2}{dt}$ equation, $p\gamma_1 I_1(t)$ enter the I_2 class, but only *after* they have been infected with strain 1. Then

$$\mathcal{F} = \begin{bmatrix} \frac{c\lambda_1 S(t)I_1(t)}{N(t)} \\ \frac{c\lambda_2 S(t)I_2(t)}{N(t)} \end{bmatrix} \quad \text{and} \quad \mathcal{V} = \begin{bmatrix} (\mu + \gamma_1)I_1(t) \\ (\mu + \gamma_2)I_2(t) - p\gamma_1 I_1(t) \end{bmatrix}.$$

Since we have only two infected classes, $n = 2$, and our Jacobian matrices are

$$\begin{aligned} \mathbf{F} &= \begin{bmatrix} \frac{c\lambda_1 S(t)}{N(t)} & 0 \\ 0 & \frac{c\lambda_2 S(t)}{N(t)} \end{bmatrix} \Big|_{DFE} = \begin{bmatrix} c\lambda_1 & 0 \\ 0 & c\lambda_2 \end{bmatrix} \\ \mathbf{V} &= \begin{bmatrix} \mu + \gamma_1 & 0 \\ -p\gamma_1 & \mu + \gamma_2 \end{bmatrix} \Big|_{DFE} = \begin{bmatrix} \mu + \gamma_1 & 0 \\ -p\gamma_1 & \mu + \gamma_2 \end{bmatrix}. \end{aligned}$$

Then we can calculate the inverse of \mathbf{V} ,

$$\mathbf{V}^{-1} = \begin{bmatrix} \frac{1}{\mu + \gamma_1} & 0 \\ \frac{p\gamma_1}{(\mu + \gamma_1)(\mu + \gamma_2)} & \frac{1}{\mu + \gamma_2} \end{bmatrix},$$

and

$$\mathbf{FV}^{-1} = \begin{bmatrix} \frac{c\gamma_1}{\mu+\gamma_1} & 0 \\ \frac{c\gamma_2 p\gamma_1}{(\mu+\gamma_1)(\mu+\gamma_2)} & \frac{c\gamma_2}{\mu+\gamma_2} \end{bmatrix}.$$

To calculate the spectral radius of \mathbf{FV}^{-1} we find the eigenvalues of the matrix:

$$\rho(\mathbf{FV}^{-1}) = \max \left\{ \frac{c\lambda_1}{\mu + \gamma_1}, \frac{c\lambda_2}{\mu + \gamma_2} \right\}.$$

We often call $\mathcal{R}_1 = \frac{c\lambda_1}{\mu+\gamma_1}$ the reproductive number for strain 1 and $\mathcal{R}_2 = \frac{c\lambda_2}{\mu+\gamma_2}$ the reproductive number for strain 2. Then the basic reproductive number is

$$\mathcal{R}_0 = \max[\mathcal{R}_1, \mathcal{R}_2].$$

Example 4. (A three population model of West Nile virus) The following example comes from Chowell-Puente et al. (2004b). The system of equations is

$$\begin{aligned} S'_M &= \mu_M N_M - \beta_H S_M b \frac{N_H}{N_M} \frac{P_H N_M}{N_H + N_B} \frac{I_H}{N_H} - \beta_B S_M b \frac{N_B}{N_M} \frac{P_B N_M}{N_H + N_B} \frac{I_B}{N_B} - S_M \mu_M, \\ I'_M &= \beta_H S_M b \frac{N_H}{N_M} \frac{P_H N_M}{N_H + N_B} \frac{I_H}{N_H} + \beta_B S_M b \frac{N_B}{N_M} \frac{P_B N_M}{N_H + N_B} \frac{I_B}{N_B} - I_M \mu_M, \\ S'_B &= \Lambda - \beta_M S_B b \frac{N_M}{N_B} \frac{N_B}{N_H + N_B} \frac{I_M}{N_M} - S_B \mu_B, \\ I'_B &= \beta_M S_B b \frac{N_M}{N_B} \frac{N_B}{N_H + N_B} \frac{I_M}{N_M} - I_B \mu_B, \\ S'_H &= \mu_H N_H - \beta_M S_H b \frac{N_M}{N_H} \frac{N_H}{N_H + N_B} \frac{I_M}{N_M} - S_H \mu_H, \\ I'_H &= \beta_M S_H b \frac{N_M}{N_H} \frac{N_H}{N_H + N_B} \frac{I_M}{N_M} - (\mu + \theta) I_H, \\ R'_H &= \theta I_H - \mu_H R_H, \end{aligned}$$

where S_i refers to the susceptible class of species i , and I_i refers to the infected class of species i for $i = M, B, H$, mosquitoes, birds, and humans, respectively. Then P_i is the mosquito biting preference for species i , μ_i is the natural mortality rate of species i , b is the number of bites per mosquito per unit time, θ is the human recovery rate, β_M is the transmission probability from mosquito to host per bite, β_B is the transmission probability from birds to mosquito, and β_H is the transmission probability from humans to mosquito.

The disease-free equilibrium for this model is $[S_M = N_M, I_M = 0, S_B = \frac{\Lambda}{\mu_B}, I_B = 0, S_H = N_H, I_H = 0, R_H = 0]$. Next we will reorder our variables

$$[I_M, I_B, I_H, S_M, S_B, S_H, R_H]$$

and again note that we need only the infected classes. Then the new infections enter our \mathcal{F} vector as

$$\mathcal{F} = \begin{bmatrix} \beta_H S_M b \frac{N_H}{N_M} \frac{P_H N_M}{N_H + N_B} \frac{I_H}{N_H} + \beta_B S_M b \frac{N_B}{N_M} \frac{P_B N_M}{N_H + N_B} \frac{I_B}{N_B} \\ \beta_M S_B b \frac{N_M}{N_B} \frac{N_B}{N_H + N_B} \frac{I_M}{N_M} \\ \beta_M S_H b \frac{N_M}{N_H} \frac{N_H}{N_H + N_B} \frac{I_M}{N_M} \end{bmatrix},$$

and our \mathcal{V} vector is the remaining terms,

$$\mathcal{V} = \begin{bmatrix} I_M \mu_M \\ I_B \mu_B \\ (\mu + \theta) I_H \end{bmatrix}.$$

Then our Jacobian matrices are

$$\mathbf{F}|_{DFE} = \begin{bmatrix} 0 & \frac{\beta_B b P_B N_M}{N_H + \frac{\Lambda}{\mu_B}} & \frac{\beta_H b P_H N_M}{N_H + \frac{\Lambda}{\mu_B}} \\ \frac{\beta_M b \Lambda}{\mu_B (N_H + \frac{\Lambda}{\mu_B})} & 0 & 0 \\ \frac{\beta_M b N_H}{N_H + \frac{\Lambda}{\mu_B}} & 0 & 0 \end{bmatrix},$$

$$\mathbf{V}|_{DFE} = \begin{bmatrix} \mu_M & 0 & 0 \\ 0 & \mu_B & 0 \\ 0 & 0 & \mu_H + \theta \end{bmatrix},$$

and

$$\mathbf{V}^{-1} = \begin{bmatrix} \frac{1}{\mu_M} & 0 & 0 \\ 0 & \frac{1}{\mu_B} & 0 \\ 0 & 0 & \frac{1}{\mu_H + \theta} \end{bmatrix},$$

$$\mathbf{FV}^{-1} = \begin{bmatrix} 0 & \frac{\beta_B b P_B N_M}{(N_H + \frac{\Lambda}{\mu_B}) \mu_B} & \frac{\beta_H b P_H N_M}{(N_H + \frac{\Lambda}{\mu_B}) (\mu_H + \theta)} \\ \frac{\beta_M b \Lambda}{\mu_B (N_H + \frac{\Lambda}{\mu_B}) \mu_M} & 0 & 0 \\ \frac{\beta_M b N_H}{N_H + \frac{\Lambda}{\mu_B}} & 0 & 0 \end{bmatrix},$$

from which we calculate the eigenvalues and determine that the spectral radius is

$$\mathcal{R}_0 = \sqrt{\frac{b N_H}{\frac{\Lambda}{\mu_B} + N_H} \frac{\beta_M}{\mu_M} \frac{b P_H N_M}{\frac{\Lambda}{\mu_B} + N_H} \frac{\beta_H}{\mu_H + \theta} + \frac{b \frac{\Lambda}{\mu_B}}{\frac{\Lambda}{\mu_B} + N_H} \frac{\beta_M}{\mu_M} \frac{b P_B N_M}{\frac{\Lambda}{\mu_B} + N_H} \frac{\beta_B}{\mu_B}}$$

We have described the next generation matrix method for continuous models. There is an analogous theory for discrete systems, described in [Allen and van den Driessche (2008)].

9.10.1 A Global Asymptotic Stability Result

There are some situations in which $\mathcal{R}_0 < 1$ in which it is possible to show that the asymptotic stability of the disease - free equilibrium is global, that is, all solutions approach the disease - free equilibrium, not only those with initial values sufficiently close to this equilibrium.

We will say that a vector is nonnegative if each of its components is nonnegative, and that a matrix is nonnegative if each of its entries is non - negative. We rewrite the system (9.35) as

$$\begin{aligned} x' &= -Ax - \hat{f}(x, y), \\ y'_j &= g_j(x, y), \quad j = 1, \dots, m. \end{aligned} \tag{9.39}$$

If $\mathcal{R}_0 < 1$, we have shown that the disease - free equilibrium is asymptotically stable, and that $-A = -(F - V)$ is a non - singular M - matrix.

Theorem 9.2 (Castillo-Chavez, Feng, and Huang (2002)). *If $-A$ is a nonsingular M-matrix and $\hat{f} \geq 0$, if the assumptions on the model (9.35) made earlier in this section are satisfied, and if $\mathcal{R}_0 < 1$, then the disease-free equilibrium of (9.39) is globally asymptotically stable.*

Proof. The variation of constants formula for the first equation of (9.39) gives

$$x(t) = e^{-tA}x(0) - \int_0^t e^{-(t-s)A} \hat{f}(x(s), y(s)) ds.$$

It can be shown that $e^{-tA} \geq 0$ if $-A$ is an M-matrix, because

$$-A = B - sI$$

with $B \geq 0$,

$$e^{-tA} = e^{tB} e^{-stI} = e^{tB} e^{-st} I = e^{tB} e^{-st},$$

and $e^{tB} \geq 0$, since $B \geq 0$. This, together with the assumption that $\hat{f} \geq 0$, implies that

$$0 \leq x(t) \leq e^{-tA}x(0),$$

and since $e^{-tA}x(0) \rightarrow 0$ as $t \rightarrow \infty$ it follows that $x(t) \rightarrow 0$ as $t \rightarrow \infty$.

There are examples to show that the disease-free equilibrium may not be globally asymptotically stable if the condition $\hat{f} \geq 0$ is not satisfied.

Exercises

1. For each of the models (9.16), (9.18), and (9.19) use the next generation approach to calculate their reproduction numbers.

2. Use the next generation approach to calculate the basic reproduction number for the model (9.26) but with infectivity in the exposed class having a reduction factor ϵ .
3. Formulate an SEITR model and calculate its reproduction number.
4. For each of the examples in this section determine whether the disease-free equilibrium is globally asymptotically stable when $\mathcal{R}_0 < 1$.

9.11 Directions for Generalization

A fundamental assumption in the model (9.2) is homogeneous mixing, that all individuals are equivalent in contacts. A more realistic approach would include separation of the population into subgroups with differences in behavior. For example, in many childhood diseases the contacts that transmit infection depend on the ages of the individuals, and a model should include a description of the rate of contact between individuals of different ages. Other heterogeneities that may be important include activity levels of different groups and spatial distribution of populations. Network models may be formulated to include heterogeneity of mixing, or more complicated compartmental models can be developed.

An important question that should be kept in mind in the formulation of epidemic models is the extent to which the fundamental properties of the simple model (9.2) carry over to more elaborate models.

An epidemic model for a disease in which recovery from infection brings only temporary immunity cannot be described by the models of this chapter because of the flow of new susceptibles into the population. This effectively includes demographics in the model, and such models will be described in the next chapter.

Many of the important underlying ideas of mathematical epidemiology arose in the study of malaria begun by Sir R.A. Ross (1911). Malaria is one example of a disease with vector transmission, the infection being transmitted back and forth between vectors (mosquitoes) and hosts (humans). Other vector diseases include West Nile virus and HIV with heterosexual transmission. Vector transmitted diseases require models that include both vectors and hosts.

9.12 Some Warnings

An actual epidemic differs considerably from the idealized models (9.2) and (9.28). Some notable differences are these:

1. When it is realized that an epidemic has begun, individuals are likely to modify their behavior by avoiding crowds to reduce their contacts and by being more careful about hygiene to reduce the risk that a contact will produce infection.
2. If a vaccine is available for the disease that has broken out, public health measures will include vaccination of part of the population. Various vaccination

strategies are possible, including vaccination of health care workers and other first line responders to the epidemic, vaccination of members of the population who have been in contact with diagnosed infectives, and vaccination of members of the population who live in close proximity to diagnosed infectives.

3. Isolation may be imperfect; in-hospital transmission of infection was a major problem in the SARS epidemic.
4. In the SARS epidemic of 2002–2003, in-hospital transmission of disease from patients to health care workers or visitors because of imperfect isolation accounted for many of the cases. This points to an essential heterogeneity in disease transmission that must be included whenever there is any risk of such transmission.

9.13 Project: Discrete Epidemic Models

The discrete analogue of the continuous-time epidemic model (9.2) is

$$\begin{aligned} S_{j+1} &= S_j G_j, \\ I_{j+1} &= S_j(1 - G_j) + \sigma I_j, \\ G_j &= e^{-\beta I_j/N}, \quad j = 1, 2, \dots, \end{aligned} \tag{9.40}$$

where S_j and I_j denote the numbers of susceptible and infective individuals at time j , respectively, G_j is the probability that a susceptible individual at time j will remain susceptible to time $j + 1$, and $\sigma = e^{-\alpha}$ is the probability that an infected individual at time j will remain infected to time $j + 1$. Assume that the initial conditions are $S(0) = S_0 > 0$, $I(0) = I_0 > 0$, and $S_0 + I_0 = N$.

Exercise 1. Consider the system (9.40).

- (a) Show that the sequence $\{S_j + I_j\}$ has a limit

$$S_\infty + I_\infty = \lim_{j \rightarrow \infty} (S_j + I_j).$$

- (b) Show that

$$I_\infty = \lim_{j \rightarrow \infty} I_j = 0.$$

- (c) Show that

$$\log \frac{S_0}{S_\infty} = \beta \sum_{m=0}^{\infty} \frac{I_m}{N}.$$

- (d) Show that

$$\log \frac{S_0}{S_\infty} = \mathcal{R}_0 \left[1 - \frac{S_\infty}{N} \right],$$

$$\text{with } \mathcal{R}_0 = \frac{\beta}{1-\sigma}.$$

Next, consider the case that there are k infected stages and there is treatment in some stages, with treatment rates that can be different in different stages. Assume that selection of members for treatment occurs only at the beginning of a stage. Let $I_j^{(i)}$ and $T_j^{(i)}$ denote the numbers of infected and treated individuals respectively in stage i ($i = 1, 2, \dots, k$) at time j . Let σ_i^I denote the probability that an infected individual in the $I^{(i)}$ stage continues on to the next stage, either treated or untreated, and let σ_i^T denote the probability that an individual in the $T^{(i)}$ stage continues on to the next treated stage. In addition, of the members leaving an infected stage $I^{(i)}$, a fraction p_i enters treatment in $T^{(i+1)}$, while the remaining fraction q_i continues to $I^{(i+1)}$. Let m_i denote the fraction of infected members who go through the stage $I^{(i)}$, and n_i the fraction of infected members who go through the stage $T^{(i)}$. Then,

$$\begin{aligned} m_1 &= q_1, \quad m_2 = q_1 q_2, \quad \dots, \quad m_k = q_1 q_2 \cdots q_k, \\ n_1 &= p_1, \quad n_2 = p_1 + q_1 p_2, \quad \dots, \quad n_k = p_1 + q_1 p_2 + \dots + q_1 q_2 \cdots q_{k-1} p_k. \end{aligned}$$

The discrete system with treatment is

$$\begin{aligned} S_{j+1} &= S_j G_j, \\ I_{j+1}^{(1)} &= q_1 S_j (1 - G_j) + \sigma_1^I I_j^{(1)}, \\ T_{j+1}^{(1)} &= p_1 S_j (1 - G_j) + \sigma_1^T T_j^{(1)} \\ I_{j+1}^{(i)} &= q_i (1 - \sigma_{i-1}^I) I_j^{(i-1)} + \sigma_i^I \eta_i I_j^{(i)}, \\ T_{j+1}^{(i)} &= p_i (1 - \sigma_{i-1}^I) I_j^{(i-1)} + (1 - \sigma_{i-1}^T) T_j^{(i-1)} + \sigma_i^T T_j^{(i)}, \end{aligned} \tag{9.41}$$

$[i = 2, \dots, k, j \geq 0]$, with

$$G_j = e^{-\beta \sum_{i=1}^k (\varepsilon_i I_j^{(i)} / N + \delta_i T_j^{(i)} / N)},$$

where ε_i is the relative infectivity of untreated individuals at stage i and δ_i is the relative infectivity of treated individuals at stage i . Consider the initial conditions

$$I_0^{(1)}(0) = q_1 I_0, \quad T_0^{(1)}(0) = p_1 I_0, \quad I_0^{(i)}(0) = T_0^{(i)}(0) = 0, \quad i \geq 2, \quad S_0 + I_0 = N.$$

Exercise 2. Consider the system (9.41). Show that

$$\log \frac{S_0}{S_\infty} = \mathcal{R}_c \left[1 - \frac{S_\infty}{N} \right], \tag{9.42}$$

with

$$\mathcal{R}_c = \beta \sum_{i=1}^k \left[\frac{\varepsilon_i m_i}{1 - \sigma_i^I} + \frac{\delta_i n_i}{1 - \sigma_i^T} \right].$$

Hint. Equation (9.42) can be proved by showing the following equalities first:

$$\log \frac{S_0}{S_\infty} = \beta \sum_{i=1}^k \left[\varepsilon_i \sum_{j=1}^{\infty} \frac{I_j^{(i)}}{N} + \delta_i \sum_{j=1}^{\infty} \frac{T_j^{(i)}}{N} \right],$$

$$m_i(N - S_\infty) = (1 - \sigma_i^I) \sum_{j=0}^{\infty} I_j^{(i)}, \quad i \geq 2,$$

$$\frac{n_i(N - S_\infty)}{1 - \sigma_i^T} = \sum_{j=0}^{\infty} T_j^{(i)}, \quad i \geq 2.$$

[References: Feng (2007), Feng, Xu, & Zhao (2007).]

9.14 Project: Fitting Data for an Influenza Model

Consider an SIR model (9.2) with basic reproduction number 1.5.

1. Describe the qualitative changes in (S, I, R) as a function of time for different values of β and α with $\beta \in \{0.0001, 0.0002, \dots, 0.0009\}$, for the initial condition $(S, I, R) = (10^6, 1, 0)$.
2. Discuss the result of part (a) in terms of the basic reproductive number (what is β/γ ?). Use a specific disease such as influenza to provide simple interpretations for the different time courses of the disease for the different choices of β and γ .
3. Repeat the steps in part (a) for values of $R_0 \in \{1.75, 2, 2.5\}$, and for each value of R_0 , choose the best pair of values (β, α) that fits the slope before the first peak in the data found in the table for reported H1N1 influenza cases in México below. (Hint: normalize the data so that the peak is 1, and then multiply the data by the size of the peak in the simulations.)

9.15 Project: Social Interactions

Suppose we have a system with multiple classes of mathematical biology teachers (MBT) at time t . The classes roughly capture the MBT individual attitudes toward learning new stuff. “Reluctant” means the class of MBTs that come into the door as new hires without a disposition to learn new stuff; the positive class corresponds to those who join the MBTs with the right attitude; and the rest of the classes should be self-explanatory.

Day	Cases	Day	Cases	Day	Cases	Day	Cases	Day	Cases	Day	Cases
75	2	95	4	115	318	135	83	155	152	175	328
76	1	96	11	116	399	136	75	156	138	176	298
77	3	97	5	117	412	137	87	157	159	177	335
78	2	98	7	118	305	138	98	158	186	178	330
79	3	99	4	119	282	139	71	159	222	179	375
80	3	100	4	120	227	140	73	160	204	180	366
81	4	101	4	121	212	141	78	161	257	181	291
82	4	102	11	122	187	142	67	162	208	182	251
83	5	103	17	123	212	143	68	163	198	183	215
84	7	104	26	124	237	144	69	164	193	184	242
85	3	105	20	125	231	145	65	165	243	185	223
86	1	106	12	126	237	146	85	166	231	186	317
87	2	107	26	127	176	147	55	167	225	187	305
88	5	108	33	128	167	148	67	168	239	188	228
89	7	109	44	129	139	149	75	169	219	189	251
90	4	110	107	130	142	150	71	170	199	190	207
91	10	111	114	131	162	151	97	171	215	191	159
92	11	112	155	132	138	152	168	172	309	192	155
93	13	113	227	133	117	153	126	173	346	193	214
94	4	114	280	134	100	154	148	174	332	194	237

Table 9.2 Reported case for H1N1-pandemic in Mexico.

$R(t)$:	reluctant MBTs
$P(t)$:	positive MBTs
$M(t)$:	masterful MBTs
$U(t)$:	unchangeable (that is, negative) MBTs
$I(t)$:	inactive MBTs

Assume that $N(t) = R(t) + P(t) + M(t) + U(t) + I(t)$ and that the total number of MBTs is constant, that is, $N(t) = \frac{K}{\mu}$ for all t , where K is a constant. The model is

$$\begin{aligned}
 \frac{dP}{dt} &= qK - \beta P \frac{M}{K} + \delta R - \mu P, \\
 \frac{dR}{dt} &= (1 - q)K - (\delta + \mu)R - \alpha R, \\
 \frac{dM}{dt} &= \beta P \frac{M}{K} - (\gamma + \mu)M, \\
 \frac{dU}{dt} &= -\mu U + \alpha R, \\
 \frac{dI}{dt} &= \gamma M - \mu I,
 \end{aligned}$$

where $q, \beta, \delta, \mu, \gamma$, and α are constants and $0 \leq q \leq 1$.

1. Interpret the parameters.

2. Look at the stability of the simplest equilibrium (your choice).
3. From \mathfrak{R}_0 , discuss what would be the impact of changing parameters q , γ , and δ .
4. What are your conclusions from this model?

Chapter 10

Models for Endemic Diseases

10.1 A Model for Diseases with No Immunity

We have been studying SIR models, in which the transitions are from susceptible to infective to removed, with the removal coming through recovery with full immunity (as in measles) or through death from the disease (as in plague, rabies, and many other animal diseases). Another type of model is an SIS model in which infectives return to the susceptible class on recovery because the disease confers no immunity against reinfection. Such models are appropriate for most diseases transmitted by bacterial or helminth agents, and most sexually transmitted diseases (including gonorrhea, but not such diseases as AIDS, from which there is no recovery). One important way in which SIS models differ from SIR models is that in the former there is a continuing flow of new susceptibles, namely recovered infectives. Later in this chapter we will study models that include demographic effects, namely births and deaths, another way in which a continuing flow of new susceptibles may arise.

The simplest SIS model, due to Kermack and McKendrick (1932), is

$$\begin{aligned} S' &= -\beta SI + \gamma I, \\ I' &= \beta SI - \gamma I. \end{aligned} \quad (10.1)$$

This differs from the SIR model only in that the recovered members return to the class S at a rate γI instead of passing to the class R . The total population $S + I$ is a constant, since $(S + I)' = 0$. We call this constant N ; sometimes population size is measured using K as the unit, so that the total population size is one. We may reduce the model to a single differential equation by replacing S by $N - I$ to give the single differential equation

$$I' = \beta I(N - I) - \gamma I = (\beta N - \gamma)I - \beta I^2 = (\beta N - \gamma)I \left(1 - \frac{I}{N - \frac{\gamma}{\beta}}\right). \quad (10.2)$$

Since (10.2) is a logistic differential equation of the form

$$I' = rI \left(1 - \frac{I}{K}\right),$$

with $r = \beta N - \gamma$ and with $K = N - \gamma/\beta$, our qualitative result tells us that if $\beta N - \gamma < 0$ or $\beta N/\gamma < 1$, then all solutions of the model (10.2) with nonnegative initial values except the constant solution $I = K - \beta/\gamma$ approach the limit zero as $t \rightarrow \infty$, while if $\beta K/\gamma > 1$, then all solutions with nonnegative initial values except the constant solution $I = 0$ approach the limit $K - \gamma/\beta > 0$ as $t \rightarrow \infty$. Thus there is always a single limiting value for I , but the value of the quantity $\beta K/\gamma$ determines which limiting value is approached, regardless of the initial state of the disease. In epidemiological terms this says that if the quantity $\beta K/\gamma$ is less than one, the infection dies out in the sense that the number of infectives approaches zero. For this reason the equilibrium $I = 0$, which corresponds to $S = K$, is called the *disease-free equilibrium*. On the other hand, if the quantity $\beta K/\gamma$ exceeds one, the infection persists. The equilibrium $I = K - \gamma/\beta$, which corresponds to $S = \gamma/\beta$, is called an *endemic equilibrium*.

As we have seen in epidemic models, the dimensionless quantity $\beta K/\gamma$ is called the basic reproduction number or contact number for the disease, and it is usually denoted by \mathcal{R}_0 . In studying an infectious disease, the determination of the basic reproduction number is invariably a vital first step. The value one for the basic reproduction number defines a threshold at which the course of the infection changes between disappearance and persistence. Since βK is the number of contacts made by an average infective per unit time and $1/\gamma$ is the mean infective period, \mathcal{R}_0 represents the average number of secondary infections caused by each infective over the course of the infection. Thus, it is intuitively clear that if $\mathcal{R}_0 < 1$, the infection should die out, while if $\mathcal{R}_0 > 1$, the infection should establish itself. In more highly structured models than the simple one we have developed here, the calculation of the basic reproduction number may be much more complicated, but the essential concept remains, that of the basic reproduction number as the number of secondary infections caused by an average infective over the course of the disease. However, there is a difference from the behavior of epidemic models. Here, the basic reproduction number determines whether the infection establishes itself or dies out, whereas in the SIR epidemic model the basic reproduction number determines whether there will be an epidemic.

We were able to reduce the system of two differential equations (10.1) to the single equation (10.2) because of the assumption that the total population $S + I$ is constant. If there are deaths due to the disease, this assumption is violated, and it would be necessary to use a two-dimensional system as a model. We shall consider this in a more general context in the next section.

A model for a disease from which infectives recover with no immunity against reinfection and that includes births and deaths is

$$\begin{aligned} S' &= \Lambda(N) - \beta(N)SI - \mu S + f\alpha I, \\ I' &= \beta(N)SI - \alpha I - \mu I, \end{aligned} \tag{10.3}$$

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

It is easy to verify that

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

If we add the two equations of (10.11) and use $N = S + I$, we obtain

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

We will carry out the analysis of the SIS model only in the special case $f = 1$, so that N is the constant K . The system (10.11) is asymptotically autonomous and its asymptotic behavior is the same as that of the single differential equation

$$I' = \beta(K)I(K - I) - (\alpha + \mu)I, \quad (10.4)$$

where S has been replaced by $K - I$. But (10.4) is a logistic equation that is easily solved analytically by separation of variables or qualitatively by an equilibrium analysis. We find that $I \rightarrow 0$ if $K\beta(K) < (\mu + \alpha)$, or $\mathcal{R}_0 < 1$ and $I \rightarrow I_\infty > 0$ with

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

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

The endemic equilibrium, which exists if $\mathcal{R}_0 > 1$, is always asymptotically stable. If $\mathcal{R}_0 < 1$, the system has only the disease-free equilibrium and this equilibrium is asymptotically stable. The verification of these properties remains valid if there are no births and deaths. This suggests that a requirement for the existence of an endemic equilibrium is a flow of new susceptibles either through recovery without immunity against reinfection or through births.

Exercises

1. Modify the SIS model (10.1) to the situation in which there are two competing strains of the same disease, generating two infective classes I_1, I_2 under the assumption that coinfections are not possible. Does the model predict coexistence of the two strains or competitive exclusion?
- 2.* A communicable disease from which infectives do not recover may be modeled by the pair of differential equations

$$S' = -\beta SI, \quad I' = \beta SI.$$

Show that in a population of fixed size K such a disease will eventually spread to the entire population.

- 3.* Consider a disease spread by carriers who transmit the disease without exhibiting symptoms themselves. Let $C(t)$ be the number of carriers and suppose that carriers are identified and isolated from contact with others at a constant per capita rate α , so that $C' = -\alpha C$. The rate at which susceptibles become infected is proportional to the number of carriers and to the number of susceptibles, so that $S' = -\beta SC$. Let C_0 and S_0 be the numbers of carriers and susceptibles, respectively, at time $t = 0$.
 - (i) Determine the number of carriers at time t from the first equation.
 - (ii) Substitute the solution to part (i) into the second equation and determine the number of susceptibles at time t .
 - (iii) Find $\lim_{t \rightarrow \infty} S(t)$, the number of members of the population who escape the disease.
- 4.* Consider a population of fixed size K in which a rumor is being spread by word of mouth. Let $y(t)$ be the number of people who have heard the rumor at time t and assume that everyone who has heard the rumor passes it on to r others in unit time. Thus, from time t to time $(t+h)$, the rumor is passed on $hry(t)$ times, but a fraction $y(t)/K$ of the people who hear it have already heard it, and thus there are only $hry(t) \left(\frac{K-y(t)}{K} \right)$ people who hear the rumor for the first time. Use these assumptions to obtain an expression for $y(t+h) - y(t)$, divide by h , and take the limit as $h \rightarrow 0$ to obtain a differential equation satisfied by $y(t)$.
5. At 9 AM one person in a village of 100 inhabitants starts a rumor. Suppose that everyone who hears the rumor tells one other person per hour. Using the model of the previous exercise, determine how long it will take until half the village has heard the rumor.
- 6.* If a fraction λ of the population susceptible to a disease that provides immunity against reinfection moves out of the region of an epidemic, the situation may be modeled by a system

$$S' = -\beta SI - \lambda S, \quad I' = \beta SI - \alpha I.$$

Show that both S and I approach zero as $t \rightarrow \infty$.

10.2 The SIR Model with Births and Deaths

Epidemics that sweep through a population attract much attention and arouse a great deal of concern. We have omitted births and deaths in our description of epidemic models because the time scale of an epidemic is generally much shorter than the demographic time scale. In effect, we have used a time scale on which the number of births and deaths in unit time is negligible.

However, there are diseases that are endemic in many parts of the world and that cause millions of deaths each year. To model a disease that may be endemic we need to think on a longer time scale and include births and deaths. A reference describing the properties of many endemic diseases is Anderson and May (1991). For diseases that are endemic in some region, public health physicians would like to be able to estimate the number of infectives at a given time as well as the rate at which new infections arise. The effects of quarantine or vaccination in reducing the number of victims are of importance, just as in the treatment of epidemics. In addition, the possibility of defeating the endemic nature of the disease and thus controlling or even eradicating the disease in a population is worthy of study.

A model of Kermack and McKendrick (1932) includes births in the susceptible class proportional to total population size and a death rate in each class proportional to the number of members in the class. This model allows the total population size to grow exponentially or die out exponentially if the birth and death rates are unequal. It is applicable to such questions as whether a disease will control the size of a population that would otherwise grow exponentially. We shall return to this topic, which is important in the study of many diseases in less-developed countries with high birth rates. To formulate a model in which total population size remains bounded we could follow the approach suggested by Hethcote (1976) in which the total population size K is held constant by making birth and death rates equal. Such a model is

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

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

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

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

Let us analyze the model

$$\begin{aligned}
S' &= \Lambda - \beta SI - \mu S, \\
I' &= \beta SI - \mu I - \alpha I, \\
N' &= \Lambda - (1-f)\alpha I - \mu N,
\end{aligned} \tag{10.5}$$

where $N = S + I + R$, with a mass action contact rate, a constant number of births Λ per unit time, a proportional natural death rate μ in each class, and a rate of recovery or disease death α of infectives with a fraction f of infectives recovering with immunity against reinfection. In this model, if $f = 1$ the total population size approaches a limit $K = \Lambda/\mu$. Then K is the carrying capacity of the population. If $f < 1$, the total population size is not constant, and K again represents a carrying capacity or maximum possible population size, rather than a population size. We view the first two equations as determining S and I , and then consider the third equation as determining N once S and I are known. This is possible because N does not enter into the first two equations. Instead of using N as the third variable in this model we could have used R , and the same reduction would have been possible.

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

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

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

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

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

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

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

$$\Lambda(0) = 0, \quad \Lambda''(N) \leq 0.$$

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

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

$$\begin{aligned}
S' &= \Lambda(N) - \beta(N)SI - \mu S, \\
I' &= \beta(N)SI - \mu I - \alpha I, \\
N' &= \Lambda(N) - (1-f)\alpha I - \mu N.
\end{aligned} \tag{10.6}$$

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

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

so that $N(t)$ approaches a limiting population size K . The theory of *asymptotically autonomous systems* [Castillo-Chavez and Thieme (1993), Markus (1956), Thieme (1994), Thieme and Castillo-Chavez (1993)] implies that if N has a constant limit then the system is equivalent to the system in which N is replaced by this limit. Then the system (10.6) is the same as the system (10.5) with β replaced by the constant $\beta(K)$ and N by K , and $\Lambda(N)$ replaced by the constant $\Lambda(K) = \mu K$.

We shall analyze the model (10.5) qualitatively. In view of the remark above, our analysis will also apply to the more general model (10.6) if there are no disease deaths. Analysis of the system (10.6) with $f < 1$ is much more difficult. We will confine our study of (10.6) to a description without details.

The first stage of the analysis is to note that the model (10.5) is a properly posed problem. That is, since $S' \geq 0$ if $S = 0$ and $I' \geq 0$ if $I = 0$, we have $S \geq 0$, $I \geq 0$ for $t \geq 0$, and since $N' \leq 0$ if $N = K$, we have $N \leq K$ for $t \geq 0$. Thus the solution always remains in the biologically realistic region $S \geq 0$, $I \geq 0$, $0 \leq N \leq K$ if it starts in this region. By rights, we should verify such conditions whenever we analyze a mathematical model, but in practice this step is frequently overlooked.

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

$$I_\infty = \frac{\Lambda}{\mu + \alpha} - \frac{\mu}{\beta}. \quad (10.7)$$

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

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

The coefficient matrix of this linear system is

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

We then look for solutions whose components are constant multiples of $e^{\lambda t}$; this means that λ must be an eigenvalue of the coefficient matrix. The condition that all solutions of the linearization at an equilibrium tend to zero as $t \rightarrow \infty$ is that the real part of every eigenvalue of this coefficient matrix be negative. At the disease-free equilibrium the matrix is

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

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

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

$$\begin{bmatrix} -\beta I_\infty - \mu & -\beta S_\infty \\ \beta I_\infty & 0 \end{bmatrix}, \quad (10.8)$$

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

$$R_0 = \frac{\beta K}{\mu + \alpha} = \frac{K}{S_\infty} \quad (10.9)$$

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

There are two aspects of the analysis of the model (10.6) that are more complicated than the analysis of (10.5). The first is in the study of equilibria. Because of the dependence of $\Lambda(N)$ and $\beta(N)$ on N , it is necessary to use two of the equilibrium conditions to solve for S and I in terms of N and then substitute into the third condition to obtain an equation for N . Then by comparing the two sides of this

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

The second complication is in the stability analysis. Since (10.6) is a three-dimensional system that cannot be reduced to a two-dimensional system, the coefficient matrix of its linearization at an equilibrium is a 3×3 matrix, and the resulting characteristic equation is a cubic polynomial equation of the form

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

The *Routh–Hurwitz* conditions

$$a_1 > 0, \quad a_1a_2 > a_3 > 0$$

are necessary and sufficient conditions for all roots of the characteristic equation to have negative real part. A technically complicated calculation is needed to verify that these conditions are satisfied at an endemic equilibrium for the model (10.6).

The asymptotic stability of the endemic equilibrium means that the compartment sizes approach a steady state. If the equilibrium had been unstable, there would have been a possibility of sustained oscillations. Oscillations in a disease model mean fluctuations in the number of cases to be expected, and if the oscillations have long period, that could also mean that experimental data for a short period would be quite unreliable as a predictor of the future. Epidemiological models that incorporate additional factors may exhibit oscillations. A variety of such situations is described in [Hethcote and Levin (1989), Hethcote, Stech, and van den Driessche (1981)].

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

From the third equation of (10.5) we obtain

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

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

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

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

Numerical simulations indicate that the approach to endemic equilibrium for an SIR model is like a rapid and severe epidemic if the epidemiological and demographic time scales are very different. The same happens in the SIS model. If there are few disease deaths, the number of infectives at endemic equilibrium may be

substantial, and there may be damped oscillations of large amplitude about the endemic equilibrium. For both the SIR and SIS models we may write the differential equation for I as

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

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

Next, we will turn to some applications of SIR and SIS models.

10.3 Some Applications

10.3.1 Herd Immunity

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

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

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

For measles, epidemiological data in the United States indicate that \mathcal{R}_0 for rural populations ranges from 5.4 to 6.3, requiring vaccination of 81.5 percent to 84.1 percent of the population. In urban areas \mathcal{R}_0 ranges from 8.3 to 13.0, requiring vaccination of 88.0 percent to 92.3 percent of the population. In Great Britain, \mathcal{R}_0 ranges from 12.5 to 16.3, requiring vaccination of 92 percent to 94 percent of the population. The measles vaccine is not always effective, and vaccination campaigns are never able to reach everyone. As a result, herd immunity against measles has not been achieved (and probably never can be). Since smallpox is viewed as more serious and requires a lower percentage of the population be immunized, herd immunity was attainable for smallpox. In fact, smallpox has been eliminated; the last known case was in Somalia in 1977, and the virus is maintained now only in laboratories. The eradication of smallpox was actually more difficult than expected, because high vaccination rates were achieved in some countries but not everywhere, and the disease persisted in some countries. The eradication of smallpox was possible only after an intensive campaign for worldwide vaccination [Hethcote (1978)].

10.3.2 Age at Infection

In order to calculate the basic reproduction number \mathcal{R}_0 for a disease, we need to know the values of the contact rate β and the parameters μ, K , and α . The parameters μ, K , and α can usually be measured experimentally, but the contact rate β is difficult to determine directly. There is an indirect means of estimating \mathcal{R}_0 in terms of the life expectancy and the mean age at infection that enables us to avoid having to estimate the contact rate. In this calculation, we will assume that β is constant, but we will also indicate the modifications needed when β is a function of total population size N . The calculation assumes exponentially distributed life spans and infective periods. In fact, the result is valid so long as the life span is exponentially distributed, but if the life span is not exponentially distributed, the result could be quite different.

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

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

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

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

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

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

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

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

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

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

For our model the endemic equilibrium is

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

and this implies

$$\frac{L}{A} = \frac{\beta I_{\infty}}{\mu} = \mathcal{R}_0 - 1. \quad (10.10)$$

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

The derivation of $A = 1/\beta I_{\infty}$ is obtained from considering surviving susceptible members at each age. This is the value that would be obtained from data giving the fraction of susceptibles at each age. However, if average age at infection has the normal meaning of average age at which those people who become infected do become infected, then the calculation would be different. The susceptible population at age a is a fraction $e^{-(\mu+\beta I_{\infty})a}$ of the number of newborn members, and the incidence of new infections is $\beta I_{\infty} e^{-(\mu+\beta I_{\infty})a}$. This would lead to an average age at infection $A^* = 1/(\mu + \beta I_{\infty})$ and the relation $L/A^* = \mathcal{R}_0$.

If β is a function $\beta(N)$ of total population size, the relation (10.10) becomes

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

If disease mortality does not have a large effect on total population size, in particular if there is no disease mortality, this relation is very close to (10.10).

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

10.3.3 The Interepidemic Period

Many common childhood diseases, such as measles, whooping cough, chicken pox, diphtheria, and rubella, exhibit variations from year to year in the number of cases.

These fluctuations are frequently regular oscillations, suggesting that the solutions of a model might be periodic. This does not agree with the predictions of the model we have been using in this section; however, it would not be inconsistent with solutions of the characteristic equation, which are complex conjugate with small negative real part corresponding to lightly damped oscillations approaching the endemic equilibrium. Such behavior would look like recurring epidemics. If the eigenvalues of the matrix of the linearization at an endemic equilibrium are $-u \pm iv$, where $i^2 = -1$, then the solutions of the linearization are of the form $Be^{-ut} \cos(vt + c)$, with decreasing “amplitude” Be^{-ut} and “period” $\frac{2\pi}{v}$.

For the model (10.5) we recall from (10.7) that at the endemic equilibrium we have

$$\beta I_\infty + \mu = \mu R_0, \quad \beta S_\infty = \mu + \gamma + \alpha,$$

and from (10.8), the matrix of the linearization is

$$\begin{bmatrix} -\mu R_0 & -(\mu + \gamma + \alpha) \\ \mu(R_0 - 1) & 0 \end{bmatrix}.$$

The eigenvalues are the roots of the quadratic equation

$$\lambda^2 + \mu R_0 \lambda + \mu(R_0 - 1)(\mu + \gamma + \alpha) = 0,$$

which are

$$\lambda = \frac{-\mu R_0 \pm \sqrt{\mu^2 R_0^2 - 4\mu(R_0 - 1)(\mu + \gamma + \alpha)}}{2}.$$

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

$$\frac{-\mu R_0 \pm \sqrt{-4\mu(R_0 - 1)(\gamma + \alpha)}}{2},$$

and these are complex with imaginary part $\sqrt{\mu(R_0 - 1)(\gamma + \alpha)}$. This indicates oscillations with period approximately

$$\frac{2\pi}{\sqrt{\mu(R_0 - 1)(\gamma + \alpha)}}.$$

We use the relation $\mu(R_0 - 1) = \mu L/A$ and the mean infective period $\tau = 1/(\gamma + \alpha)$ to see that the interepidemic period T is approximately $2\pi\sqrt{A\tau}$. Thus, for example, for recurring outbreaks of measles with an infective period of 2 weeks or 1/26 year, in a population with a life expectancy of 70 years with R_0 estimated as 15, we would expect outbreaks spaced 2.76 years apart. Also, since the “amplitude” at time t is $e^{-\mu R_0 t/2}$, the maximum displacement from equilibrium is multiplied by a factor $e^{-(15)(2.76)/140} = 0.744$ over each cycle. In fact, many observations of measles outbreaks indicate less damping of the oscillations, suggesting that there may be additional influences that are not included in our simple model. To explain oscillations about the endemic equilibrium, a more complicated model is needed. One

possible generalization would be to assume seasonal variations in the contact rate. This is a reasonable supposition for a childhood disease most commonly transmitted through school contacts, especially in winter in cold climates. Note, however, that data from observations are never as smooth as model predictions and models are inevitably gross simplifications of reality that cannot account for random variations in the variables. It may be difficult to judge from experimental data whether an oscillation is damped or persistent.

10.3.4 “Epidemic” Approach to Endemic Equilibrium

In the model (10.5), the demographic time scale described by the birth and natural death rates μK and μ and the epidemiological time scale described by the rate $(\alpha + \gamma)$ of departure from the infective class may differ substantially. Think, for example, of a natural death rate $\mu = 1/75$, corresponding to a human life expectancy of 75 years, and epidemiological parameters $\alpha = 0$ and $\gamma = 25$, describing a disease from which all infectives recover after a mean infective period of $1/25$ year, or two weeks. Suppose we consider a carrying capacity $K = 1000$ and take $\beta = 0.1$, indicating that an average infective makes $(0.1)(1000) = 100$ contacts per year. Then $R_0 = 4.00$, and at the endemic equilibrium we have $S_\infty = 250.13$, $I_\infty = 0.40$, $R_\infty = 749.47$. This equilibrium is globally asymptotically stable and is approached from every initial state.

However, if we take $S(0) = 999$, $I(0) = 1$, $R(0) = 0$, simulating the introduction of a single infective into a susceptible population, and solve the system numerically, we find that the number of infectives rises sharply to a maximum of 400 and then decreases to almost zero in a period of 0.4 year, or about 5 months. In this time interval the susceptible population decreases to 22 and then begins to increase, while the removed (recovered and immune against reinfection) population increases to almost 1000 and then begins a gradual decrease. The size of this initial “epidemic” could not have been predicted from our qualitative analysis of the system (10.5). On the other hand, since μ is small compared to the other parameters of the model, we might consider neglecting μ , replacing it by zero in the model. If we do this, the model reduces to the simple Kermack–McKendrick epidemic model (without births and deaths) of Section 9.2.

If we follow the model (10.5) over a longer time interval, we find that the susceptible population grows to 450 after 46 years, then drops to 120 during a small epidemic with a maximum of 18 infectives, and exhibits widely spaced epidemics decreasing in size. It takes a very long time before the system comes close to the endemic equilibrium and remains close to it. The large initial epidemic conforms to what has often been observed in practice when an infection is introduced into a population with no immunity, such as the smallpox inflicted on the Aztecs by the invasion of Cortez.

If we use the model (10.5) with the same values of β , K , and μ , but take $\alpha = 25$, $\gamma = 0$ to describe a disease fatal to all infectives, we obtain very similar results. Now the total population is $S + I$, which decreases from an initial size of 1000 to a minimum of 22 and then gradually increases and eventually approaches its equilibrium size of 250.53. Thus, the disease reduces the total population size to one-fourth of its original value, suggesting that infectious diseases may have large effects on population size. This is true even for populations that would grow rapidly in the absence of infection, as we shall see in the next section.

10.3.5 The SIS Model with Births and Deaths

In order to describe a model for a disease from which infectives recover with immunity against reinfection and that includes births and deaths as in the model (10.5), we may modify the model (10.5) by removing the equation for R' and moving the term αI describing the rate of recovery from infection to the equation for S' . This gives the model

$$\begin{aligned} S' &= -\beta SI + \mu(K - S) + \alpha I, \\ I' &= \beta SI - \alpha I - \mu I - \gamma I, \end{aligned} \tag{10.11}$$

describing a population with a constant number of births μK per unit time, a proportional death rate μ in each class, and a fraction γ of infectives dying from infection and a fraction α of infectives recovering with no immunity against reinfection. In this model, if $\gamma > 0$, the total population size is not constant and K represents a *carrying capacity*, or maximum possible population size, rather than a constant population size. The analysis of the model (10.11) is very similar to that of the *SIR* model (10.5), except that there is no equation for R' to be eliminated.

The only difference is the additional term αI in the equation for S' , and this does not change any of the qualitative results. As in the *SIR* model we have a basic reproductive number

$$\mathcal{R}_0 = \frac{\beta K}{\mu + \gamma + \alpha} = \frac{K}{S_\infty},$$

and if $\mathcal{R}_0 < 1$, the disease-free equilibrium $S = K$, $I = 0$ is asymptotically stable, while if $\mathcal{R}_0 > 1$, there is an endemic equilibrium (S_∞, I_∞) with $\beta S_\infty = \mu + \gamma + \alpha$ and I_∞ given by (10.7), which is asymptotically stable. There are, however, differences that are not disclosed by the qualitative analysis. If the epidemiological and demographic time scales are very different, for the *SIR* model we observed that the approach to endemic equilibrium is like a rapid and severe epidemic. The same happens in the *SIS* model, especially if there is a significant number of deaths due to disease. If there are few disease deaths, the number of infectives at endemic equilibrium may be substantial, and there may be oscillations of large amplitude about the endemic equilibrium.

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

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

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

Exercise

1. Recurrent outbreaks of measles and other childhood diseases have previously been explained by an interaction between intrinsic epidemiological forces generating dampened oscillations and seasonal and/or stochastic excitation. The following model shows that isolation or quarantine (i.e., sick individuals stay at home and have a reduced infective impact) can create self-sustained oscillations.

In the model considered here the population is divided into susceptibles (S), infectives (I), isolated or quarantined individuals (Q), and recovered individuals (R), for whom permanent immunity is assumed. Let N denote the total population, and let $A = S + I + R$ denote the active (nonisolated) individuals. The model takes the form

$$\begin{aligned} \frac{dS}{dt} &= \mu N - \mu S - \sigma S \frac{I}{A}, \\ \frac{dI}{dt} &= -(\mu + \gamma)I + \sigma S \frac{I}{A}, \\ \frac{dQ}{dt} &= -(\mu + \xi)Q + \gamma I, \\ \frac{dR}{dt} &= -\mu R + \xi Q, \\ A &= S + I + R. \end{aligned} \tag{10.12}$$

All newborns are assumed to be susceptible. μ is the per capita mortality rate, σ is the per capita infection rate of an average susceptible individual provided everybody else is infected, γ is the rate at which individuals leave the infective class, and ξ is the rate at which individuals leave the isolated class; all are positive constants.

- (i) Show that the total population size N is constant.
- (ii) Give the meanings of $1/\mu$, $1/\gamma$, $1/\xi$ and their units.
- (iii) Rescale the model by: $\tau = \sigma t, u = S/A, y = I/A, q = Q/A, z = R/A$. Rearrange your new model as follows:

$$\begin{aligned} \dot{y} &= y(1 - v - \theta - y - z + \theta y - (v + \zeta)q) \\ \dot{q} &= (1 + q)(\theta y - (v + \zeta)q) \\ \dot{z} &= \zeta q - vz + z(\theta y - (v + \zeta)q). \end{aligned} \tag{10.13}$$

- Express the new parameters in terms of the old parameters. Check that all the new parameters and variables are dimensionless.
- (iv) Study the stability of the equilibrium point $(0, 0, 0)$ and derive a basic reproductive number R_0 .
 - (v) Use a computer algebra system to demonstrate that (10.13) has periodic trajectories. Use the parameter values $\nu = 0.0002$, $\theta = 0.0156$, and ξ close to $\theta^2(1 - \theta)$. You also need to choose proper initial values.

10.4 Temporary Immunity

In the *SIR* models that we have studied, it has been assumed that the immunity received by recovery from the disease is permanent. This is not always true, since there may be a gradual loss of immunity with time. In addition, there are often mutations in a virus, and as a result the active disease strain is sufficiently different from the strain from which an individual has recovered, that the immunity received may wane.

Temporary immunity may be described by an *SIRS* model in which a rate of transfer from *R* to *S* is added to an *SIR* model. For simplicity, we confine our attention to epidemic models, without including births, natural deaths, and disease deaths, but the analysis of models including births and deaths would lead to the same conclusions. Thus we begin with a model

$$\begin{aligned} S' &= -\beta SI + \theta R, \\ I' &= \beta SI - \alpha I, \\ R' &= \alpha I - \theta R, \end{aligned}$$

with a proportional rate θ of loss of immunity.

Since $N' = (S + I + R)' = 0$, the total population size N is constant, and we may replace R by $N - S - I$ and reduce the model to a two-dimensional system

$$\begin{aligned} S' &= -\beta SI + \theta(N - S - I), \\ I' &= \beta SI - \alpha I. \end{aligned} \tag{10.14}$$

Equilibria are solutions of the system

$$\begin{aligned} \beta SI + \theta S + \theta I &= \theta N, \\ \alpha I + \theta S + \theta I &= \theta N, \end{aligned}$$

and there is a disease-free equilibrium $S = \alpha/\beta$, $I = 0$. If $\mathcal{R}_0 = \beta N/\alpha > 1$, there is also an endemic equilibrium with

$$\beta S = \alpha, \quad (\alpha + \theta)I = \theta(N - S).$$

The matrix of the linearization of (10.14) at an equilibrium (S, I) is

$$A = \begin{bmatrix} -(\beta I + \theta) & -(\beta S + \theta) \\ \beta I & \beta S - \alpha \end{bmatrix}.$$

At the disease-free equilibrium, A has the sign structure

$$\begin{bmatrix} - & - \\ + & \beta N - \alpha \end{bmatrix}.$$

This matrix has negative trace and positive determinant if and only if $\beta N < \alpha$, or $\mathcal{R}_0 < 1$. At an endemic equilibrium, the matrix has sign structure

$$\begin{bmatrix} - & - \\ + & 0 \end{bmatrix}.$$

and thus always has negative trace and positive determinant. We see from this that as in other models studied in this chapter, the disease-free equilibrium is asymptotically stable if and only if the basic reproduction number is less than 1 and the endemic equilibrium, which exists if and only if the basic reproduction number exceeds 1, is always asymptotically stable. However, it is possible for a different SIRS model to have quite different behavior.

We consider an SIRS model, that assumes a constant period of temporary immunity following recovery from the infection in place of an exponentially distributed period of temporary immunity. It will turn out that the endemic equilibrium for this model may be unstable, thus giving an example of a generalization that leads to new possibilities for the behavior of a model.

We add the assumption that there is a temporary immunity period of fixed length ω , after which recovered infectives revert to the susceptible class. The resulting model is described by the system

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

Since $N = S + I + R$ is constant, we may discard the equation for R and use a two-dimensional model

$$\begin{aligned} S'(t) &= -\beta S(t)I(t) + \alpha I(t - \omega), \\ I'(t) &= \beta S(t)I(t) - \alpha I(t). \end{aligned} \tag{10.15}$$

Equilibria are given by $I = 0$ or $\beta S = \alpha$. There is a disease-free equilibrium $S = N$, $I = 0$. There is also an endemic equilibrium for which $\beta S = \alpha$. However, the two equations for S and I give only a single equilibrium condition. To determine the endemic equilibrium (S_∞, I_∞) we must write the equation for R in the integrated form

$$R(t) = \int_{t-\omega}^t \alpha I(x) dx$$

to give $R_\infty = \omega \alpha I_\infty$. We also have $\beta S_\infty = \alpha$, and from $S_\infty + I_\infty + R_\infty = N$ we obtain

$$\beta I_\infty = \frac{\beta N - \alpha}{1 + \omega \alpha}.$$

The characteristic equation at an equilibrium is the condition that the linearization at the equilibrium have a solution whose components are constant multiples of $e^{\lambda t}$. In the ordinary differential equation case this is just the equation that determines the eigenvalues of the coefficient matrix, a polynomial equation, but in the general case it is a transcendental equation. *The result on which our analysis depends is that an equilibrium is asymptotically stable if all roots of the characteristic equation have negative real part, or equivalently that the characteristic equation have no roots with real part greater than or equal to zero.*

To linearize about an equilibrium (S_∞, I_∞) of (10.15) we substitute

$$S(t) = S_\infty + u(t), \quad I(t) = I_\infty + v(t),$$

and neglect the quadratic term, giving the linearization

$$\begin{aligned} u'(t) &= -\beta I_\infty u(t) - \beta S_\infty v(t) + \alpha v(t - \omega), \\ v'(t) &= \beta I_\infty u(t) + \beta S_\infty v(t) - \alpha v(t). \end{aligned}$$

The characteristic equation is the condition on λ that this linearization have a solution

$$u(t) = u_0 e^{\lambda t}, \quad v(t) = v_0 e^{\lambda t},$$

and this is

$$\begin{aligned} (\beta I_\infty + \lambda) u_0 + (\beta S_\infty - \alpha e^{-\lambda \omega}) v_0 &= 0, \\ \beta I_\infty u_0 + (\beta S_\infty - \alpha - \lambda) v_0 &= 0, \end{aligned}$$

or

$$\det \begin{bmatrix} \lambda + \beta I_\infty & \beta S_\infty - \alpha e^{\lambda \omega} \\ \beta I_\infty & \beta S_\infty - \alpha - \lambda \end{bmatrix}.$$

This reduces to

$$\beta \alpha I_\infty \frac{1 - e^{-\omega \lambda}}{\lambda} = -[\lambda + \beta S_\infty + \beta I_\infty - \alpha]. \quad (10.16)$$

At the disease-free equilibrium $S_\infty = N$, $I_\infty = 0$, this reduces to a linear equation with a single root $\lambda = \beta N - \alpha$, which is negative if and only if $R_0 = \beta N / \alpha < 1$.

We think of ω and N as fixed and consider β and α as parameters. If $\alpha = 0$, the equation (10.16) is linear and its only root is $-\beta I_\infty < 0$. Thus, there is a region in the (β, α) parameter space containing the β -axis in which all roots of (10.16) have negative real part. In order to find how large this stability region is we make use of the fact that the roots of (10.16) depend continuously on β and α . A root can

move into the right half-plane only by passing through the value zero or by crossing the imaginary axis as β and α vary. Thus, the stability region contains the β -axis and extends into the plane until there is a root $\lambda = 0$ or until there is a pair of pure imaginary roots $\lambda = \pm iy$ with $y > 0$. Since the left side of (10.16) is positive and the right side of (10.16) is negative for real $\lambda \geq 0$, there can not be a root $\lambda = 0$.

The condition that there is a root $\lambda = iy$ is

$$\alpha\beta I_\infty \frac{1 - e^{-i\omega\alpha}}{iy} = -(\beta I_\infty + iy), \quad (10.17)$$

and separation into real and imaginary parts gives the pair of equations

$$\alpha \frac{\sin \omega y}{y} = -1, \quad \alpha\beta I_\infty \frac{1 - \cos \omega y}{y} = y. \quad (10.18)$$

To satisfy the first condition it is necessary to have $\omega\alpha > 1$, since $|\sin \omega y| \leq |\omega y|$ for all y . This implies, in particular, that the endemic equilibrium is asymptotically stable if $\omega\alpha < 1$. In addition, it is necessary to have $\sin \omega y < 0$. There is an infinite sequence of intervals on which $\sin \omega y < 0$, the first being $\pi < \omega y < 2\pi$. For each of these intervals the equations (10.18) define a curve in the (β, α) plane parametrically with y as parameter. The region in the plane below the first of these curves is the region of asymptotic stability, that is, the set of values of β and α for which the endemic equilibrium is asymptotically stable. This curve is shown for $\omega = 1$, $N = 1$ in Figure 10.1. Since $\mathcal{R}_0 = \beta/\alpha > 1$, only the portion of the (β, α) plane below the line $\gamma = \beta$ is relevant.

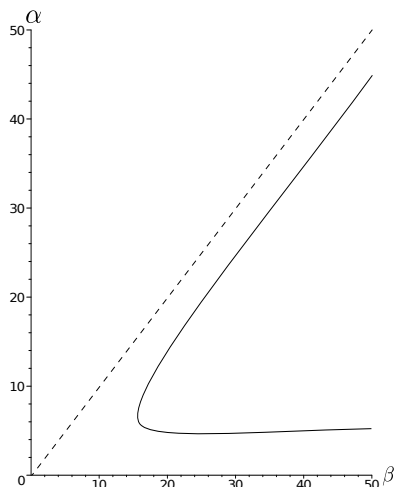


Fig. 10.1 Region of asymptotic stability for endemic equilibria ($\omega = 1$, $N = 1$).

The new feature of the model of this section is that the endemic equilibrium is not asymptotically stable for all parameter values. What is the behavior of the model if the parameters are such that the endemic equilibrium is unstable? A plausible suggestion is that since the loss of stability corresponds to a root $\lambda = iy$ of the characteristic equation, there are solutions of the model behaving like the real part of e^{iyt} , that is, there are periodic solutions. This is exactly what does happen according to a very general result called the Hopf bifurcation theorem, which says that when roots of the characteristic equation cross the imaginary axis a stable periodic orbit arises.

From an epidemiological point of view, periodic behavior is unpleasant. It implies fluctuations in the number of infectives, which makes it difficult to allocate resources for treatment. It is also possible for oscillations to have a long period. This means that if data are measured over only a small time interval the actual behavior may not be displayed. Thus, the identification of situations in which an endemic equilibrium is unstable is an important problem.

Exercises

1. Verify that the basic reproduction number of the SIRS model (10.14) is $\beta N/\alpha$.
2. Verify that the basic reproduction number of the model (10.15) is $\beta N/\alpha$.
3. If we add vaccination that reduces susceptibility by a factor σ at a rate φ to the model (10.14), we obtain a system

$$\begin{aligned} S' &= -\beta SI - \varphi S + \theta R, \\ V' &= \varphi S - \sigma\beta VI, \\ I' &= \beta SI + \sigma\beta VI - \alpha I, \\ R' &= \alpha I - \theta R. \end{aligned} \tag{10.19}$$

Determine the basic reproduction number of the model (10.19) and find all endemic equilibria.

10.5 Diseases as Population Control

Many parts of the world experienced very rapid population growth in the eighteenth century. The population of Europe increased from 118 million in 1700 to 187 million in 1800. In the same time period, the population of Great Britain increased from 5.8 million to 9.15 million, and the population of China increased from 150 million to 313 million [McNeill (1976)]. The population of the English colonies in North America grew much more rapidly than this, aided by substantial immigration from England, but the native population, which had been reduced to one-tenth of their previous size by disease following the early encounters with Europeans and Euro-

pean diseases, grew even more rapidly. While some of these population increases may be explained by improvements in agriculture and food production, it appears that an even more important factor was the decrease in the death rate due to diseases. Disease death rates dropped sharply in the eighteenth century, partly from better understanding of the links between illness and sanitation and partly because the recurring invasions of bubonic plague subsided, perhaps due to reduced susceptibility. One plausible explanation for these population increases is that the bubonic plague invasions served to control the population size, and when this control was removed, the population size increased rapidly.

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

The SIR model with births and deaths of Kermack and McKendrick (1932) includes births in the susceptible class proportional to population size and a natural death rate in each class proportional to the size of the class. Let us analyze a model of this type with birth rate r and a natural death rate $\mu < r$. For simplicity we assume that the disease is fatal to all infectives with disease death rate α , so that there is no removed class and the total population size is $N = S + I$. Our model is

$$\begin{aligned} S' &= r(S+I) - \beta SI - \mu S \\ I' &= \beta SI - (\mu + \alpha)I \end{aligned} \tag{10.20}$$

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

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

which leads to

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

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

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

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

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

$$\begin{aligned} S' &= rS - \beta SI - \mu S = S(r - \mu - \beta I), \\ I' &= \beta SI - (\mu + \alpha)I = I(\beta S - \mu - \alpha), \end{aligned} \quad (10.21)$$

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

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

to obtain the separable first-order differential equation

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

Separation of variables gives

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

Integration gives the relation

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

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

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

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

This model is the same as the simple epidemic model of Section 9.2 except for the birth and death terms, and in many examples the time scale of the disease is

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

Exercises

1. Consider the model (10.21).
 - (i) Show that there are always two equilibria, an extinction equilibrium $(0,0)$ and a coexistence equilibrium with

$$\beta S = \mu + \alpha, \quad \beta I = r - \mu.$$

- (ii) Show that both equilibria are unstable, in fact saddle points.

10.6 Parameter Estimation: Ordinary Least Squares

10.6.1 Connecting Models to Data

This book is concerned primarily with theoretical models for natural phenomena. Such models necessarily contain parameters whose values must be estimated in order to make it possible to compare model predictions with real-life data. The challenges of connecting models and data and the validation of models are critically important in science. In fact, one of the first recorded modeling contributions in the field of epidemiology was that of Daniel Bernoulli (1766), which focused on the increase in the average life expectancy generated by elimination of a lethal disease, a study related directly to data on a single smallpox outbreak, an intellectual contribution that has been analyzed in detail and expanded by Dietz and Heesterbeek (2002).

In 2004 Sally Blower observed that

In the current debate concerning whether the United States population should be vaccinated against smallpox (in order to prepare for a possible terrorist attack) critics of mass vacci-

nation have argued that the risks associated with widespread usage of the current smallpox vaccine would outweigh the potential benefits. Over 200 years ago, opponents of smallpox variolation in the 18th century used the same arguments. They argued that: (i) inoculation was risky because artificial smallpox could cause mortality, and (ii) inoculation programs could increase the transmission of smallpox (because individuals inoculated with artificial smallpox could transmit smallpox). Bernoulli used mathematical reasoning to counter both arguments.

[*ibid*, page 285]. She further observes that as we do today, for example,

Bernoulli conducted a [rudimentary] form of sensitivity analysis by changing assumptions and parameter values and then presenting a series of analyses. In Bernoulli's first series of calculations he assumed that 100% of newborns would be inoculated, that inoculation would induce complete immunity to infection by the wild-type strains of smallpox, and that inoculation would carry no risks. Under these assumptions he calculated that an individual's expectation of life at birth would increase from 26 years 7 months to 29 years 9 months. Bernoulli then repeated his calculations including the assumption that one individual out of every 200 inoculated individuals would die as the result of artificial smallpox.

[*ibid*, 286-287]. Blower concludes that:

It is not clear how influential Bernoulli's paper was in influencing public health policy, but it remains a classic paper as it was the first known mathematical analysis that was used to try to influence public health policy.

[*ibid*, 287].

Efforts to connect models to data increased dramatically with the onset of the HIV epidemic, with emphasis on the estimation of the incubation period distribution for HIV ("back calculation") in efforts to estimate the number of individuals with asymptomatic HIV infections [Anderson et al. (1989), Bailey (1988), Brookmeyer and Gail (1988), Cox and Medley (1989), Hyman and Stanley (1988), Isham (1989), Lagakos et al. (1988), Medley et al. (1987), Wilkie (1989), Castillo-Chavez (1989)]. Additional efforts to connect models to data for the purposes of forecasting and assisting public health decisions gained additional impetus with the emergence of SARS in 2002 [Chowell et al. 2003, Chowell et al. (2004a), Bettencourt et al. (2007, 2006), Gumel et al. (2004)] and the threat of bioterrorism after the tragic events of September 11, 2001 in the United States [Banks and Castillo-Chavez (2003), and Zeng et al. (2011)]. More recently, seasonal and pandemic influenza have been important drivers of the theory giving rise to a large number of articles such as [Nishiura et al. (2009, 2010, 2011)]. Three volumes with a large number of contributions that involve connecting models to data include [Gumel et al. (2006), Chowell et al. (2009), and Castillo-Chavez and Chowell (2011)]. Methods that connect statistical theory and dynamical systems have been refined to improve our ability to connect model and epidemiological data [Ortiz et al. (2011), Sutton et al. (2010a, 2010b, 2008), Cintron-Arias et al. (2008), Shim et al. (2006)]. Next, we outline the "essentials" of the most typical parameter estimation approach, ordinary least squares (OLS) estimation, in the context of the classical $S - I_R$ epidemiological model.

10.6.2 Ordinary Least Squares (OLS) Estimation

The formulation of the standard ordinary least squares formulation usually involves two models: mathematical and statistical. The mathematical model used here is a compartmental mathematical model, which is given in terms of a nonlinear system of ordinary differential equations involving a parameter vector θ . Specifically, we have

$$\frac{dx}{dt} = g(x(t; \theta); \theta), \quad (10.22)$$

where $x(t; \theta) \in \mathbb{R}^n$ denotes the state variable vector at time t and $\theta \in \mathbb{R}^p$ denotes the parameter vector.

The statistical model linked to the process generated by the dynamical system is formulated under the assumption that the model output and associated random deviations (measurement error) are captured by the random variables

$$Y_j = z(t_j; \theta_0) + \mathcal{E}_j \quad \text{for } j = 1, \dots, n, \quad (10.23)$$

where $z(t_j; \theta_0)$ denotes the output of the mathematical model. Usually the model output is a functional of the state variable vector, that is, $z(t; \theta) = \mathcal{F}(x(t; \theta))$, and in equation (10.23) the model output is evaluated at $\theta = \theta_0$, the “true” parameter vector. The random variables \mathcal{E}_j model the random deviations away from $z(t, \theta_0)$ and are assumed to satisfy

1. \mathcal{E}_j 's are independent and identically distributed random variables;
2. $E[\mathcal{E}_j] = 0$ for every j ;
3. $\text{var}(\mathcal{E}_j) = \sigma_0^2 < \infty$, for every j .

The ordinary least squares problem arises from efforts to minimize $[Y_j - z(t_j; \theta)]^2$ over the set of parameter vectors θ constrained by a prespecified feasible region, here denoted by Θ . The minimizer is a random variable, called the estimator θ_{OLS} and given in this context by

$$\theta_{OLS} = \underset{\theta \in \Theta}{\operatorname{argmin}} \sum_{j=1}^n [Y_j - z(t_j; \theta)]^2. \quad (10.24)$$

The usefulness of the estimator θ_{OLS} derives from its statistical properties. A classical asymptotic result, in the spirit of a central limit theorem, establishes that for n sufficiently large, this estimator has a p -multivariate normal sampling distribution. In other words,

$$\theta_{OLS} \sim \mathcal{N}_p(\theta_0, \Sigma_0),$$

where $\Sigma_0 = n^{-1} \sigma_0^2 \Omega_0^{-1}$ and $\Omega_0 = \lim_{n \rightarrow \infty} \frac{1}{n} \chi(\theta_0)^T \chi(\theta_0)$, provided this limit exists and the matrix Ω_0 is nonsingular. The $p \times p$ matrix Σ_0 is the covariance matrix of the θ_{OLS} estimator, while the matrix $\chi(\theta_0)$ is $n \times p$ and is called the sensitivity matrix, with its (i, j) entry defined by

$$\chi_{ij}(\theta) = \frac{\partial z(t_i; \theta)}{\partial \theta_j}, \quad 1 \leq i \leq n, \quad 1 \leq j \leq p.$$

[Banks and Tran (2009)]. The theoretical quantities θ_0 , σ_0^2 , and Σ_0 are in general unknown. In practice, one has only the data associated with a single realization y_i of the observation process Y_i (for $i = 1, \dots, n$) and has no option but to solve the minimization problem, that is, the computation of an estimate for $\hat{\theta}_{OLS}$ under these conditions. That is, we carry out the following minimization process:

$$\hat{\theta}_{OLS} = \operatorname{argmin}_{\theta \in \Theta} \sum_{j=1}^n [y_j - z(t_j; \theta)]^2.$$

The estimate obtained is then used to approximate the error assuming constant variance σ_0^2 via the approximation

$$\sigma_0^2 \approx \hat{\sigma}_{OLS}^2 = \frac{1}{n-p} \sum_{j=1}^n [y_j - z(t_j; \hat{\theta}_{OLS})]^2.$$

The covariance matrix Σ_0 can also be approximated using $\hat{\theta}_{OLS}$ and $\hat{\sigma}_{OLS}^2$ as follows:

$$\Sigma_0 \approx \hat{\Sigma}_{OLS} = \hat{\sigma}_{OLS}^2 [\chi(\hat{\theta}_{OLS})^T \chi(\hat{\theta}_{OLS})]^{-1}.$$

We proceed to apply the above OLS methodology to a well-known compartmental model and an influenza data set. In 1978, an outbreak of influenza was reported in a boarding school for boys in the United Kingdom [Communicable Disease Surveillance Centre (1978)]. The single outbreak is modeled using the classical SIR compartmental model given by the following set of nonlinear differential equations:

$$\begin{aligned} \frac{dS}{dt} &= -\beta S \frac{I}{N}, \\ \frac{dI}{dt} &= \beta S \frac{I}{N} - \gamma I, \\ \frac{dR}{dt} &= \gamma I, \\ N &= S(t) + I(t) + R(t), \\ S(0) &= S_0, \\ I(0) &= I_0. \end{aligned}$$

The dataset reported in [Communicable Disease Surveillance Centre (1978)] corresponds to the prevalence of influenza, and therefore, the output in this case is modeled by

$$z(t; \theta) = I(t; \theta),$$

from which the statistical model for the observation process becomes

$$Y_j = I(t_j; \theta_0) + \mathcal{E}_j \quad \text{for } j = 1, \dots, n. \quad (10.25)$$

In this case, it was reported that $N = 763$, and so the conditions at the start of this outbreak can be assumed to be $S_0 = 762$ and $I_0 = 1$. The parameter vector to be estimated involves the transmission coefficient β and the recovery rate γ , that is, $\theta = (\beta, \gamma)$. The data set in [Communicable Disease Surveillance Centre (1978)] is denoted by y_i for $i = 1, \dots, 12$, and it is used to compute

$$\hat{\theta}_{OLS} = \operatorname{argmin}_{\theta \in \Theta} \sum_{j=1}^{12} [y_j - z(t_j; \theta)]^2.$$

The minimization leads to an estimate of $\theta = (\beta, \gamma)$. The minimization can be carried out in multiple ways. There are in fact several optimization algorithms (for example, Nelder–Mead simplex), and, for example, the computing software Matlab (Mathworks, Inc) includes the following appropriate optimization routines `fminsearch`, `fmincon`, `lsqcurvefit`, `lsqnonlin`.

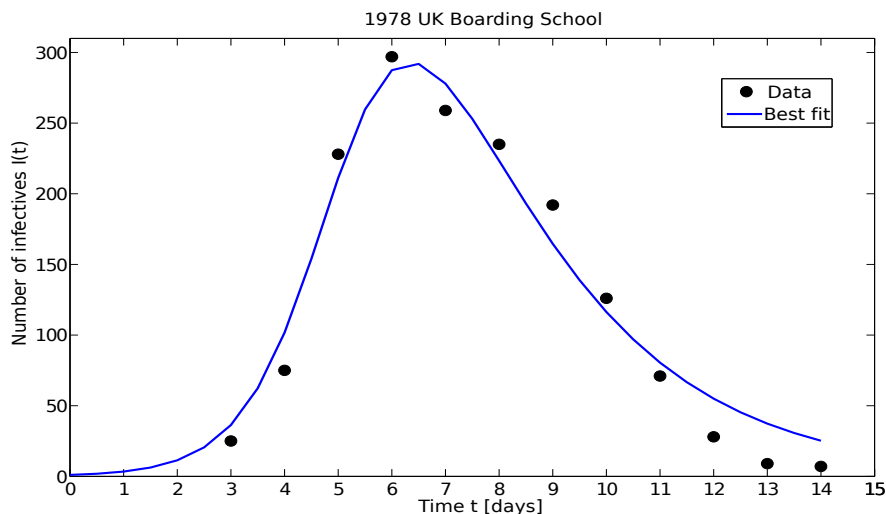


Fig. 10.2 Prevalence of an influenza outbreak in a boarding school for boys, UK [Communicable Disease Surveillance Centre (1978)]. Total population size, $N = 763$, initial number of susceptible, $S_0 = 762$, and initial number of infectives, $I_0 = 1$. Parameter estimates with one standard error: $\hat{\beta} = 1.6682 \pm 0.0294$ (days^{-1}), $\hat{\gamma} = 0.4417 \pm 0.0177$ (days^{-1}).

Figure 10.2 displays the prevalence data (circles) and the best fit solution $I(t; \hat{\theta}_{OLS})$ (solid curve) versus time generated by the solution of the dynamical system and the optimization process. The results look “good”, but the uncertainty in the estimation can be quantified, and here is where statistical theory comes in [Banks and Tran (2009)]. One way of quantifying uncertainty requires the computation of $\hat{\Sigma}_{OLS}$. In this illustration the sensitivities (entries of $\hat{\Sigma}_{OLS}$) are defined as

$$\begin{aligned}\frac{\partial z}{\partial \beta}(t_j; \theta) &= \frac{\partial I}{\partial \beta}(t_j; \theta), \\ \frac{\partial z}{\partial \gamma}(t_j; \theta) &= \frac{\partial I}{\partial \gamma}(t_j; \theta).\end{aligned}$$

Numerical values are computed by solving

$$\begin{aligned}\frac{d}{dt}x(t) &= g(x(t; \hat{\theta}); \hat{\theta}), \\ \frac{d}{dt}\phi(t) &= \frac{\partial g}{\partial x}\phi(t) + \frac{\partial g}{\partial \theta},\end{aligned}$$

where $\phi(t) = \frac{\partial x}{\partial \theta}(t; \theta)$,

$$\frac{\partial g}{\partial x} = \begin{bmatrix} -\hat{\beta} \frac{I}{N} & -\hat{\beta} \frac{S}{N} & 0 \\ \hat{\beta} \frac{I}{N} & \hat{\beta} \frac{S}{N} - \hat{\gamma} & 0 \\ 0 & \hat{\gamma} & 0 \end{bmatrix},$$

and

$$\frac{\partial g}{\partial \theta} = \begin{bmatrix} -S \frac{I}{N} & 0 \\ S \frac{I}{N} & -I \\ 0 & I \end{bmatrix}.$$

In this case (influenza data), the OLS estimation of the covariance matrix is

$$\hat{\Sigma}_{OLS} = \hat{\sigma}^2 [\chi(\hat{\theta}_{OLS})^T \chi(\hat{\theta}_{OLS})]^{-1} = \begin{bmatrix} 0.8615 \times 10^{-3} & 0.1946 \times 10^{-3} \\ 0.1946 \times 10^{-3} & 0.3140 \times 10^{-3} \end{bmatrix},$$

with the standard errors calculated by taking the square root of the diagonal entries in $\hat{\Sigma}_{OLS}$. The estimates, within one standard error, are therefore given by

$$\hat{\beta} = 1.6682 \pm 0.0294 \text{ (days}^{-1}\text{)}$$

and

$$\hat{\gamma} = 0.4417 \pm 0.0177 \text{ (days}^{-1}\text{)}.$$

Estimation can go further. For example, it is often important to estimate the effective reproductive number $\mathcal{R}(t)$, defined by the expression

$$\mathcal{R}(t) \equiv \mathcal{R}(t; \theta) = \frac{\beta}{\gamma} \frac{S(t; \theta)}{N}.$$

In order to estimate $\mathcal{R}(t)$, we rewrite the equation for the infective population (10.25) as follows:

$$\frac{dI}{dt} = \gamma(\mathcal{R}(t) - 1)I.$$

Here, $\mathcal{R}(t) = \frac{S}{N}\mathcal{R}_0$ is the effective reproduction number and $\mathcal{R}_0 = \beta/\gamma$ is the basic reproductive number). We have $\mathcal{R}(t) \leq \mathcal{R}_0$, with the upper bound, the basic reproductive number, being achieved only when the entire population is susceptible.

Days	Cases
3	25
4	75
5	228
6	297
7	259
8	235
9	192
10	126
11	71
12	28
13	9
14	7

Table 10.1 Influenza prevalence, 1978 UK Boarding School.

Here, $\mathcal{R}(t)$ is defined as the product of the transmission rate and the average infectious period, that is, the effective reproductive number gives the average number of secondary infections caused by a single infective individual, at a given susceptible fraction. The prevalence of infection increases or decreases according to whether $\mathcal{R}_0(t)$ is greater than or less than one. Since there is no replenishment of the susceptible population in this model, $\mathcal{R}_0(t)$ decreases over the course of an outbreak as susceptible individuals become infected.

Using $\hat{\theta}_{OLS}$ we can compute the relevant point-estimate (without uncertainty bounds) curve given by

$$\mathcal{R}(t; \hat{\theta}_{OLS}) = \frac{\hat{\beta}}{\hat{\gamma}} \frac{S(t; \hat{\theta}_{OLS})}{N}.$$

Figure 10.3 displays $\mathcal{R}(t; \hat{\theta}_{OLS})$ versus time t .

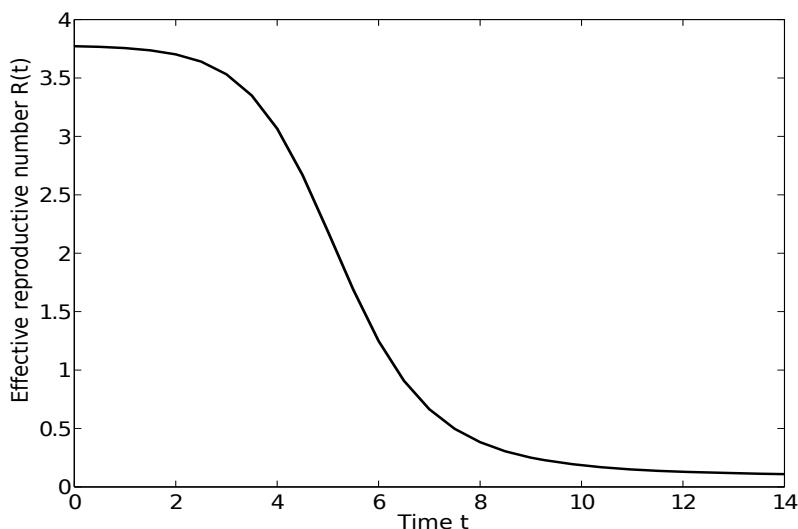


Fig. 10.3 Effective reproductive number $\mathcal{R}(t; \hat{\theta}_{OLS})$ versus time t . This curve is a point-estimate curve calculated using the estimate $\hat{\theta}_{OLS} = (\hat{\beta}, \hat{\gamma})$.

10.7 Possible Extensions

As we have seen, in a large variety of models, the behavior when $\mathcal{R}_0 < 1$ is different from the behavior when $\mathcal{R}_0 > 1$. More precisely, as \mathcal{R}_0 increases through 1 there is an exchange of stability between the disease-free equilibrium and the endemic equilibrium (which is negative as well as unstable and thus biologically meaningless if $\mathcal{R}_0 < 1$). There is a bifurcation, or change in equilibrium behavior, at $\mathcal{R}_0 = 1$, but the equilibrium infective population size depends continuously on \mathcal{R}_0 . Such a transition is called a forward, or transcritical, bifurcation. It is also possible, as we have seen in an SIRS model, that the endemic equilibrium for $\mathcal{R}_0 > 1$ may be unstable, depending on the distribution of infective periods.

However, it would be a serious mistake to assume that this normal situation is universal. It has been noted [Dushoff, Huang and Castillo-Chavez(1998), Haderler and Castillo-Chavez (1995), Haderler and van den Driessche (1997), Kribs-Zaleta and Velasco-Hernandez (2000)] that in epidemic models with multiple groups and asymmetry between groups or multiple interaction mechanisms, it is possible to have a very different bifurcation behavior at $\mathcal{R}_0 = 1$. There may be multiple positive endemic equilibria for values of $\mathcal{R}_0 < 1$. Typically, there is an interval of values of \mathcal{R}_0 on which there are two asymptotically stable equilibria, one disease-free and one endemic, and an unstable endemic equilibrium between them. Such a situation is called a backward bifurcation at $\mathcal{R}_0 = 1$.

The qualitative behavior of an epidemic system with a backward bifurcation differs from that of a system with a forward bifurcation in at least three important ways.

If there is a forward bifurcation at $\mathcal{R}_0 = 1$, it is not possible for a disease to invade a population if $\mathcal{R}_0 < 1$, because the system will return to the disease-free equilibrium $I = 0$ if some infectives are introduced into the population. On the other hand, if there is a backward bifurcation at $\mathcal{R}_0 = 1$ and enough infectives are introduced into the population to put the initial state of the system above the unstable endemic equilibrium with $\mathcal{R}_0 < 1$, the system will approach the asymptotically stable endemic equilibrium.

Other differences are observed if the parameters of the system change to produce a change in \mathcal{R}_0 . With a forward bifurcation at $\mathcal{R}_0 = 1$ the equilibrium infective population remains zero so long as $\mathcal{R}_0 < 1$ and then increases continuously as \mathcal{R}_0 increases. With a backward bifurcation at $\mathcal{R}_0 = 1$, the equilibrium infective population size also remains zero so long as $\mathcal{R}_0 < 1$ but then jumps to the positive endemic equilibrium as \mathcal{R}_0 increases through 1. In the other direction, if a disease is being controlled by means that decrease \mathcal{R}_0 , it is sufficient to decrease \mathcal{R}_0 to 1 if there is a forward bifurcation at $\mathcal{R}_0 = 1$, but it is necessary to bring \mathcal{R}_0 well below 1 if there is a backward bifurcation.

We have been assuming homogeneous mixing of members of the population being studied, and this is certainly unrealistically simple. Members of the population may differ, for example, in rate of contact or in location. In the study of sexually transmitted diseases, differences in activity levels are important aspects. Contact rates may be age-dependent, and this would suggest the use of age-structured models. Spatial dependence may take two forms, the local diffusion of members of the population, which would lead to partial differential equations of reaction–diffusion types, and travel between communities, which would lead to patch or metapopulation models.

Models incorporating one or more of these kinds of heterogeneity can be developed and analyzed. Inevitably, their analysis involves more structure, equations, and parameters, as well as more sophisticated mathematical methods.

There are other modes of transmission of communicable diseases that can be described by compartmental models. Some infections can be transmitted *vertically*, for example, from mother to daughter prior to birth [Busenberg and Cooke (1993)]. Another form is transmission by a *vector*. For example, malaria is transmitted back and forth between humans and mosquitoes. Thus an infected mosquito may bite a human and thus infect the human. An uninfected mosquito may bite an infected human and become infected, but infection is not transmitted directly from human to human or from mosquito to mosquito [Ross (1911)]. Sexually transmitted diseases that are transmitted by heterosexual contact are also examples of vector transmission, with male and females playing the roles of the two species.

10.8 Project: Pulse Vaccination

Consider an SIR model $\dot{S} = \mu K - \mu S - \beta SI$. For measles, typical parameter choices are $\mu = 0.02$, $\beta = 1800$, $\alpha = 100$, $K = 1$ (to normalize carrying capacity to 1) [Engbert and Drepper (1984)].

Question 1.

Show that for these parameter choices $R_0 \approx 18$ and to achieve herd immunity would require vaccination of about 95 percent of the susceptible population.

In practice, it is not possible to vaccinate 95 percent of a population because not all members of the population would come to be vaccinated and not all vaccinations are successful. One way to avoid recurring outbreaks of disease is “pulse vaccination” [Agur, Mazor, Anderson, and Danon (1993), Shulgin, Stone, and Agur (1998), Stone, Shulgin, and Agur (2000)]. The basic idea behind pulse vaccination is to vaccinate a given fraction p of the susceptible population at intervals of time T with T (depending on p) chosen to ensure that the number of infectives remains small and approaches zero. In this project we will give two approaches to the calculation of a suitable function $T(p)$.

The first approach depends on the observation that I decreases so long as $S < \Gamma < (\mu + \gamma)/\beta$. We begin by vaccinating $p\Gamma$ members, beginning with $S(0) = (1 - p)\Gamma$. From (8.7),

$$S' = \mu K - \mu S - \beta SI \geq \mu K - \mu S.$$

Then $S(t)$ is greater than the solution of the initial value problem

$$S' = \mu K - \mu S, \quad S(0) = (1 - p)\Gamma.$$

Question 2.

Solve this initial value problem and show that the solution obeys

$$S(t) < \Gamma \text{ for } 0 \leq t < \frac{1}{\mu} \log \frac{K - (1 - p)\Gamma}{K - \Gamma}.$$

Thus a suitable choice for $T(p)$ is

$$T(p) = \frac{1}{\mu} \log \frac{K - (1 - p)\Gamma}{K - \Gamma} = \frac{1}{\mu} \log \left[1 + \frac{p\Gamma}{K - \Gamma} \right].$$

Calculate $T(p)$ for $p = m/10$ ($m = 1, 2, \dots, 10$).

The second approach is more sophisticated. Start with $I = 0$, $S' = \mu K - \mu S$. We let $t_n = nT$ ($n = 0, 1, 2, \dots$) and run the system for $0 \leq t \leq t_1 = T$. Then we let $S_1 = (1 - p)S(t_1)$. We then repeat, i.e., for $t_1 \leq t \leq t_2$, $S(t)$ is the solution of $S' = \mu K - \mu S$, $S(t_1) = S_1$, and $S_2 = (1 - p)S_1$. We obtain a sequence S_n in this way.

Question 3.

Show that

$$S_{n+1} = (1-p)K(1 - e^{-\mu T}) + (1-p)S_n e^{-\mu T}$$

and for $t_n \leq t \leq t_{n+1}$,

$$S(t) = K \left[1 - e^{-\mu(t-t_n)} \right] + S_n e^{-\mu(t-t_n)}.$$

Question 4.

Show that the solution is periodic if

$$S_{n+1} = S_n = S^* \quad (n = 0, 1, 2, \dots)$$

with

$$S^* = K \left[1 - \frac{pe^{\mu T}}{e^{\mu T} - (1-p)} \right]$$

and that the periodic solution is

$$S(t) = \begin{cases} K \left[1 - \frac{pe^{\mu T}}{e^{\mu T} - (1-p)} e^{-\mu(t-t_n)} \right] & \text{for } t_n \leq t \leq t_{n+1}, \\ S^* & \text{for } t = t_{n_1}, \end{cases}$$

$$I(t) = 0.$$

It is possible to show by linearizing about this periodic solution that the periodic solution is asymptotically stable if

$$\frac{1}{T} \int_0^T S(t) dt < \frac{\mu + \xi}{\beta}.$$

If this condition is satisfied, the infective population will remain close to zero.

Question 5.

Show that this stability condition reduces to

$$\frac{K(\mu T - p)(e^{\mu T} - 1) + pK\mu T}{\mu T [e^{\mu} - (1-p)]} < \frac{\mu + \xi}{\beta}.$$

Question 6.

Use a computer algebra system to graph $T(p)$, where T is defined implicitly by

$$\frac{K(\mu T - p)(e^{\mu T} - 1) + pK\mu T}{\mu T [e^{\mu} - (1-p)]} = \frac{\mu + \xi}{\beta}.$$

Compare this expression for T with the one obtained earlier in Question 2 in this project. A larger estimate for a safe value of T would save money by allowing less frequent vaccination pulses.

10.9 Project: A Model with Competing Disease Strains

We model a general discrete-time SIS model with two competing strains in a population with discrete and nonoverlapping generations. This model arises from a particular discretization in time of the corresponding SIS continuous-time stochastic model for two competing strains.

State variables

S_n population of susceptible individuals in generation n

I_n^1 population of infected individuals with strain 1 in generation n

I_n^2 population of infected individuals with strain 2 in generation n

T_n total population in generation n

f recruitment function

Parameters

μ per capita natural death rate

γ_i per capita recovery rate for strain i

α_i per capita infection rate for strain i

Construction of the model equations: The model assumes that (i) the disease is not fatal; (ii) all recruits are susceptible and the recruitment function depends only on T_n ; (iii) there are no coinfections; (iv) death, infections, and recoveries are modeled as Poisson processes with rates μ, α_i, γ_i ($i = 1, 2$); (v) the time step is measured in generations; (vi) the populations change only because of “births” (given by the recruitment function), deaths, recovery, and infection of a susceptible individual for each strain; (vii) individuals recover but do not develop permanent or temporary immunity, that is, they immediately become susceptible again.

By assumption we have that the probability of k successful encounters is a Poisson distribution, which in general has the form $p(k) = e^{-\beta} \beta^k / k!$, where β is the parameter of the Poisson distribution. In our context, only one success is necessary. Therefore, when there are no successful encounters, the expression $p(0) = e^{-\beta}$ represents the probability that a given event does not occur. For example, the probability that a susceptible individual does not become infective is $\text{Prob}(\text{not being infected by strain } i) = e^{-\alpha_i I_n^i}$, and, $\text{Prob}(\text{not recovering from strain } i) = e^{-\gamma_i I_n^i}$. Hence, $\text{Prob}(\text{not being infected}) = \text{Prob}(\text{not being infected by strain 1}) \text{Prob}(\text{not being infected by strain 2}) = e^{-\alpha_1 I_n^1} e^{-\alpha_2 I_n^2}$.

Now the probability that a susceptible does become infected is given by $1 - e^{-\alpha_i I_n^i}$. Then, $\text{Prob}(\text{infected by strain } i) = \text{Prob}(\text{infected}) \cdot \text{Prob}(\text{infected by strain } i \mid \text{infected}) = (1 - e^{-(\alpha_1 I_n^1 + \alpha_2 I_n^2)}) \frac{\alpha_i I_n^i}{\alpha_1 I_n^1 + \alpha_2 I_n^2}$.

(a) Using the above discussion, show that the dynamics are governed by the system

$$\begin{aligned} S_{n+1} &= f(T_n) + S_n e^{-\mu} e^{-(\alpha_1 I_n^1 + \alpha_2 I_n^2)} + I_n^1 e^{-\mu} (1 - e^{-\gamma_1}) + I_n^2 e^{-\mu} (1 - e^{-\gamma_2}), \\ I_{n+1}^1 &= \frac{\alpha_1 S_n I_n^1}{\alpha_1 I_n^1 + \alpha_2 I_n^2} e^{-\mu} (1 - e^{-(\alpha_1 I_n^1 + \alpha_2 I_n^2)}) + I_n^1 e^{-\mu} e^{-\gamma_1}, \\ I_{n+1}^2 &= \frac{\alpha_2 S_n I_n^2}{\alpha_1 I_n^1 + \alpha_2 I_n^2} e^{-\mu} (1 - e^{-(\alpha_1 I_n^1 + \alpha_2 I_n^2)}) + I_n^2 e^{-\mu} e^{-\gamma_2}. \end{aligned} \quad (10.26)$$

(b) Show that

$$T_{n+1} = f(T_n) + T_n e^{-\mu},$$

where

$$T_n = S_n + I_n^1 + I_n^2. \quad (10.27)$$

This equation is called the *demographic equation*. It describes the total population dynamics.

(c) If we set $I_{n+1}^1 = I_{n+1}^2 = 0$, then model (10.26) reduces to the demographic model

$$T_{n+1} = f(T_n) + T_n e^{-\mu}.$$

Check that this is the case.

(d) Study the disease dynamics at a demographic equilibrium, that is, at a point where $T_\infty = T_\infty e^{-\mu} + f(T_\infty)$. Substitute $S_n = T_\infty - I_n^1 - I_n^2$ where T_∞ is a stable demographic equilibrium, that is, assume $T_0 = T_\infty$ to get the following equations:

$$\begin{aligned} I_{n+1}^1 &= \frac{\alpha_1 I_n^1}{\alpha_1 I_n^1 + \alpha_2 I_n^2} (T_\infty - I_n^1 - I_n^2) e^{-\mu} \left(1 - e^{-(\alpha_1 I_n^1 + \alpha_2 I_n^2)} \right) + I_n^1 e^{-\mu} e^{-\gamma_1}, \\ I_{n+1}^2 &= \frac{\alpha_2 I_n^2}{\alpha_1 I_n^1 + \alpha_2 I_n^2} (T_\infty - I_n^1 - I_n^2) e^{-\mu} \left(1 - e^{-(\alpha_1 I_n^1 + \alpha_2 I_n^2)} \right) + I_n^2 e^{-\mu} e^{-\gamma_2}. \end{aligned} \quad (10.28)$$

System (10.28) describes the dynamics of a population infected with the two strains at a demographic equilibrium.

Show that in system (10.28), if $R_1 = \frac{e^{-\mu} T_\infty \alpha_1}{1 - e^{-(\mu + \gamma_1)}} < 1$ and $R_2 = \frac{e^{-\mu} T_\infty \alpha_2}{1 - e^{-(\mu + \gamma_2)}} < 1$, then the equilibrium point $(0, 0)$ is asymptotically stable.

(e) Interpret biologically the numbers R_i , $i = 1, 2$.

(f) Consider $f(T_n) = \Lambda$, where Λ is a constant. Show that

$$T_{n+1} = \Lambda + T_n e^{-\mu}$$

and that

$$T_{\infty} = \frac{\Lambda}{1 - e^{-\mu}}.$$

- (g) Consider $f(T_n) = rT_n(1 - T_n)/k$, and show that in this case the total population dynamic is given by,

$$T_{n+1} = rT_n \left(1 - \frac{T_n}{k}\right) + T_n e^{-\mu}$$

and that the fixed points are

$$T_n^* = 0, \quad T_n^{**} = \frac{k(r + e^{-\mu} - 1)}{r},$$

whenever $r + e^{-\mu} > 1$.

- (h) Assume that one of the strains is missing, that is, let $I_n^i = 0$ for either $i = 1$ or 2 . Equation (10.28) reduces to

$$I_{n+1} = (T_{\infty} - I_n)e^{-\mu}(1 - e^{\alpha_1 I_n}) + I_n e^{-(\mu+\gamma)}.$$

Establish necessary and sufficient conditions for the stability and/or instability of boundary equilibria for the system (10.28). Compare your results with simulations of the system (10.28) and of the full system (10.26).

- (i) Does the system (10.28) have endemic ($I_1^* > 0, I_2^* > 0$) equilibria?
(j) Simulate the full system (10.26) when the demographic equation is in the period doubling regime. What are your conclusions?

References: Perez-Velazquez (2000) and Castillo-Chavez, Huang, and Li (1996a, 1997). (2000).

10.10 Project: An Epidemic Model in Two Patches

Consider the following SIS model with dispersion between two patches, Patch 1 and Patch 2, where in Patch $i \in \{1, 2\}$ at generation t , $S_i(t)$ denotes the population of susceptible individuals; $I_i(t)$ denotes the population of infecteds assumed infectious; $T_i(t) \equiv S_i(t) + I_i(t)$ denotes the total population size. The constant dispersion coefficients D_S and D_I measure the probability of dispersion by the susceptible and infective individuals, respectively. Observe that we are using a different notation from what we have used elsewhere, writing variables as a function of t rather than using a subscript for the independent variable in order to avoid needing double subscripts:

$$\begin{aligned} S_1(t+1) &= (1 - D_S)\tilde{S}_1(t) + D_S\tilde{S}_2(t), \\ I_1(t+1) &= (1 - D_I)\tilde{I}_1(t) + D_I\tilde{I}_2(t), \\ S_2(t+1) &= D_S\tilde{S}_1(t) + (1 - D_S)\tilde{S}_2(t), \\ I_2(t+1) &= D_I\tilde{I}_1(t) + (1 - D_I)\tilde{I}_2(t), \end{aligned}$$

where

$$\begin{aligned}\tilde{S}_i(t) &= f_i(T_i(t)) + \gamma_i S_i(t) \exp\left(\frac{-\alpha_i I_i(t)}{T_i(t)}\right) + \gamma_i I_i(t)(1 - \sigma_i), \\ \tilde{I}_i(t) &= \gamma_i(1 - \exp\left(\frac{-\alpha_i I_i(t)}{T_i(t)}\right))S_i(t) + \gamma_i \sigma_i I_i(t)\end{aligned}$$

and

$$0 \leq \gamma_i, \sigma_i, \alpha_i, D_S, D_I \leq 1.$$

Let

$$f_i(T_i(t)) = T_i(t) \exp(r_i - T_i(t)),$$

where r_i is a positive constant.

(a) Using computer explorations, determine whether it is possible to have a globally stable disease-free equilibrium on a patch (without dispersal) where the full system with dispersal has a stable endemic equilibrium. Do you have a conjecture?

(b) Using computer explorations determine whether it is possible to have a globally stable endemic equilibrium on a patch (without dispersal) where the full system with dispersal has a stable disease-free equilibrium. Do you have a conjecture?

References: Gonzalez, Sanchez and Saenz (2000), Arreola, Crossa, and Velasco (2000), Castillo-Chavez and Yakubu (2000a, 2000c).

10.11 Project: Population Growth and Epidemics

When one tries to fit epidemiological data over a long time interval to a model, it is necessary to include births and deaths in the population. Throughout the book we have considered population models with birth and death rates that are constant in time. However, population growth often may be fit better by assuming a linear population model with a time-dependent growth rate, even though this does not have a model-based interpretation. There could be many reasons for variations in birth and death rates; we could not quantify the variations even if we knew all of the reasons. Let $r(t) = \frac{dN}{dt}/N$ denote the time-dependent per capita growth rate. To estimate $r(t)$ from linear interpolation of census data, proceed as follows:

1. Let N_i and N_{i+1} be the consecutive census measurements of population size taken at times t_i and t_{i+1} respectively. Let $\Delta N = N_{i+1} - N_i$, $\Delta t = t_{i+1} - t_i$, and $\delta N = N(t + \delta t) - N(t)$.
2. If $t_i \leq t \leq t_{i+1}$, $\frac{\Delta N}{\Delta t} = \frac{\delta N}{\delta t}$, then we make the estimate $r(t) \approx \frac{\Delta N}{\Delta t N(t)}$.
3. A better approximation is obtained by replacing $N(t)$ by $N(t + \delta t/2)$. Why? Show that in this case, $r(t) \approx (\frac{\delta t}{2} + \frac{N(t)\Delta t}{\Delta N})^{-1}$.

Question 1.

Use the data of [Table 10.2](#) to estimate the growth rate $r(t)$ for the population of the USA.

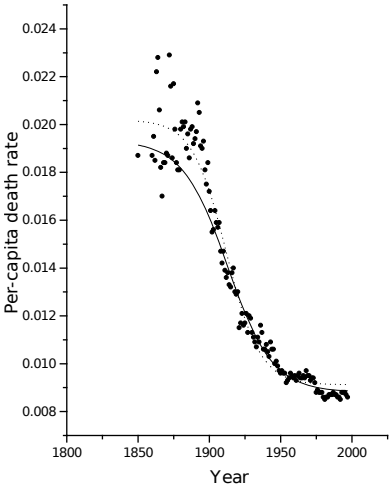


Fig. 10.4 Observed death rate (●) and the best fit obtained with the function (10.29).

1700	250,888	1800	5,308,483	1900	75,994,575
1710	331,711	1810	7,239,881	1910	91,972,266
1720	466,185	1820	9,638,453	1920	105,710,620
1730	629,445	1830	12,866,020	1930	122,775,046
1740	905,563	1840	17,069,453	1940	131,669,275
1750	1,170,760	1850	23,192,876	1950	151,325,798
1760	1,593,625	1860	31,443,321	1960	179,323,175
1770	2,148,076	1870	39,818,449	1970	203,302,031
1780	2,780,369	1880	50,155,783	1980	226,542,199
1790	3,929,214	1890	62,947,714	1990	248,718,301
—	—	—	—	2000	274,634,000

Table 10.2 Population data growth for the USA

Figure 10.4 shows the time evolution of the USA mortality rate. This mortality rate is fit well by

$$\mu = \mu_0 + \frac{\mu_0 - \mu_f}{1 + e^{(t-t'_{1/2})/\Delta'}} \tag{10.29}$$

with $\mu_0 = 0.01948$, $\mu_f = 0.008771$, $t'_{1/2} = 1912$, and $\Delta' = 16.61$. Then the “effective birth rate” $b(t)$ is defined as the real birth rate plus the immigration rate.

Question 2.

Estimate $b(t)$ using $r(t) = b(t) - \mu(t)$, with $r(t)$ found in Question 1.

Consider an SEIR disease transmission model. We assume that:

- (a) An average infective individual produces β new infections per unit of time when all contacts are with susceptibles but that otherwise, this rate is reduced by the ratio S/N .
- (b) Individuals in the exposed class E progress to the infective class at the per capita rate k .
- (c) There is no disease-induced mortality or permanent immunity, and there is a mean infective period of $1/\gamma$.

We define $\gamma = r + \mu$. The model becomes:

$$\begin{aligned}
 \frac{dS}{dt} &= bN - \mu S - \beta S \frac{I}{N}, \\
 \frac{dE}{dt} &= \beta S \frac{I}{N} - (k + \mu)E, \\
 \frac{dI}{dt} &= kE - (r + \mu)I, \\
 \frac{dR}{dt} &= rI - \mu R..
 \end{aligned} \tag{10.30}$$

Question 3.

- (a) Show that the mean number of secondary infections (belonging to the exposed class) produced by one infective individual in a population of susceptibles is $Q_0 = \beta/\gamma$.
- (b) Assuming that k and μ are time-independent, show that R_0 is given by $Q_0 f$, where $f = k/(k + \mu)$. What is the epidemiological interpretation of $Q_0 f$?

The usual measure of the severity of an epidemic is the incidence of infective cases. The incidence of infective cases is defined as the number of new infective individuals per year. If we take one year as the unit of time, the incidence of infective cases is given approximately by kE . The incidence rate of infective cases per 100,000 population is given approximately by $10^5 kE/N$.

Tuberculosis (TB) is an example of a disease with an exposed (noninfective) stage. Infective individuals are called active TB cases. Estimated incidence of active TB in the USA was in a growing phase until around 1900 and then experienced a subsequent decline. The incidence rate of active TB exhibited a declining trend from 1850 (See [Table 10.3](#) and [Figure 10.5](#)). The proportion of exposed individuals who survive the latency period and become infective is $f = \frac{k}{k + \mu}$. The number f will be used as a measure of the risk of developing active TB by exposed individuals.

Question 4.

Assume that mortality varies according to the expression (10.29), and that the value of b found in Question 2 is used. Set $\gamma = 1 \text{ years}^{-1}$ and $\beta = 10 \text{ years}^{-1}$, both constant through time. Simulate TB epidemics starting in 1700 assuming constant values for f . Can you reproduce the observed trends ([Table 10.3](#))?

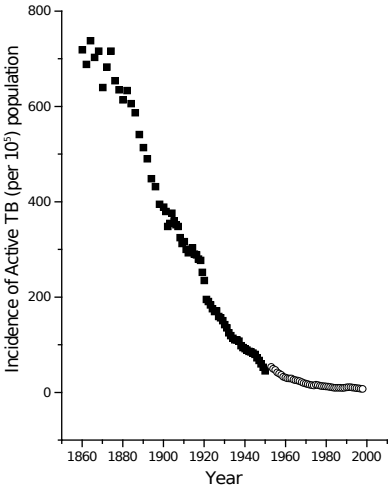


Fig. 10.5 Incidence of active TB

Year	Incidence rate	Incidence	Year	Incidence rate	Incidence
1953	53	84304	1976	15	32105
1954	49.3	79775	1977	13.9	30145
1955	46.9	77368	1978	13.1	28521
1956	41.6	69895	1979	12.6	27769
1957	39.2	67149	1980	12.3	27749
1958	36.5	63534	1981	11.9	27337
1959	32.5	57535	1982	11	25520
1960	30.8	55494	1983	10.2	23846
1961	29.4	53726	1984	9.4	22255
1962	28.7	53315	1985	9.3	22201
1963	28.7	54042	1986	9.4	22768
1964	26.6	50874	1987	9.3	22517
1965	25.3	49016	1988	9.1	22436
1966	24.4	47767	1989	9.5	23495
1967	23.1	45647	1990	10.3	25701
1969	19.4	39120	1992	10.5	26673
1970	18.3	37137	1993	9.8	25287
1971	17.1	35217	1994	9.4	24361
1972	15.8	32882	1995	8.7	22860
1973	14.8	30998	1996	8	21337
1974	14.2	30122	1997	7.4	19885
1975	15.9	33989	1998	6.8	18361

Table 10.3 Reported incidence and incidence rate (per 100,000 population) of active TB.

It is not possible to obtain a good fit of the data of Table 10.3 to the model (10.30). It is necessary to use a refinement of the model that includes time-dependence in the parameters, and the next step is to describe such a model. The risk of progression

to active TB depends strongly on the standard of living. An indirect measure of the standard of living can be obtained from the life expectancy at birth. The observed life expectancy for the USA is approximated well by the sigmoid shape function

$$\tau = \tau_f + \frac{(\tau_0 - \tau_f)}{1 + \exp[(t - t_{1/2})/\Delta]}, \quad (10.31)$$

shown in Figure 10.6. Here τ_0 and τ_f are asymptotic values for life expectancy; $t_{1/2} = 1921.3$ is the time by which life expectancy reaches the value $(\tau_0 + \tau_f)/2$; and $\Delta = 18.445$ determines the width of the sigmoid.

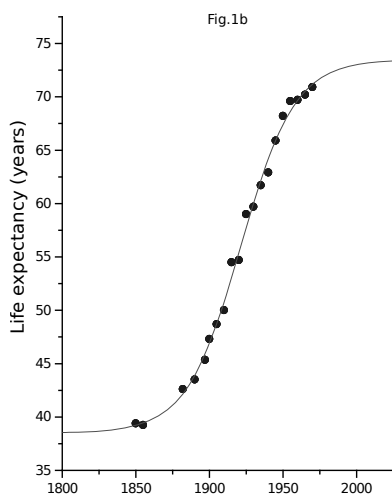


Fig. 10.6 Observed average life expectancy at birth (●) and its best fit (continuous line) using expression (10.31).

Assume that the risk f varies exactly like life expectancy, that is, assume that f is given by

$$f(t) = f_f + \frac{(f_i - f_f)}{1 + \exp[(t - t_{1/2})/\Delta]}. \quad (10.32)$$

We refine the model (10.30) by replacing the parameter k by the variable expression $\mu f(t)/(1 - f(t))$ and $k + \mu$ by $\mu/(1 - f(t))$, obtained from the relation $f = k/(k + \mu)$. Since the time scale of the disease is much faster than the demographic time scale, the recovery rate r is approximately equal to γ . This gives the model

$$\begin{aligned}
\frac{dS}{dt} &= b(t)N - \mu(t)S - \beta S \frac{I}{N}, \\
\frac{dE}{dt} &= \beta S \frac{I}{N} - \frac{\mu(t)}{1-f(t)}E, \\
\frac{dI}{dt} &= \frac{\mu(t)f(t)}{1-f(t)}E - \gamma I, \\
\frac{dR}{dt} &= \gamma I - \mu(t)R..
\end{aligned} \tag{10.33}$$

Question 5.

Simulate TB epidemics starting in 1700 using the model (10.33) with $\gamma = 1 \text{ years}^{-1}$ and $\beta = 10 \text{ years}^{-1}$, both constant, and with $\mu(t)$ given by (10.29) and $f(t)$ given by (10.32). Find values of f_0 and f_f for which an accurate reproduction of the observed TB trends (Table 10.3) is achieved.

References: Aparicio, Capurro, and Castillo-Chavez (2000a, 2000b, 2001a, 2001d); U.S. Bureau of the Census (1975, 1980, 1991, 1999); Castillo-Chavez and Feng (1997, 1998a, 1998b); Feng, Castillo-Chavez, and Capurro (2000); Feng, Huang, and Castillo-Chavez (2001).

10.12 Project: Estimating Parameters for Leishmaniasis

Leishmaniasis is a vector-borne disease caused by a protozoan parasites and transmitted by the bite of certain species of sand fly (referred as “vector”). Leishmaniasis is found in many tropical and subtropical countries. The most serious and potentially fatal, if left untreated, form of leishmaniasis is a “visceral” form. The Indian state of Bihar is one of the major foci of visceral leishmaniasis (VL) in the world.

Two mathematical models of the spread of VL in Bihar are shown in Figure 10.7. The models (variables defined in Table 10.12) incorporate the possibility that individuals seek treatment at private (T) or public (G) health facilities. The treatment of individuals at private health facilities results in *underreporting* of cases and deaths in the state, since private practitioners are not required by law to report cases and deaths. Reported-incidence and -mortality data for 2003 can be used to obtain estimates of model-dependent underreporting levels ($1 - p$). Moreover, the transmission coefficients (β in Model I and λ_h and λ_v in Model II) are also unknown for this disease.

Model I can be written as

$$\begin{aligned}
S'(t) &= \mu N - \beta S(t) \frac{I(t)}{N} - \mu S(t), \\
I'(t) &= \beta S(t) \frac{I(t)}{N} - (\eta + \mu) I(t), \\
G'(t) &= p\eta I(t) - (\alpha_1 + \mu) G(t), \\
T'(t) &= (1-p)\eta I(t) - (\alpha_2 + \mu) T(t), \\
R'(t) &= \alpha_1 G(t) + \alpha_2 T(t) - \mu R(t).
\end{aligned}$$

The reproduction number for Model I is

$$\mathcal{R}_c^I = \frac{\beta}{\mu + \eta}.$$

Aim 1: Estimate parameter(s) p and β . The incidence-underreporting level is then given by $(1-p) \times 100\%$.

Aim 2: Use these estimated parameters to estimate the *controlled* (since treatment modifies the infectious period) reproduction number (R_c).

In Model II, the rate of infection for susceptible humans is modeled by $F_h(t) = \lambda_h \frac{Z(t)}{N_v(t)}$ with $\lambda_h \equiv mC\beta_{hv}$. Here, m is the per capita average number of sand flies (assumed constant), C is the mean rate of bites per sand fly, β_{hv} is the transmission “probability” per bite from an infectious sand fly, and $\frac{Z}{N_v}$ is the proportion of infectious sand flies in the vector population. The proportion of bites of susceptible sand flies on infectious humans is modeled by $\frac{I}{N_h}$, that is, sand flies bite the host population at random. The infection rate of susceptible sand flies is $F_v(t) = \lambda_v \frac{I(t)}{N_h(t)}$ with $\lambda_v = C\beta_{vh}$ where β_{vh} is the transmission “probability” per infectious human bite by a susceptible sand fly.

Model II is given by

$$\begin{aligned}
S'(t) &= \Lambda - \lambda_h S(t) \frac{Z(t)}{N_v(t)} - \mu_h S(t), \quad X'(t) = \mu_v N_v - \lambda_v X(t) \frac{I_h(t)}{N_h(t)} - \mu_v X(t), \\
I'(t) &= \lambda_h S(t) \frac{Z(t)}{N_v(t)} - (\delta_1 + \eta + \mu_h) I(t), \quad Z'(t) = \lambda_v X(t) \frac{I_h(t)}{N_h(t)} - \mu_v Z(t), \\
G'(t) &= p\eta I(t) - (\delta_2 + \alpha_1 + \mu_h) G(t), \\
T'(t) &= (1-p)\eta I(t) - (\delta_2 + \alpha_2 + \mu_h) T(t), \\
R'(t) &= \alpha_1 G(t) + \alpha_2 T(t) - \mu_h R(t),
\end{aligned}$$

In Model II, the rate of infection for susceptible humans can be modeled by $F_h(t) = \lambda_h \frac{Z(t)}{N_v(t)}$ (see Table 10.12 for definitions) with $\lambda_h \equiv mC\beta_{hv}$. Here, m is the per capita average number of sand flies (assumed constant), C is the mean rate of bites per sand fly, β_{hv} is the transmission “probability” per bite from an infectious sand fly, and $\frac{Z}{N_v}$ is the proportion of infectious sand flies in the vector population. The proportion of bites of susceptible sand flies on infectious humans is modeled by $\frac{I}{N_h}$, that is, sand flies bite the host population at random. The infection rate of susceptible sand flies is $F_v(t) = \lambda_v \frac{I(t)}{N_h(t)}$ with $\lambda_v = C\beta_{vh}$ where β_{vh} is the transmis-

sion “probability” per infectious human bite by a susceptible sand fly. Even though we consider a systematic sub-representation of λ_h and λ_v , we do not estimate their underrepresentation parameters. That is, we estimate transmission coefficients λ_h and λ_v as a lumped parameters and do not estimate m , C β_{hv} and β_{vh} explicitly. The reproduction number for Model II is

$$\mathcal{R}_c^{II} = \sqrt{\left(\frac{\lambda_v}{\mu_h + \delta_1 + \eta}\right) \left(\frac{\lambda_h}{\mu_v}\right)}.$$

Estimates of the *controlled* (since treatment modifies the infectious period) reproduction number (\mathcal{R}_c) and the proportion of reported cases (p) are generated from the models. *Berkeley Madonna* or *Matlab* (built-in routine `lsqcurvefit`) curve-fitting tools can be used to estimate model parameters. As an example, parameters estimates for Model I are shown in Table 10.12. The “best” fit of the model to the cumulative number of reported cases (official data; Table 10.12) using a least squares fitting procedure is used to generate the unknown parameter estimates. However, for Model II, multiple data (cases and deaths) can be used to estimate parameters.

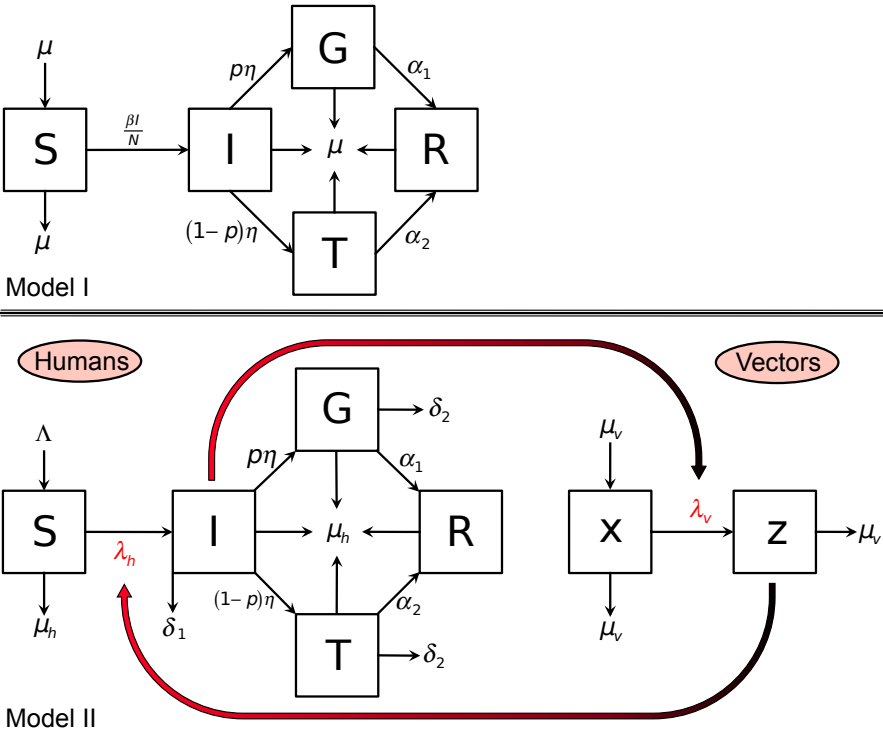


Fig. 10.7 Flow chart: Arrows in Model II indicate interactions between vector (sand fly) and host (human) populations.

Variable	Definition
S	Number of human susceptibles
I	Number of infected humans
G	Number of humans undergoing treatment at public health facilities
G	Number of humans undergoing treatment at private health facilities
R	Number of recovered individuals
N_h	Total population of humans
X	Number of susceptible vectors
Z	Number of infected vectors
N_v	Total vector population size

Table 10.4 Model variables.

Para.	Definition	Point estimate	Ref.
β	Transmission coefficient	0.002 /month	Estimated
p	Proportion of infected using public health clinics	0.24	Estimated
μ	Natural mortality	0.00138 / month	Census, 2001
η	Treatment per capita rate	0.25 / month	Sud et al., 2004
α_1	Per capita recovery rate for G class individuals	1.32 / month	Mubayi et al., 2010
α_2	Per capita recovery rate for T class individuals	0.65 / month	Mubayi et al., 2010
Λ	Human susceptibles recruitment rate	7224 people/month	Mubayi et al., 2010
δ_1	Disease related mortality in class I	(38.5%)	Zerpa et al., 2003
δ_2	Disease related mortality in treatment (G and T) classes	(10%)	Bora, (1999)
$\mu_h (\mu_v)$	Natural mortality per capita rate in humans (vectors)	0.00138 (2.13) /month	Mubayi et al., (2010)
$\lambda_h (\lambda_v)$	Transmission coefficients for humans (vectors)	2.1 (1.5) / month	Estimated

Table 10.5 Parameters of the models.

In the models, the cumulative reported cases from time t_0 to t are given by

$$C(t) = \int_{t_0}^t p\eta I(\tau)d\tau \approx C(t_k) = \sum_{n=1}^k p\eta I_n$$

where $k \in \{1, 2, \dots, 12\}$, representing 12 months. Corresponding estimates of $C(t_k)$ (i.e., \tilde{C}_k) can be found from the data given in Table 10.12. Hence, the estimation problem, for example in the case of Model I, is then to find optimal values ($\hat{\beta}$ and \hat{p}) of β and p such that

$$(\hat{\beta}, \hat{p}) = \min_{(\beta, p)} \sum_{k=1}^{12} [C(t_k) - \tilde{C}_k]^2.$$

Time	Incidence rate per 10,000 individuals ^a	VL mortality rate per 10,000 infected ^b
Jan-03	0.0638	0.0008
Feb-03	0.1022	0.0007
Mar-03	0.1233	0.0031
Apr-03	0.1595	0.0009
May-03	0.1054	0.0020
Jun-03	0.1235	0.0022
Jul-03	0.1585	0.0013
Aug-03	0.2028	0.0034
Sep-03	0.1567	0.0032
Oct-03	0.1460	0.0021
Nov-03	0.1289	0.0009
Dec-03	0.1453	0.0010

Table 10.6 2003 Visceral leishmanias reported data from Indian state of Bihar.

^aThis rate is computed by dividing the number of new cases occurring in a particular month by the 2003 population of Bihar (2003 population of Bihar was 86,396,255) and then multiplying the resultant number by 10,000.

^bThis rate is computed by dividing the number of deaths occurring in a particular month by the 2003 population of Bihar (2003 population of Bihar was 86,396,255), and then multiplying the resultant number by 10,000.

Similarly, if mortality data ($\tilde{D}_k, k \in \{1, 2, \dots, 12\}$) are to be used then the cumulative number of deaths from the model is given by

$$D(t) = \int_{t_0}^t \delta_2 G(\tau) d\tau \approx D(t_k) = \sum_{n=1}^k \delta_2 G_n.$$

Assignment:

1. Use Model II and incidence data in Table 10.12 to estimate $\lambda_h, \lambda_v, \delta_1$, and p .
2. Compute underreporting incidence levels and estimate R_c^{II} using model estimates from step (1).
3. Repeat the first two steps using both data sets (i.e., incidence and mortality data sets in Table 10.12).

References: Bora (1999), Census of India (2001), Desjeux (2004), Mubayi et al (2010), Murray et al.(2005), Sud (2004), Zerpa (2003).

10.13 Project: Invasive Pneumococcal Disease Surveillance Data

We consider for this example invasive diseases (in contrast to respiratory infections including the common ear infection) caused by the bacterium *Streptococcus pneumoniae*. These most notably include pneumonia, meningitis, and bacteremia. *S. pneumoniae*, or the pneumococcus, is commonly part of the normal flora of healthy individuals, and is spread through casual contacts via respiratory droplets. Colonization of the nasopharyngeal region is asymptomatic and typically reversed in a

couple of weeks (it lasts longer in older, younger, and other immunocompromised individuals). A successful infection occurs only in the event that pneumococci are able to spread to and colonize another part of the body, such as the lungs, ear, or blood, etc.

While there are some vaccines routinely used against these infections, it is a lively research area with many open questions. Of the vaccines available, the older one protect against infection only, leaving the colonization phase unaffected, and targets the 23 most common of the over 90 *serotypes* (analogous to *strains* of a virus). The somewhat inefficient protection it provides to juvenile and older individuals is not permanent, and children, one of the most affected groups, are unresponsive to the vaccine. Another vaccine, targeting seven common serotypes, is very effective against invasive infection, particularly in children, and may provide protection against colonization. However, whether this aspect of protection is beneficial is under investigation, since it may also provide selective pressure for previously uncommon, yet more invasive, serotypes not targeted by the vaccine. A model of these dynamics was developed in [Sutton, Banks, and Castillo-Chavez (2008)], and is discussed here. See this paper and references therein for more information on the epidemiology of these infections and vaccines developed to target them.

A schematic of pneumococcal infection dynamics including vaccination is shown in Figure 10.8. The model equations are given by

$$\begin{aligned}
 \frac{dS}{dt} &= \lambda - \beta S \frac{E + E_V + I + I_V}{N} + \alpha E + \gamma I - \phi S + \rho S_V - \mu S, \\
 \frac{dE}{dt} &= \beta S \frac{E + E_V + I + I_V}{N} - \alpha E - l\kappa(t)E - \phi E + \rho E_V - \mu E, \\
 \frac{dS_V}{dt} &= \phi S - \varepsilon \beta S_V \frac{E + E_V + I + I_V}{N} + \alpha E_V + \gamma I_V - \rho S_V - \mu S_V, \\
 \frac{dE_V}{dt} &= \varepsilon \beta S_V \frac{E + E_V + I + I_V}{N} - \alpha E_V + \phi E - \rho E_V - \delta \kappa(t)E_V - \mu E_V, \\
 \frac{dI}{dt} &= l\kappa(t)E - (\gamma + \eta + \mu)I, \\
 \frac{dI_V}{dt} &= \delta \kappa(t)E_V - (\gamma + \eta + \mu)I_V,
 \end{aligned} \tag{10.34}$$

where the infection rate $\kappa(t)$ is the oscillatory function $\kappa(t) = \kappa_0 \{1 + \kappa_1 \cos[\omega(t - \tau)]\}$.

1. Give interpretations for the parameters ϕ , ρ , ε , and δ .

The simplified model excluding vaccination is given by

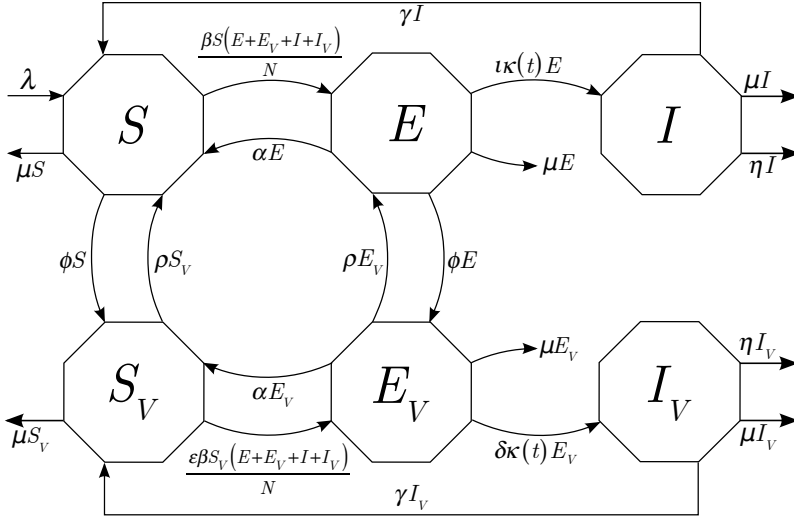


Fig. 10.8 Pneumococcal infection dynamics with vaccination.

$$\begin{aligned}
 \frac{dS}{dt} &= \Lambda - \beta S \frac{E+I}{N} + \alpha E + \gamma I - (\phi + \mu)S, \\
 \frac{dE}{dt} &= \beta S \frac{E+I}{N} - (\alpha E + \kappa(t) + \mu)E, \\
 \frac{dI}{dt} &= \kappa(t)E - (\gamma + \eta + \mu)I.
 \end{aligned} \tag{10.35}$$

2. Show that the reproductive number for the model without vaccination is $\mathcal{R}_0 = \frac{\beta}{\alpha + \kappa + \mu} + \frac{\kappa}{\alpha + \kappa + \mu} \frac{\beta}{\gamma + \eta + \mu}$. Give an interpretation.

The total population in the model is not constant, with case fatality rate $\eta > 0$, as is clear by considering the equation $\frac{dN}{dt} = \lambda - \mu N - \eta(I + I_V)$, where $N = S + E + S_V + E_V + I + I_V$. But in the absence of infection, the total population $N(t)$ approaches $\frac{\lambda}{\mu}$.

3. Rescale the (full) model equations so that the state variables

$$X(t) = (S(t), E(t), S_V(t), E_V(t), I(t), I_V(t))^T$$

now represent proportions of the total population,

$$x(t) = (s(t), e(t), s_v(t), e_v(t), i(t), i_v(t))^T.$$

Hint: $\frac{dx}{dt} \neq \frac{1}{N} \frac{dX}{dt}$, since $N(t) \neq N$ constant.

4. Show that the proportions of population unvaccinated and vaccinated are $\frac{\rho + \mu}{\rho + \mu + \phi}$ and $\frac{\phi}{\rho + \mu + \phi}$, respectively.

Many model parameters are not usually available in typical sources such as census data, as is usually the case with epidemiological models, and further are not directly measurable from surveillance data. However, to study prevention strategies, such as vaccination, in a specific population, we need to know all parameter values. In [Sutton, Banks, and Castillo-Chavez (2008)], the authors used surveillance data collected by the Australian government, available at [National Notifiable Diseases, Roche et al. (2004), Roche et al. (2006), Roche et al. (2007)], and used a least squares approach to estimate pertinent parameters. In the simplest case, if one has only one type of observation d_j at times t_j , where $j \in [1, \dots, n]$, that is representative of model quantities $f(t_j)$, we can obtain estimates to model parameters $\hat{\theta}$ (commonly θ is a vector of all desired parameters) by minimizing the objective functional

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \sum_{j=1}^n (d_j - f(t_j))^2. \quad (10.36)$$

In the above, Θ is the space of all possible values the model parameters may take, or the feasible parameter space. For more information on how the surveillance data were handled with the full model including vaccination, see [Sutton, Banks, and Castillo - Chavez (2008)], and for a discussion of the mathematical and statistical aspects of parameter estimation and model comparison as techniques in the field of inverse problems, see [Banks et al. (2009)].

5. Determine the form of $f(t_j)$ if the data d_j are the number of infections in a population at a given time t_j . What is the form of $f(t_j)$ if the data d_j are reported as the number of cases reported during a given time interval t_{j-1} to t_j ?

Below is a sample Matlab program to estimate parameters $\theta = (\beta, \gamma)^T$ from “data” (data were actually generated from a forward solution of the model run with the true parameters) in which the number of infections during a week is reported for 14 weeks.

The code below fits the sample case notification data using a basic SIR model

$$\begin{aligned} \dot{S} &= -\beta SI, \\ \dot{I} &= \beta SI - \gamma I, \\ \dot{R} &= \gamma I, \end{aligned}$$

using the initial guess $(0.00202, 0.0022)^T$ for parameters $(\beta, \gamma)^T$. The data here are without any noise, or *error*. That is, these data were calculated from solutions of the above SIR model exactly. Of course, there is always some error in actual data, since even a very good model is only an approximation of the epidemiological/physical processes occurring that generate the observations.

```

clear all
tic;

data = [2.4788; 8.4667; 27.4354; 75.6336; 140.4773; 137.2901; 71.5112; 25.5687; 7.8727;
2.3170; 0.6729; 0.1954; 0.0567; 0.0165]; % enter data

n = numel(data);
t_wndow = (0:n)'; %n+1 time points
IC = [500, 1, 0, 0];

%% setting initial model parameter values and ranges %%%
theta0 = [0.0018; 0.0022];
lb = [0.0005; 0.0005];
ub = [0.01; 0.01];
theta_hat = theta0;
%anonymous function to be able to pass additional arguments to the
%objective functional
anonym = @(theta_hat,t_wndow)obj_fcn(theta_hat,t_wndow,n,IC);

options=optimset('Display','off','TolFun',1e-24,'TolX',1e-24,'LargeScale','on');

% calculating estimated parameters theta_hat
[ theta_hat ]= lsqcurvefit(anonym,theta_hat,t_wndow,data,lb,ub,options);
%model solution with estimated parameters
[t z_h]=ode45(@simple_SIR,t_wndow,IC,[],theta_hat);
% calculating new infections as predicted by the 'fitted' model
% solution with estimated parameters
model = z_h(2:n+1,4)-z_h(1:n,4);
t_data = 1:numel(data);
figure(1);
plot(t_data,model,t_data,data,'.')
title('Model fit to weekly cases reported')
xlabel('t (weeks)')
ylabel('Cases')

fprintf('The estimated parameters are beta = %d and gamma = %d',theta_hat(1),theta_hat(2))
function f=simple_SIR(t,y,theta)
f=zeros(4,1);
f(1)=-theta(1)*y(1)*y(2);
f(2)=theta(1)*y(1)*y(2) - theta(2)*y(2);
f(3)=theta(2)*y(2); f(4) = theta(1)*y(1)*y(2);
function F=obj_fcn(p,t,n,IC)
[t,z]=ode45(@simple_SIR,t,IC,[],p);
F = zeros(n,1); F(1:n) = z(2:n+1,4)-z(1:n,4);

```

6. After running the sample code to estimate the parameter values from the “noiseless” data, use it to estimate the same parameters, starting from the same initial guess $(0.00202, 0.0022)^T$, using the data sets below with an increasing amount of error (1%, 5%, 10%) in the observations. Interpret your results for the parameter estimates as the observational error in the data increases. It may be that one parameter is more reliably estimated from these data than the other. If that

is the case, suggest what type of data may be more informative for that parameter.

Case notification data to be used to estimate parameters using a simple SIR model, with noise added as denoted in the column headings below:

1% error	5% error	10% error
2.7341	3.0036	3.6188
8.3935	7.0609	9.5840
27.3911	29.0218	24.3466
76.1656	73.5853	75.5263
140.9805	138.5686	139.8885
137.7963	135.8446	139.5319
71.7510	66.2536	75.4156
25.1375	28.1372	29.5304
8.1290	8.4535	4.7884
2.8993	0.9690	2.5933
0.8475	3.1198	0
0.5649	0	0
0.3163	0	0.0322
0	0	5.4901

7. Adapt the sample code to estimate parameters $\theta = (\beta, \kappa_0, \kappa_1)^T$ with the two sets of monthly case notifications:

Jan	95	104	Jan	113	169
Feb	93	127	Feb	86	99
Mar	120	77	Mar	85	84
Apr	167	183	Apr	112	125
May	221	227	May	158	154
Jun	267	242	Jun	211	209
Jul	293	284	Jul	258	286
Aug	292	297	Aug	284	311
Sep	291	329	Sep	284	311
Oct	262	329	Oct	257	270
Nov	214	266	Nov	212	189
Dec	160	134	Dec	158	172

Use the full model (including vaccination) for pneumococcal disease dynamics. The other parameter values and initial conditions needed are given in the table below.

α	1.3555	$S(0)$	16,576,528
ω	$\frac{\pi}{6}$	$E(0)$	2,453,326
τ	8	$S_V(0)$	435,002
l	0.05	$E_V(0)$	64,380
γ	1	$I(0)$	33
η	0.0787	$I_V(0)$	5
λ	25,000		
μ	0.0003		
ϕ	0.0007579		
ρ	0.02741		
ε	0.953		
δ	0.53551		

8. In 2005, the studied population was prescribed a new vaccine nationally, the long-term effects of which were not clear at the time. Vaccine companies publish estimates of the efficacy of released vaccines, but these may vary depending on the ability of the immunized population to mount a response to them. The effect of this vaccine on the colonization susceptibility was unclear. Estimate the vaccine efficacies ε, δ from the data below, assuming unchanged loss of protection ρ and an increased vaccination rate of $\phi = 0.007984$. Data of type d^1 are monthly reported cases of (total) infections, and type d^2 are vaccinated cases. Use the state variable values from the end (last time point) of the “best fit” solution to the answer to the previous question as initial conditions.

	d^1	d^1		d^2	d^2
Jan	68	134	Jan	11	61
Feb	107	119	Feb	24	52
Mar	35	180	Mar	30	101
Apr	159	230	Apr	34	105
May	206	319	May	65	174
Jun	226	437	Jun	79	220
Jul	291	486	Jul	122	259
Aug	326	494	Aug	112	268
Sep	389	441	Sep	114	237
Oct	326	325	Oct	104	209
Nov	169	287	Nov	58	159
Dec	222	233	Dec	88	125

9. The reproductive number for the full model with vaccination \mathcal{R}_ϕ is given by

$$\begin{aligned} \mathcal{R}_\phi = & \frac{\beta s^0(\rho + \alpha + \delta\kappa + \mu + \phi) + \varepsilon\beta s_v^0(\rho + \phi + \alpha + \kappa_\mu)}{(\phi + \alpha + \kappa + \mu)(\rho + \alpha + \delta\kappa + \mu) - \phi\rho} \\ & + \frac{\beta s^0\kappa(\rho + \alpha + \delta\kappa + \mu + \delta\phi) + \varepsilon\beta s_v^0(\kappa\rho + \delta\kappa(\phi + \alpha + \kappa + \mu))}{(\gamma + \eta + \mu)[(\phi + \alpha + \kappa + \mu)(\rho + \alpha + \delta\kappa + \mu) - \phi\rho]}, \end{aligned} \quad (10.37)$$

where $\kappa \approx \kappa_0$ and s^0 and s_v^0 are the proportions of the population unvaccinated and vaccinated, respectively, at the disease-free equilibrium. With the vaccine efficacy parameters just obtained, what can be said about the possibility of effectively vaccinating this population against IPD with the given vaccines?

References: Banks et al. (2009), National Notifiable Diseases Surveillance System, Roche et al. (2004), Roche et al. (2006), Roche et al. (2007), Sutton, Banks and Castillo-Chavez (2008).

Epilogue

On Mathematical and Theoretical Biology

This book attempts to bridge the gap between mathematics and population biology. It is intended to show students of biology how to apply mathematics to the study of some questions of importance to population biology and to introduce modeling in the natural sciences to students of mathematics. It may also be used as a reference on mathematical methods for working biological scientists.

For the most part, we have given little description of the background of the subject, and we urge the reader to explore the history of population ecology in such sources as Kingsland (1985). We also suggest exploring Real and Brown (1991), a collection of 40 classic papers in ecology over the period 1887 to 1974.

Naturally, there are many topics in the mathematics of population biology that have been omitted from this book. An important omission involves the use of stochastic models. In any real-life situation there are random effects. If population sizes are large, these effects are often small enough to be ignored, but when population sizes are small, then their use is essential. Some references for such models are Durrett and Levin (1994), Nisbet and Gurney (1982), Pollard (1973), and Renshaw (1991).

There is also no consideration of structured models that take into account nonlinear birth, death, and infection processes. Some references in this area are Castillo-Chavez (1987), Diekmann and Metz (1986), Gurney and Nisbet (1998), Gurtin and McCamy (1974), and Hoppensteadt (1975).

There has been a great deal of work in the mathematical theory of epidemics; see for example [Anderson and May (1991), Brauer, van den Driessche, and Wu (2008), Castillo-Chavez (1989), Castillo-Chavez, Blower, van den Driessche, Kirschner, and Yakubu (2001a, 2001b), Diekmann and Heesterbeek (2000), Diekmann, Heesterbeek, and Metz (1990), Haderl (1989a, 1989b, 1992, 1993), Murray (1989), Thieme (2003), and Thieme and Castillo-Chavez (1989, 1993)] for a few of the directions of current interest.

Because both of us are currently interested in the study of disease dynamics we have emphasized epidemiological systems throughout the book. In fact, we are planning to write a book on the subject that will go well beyond the introduction in Chapters 9 and 10 of this book. Mathematical and theoretical epidemiology have experienced a great deal of growth over the last two decades due to the emergence and reemergence of diseases like tuberculosis and AIDS, and epidemic disease outbreaks such as SARS (2002–2003), recurring threats of avian influenza, and the H1N1 influenza pandemic of 2009. In our discussion of disease dynamics we have shown our biases. We have focused on the impact at the population (organizational) level and have ignored the outstanding research that is being carried out in immunology. Fortunately, many of the techniques we have illustrated in the context of population dynamics and epidemiology are also of use in population genetics,

mathematical physiology, immunology, and other areas of biology. A relatively new development in the modeling of disease transmission has been the study of networks; see for example, [Newman (2002, 2003), and Strogatz (2001)]. This subject is developing rapidly, and promises to be one of the key areas of mathematical epidemiology.

In this book little has been said about evolution, units of selection, levels of aggregation, and scales (temporal or spatial). Consideration of these factors in the study of population biology and epidemiology is critical. In the study of a disease such as influenza, multiple scales come into play. We have fast (disease dynamics), slow (host demography), and super slow (evolution of cross immunity) time scales. Influenza spreads locally (as in schools) and globally. Influenza epidemic waves move across cities, countries, and continents. Locally they are driven by age structure in contact rates and by public transportation, while globally they may be driven by train or airplane flow. We hope that the mathematical techniques and modeling approaches of this book will be useful for the interdisciplinary groups of scientists working on the types of challenges posed by influenza epidemics.

We may summarize our goals by saying that if our book facilitates communication between biologists and mathematicians, then we will feel that we have made a contribution to science.

Answers to Selected Exercises

Chapter 1

Section 1.1

1. 106.
3. 435.
5. 30.
7. (i) $x = 10e^{ct}$.
(ii) $c = 0.0347$.
(iii) 60 hours, 66.4 hours.

Section 1.2

1. 75.
3. 9.758×10^7 kg, 1.547 years.
- 9b For spherical cow, $MR = kW^{3/2}$; for cubical cow, $MR = k_c W^{3/2}$ with $k = k_c \pi / 6^{1/3}$.
- 10 (a) $W = (-t/12 + 12.57)^4$ (b) About 24 days.

Section 1.4

1. $x = 0$ unstable, $x = K$ asymptotically stable (if $r > 0$, $K > 0$).
3. $x = 0$ unstable, $x = K$ asymptotically stable (if $r > 0$, $K > 0$).
5. $x = 0$ unstable, $x = K(1 + \log \frac{r}{d})$ asymptotically stable (if $r > 0$, $K > 0$).
7. All values of $y(0)$ less than the solution of $e^{-y} = 2y$.

9. (i) If $e > \beta$, equilibrium 0 is asymptotically stable and equilibrium $1 - \frac{e}{\beta} < 0$ is unstable. If $e < \beta$, equilibrium 0 is unstable and equilibrium $1 - \frac{e}{\beta} > 0$ is asymptotically stable.
 (ii) The only equilibrium is 0, asymptotically stable if $a > 0$ and unstable if $a < 0$.
11. (i) Unstable.
 (ii) Asymptotically stable.
 (iii) Asymptotically stable.
15. For $H = 10$, there is one equilibrium ($V \approx 22$); for $H = 20$, there are two equilibria ($V \approx 16, V \approx 2$); for $H = 30$, there is only one equilibrium ($V \approx 1$). A large herd depletes vegetation discontinuously.

Section 1.5

1. Critical harvest rate 4800, equilibrium with harvest of 3000 per year is 156,900.
 3. 7066
 5. $\frac{rK}{e}$.
 7. Maximum yield is the value of $rx e^{1-\frac{x}{K}} - dx$, with x defined by $re^{1-\frac{x}{K}} \left[1 - \frac{x}{K}\right] = d$; cannot be evaluated explicitly.

Section 1.6

1. Lake would be eutrophic.
 3. Dump would move lake from current equilibrium past unstable equilibrium to eutrophic equilibrium.

Section 1.7

1. (i) Mean is $\frac{b^2 - a^2}{2}$.
 (ii) Cumulative distribution function is

$$f(t) = \begin{cases} 0, & \text{if } t \leq a, \\ \frac{t-a}{b-a} & \text{if } a < t < b, \\ 1, & \text{if } t \geq b. \end{cases}$$

3. Mean is $\frac{\alpha_1 + \alpha_2}{\alpha_1 \alpha_2}$. Probability density function is $\frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} (e^{-\alpha_1 t} - e^{-\alpha_2 t})$. If $\alpha_1 = \alpha_2 = \alpha$, probability density function is $\alpha^2 t e^{-\alpha t}$.

Chapter 2

Section 2.1

1. $x_n = 2^{1-n}$.
3. $x_{2n} = \frac{1}{2^n}, x_{2n+1} = -\frac{1}{2^n}$.
5. $x_{2n} = r^n, x_{2n+1} = -r^n$.
9. 0, 1, 1, 2, 3, 5, 8, 13.

Section 2.2

1. Figure 10.9.

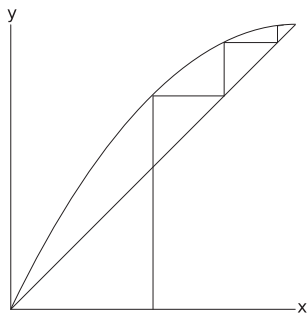


Fig. 10.9

3. Figure 10.10.

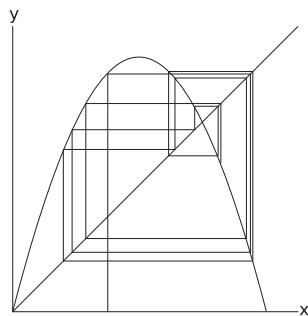


Fig. 10.10

5. x_n approaches 1000 if $R = 2$.

Section 2.3

1. If $r^2 - 4A < 0$, the only equilibrium is 0 (asymptotically stable). If $r^2 - 4A \geq 0$, equilibria are 0, $\frac{r + \sqrt{r^2 - 4A}}{2}$ (asymptotically stable) and $\frac{r - \sqrt{r^2 - 4A}}{2}$ (unstable).
3. $x = 0$ (stable if $r < 1$), $x = \frac{1}{\alpha} \left(r^{\frac{1}{\beta}} - 1 \right)$, asymptotically stable if $\left(r^{\frac{1}{\beta}} - 1 \right) (\beta - 2) < 2$.
5. $x = 0$ (unstable), $x = 1$ (asymptotically stable).
7. (b) $p < 1 - \frac{1}{e} = 0.632$.
9. (i) $x = 0$, $x = 2$ asymptotically stable, $x = 1$ unstable.
(ii) $\sqrt{\frac{9}{8}} - 1 = 0.06 < a < 1$.
14. $x = \frac{1}{2}$ is asymptotically stable if $0 \leq \alpha < \frac{1}{2\pi}$. Periodic orbit for $\alpha = \frac{1}{4}$ is $\left\{ \frac{1}{4}, \frac{3}{4}, \frac{1}{4}, \dots \right\}$.
15. $x = 0$ (asymptotically stable if $a < 1$). $x = \log a$ (asymptotically stable if $1 < a < e^2$). Period doubling bifurcation appears when $a = e^2$. Population goes extinct if $a < 1$.

Section 2.4

1. $r = e$.
5. Fixed points 0, $\frac{5}{2}$, $\frac{7}{12}$, 1 all unstable. Cycle of period 2 with initial value $\frac{1}{4}$.
9. $r = e$.

Section 2.5

1. $x = 0$ is asymptotically stable for $0 < \alpha < 1$, $x = \frac{\log \alpha}{\beta}$ is asymptotically stable for $1 < \alpha < e^2$.
3. $x = 0$ is asymptotically stable for $1 < \alpha < 1$, $x = \frac{\log \alpha}{\beta}$ is asymptotically stable for $\alpha > 1$.

Section 2.6

1. $x_{n+1} = 6x_n$; (2,3), (6,6), (12,18).
3. $(-1, 1, 1, -1, -1, 1, 1, \dots)$.
5. (0,0) and $\left(\frac{a}{a-1} \frac{\log a}{bc}, \frac{1}{b} \log a \right)$ [if $a > 1$], both unstable.

Section 2.7

- Equilibria are solutions of $y = \alpha B(y) - D(y)$, with x given by $B(y)$. Asymptotic stability conditions are $|D'(y)| < 1 - \alpha B'(y)$, $-\alpha B'(y) < 1$.

Section 2.8

- $A = 104.85, L = 179.32, P = 143.46$.
 - $A = 2.75, L = 13.06, P = 10.45$.

Chapter 3

Section 3.2

$$6. x = \frac{K \pm \sqrt{K^2 - \frac{4KH}{r}}}{2}.$$

Section 3.3

- $x = 0$ asymptotically stable for all T if $r < dA$. $x = \frac{r - Ad}{d}$, asymptotically stable for all T if $r > dA$.
- $x = 0$ is unstable for all T , $x = 3$ is asymptotically stable if $\sec z < -2$, where $z = -T \tan z$.
 - $x = 0$ is unstable for all T , $x = 3 - \log(1 + p)$ is asymptotically stable if $\log(1 + p) - 2 > \sec z$, where $z = -(1 + p)T \tan z$.

Section 3.4

- $x = 0$ asymptotically stable for all T if $r < \pi Ad$, $x = \frac{r}{\pi d} - A$ asymptotically stable for all T if $r > \pi Ad$ [for both parts (a) and (b)].
- $x = 0$ asymptotically stable for all T if $r < \pi d$. $x = \log \frac{r}{\pi d}$, existing only if $r > \pi d$, asymptotically stable for all T .

Section 3.5

1. Positive equilibrium x_∞ is asymptotically stable for all T if it exists, which happens if $H < \max x(e^{3-x} - 1) = 6.45$.

Chapter 4

Section 4.3

1. $u' = u - v, v' = u + v$ at $(1, 1)$.
3. $u' = v, v' = 2u + v$ at $(1, -1)$ and $u' = v, v' = -2u + v$ at $(-1, 1)$.
5. No equilibrium.
7. $u' = \lambda u, v' = \mu v$ at $(0, 0)$, $u' = -ax_\infty u - by_\infty v, v' = -cx_\infty u - dy_\infty v$ at $x_\infty = \frac{d\lambda - b\mu}{ad - bc}, y_\infty = \frac{a\mu - c\lambda}{ad - bc}$.
9. $(0, 0)$ with community matrix $\begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$, asymptotically stable. $(1, 1)$ with community matrix $\begin{bmatrix} -1 & 1 \\ \frac{8}{5} & -1 \end{bmatrix}$, unstable. $(4, 4)$ with community matrix $\begin{bmatrix} -1 & 1 \\ \frac{2}{5} & -1 \end{bmatrix}$, asymptotically stable.
11. (i) $S' = -\frac{\beta SI}{N} + \gamma(N - I - S)$ $I' = \frac{\beta SI}{N} - \gamma I$ [N constant].
(ii) Disease-free equilibrium $S = N, I = 0$ is asymptotically stable if $\beta < \nu$.
Endemic equilibrium with $\beta S = \nu N$ is asymptotically stable if it exists ($\beta > \nu$).
13. (i) Λ -people/time, β, μ , and γ -1/time.
(ii) $N' = \Lambda - \mu N$.
(iii) $N(t) = \frac{\Lambda}{\mu}(1 - e^{-\mu t}) + N(0)e^{-\mu t}$.
(v) $R_0 = \frac{\beta}{\mu + \gamma}$, equilibrium $S = K, I = 0$ is asymptotically stable if $R_0 < 1$ and equilibrium $S = \frac{\mu + \gamma}{\beta}$ is asymptotically stable if $R_0 > 1$.
15. (a) $\begin{bmatrix} -\mu & -\beta & 0 \\ 0 & \beta - (\mu + \gamma) & 0 \\ 0 & \gamma & -\mu \end{bmatrix}$, with eigenvalues $-\mu, -\mu, \beta - (\mu + \gamma)$.
(c) $R_0 = \frac{\beta}{\mu + \gamma}$, same as for (4.9).
(d) Disease-free equilibrium is asymptotically stable if $R_0 < 1$.

Section 4.4

1. $(1, 1)$ unstable.
3. $(-1, -1)$ unstable, $(1, 1)$ unstable.
5. No equilibrium.
7. Orbits depend on values of the parameters; solution is given in Section 5.1.
9. Orbits depend on the values of the parameters; solution is given in Section 5.2.

11. (b) $y \approx \frac{\alpha N}{\gamma - \alpha} (e^{-\alpha t} - e^{-\gamma t})$.
 (c) $y \approx \frac{N}{1000}$ for $t = 6.915$.
17. Competitive system; all solutions approach $(\frac{2}{3}, \frac{2}{3})$.
19. $(0, 0)$ has matrix $\begin{bmatrix} -0 & -1 \\ -1 & -0 \end{bmatrix}$, (saddle point). $(1, 1)$ has matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (unstable node).

Section 4.5

- There are three equilibria $(0, 0)$, $(20, 0)$, and $(10, 30)$. All of them are unstable. There is a stable limit cycle around $(10, 30)$. Two species coexist with oscillation.
- There are three equilibria $(0, 0)$, $(20, 0)$, and $(2, 32.4)$. All of them are unstable. There is a stable limit cycle around $(2, 42.4)$. Two species coexist with oscillation.
- There are two equilibria $(0, 0)$, $(20, 0)$; $(20, 0)$ is asymptotically stable. Thus the predator species goes to extinction.
- The equilibrium $(60, 20)$ is asymptotically stable and all trajectories approach it. Two species coexist.
- The equilibrium $(0, 45)$ is asymptotically stable and all trajectories approach it. The x -species goes to extinction and the y -species wins the competition.
- The equilibrium $(0, 16)$ is asymptotically stable and all trajectories approach it. The x -species goes to extinction and the y -species wins the competition.
- If $n < 8$, steady state is an asymptotically stable, a spiral point if $n < \sqrt{48}$ and a node if $\sqrt{48} < n < 8$. If $n > 8$, steady state is a saddle point.

Chapter 5

Section 5.1

- The equilibrium $(60, 20)$ is asymptotically stable and all trajectories approach it. Two species coexist.
- The equilibrium $(0, 45)$ is asymptotically stable and all trajectories approach it. The x -species goes to extinction and the y -species wins the competition.

5. The equilibrium $(0, 16)$ is asymptotically stable and all trajectories approach it. The x -species goes to extinction and the y -species wins the competition.

Section 5.2

1. There are three equilibria $(0, 0)$, $(20, 0)$, and $(10, 30)$. All of them are unstable. There is a stable limit cycle around $(10, 30)$. Two species coexist with oscillation.
3. There are three equilibria $(0, 0)$, $(20, 0)$, and $(2, 32.4)$. All of them are unstable. There is a stable limit cycle around $(2, 42.4)$. Two species coexist with oscillation.
5. There are two equilibria $(0, 0)$, $(20, 0)$. $(20, 0)$ is asymptotically stable. Thus the predator species goes to extinction.

Section 5.3

1. [Figure 10.11.](#)

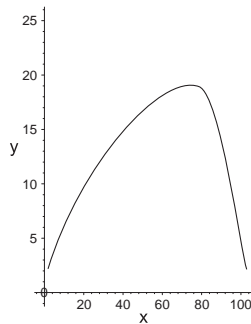


Fig. 10.11

Section 5.5

1. There are two equilibria $(0,0)$ and $(120,70)$. The first is unstable and the second is asymptotically stable.
3. There are three equilibria $(0,0)$, $(0,10)$, and $(4 + \sqrt{34}, 5 + \sqrt{34})$. The first two are unstable and the third is asymptotically stable.
5. There are three equilibria $(0,0)$, $(0,M)$, and $(K,0)$. They are all unstable. If $ab \geq 1$, orbits are unbounded. If $ab < 1$, a stable equilibrium $(x_\infty, y_\infty) = \left(\frac{K+aM}{1-ab}, \frac{Mb+K}{1-ab} \right)$ appears and all trajectories approach this equilibrium.

Chapter 6

Section 6.1

1. When $E = 0$, the equilibrium $(x_\infty, y_\infty) = (60, 20)$ is asymptotically stable and thus the two species coexist. When $E > 0$, the stable equilibria are on the line $x + y = 80$, and when $E = 120$, it coalesces to $(0, 80)$. Therefore, harvesting decreases the x -species size and increases the y -species, and eventually moves coexistence to x -extinction.
3. Without harvesting, the x -species goes to extinction and the y -species wins the competition. Harvesting the x -species will speed up extinction of the x -species.
5. Four equilibria $O_1 = (40 + \sqrt{40^2 - H}, 0)$, $O_2 = (40 - \sqrt{40^2 - H}, 0)$, $O_3 = \left(\frac{60 + \sqrt{60^2 - 6H}}{2}, 40 - \frac{60 + \sqrt{60^2 - 6H}}{6} \right)$ and $O_4 = \left(\frac{60 - \sqrt{60^2 - 6H}}{2}, 40 - \frac{60 - \sqrt{60^2 - 6H}}{6} \right)$. O_1 and O_2 are always unstable; O_4 is unstable and O_3 is stable. With harvesting of the x -species, as H increases O_3 and O_4 move along the line $x + 3y = 120$ until they coalesce when $H = 60$, resulting in coexistence at $(30, 30)$.
7. There is no equilibrium with positive population sizes. The x -species becomes extinct in finite time.

Section 6.2

1. $H_c = H^* = \frac{5}{6}$.
3. $H_c = 5.787$, $H_1^* = 3.7037$, $H_2^* = 5.5556$. H must be in the interval $(0, H_1^*)$ or in the interval (H_2^*, H_c) .
5. $H_c = 8.8889$. Trace of community matrix is positive, meaning that equilibrium is unstable. There exists a stable limit cycle.

Section 6.3

1. If we repeat the process 5 times, the estimated average instantaneous harvest is 34.5526. If we repeat the process 10 times, the estimated average instantaneous harvest is 34.2373. If we repeat the process 15 times, the estimated average instantaneous harvest is 34.1638. If we repeat the process 100 times, the estimated average instantaneous harvest is 34.0324. If we repeat the process 500 times, the estimated average instantaneous harvest is 34.0318.

Section 6.4

1. $Y(E) = \frac{aE}{E+b}$, always less than a and approaches a as $E \rightarrow \infty$. However, maximum fish population size is $\frac{a}{b}$. Thus the yield cannot exceed $\frac{a}{b}$.
3. E_{MSY} is the unique solution of the equation $\log\left(\frac{r}{d+E}\right) - \frac{E}{d+E} = 0$ for E in the interval $(0, r-d)$. The maximum sustainable yield $Y(E_{MSY}) = \frac{E_{MSY}^2}{d+E_{MSY}}$.

Section 6.5

1. $F(y) = ry \log \frac{K}{y}$, $h = Ey$, $C(y) = \frac{c}{y}$; y^* is the unique solution to the equation $rp \log \frac{K}{y} - (\delta + r)(p - \frac{c}{y}) = 0$. If $\frac{K}{e} < \frac{c}{p}$, $y^* < \frac{K}{e} = y_{MSY}$, implying overfishing.

Chapter 7

Section 7.1

1. $\lambda_0 = 0.9491$, $v_0 = [0.8021, 0.4534, 0.1979]$,
stable age distribution $[0.5131, 0.3604, 0.1226]$.
3. $\lambda_0 = 1.6826$, $v_0 = [0.9552, 0.2838, 0.0843]$,
stable age distribution $[0.7281, 0.2145, 0.0637]$.
5. Leslie matrix is

$$A = \begin{bmatrix} 1 & 3 \\ 2/3 & 0 \end{bmatrix},$$

$\lambda_0 = 2$, $v_0 = [0.9487, 0.3162]$, the stable age distribution is $[0.75, 0.25]$. The total population $P_n \approx c2^n$ when n is large.

Section 7.3

- 1(d). $B(t) = \left(\beta \int_0^\infty \phi(s) ds \right) e^{(\beta - \mu)t}$.
4. $P(t) = \varphi_0(a) \frac{\pi(a+t)}{\pi(a)} + \int_0^t P(t-a) \pi(a) da$.

Chapter 8

Section 8.2

4. $p'_1 = cp_1p_2 - (d+e)p_1$, $p'_2 = \lambda + (d-\lambda)p_1 - (e+\lambda)p_2 - cp_1p_2$.
Equilibria are $(0, \lambda/(e+\lambda))$, $(\lambda/(\lambda+e) - \lambda(d+e)/c, 0)$.
- (ii) Threshold quantity is

$$\frac{c\lambda}{(e+\lambda)(d+e)}.$$

Section 8.4

2. Unique equilibrium

$$y_1 = \frac{\mu + d + 2\sigma_1}{(\mu + d)(\mu + \sigma_1) + \mu\sigma_2} \Lambda, \quad y_2 = \frac{\mu + 2\sigma_1}{(\mu + d)(\mu + \sigma_1) + \mu\sigma_2} \Lambda,$$

asymptotically stable.

Section 8.6

- $u_n(x, t) = B_n e^{-D(\frac{2(n-1)\pi}{2L})^2 t} \sin \frac{2(n-1)\pi}{2L} x$.
 $u(x, t) = \sum_{n=1}^{\infty} B_n e^{-D(\frac{2(n-1)\pi}{2L})^2 t} \sin \frac{2(n-1)\pi}{2L} x$.
- $u_n(x, t) = (A_n \cos \lambda_n x + B_n \sin \lambda_n x) e^{-D(\lambda_n)^2 t}$
 where λ_n are roots of $\beta \tan \lambda L = -\lambda$,
 $u(x, t) = A_0 + \sum_{n=1}^{\infty} (A_n \cos \lambda_n x + B_n \sin \lambda_n x) e^{-D(\lambda_n)^2 t}$,
 $u(x, 0) = f(x) = A_0 + \sum_{n=1}^{\infty} (A_n \cos \lambda_n x + B_n \sin \lambda_n x)$.
- From Problem 1, $u(x, t) = \sum_{n=1}^{\infty} B_n e^{-D(\frac{2(n-1)\pi}{2L})^2 t} \sin \frac{2(n-1)\pi}{2L} x$, $u(x, 0) = x = \sum_{n=1}^{\infty} B_n \sin \frac{2(n-1)\pi}{2L} x$, $B_n = \int_0^L x \sin \frac{2(n-1)\pi}{2L} x dx = \frac{8L}{2(n-1)\pi}$.
- $u_n(x, t) = B_n e^{-D(\frac{2(n-1)\pi}{2L})^2 t} \sin \frac{2(n-1)\pi}{2L} x$, $u(x, t) = \sum_{n=1}^{\infty} B_n e^{-D(\frac{2(n-1)\pi}{2L})^2 t} \sin \frac{2(n-1)\pi}{2L} x$,
 $u(x, 0) = f(x) = \sum_{n=1}^{\infty} B_n \sin \frac{2(n-1)\pi}{2L} x$, $B_n = \int_0^L f(x) \sin \frac{2(n-1)\pi}{2L} x dx$
- Let $u(x, t) = v(x) + w(x, t)$ where $v(x) = c_1 + c_2 x$ with $v(0) = c_1 = u_1$ and $v'(0) = c_2 = 0$. Therefore, $v(x) = u_1$ and $u(x, 0) = v(0) + w(x, 0) = f(x)$, leading to $w(x, 0) = f(x) - u_1$. Now, $w(x, t)$ is a solution to Problem 1; $w(x, t) = \sum_{n=1}^{\infty} B_n e^{-D(\frac{2(n-1)\pi}{2L})^2 t} \sin \frac{2(n-1)\pi}{2L} x$, $B_n = \int_0^L (f(x) - u_1) \sin \frac{2(n-1)\pi}{2L} x dx$.
- $X(x) = A \cos \lambda x + B \sin \lambda x$ with $X(0) = 0$ and $X'(L) = -\alpha X(L)$
 $X_n(x) = B_n \sin \lambda_n x$ where λ_n are roots of $\alpha \tan \lambda L = -\lambda$. Therefore, $u(x, t) = \sum_{n=1}^{\infty} B_n \sin \lambda_n x e^{-D(\lambda_n)^2 t}$.
- $u(x, 0) = u_0 + u_0 \cos \frac{\pi x}{L}$.

Section 8.7

$$3. u(x, t) = \int_0^{\infty} u_0 \frac{1}{\sqrt{4D\pi t}} e^{-\frac{|x-y|^2}{4Dt}} dy.$$

Section 8.8

$$1. u(x, t) = \int_{-\infty}^{\infty} f(y) \frac{1}{\sqrt{4D\pi t}} e^{-\frac{|x-y|^2}{4Dt}} dy + \int_0^{\infty} g(s) \frac{1}{\sqrt{4D\pi(t-s)}} e^{-\frac{|x|^2}{4D(t-s)}} ds.$$

Chapter 9

Section 9.2

1. 1.4416.
3. 2.827%.
5. 17.56%.
7. 58.33%.
10. $0.9125 \text{ foxes/km}^2$.
11. 23.1%.
13. No.

Section 9.3

3. Epidemic probability is 0.583.

Section 9.6

1. $S'(t) = -\beta S(t)I(t), \quad I'(t) = I'_0(t) + \beta S(t)I(t) - \beta S(t - \sigma)I(t - \sigma).$

Section 9.9

1. (i) Model

$$\begin{aligned}
 \dot{S} &= -\beta(I + \delta A)S - \nu S, \\
 \dot{S}_v &= -\varepsilon\beta(I + \delta A)S_v + \nu S - \phi S_v, \\
 \dot{P} &= \phi S_v, \\
 \dot{L} &= \beta(I + \delta A)S + \varepsilon\beta(I + \delta A)S_v - \kappa L, \\
 \dot{I} &= p\kappa L - \gamma I, \\
 \dot{A} &= (1 - p)\kappa L - \eta A \\
 \dot{R} &= \eta A + \gamma I,
 \end{aligned}$$

where S_v is the vaccinated class, P is the protected class, ν is the vaccination rate, and ϕ is the rate of advancement to protection from vaccination. Yes, the population is constant. No, there are no endemic equilibria. This can be seen since S tends to zero, causing S_v to tend toward zero, causing L to tend to zero, which ultimately causes I to tend to zero as t goes to infinity.

- (ii) As the vaccination rate is increased, the number of infected individuals will decrease. Yes, vaccination can prevent an outbreak.

Section 9.10

3. Model

$$\begin{aligned}\dot{S} &= -\beta(I + \delta T)S, \\ \dot{E} &= \beta(I + \delta T)S - \kappa E, \\ \dot{I} &= \kappa E - (\alpha + \gamma)I, \\ \dot{T} &= \gamma I - \eta T, \\ \dot{R} &= \eta T + \alpha I.\end{aligned}$$

Then

$$\mathcal{R}_0 = \frac{\beta N}{\alpha + \gamma}.$$

Chapter 10

Section 10.1

3. (i) $C = C_0 e^{-\alpha t}$. (ii) $S_0 e^{\frac{\beta C_0}{\alpha}(e^{-\alpha t} - 1)}$. (iii) $\lim_{t \rightarrow \infty} S(t) = S_0 e^{-\frac{\beta}{\alpha} C_0}$.
5. About 4 1/2 hours.

Section 10.3

1. (i) Check $\frac{dN}{dt} = 0$. (ii) $\frac{1}{\mu}$ average life span; $\frac{1}{\gamma}$ average infectious period; $\frac{1}{\xi}$ average quarantine period. Their units are the same, namely time. (iii) $\nu = \frac{\mu}{\sigma}$, $\theta = \frac{\gamma}{\sigma}$, $\zeta = \frac{\xi}{\sigma}$. (iv) $\mathcal{R}_0 = \frac{\sigma}{\mu + \gamma}$, $\mathcal{R}_0 \leq 1$, $(0, 0, 0)$ is asymptotically stable. $\mathcal{R}_0 > 1$, $(0, 0, 0)$ is unstable.

Section 10.4

3. There is a disease-free equilibrium $S = 0$, $V = N$, $I = 0$. To find endemic equilibria, substitute

$$S + \frac{\alpha I}{\varphi + \beta I}, \quad \sigma \beta V = \frac{\alpha \varphi}{\varphi + \beta I},$$

into

$$(\alpha + \theta)I + \theta S + \theta V = \theta N$$

to obtain a quadratic equation for I that may have either one positive root or two positive roots.

References

1. Adler, F., L. Smith, and C. Castillo-Chavez (1989) A distributed delay model for the local population dynamics of a parasitoid-host system, *Mathematical Approaches to Ecological and Environmental Problem Solving* (Castillo-Chavez C., S. A. Levin, and C. Shoemaker, eds.) *Lecture Notes in Biomathematics* 81, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong, pp. 152-62.
2. Agur, Z.L., G. Mazor, R. Anderson, and Y. Danon (1993) Pulse mass measles vaccination across age cohorts, *Proc. Nat. Acad. Sci.*, **90**:11698-11702.
3. Allee, W.C. (1931) *The Social Life of Animals*, Heinemann, London.
4. Anderson, R.M. (1974) Mathematical models of host-helminth parasite interactions. In: *Ecological Stability*, (M. B. Usher and M. H. Williamson, eds.), Chapman and Hall, London, pp. 43-69.
5. Anderson, R.M., ed. (1982) *Population Dynamics of Infectious Diseases: Theory and Applications*, Chapman and Hall, London-New York.
6. Anderson, R.M., D. R. Cox, and H. C. Hillier (1989) Epidemiological and statistical aspects of the AIDS epidemic: introduction, *Phil. Trans. Roy. Soc. Lond. B* **325**:39-44.
7. Anderson, R.M., H.C. Jackson, R.M. May, and A.M. Smith (1981) Population dynamics of fox rabies in Europe, *Nature*, **289**:765-771.
8. Anderson, R.M. and R.M. May (1982) Co-evolution of host and parasites, *Parasitology*, **85**:411-426.
9. Anderson, R.M. and R.M. May (1983) Vaccination against rubella and measles: Quantitative investigations of different policies, *J. Hygiene*, **90**:259-325.
10. Anderson, R.M. and R.M. May (1991) *Infectious Diseases of Humans*, Oxford Science Publications, Oxford.
11. Andrewartha, H.G. and L.C. Birch (1954) *The Distribution and Abundance of Animals*, Univ. of Chicago Press, Chicago.
12. Aparicio, J.P., A. Capurro, and C. Castillo-Chavez (2000a) Markers of disease evolution: the case of tuberculosis, *J. Theor. Biol.*, **215**: 227-238.
13. Aparicio, J.P., A. Capurro, and C. Castillo-Chavez (2000b) On the rise and fall of tuberculosis. *Department of Biometrics, Cornell University, Technical Report Series*, BU-1503-M.
14. Aparicio J.P., A.F. Capurro, and C. Castillo-Chavez (2000b) On the fall and rise of tuberculosis. *Department of Biometrics, Cornell University, Technical Report Series*, BU-1477-M.
15. Aparicio J.P., A. Capurro, and C. Castillo-Chavez (2001a). Frequency Dependent Risk of Infection and the Spread of Infectious Diseases. In: *Mathematical Approaches for Emerging and Reemerging Infectious Diseases : Models, Methods and Theory*, Edited by Castillo-Chavez, C. with S. Blower, P. van den Driessche, D. Kirschner, and A.A. Yakubu, Springer-Verlag (2002), pp. 341-350.
16. Aparicio, J.P., A. Capurro, and C. Castillo-Chavez (2001b) On the long-term dynamics and re-emergence of tuberculosis. In: *Mathematical Approaches for Emerging and Reemerging*

- Infectious Diseases: Models, Methods and Theory*, Edited by Castillo-Chavez, C. with S. Blower, P. van den Driessche, D. Kirschner, and A.A. Yakubu, Springer-Verlag, (2002) pp. 351–360.
17. Arino, J., F. Brauer, P. van den Driessche, J. Watmough and J. Wu (2007) A final size relation for epidemic models, *Math. Biosci. & Eng.* **4**: 159–176.
 18. Arino, J., F. Brauer, P. van den Driessche, J. Watmough and J. Wu (2006) Simple models for containment of a pandemic, *J. Roy. Soc. Interface*, **3**: 453–457.
 19. Armstrong, R.A. and R. McGehee (1980) Competitive exclusion, *Amer. Naturalist*, **115**:151–170.
 20. Aron, J.L. and R.M. May (1982) The population dynamics of malaria. In: *Population Dynamics of Infectious Disease*, R.M. Anderson, ed., 139–179, Chapman and Hall, London.
 21. Arreola R., A. Crossa, and M.C. Velasco (2000) Discrete-time SEIS models with exogenous re-infection and dispersal between two Patches, *Department of Biometrics, Cornell University, Technical Report Series*, BU-1533-M.
 22. Arriola P., I. Mijares-Bernal, J.A. Ortiz-Navarro, H. Campus, and R.A. Senz (2000) Dynamics of the spruce budworm population under the action of predation and insecticides, *Department of Biometrics, Cornell University, Technical Report Series*, BU-1417-M.
 23. Ayala, F.J., M.E. Gilpin, and J.G. Ehrenfeld (1973) Competition between species: Theoretical models and experimental tests, *Theoretical Pop. Biol.*, **4**:331–356.
 24. Bailey, N.T.J. (1975) *The Mathematical Theory of Infectious Diseases and Its Applications*, Griffin, London.
 25. Bailey, N.T.J. (1988) Statistical problems in the modelling and prediction of HIV/AIDS, *Aust. J. Stat.* **30A**:41–55.
 26. Banks H.T. and C. Castillo-Chavez (eds.) (2003) *Bioterrorism: Mathematical Modeling Applications in Homeland Security*. SIAM Series Frontiers in Applied Mathematics **28**.
 27. bibitemBanks2009 Banks, H.T., M. Davidian, J. R. Samuels, Jr., and K. L. Sutton (2009), An inverse problem statistical methodology summary, in *Mathematical and Statistical Estimation Approaches in Epidemiology*, (G. Chowell, C. Castillo-Chavez, J.M. Hyman, and Luis Bettencourt, eds.), Springer, Berlin Heidelberg New York: 249–302.
 28. Banks, H.T. and H.T. Tran (2009), *Mathematical and Experimental Modeling of Physical and Biological Processes*, CRC Press, Boca Raton.
 29. Bansal, S., J. Read, B. Pourbohloul, and L.A. Meyers (2010) The dynamic nature of contact networks in infectious disease epidemiology, *J. Biol. Dyn.*, **4**: 478–489.
 30. Barbour, A.D. (1978) Macdonald's model and the transmission of bilharzia, *Trans. Roy. Soc. Trop. Med. Hygiene*, **72**:6–15.
 31. Barrera J.H., A. Cintron-Arias, N. Davidenko, L.R. Denogean, and S.R. Franco-Gonzalez (2000) Dynamics of a two-dimensional discrete-time SIS model, *Department of Biometrics, Cornell University, Technical Report Series*, BU-1518-M.
 32. Bélair, J. and L. Glass (1983) Self similarity in periodically forced oscillators, *Phys. Lett.*, **96A**:113–116.
 33. Beck, K. (1984) Co-evolution: Mathematical aspects of host-parasite interactions, *J. Math. Biol.*, **19**:63–77.
 34. Bellman, R.E. and K.L. Cooke (1963) *Differential-Difference Equations*, Academic Press, New York.
 35. Bendixson, T. (1901) Sur les courbes définies par des équations différentielles, *Acta. Math.*, **24**:1–88.
 36. Berman, A. and R.J. Plemmons (1970) *Nonnegative Matrices in the Mathematical Sciences*, Academic Press (1970); Classics in Applied Mathematics, SIAM, Philadelphia (1994).
 37. Bernardelli, H. (1941) Population waves, *J. Burma Res. Soc.*, **31**:1–18.
 38. Bernoulli, D. (1766) Essai d'une nouvelle analyse de la mortalité causée par la petite verole, *Mem. Math. Phys. Acad. R. Sci. Paris*:1–45.
 39. Bettencourt, L.M.A., A. Cintron-Arias, D.I. Kaiser, and C. Castillo-Chavez (2006) The power of a good idea: quantitative modeling of the spread of ideas from epidemiological models, *Physica A* **364**: 513–536.

39. Bettencourt, L.M. A., R. M. Ribeiro, G. Chowell, T. Lant, C. Castillo-Chavez (2007) Towards real time epidemiology: data assimilation, modeling and anomaly detection of health surveillance data streams, *Intelligence and security informatics: Biosurveillance* (D. Zeng, D. Gotham, K. Komatsu, C. Lynch, Eds.) . Proceedings of the 2nd NSF Workshop, Biosurveillance: 79-90. Lecture Notes in Computer Science. New Brunswick, NJ: Springer-Verlag Berlin.
40. Beverton, R.J. and S.J. Holt (1956) The theory of fishing. In *Sea Fisheries; Their Investigation in the United Kingdom*, M. Graham, ed., 372-441, Edward Arnold, London.
41. Blower, S. (2004) An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it, by Daniel Bernoulli; reviewed by Sally Blower, *Rev. Med. Virol.*, **14**: 275-288.
42. Blythe, S.P. and C. Castillo-Chavez (1990) Scaling law of sexual activity, *Nature*, **344**:202.
43. Blythe, S.P., S. Busenberg, and C. Castillo-Chavez (1995) Affinity and paired-event probability, *Math. Biosci.*, **128**:265-84.
44. Blythe, S.P., C. Castillo-Chavez, J. Palmer, and M. Cheng (1991) Towards a unified theory of mixing and pair formation, *Math. Biosci.*, **107**:379-405.
45. Bora, D. (1999) Epidemiology of visceral leishmaniasis in India. *Natl Med J India* **12**: 62-68.
46. Brauer, F. (1976) Perturbations of the nonlinear renewal equation, *Adv. in Math.*, **22**: 32-51.
47. Brauer, F. (1986) Coexistence and survival of invading species. In *Mathematical Ecology*, T.G. Hallam, S.A. Levin, and L.J. Gross eds., 605-616, World Scientific Press, Singapore.
48. Brauer, F. (1987a) Harvesting in delayed recruitment population models. In *Oscillation, Bifurcation, and Chaos*, CMS Conference Proceedings **8**, 317-327, AMS, Providence.
49. Brauer, F. (1987b) A class of Volterra integral equations arising in delayed-recruitment population models, *Nat. Res. Modeling*, **2**:259-278.
50. Brauer, F. (1989) Multispecies interactions and coexistence. In *Differential equations and Applications*, A.R. Aftabizadeh, ed., 91-96, Ohio University Press, Athens.
51. Brauer, F. and J.A. Nohel (1969) *The Qualitative Theory of Ordinary Differential Equations*, W. A. Benjamin, New York; reprinted by Dover, New York 1989.
52. Brauer, F. and D.A. Sánchez (1975) Constant rate population harvesting: equilibrium and stability, *Theor. Pop. Biol.*, **8**:12-30.
53. Brauer, F. and A.C. Soudack (1979) Stability regions and transition phenomena for harvested predator-prey systems, *J. Math. Biol.*, **7**:319-337.
54. Brauer, F. and A.C. Soudack (1985) Optimal harvesting in predator-prey systems, *Int. J. Control*, **41**:111-128.
55. Brauer, F., A.C. Soudack, and H.S. Jarosch (1976) Stabilization and destabilization of predator-prey systems under harvesting and nutrient enrichment, *Int. J. Control*, **23**:553-573.
56. Brauer, F., P. van den Driessche and J. Wu, eds. (2008) *Mathematical Epidemiology*, Lecture Notes in Mathematics, Mathematical Biosciences subseries 1945, Springer, Berlin - Heidelberg - New York.
57. Bremermann, H.J. and J. Pickering (1983) A game-theoretical model of parasite virulence, *J. Theor. Biol.*, **100**:411-426.
58. Bremermann, H.J. and H.R. Thieme (1989) A competitive exclusion principle for pathogen virulence, *J. Math. Biol.*, **27**:179-190.
59. Britton, N.F. (2002) *Essential Mathematical Biology* Springer, Berlin - Heidelberg - New York.
60. Brookmeyer, R. and M. H. Gail (1988) A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic, *J. Am. Stat. Assoc.*, **83**:301-308.
61. Bulmer, M.G. (1976) The theory of predator-prey oscillations, *Theor. Pop. Biol.*, **9**:137-150.
62. Busenberg, S. and C. Castillo-Chavez. (1989) Interaction, pair formation and force of infection terms in sexually transmitted diseases. In *Mathematical and Statistical Approaches to AIDS Epidemiology*, Lect. Notes Biomath. **83**, C. Castillo-Chavez, ed., 289-300, Springer-Verlag, Berlin-Heidelberg-New York.
63. Busenberg, S. and C. Castillo-Chavez (1991) A general solution of the problem of mixing of subpopulations, and, its application to risk- and age-structured epidemic models, *IMA J. Math. Appl. Med. Biol.*, **8**:1-29.

64. Busenberg, S. and K.L. Cooke (1993) *Vertically Transmitted Diseases: Models and Dynamics*, Biomathematics **23**, Springer-Verlag, Berlin-Heidelberg-New York.
65. Butler, G.J., H.I. Freedman, and P. Waltman (1986) Uniformly persistent systems, *Proc. Am. Math. Soc.*, **96**:425-430.
66. Butler, G.J., S.B. Hsu, and P. Waltman (1983) Coexistence of competing predators in a chemostat, *J. Math. Biol.*, **17**:133-151.
67. Butler, G.J. and P. Waltman (1986) Persistence in dynamical systems, *J. Diff. Equations*, **86**:225-263.
68. Callaway, D.S., M.E.J. Newman, S.H. Strogatz, D.J. Watts (2000) Network robustness and fragility: Percolation on random graphs, *Phys. Rev. Letters*, **85**, 5468-5471 (2000).
69. Cannon, J.R. (1984) *The One-Dimensional Heat Equation*, Encyclopedia of Mathematics and Its Applications **32**, Addison-Wesley, Reading MA.
70. Carpenter, S.R., D. Ludwig, and W.A. Brock (1999) Management of eutrophication for lakes subject to potentially reversible change, *Ecological Applications*, **9**:751-771.
71. Castillo-Chavez, C. (1987a) Linear character dependent models with constant time delay in population dynamics, *Int. J. Math. Modeling*, **9**:821-836.
72. Castillo-Chavez, C. (1987b) Non-linear character-dependent models with constant time delay in population dynamics *J. Math. Analysis and Applications*, **128**(1):1-29.
73. Castillo-Chavez, C., ed. (1989) *Mathematical and Statistical Approaches to AIDS Epidemiology*, Lect. Notes Biomath. **83**, Springer-Verlag, Berlin-Heidelberg-New York.
74. Castillo-Chavez, C., S. Blower, P. van den Driessche, D. Kirschner, and A.A. Yakubu, eds. (2001a) *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods and Theory*, Volume I, Springer-Verlag, Berlin-Heidelberg-New York (2002).
75. Castillo-Chavez, C. with S. Blower, P. van den Driessche, D. Kirschner, and A.A. Yakubu, eds. (2001b) *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods and Theory*, Volume II, Springer-Verlag, Berlin-Heidelberg-New York (2002).
76. Castillo-Chavez, C. and S. Busenberg (1991) On the solution of the two-sex problem. In *Proceedings of the International Conference on Differential Equations and Applications to Biology and Population Dynamics*, S. Busenberg and M. Martelli, eds., Lect. Notes Biomath. **92**, 80-98, Springer-Verlag, Berlin-Heidelberg-New York.
77. Castillo-Chavez, C., S. Busenberg and K. Gerow (1991) Pair formation in structured populations. In *Differential Equations with Applications in Biology, Physics and Engineering*, J. Goldstein, F. Kappel, and W. Schappacher, eds., 47-65, Marcel Dekker, New York.
78. Castillo-Chavez, C and G. Chowell (eds.) (2011) *Math. Biosc. & Eng. special issue: Mathematical Models, Challenges, and Lessons Learned from the 2009 A/H1N1 Influenza Pandemic* **8**.
79. Castillo-Chavez, C., K. Cooke, W. Huang, and S.A. Levin (1989a) The role of long incubation periods in the dynamics of HIV/AIDS. Part 1: Single Populations Models, *J. Math. Biol.*, **27**:373-98.
80. Castillo-Chavez, C., K.L. Cooke, W. Huang, and S. A. Levin (1989b) Results on the dynamics for models for the sexual transmission of the human immunodeficiency virus, *Applied Math. Letters*, **2**(4):327-331.
81. Castillo-Chavez, C., and Z. Feng (1997) To treat or not to treat; the case of tuberculosis, *J. Math. Biol.*, **35**: 629-656.
82. Castillo-Chavez, C., and Z. Feng (1998a) Global stability of an age-structure model for TB and its applications to optimal vaccination strategies, *Mathematical Biosciences*, **151**:135-154.
83. Castillo-Chavez, C. and Z. Feng (1998b) Mathematical models for the disease dynamics of tuberculosis, *Advances in Mathematical Population Dynamics - Molecules, Cells, and Man* (O. Arino, D. Axelrod, M. Kimmel, (eds)): 629-656, World Scientific Press, Singapore.
84. Castillo-Chavez, Z. Feng, and W. Huang (2002) On the computation of \mathcal{R}_0 and its role on global stability. In *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods and Theory*, Volume I, (Castillo-Chavez, C., S. Blower, P. van

- den Driessche, D. Kirschner, and A.A. Yakubu, eds.):229-256, Springer-Verlag, Berlin-Heidelberg-New York.
85. Castillo-Chavez, C., H.W. Hethcote, V. Andreasen, S.A. Levin & Wei-min Liu (1988) Cross-immunity in the dynamics of homogeneous and heterogeneous populations. In *Mathematical Ecology*, T.G. Hallam, L.G. Gross, and S.A. Levin, eds., 303-316, World Scientific Publishing Co., Singapore.
 86. Castillo-Chavez, C., H.W. Hethcote, V. Andreasen, S.A. Levin and Wei-min Liu (1989) Epidemiological models with age structure, proportionate mixing, and cross-immunity, *J. Math. Biol.*, **27**:233-258.
 87. Castillo-Chavez, C., and W. Huang (1995) The logistic equation revisited: the two-sex case, *Math. Biosci.*, **128**:299-316.
 88. Castillo-Chavez, C., W. Huang, and J. Li (1996a) Competitive exclusion in gonorrhea models and other sexually-transmitted diseases, *SIAM J. Appl. Math.*, **56**(2):494-508.
 89. Castillo-Chavez, C., W. Huang, and J. Li, (1996b) On the existence of stable pairing distributions, *J. Math. Biol.*, **34**:413-441.
 90. Castillo-Chavez, C., W. Huang, and J. Li (1997) The effects of females' susceptibility on the coexistence of multiple pathogen strains of sexually-transmitted diseases, *Journal of Mathematical Biology*, **35**:503-522.
 91. Castillo-Chavez, C. and S-F Hsu Schmitz, (1993) On the evolution of marriage functions: It takes two to tango. In *Structured Population Models in Marine, Terrestrial, and Freshwater Systems*, S. Tuljapurkar and H. Caswell, eds., 533-553, Chapman and Hall, New York.
 92. Castillo-Chavez, C., S.A. Levin, and C. Shoemaker, eds. (1989) *Mathematical Approaches to Ecological and Environmental Problem Solving, Lecture Notes in Biomathematics 81*, Springer-Verlag, Berlin-Heidelberg-New York.
 93. Castillo-Chavez, C., S-F Shyu, G. Rubin, and D. Umbauch, (1992) On the estimation problem of mixing/pair formation matrices with applications to models for sexually-transmitted diseases. In *AIDS Epidemiology: Methodology Issues*, N.P. Jewell, K. Dietz, and V.T. Farewell, eds., 384-402, Birkhäuser, Boston.
 94. Castillo-Chavez, C., and H.R. Thieme (1993) Asymptotically autonomous epidemic models. In *Mathematical Population Dynamics: Analysis of Heterogeneity, Vol. 1, Theory of Epidemics*, O. Arino, D. Axelrod, M. Kimmel, and M. Langlais, eds., 33-50, Wuerz, Winnipeg.
 95. Castillo-Chavez, C., J.X. Velasco-Hernández, and S. Fridman (1994) Modeling Contact Structures in Biology. In *Frontiers of Theoretical Biology*, S.A. Levin, ed., Lect. Notes Biomath. **100**, 454-491, Springer-Verlag, Berlin-Heidelberg-New York.
 96. Castillo-Chavez C., and A.A. Yakubu (2000a) Discrete-time nonlinear pair formation models with geometric solutions, *Department of Biometrics, Cornell University, Technical Report Series*, BU-1496-M.
 97. Castillo-Chavez C., and A.A. Yakubu (2000b) Epidemics models on attractors, *Contemporary Mathematics*, AMS, **284**: 23–42.
 98. Castillo-Chavez C., and A.A. Yakubu (2000c) Nonlinear mating models for population with non-overlapping generations, *Department of Biometrics, Cornell University, Technical Report Series*, BU-1497-M.
 99. Castillo-Chavez, C. and A.A. Yakubu (2001a). Discrete-time S-I-S models with complex dynamics, *Nonlinear Analysis: Theory Methods and Applications*,: 4753–4762..
 100. Castillo-Chavez, C. and A.A. Yakubu (2001b). Discrete-time S-I-S models with simple and complex dynamics. In: *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods and Theory*, C. Castillo-Chavez with S. Blower, P. van den Driessche, D. Kirschner, and A.A. Yakubu, eds. Springer-Verlag, Berlin-Heidelberg-New York: pp. 153–163..
 101. Castillo-Chavez, C. and A.A. Yakubu (2001c). Intra-specific competition, dispersal and disease dynamics in discrete-time patchy environments. In: *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods and Theory*, C. Castillo-Chavez with S. Blower, P. van den Driessche, D. Kirschner, and A.A. Yakubu, eds. Springer-Verlag, Berlin-Heidelberg-New York: pp. 165–181.

102. Castillo-Chavez C., and A.A. Yakubu (2001d) Dispersal,disease and life history evolution, *Math. Biosc.*, **173**: 35–53.
103. Castillo-Garsow C., G. Jordan-Salivia, and A. Rodriguez-Herrera (2000) Mathematical models for the dynamics of tobacco use, recovery, and relapse, *Department of Biometrics, Cornell University, Technical Report Series*, BU-1505-M.
104. (2001) Census of India, 2001. Statistics of Bihar. Census of India.
105. Chowell, G., P. W. Fenimore, M. A. Castillo-Garsow and C. Castillo-Chavez (2003) SARS outbreaks in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism, *J. Theor. Biol.* **224**: 1-8.
106. Chowell, G., C. Castillo-Chavez, P.W. Fenimore, C. Kribs-Zaleta, L. Arriola, and J.M. Hyman (2004a) Model parameters and outbreak control for SARS, *Emerging Infectious Diseases* **10**: 1258-1263.
107. Chowell-Puente, D., P. Delgado, D. Perez, C. Tapia (2004b), The Impact of Mosquito-Bird Interaction on the Spread of West Nile Virus to Human Populations, *Department of Biometrics, Cornell University, Technical Report Series*, BU-01-8M.
108. Chowell, G., J.M. Hyman, L.M.A. Bettencourt, and C. Castillo-Chavez (Eds.) (2009) *Mathematical and Statistical Estimation Approaches in Epidemiology*. Springer.
109. Cintron-Arias, A., C. Castillo-Chavez, L. M. Bettencourt, A. L. Lloyd and H.T. Banks (2009) Estimation of the effective reproductive number from disease outbreak data, *Math. Biosc. & Eng.* **6**:261-282.
110. Clark, C.W. (1990) *Mathematical Bioeconomics: The Optimal Management of Renewable Resources*, 2nd ed., Wiley, New York.
111. Cohen, J.E. (1995) *How Many People Can the Earth Support?* W. W. Norton and Company, New York-London.
112. Communicable Disease Surveillance Centre (1978) Influenza in a boarding school, *British Medical Journal*, **1**: 587.
113. Cooke, K.L. and P. van den Driessche (1986) On zeroes of some transcendental functions, *Funkcialaj Ekvacioj*, **29**:77-90.
114. Cooke, K.L. and J. A. Yorke (1973) Some equations modeling growth processes and gonorrhea epidemics, *Math. Biosci.*, **58**: 93-109.
115. Costantino, R.F. and R. A. Desharnais (1991) *Population Dynamics and the 'Tribolium' Model: Genetics and Demography*. Mono. on Theor. Appl. Gen. V. 13, Springer-Verlag, Berlin.
116. Costantino, R.F., R.A. Desharnais, J.M. Cushing, and B. Dennis (1995) Experimentally induced transitions in the dynamic behavior of insect populations, *Nature*, **375**:227-230.
117. Costantino, R.F., R.A. Desharnais, J.M. Cushing, and B. Dennis (1995) Chaotic dynamics in an insect population, *Science*, **275**:389-391.
118. Cox, D.R. and G. F. Medley (1989) A process of events with notification delay and the forecasting of AIDS, *Phil. Trans. Roy. Soc. Lond. B*, **325**: 135-145.
119. Craddock, M. (2009) Fundamental solutions, transition densities and the integration of Lie symmetries, *J. Diff. Eqns.*, **207**: 2538-2560.
120. Crane, J. (1991) The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing, *Amer. J. Sociology*, **96**:1226-1259.
121. Cushing, J.M. (1998) *An Introduction to Structured Population Dynamics*, CBMS-NSF Regional Conference Series in Applied Mathematics **71**, SIAM, Philadelphia.
122. Cushing, J.M. and Jia Li (1991) Juvenile versus adult competition, *J. Math. Biol.*, **29**: 457–473.
123. Dean, A.M. (1983) A simple model of mutualism, *Amer. Naturalist*, **121**:409-417.
124. DeRoos, A., O. Diekmann, and J. A. J. Metz (1992) Studying the dynamics of structured population models: A versatile technique and its application to Daphnia, *Amer. Naturalist*, **139**:123-147.
125. (2004) Leishmania: Current situation and new perspectives, *Comparative Immunology, Microbiology and Infectious Diseases*, **27**: 305-318.
126. Devaney, R.L. (1992) *A First Course in Chaotic Systems: Theory and Experiment*, Addison-Wesley, Reading, Mass.

127. Diekmann, O., K. Dietz, and J.A.P. Heesterbeek (1991) The basic reproduction ratio for sexually transmitted diseases, Part 1: Theoretical considerations, *Math. Biosci.*, **107**: 325–339.
128. Diekmann O. and J.A.P. Heesterbeek (2000). *Mathematical epidemiology of infectious diseases: Model building, analysis and interpretation*, John Wiley and Sons, Ltd.
129. Diekmann, O., J.A.P. Heesterbeek, and J.A.J. Metz (1990) On the definition and the computation of the basic reproductive ratio R_0 in models for infectious diseases in heterogeneous populations, *J. Math. Biol.*, **28**:365–382.
130. Dietz, K. (1979) Epidemiologic interference of virus populations, *J. Math. Biol.*, **8**:291–300.
131. Dietz, K. (1982) Overall patterns in the transmission cycle of infectious disease agents. In: R.M. Anderson, R.M. May (eds) *Population Biology of Infectious Diseases*, Life Sciences Research Report **25**, Springer-Verlag, Berlin-Heidelberg-New York: pp. 87–102.
132. Dietz, K. (1988) The first epidemic model: A historical note on P. D. En'ko, *Australian J. Stat.*, **30**:56–65.
133. Dietz, K. (1988a) On the transmission dynamics of HIV, *Math. Biosci.* **90**:397–414.
134. Dietz, K. and K.P. Haderl (1988) Epidemiological models for sexually transmitted diseases, *J. Math. Biol.*, **26**:1–25.
135. Dietz, K. and J.A.P. Heesterbeek (1992) Daniel Bernoulli's epidemiological model revisited *Math. Biosci.* **180**: 1–21.
136. Dietz, K. D. Schenzle (1985) Proportionate mixing models for age-dependent infection transmission, *J. Math. Biol.*, **22**:117–120.
137. Dobson, A.P. (1988) The population biology of parasite-induced changes in host behavior. *The Quarterly Review of Biology*, **63**:139–165.
138. Dodson, S.I. (1974) Zooplankton competition and predation: An experimental test of the size-efficiency hypothesis, *Ecology*, **55**:605–613.
139. Dulac, H. (1934) Points Singulieres des Équations Differentielles, *Mém. Sci. Math.*, Fasc. **61**, Gauthier-Villars, Paris.
140. Durrett, R. and S. A. Levin (1994) Stochastic spatial models: A user's guide to ecological applications, *Phil. Trans. Roy. Soc. B*, **343**:329–350.
141. Dushoff, J., W. Huang, and C. Castillo-Chavez (1998) Backward bifurcations and catastrophe in simple models of fatal diseases, *J. of Math. Biol.*, **36**:227–248.
142. Dwyer, G., S.A. Levin, and L. Buttel (1990) A simulation model of the population dynamics and evolution of myxomatosis, *Ecological Monographs*, **60**:423–447.
143. Edelstein - Keshet, L. (1986) Mathematical theory for plant - herbivore systems, *J. Math. Biol.*, **24**: 25–58.
144. Edelstein-Keshet, L. (1988) *Mathematical Models in Biology*, Random House, New York. [reprinted as SIAM Classics in Applied Mathematics **46** (2005)].
145. Elaydi, S. (1996) *An Introduction to Difference Equations*, Springer-Verlag, Berlin-Heidelberg-New York.
146. Ellner, S., R. Gallant, and J. Theiler (1995) Detecting nonlinearity and chaos in epidemic data. In *Epidemic Models: Their Structure and Relation to Data*, D. Mollison, ed., Cambridge University Press, Cambridge, 229–247.
147. Elton, C.S. and M. Nicholson (1942) The ten year cycle in numbers of lynx in Canada, *J. Animal Ecology* **11**:215–244.
148. Engbert, R. and F. Drepper (1994) Chance and chaos in population biology-models of recurrent epidemics and food chain dynamics, *Chaos, Solutions & Fractals*, **4**(7):1147–1169.
149. En'ko, P.D. (1889) On the course of epidemics of some infectious diseases, *Vrach. St. Petersburg*:1008–1010, 1039–1042, 1061–1063.
150. Erdős, P. and A. Rényi (1959) On random graphs, *Publicationes Mathematicae*, **6**: 290–297.
151. Erdős, P. and A. Rényi (1960) On the evolution of random graphs, *Pub. Math. Inst. Hung. Acad. Science*, **5**: 17–61.
152. Erdős, P. and A. Rényi (1961) On the strengths of connectedness of a random graph, *Acta Math. Scientiae Hung.*, **12**: 261–267.
153. Evans, A.S. (1982) *Viral Infections of Humans*, 2nd ed., Plenum Press, New York.

154. Ewald, W.P. (1994) *The Evolution of Infectious Disease*, Oxford University Press, Oxford.
155. Feigenbaum, M.J. (1980) The metric universal properties of period doubling bifurcations and the spectrum for a route to turbulence. In *Nonlinear Dynamics*, R.H.G. Helleman, ed., *Ann. N.Y. Acad. Science*, **357**:330-336.
156. Feller, W. (1941) On the integral equation of renewal theory, *Ann. Math. Stat.*, **12**:243-267.
157. Feng, Z. (2007) Final and peak epidemic sizes for *SEIR* models with quarantine and isolation, *Math. Biosc. & Eng.* **4**: 675-686.
158. Feng, Z., C. Castillo-Chavez, and A. Capurro (2000) A model for TB with exogenous re-infection, *Journal of Theoretical Population Biology*, **5**:235-247.
159. Feng, Z., W. Huang and C. Castillo-Chavez (2001). On the role of variable latent periods in mathematical models for tuberculosis, *Journal of Dynamics and Differential Equations*, **13**: 425-452.
160. Feng, Z. D. Xu & W. Zhao (2007), Epidemiological models with non-exponentially distributed disease stages and applications to disease control, *Bull. Math. Biol.* **69**:1511-1536,
161. Fenner, F. and K. Myers (1978) Myxoma virus and myxomatosis in retrospect: the first quarter century of a new disease. In *Viruses and the Environment*, J.I. Cooper and F.O. MacCallum, eds., 539-570, Academic Press, London.
162. Fenner, F. and F.N. Ratcliffe (1965) *Myxomatosis*, Cambridge University Press, Cambridge.
163. Fibonacci, L. (1202) *Tipographia, delle Scienze Matematica*, Roma.
164. Fitzhugh, R. (1961) Impulses and physiological states in theoretical models of nerve membrane, *Biophys. Journ.*, **1**:445-466.
165. Foster, A., E. Hendryx, A. Murillo, M. Salas, E.J. Morals-Butler, S. Suslov, and M. Herrera-Valdez Extensions of the cable equation incorporating spatial dependent variations in nerve cell diameter (MTBI-07-01M)
166. Franke, J.E. and A.-A. Yakubu (1995) Extinction in systems of bobwhite quail populations, *Can. App. Math. Quarterly*, **3**:173-201.
167. Fredrickson, A.G. (1971) A mathematical theory of age structure in sexual populations: Random mating and monogamous marriage models, *Math. Biosci.*, **20**:117-143.
168. Freedman, H.I., G.J. Butler, and P. Waltman P. (1986) Uniformly persistent systems, *Proc. Amer. Math. Soc.*, **96**:425-430.
169. Freedman, H.I. and P. Waltman (1984) Persistence in models of three interacting predator-prey populations, *Math. Biosc.*, **68**:213-231.
170. Gambino, J., M.V. Martinez-Martinez, K. Salau, E.L. Soho, D.E. Hiebeler, F. Sanchez, and D. Murillo (2007), Cats protecting birds revisited with a spatial approach *MTBI-04-11M*.
171. Gause, G.F. (1934a) *The Struggle for Existence*, Williams and Wilkins, Baltimore.
172. Gause, G.F. (1934b) Experimental demonstration of Volterra's periodic oscillation in the numbers of animals, *J. Experimental Biology*, **11**:44-48.
173. Gilpin, M.E. (1973) Do hares eat lynx?, *Amer. Naturalist*, **107**:727-730.
174. Gladwell, M. (2000) *The Tipping Point*, Little, Brown and Co., Boston.
175. Glass, L. and M.C. Mackey (1979) Pathological conditions resulting from instabilities in physiological control systems, *Ann. N. Y. Acad. Science*, **316**:214-235.
176. Glass, L. and M.C. Mackey (1988) *From Clocks to Chaos: The Rhythms of Life*, Princeton University Press, Princeton, N. J.
177. Gompertz, B. (1825) On the nature of the function expressing the law of human mortality, *Phil. Trans.*, **115**:513-585
178. Gumel A., C. Castillo-Chavez, D.P. Clemence, and R.E. Mickens (eds.)(2006) *Mathematical Studies on Human Disease Dynamics* (A. Gumel, C. Castillo-Chavez, D.F. Clemence, and R. E. Mickens, eds.), Am. Math. Soc. **410**.
179. Gumel, A., S. Ruan, T. Day, J. Watmough, P. van den Driessche, F. Brauer, D. Gabrielson, C. Bowman, M.E. Alexander, S. Ardal, J. Wu, and B.M. Sahai (2004) Modeling strategies for controlling SARS outbreaks based on Toronto, Hong Kong, Singapore and Beijing experience, *Proc. Roy. Soc. London*, **271**: 2223-2232.
180. Gurney, W.S.C., S.P. Blythe, and R.M. Nisbet (1980) Nicholson's blowflies revisited, *Nature*, **287**:17-21.

181. Gurney, W.S. and R. M. Nisbet (1998) *Ecological Dynamics*, Oxford Univ. Press, Oxford.
182. Gurtin, M.L. and R.C. MacCamy (1974) Nonlinear age dependent population dynamics, *Arch. Rat. Mech. Analysis*, **54** (3):281-300.
183. Haderler, K.P. (1984). Integral equations with discrete parasites: hosts with a Lotka birth law, In: *Conf. Proc. Autumn Course on Math. Ecology, Trieste, 1982*, (S. Levin and T. Hallam, eds), 356-365, *Lecture Notes in Biomathematics*, **54** Springer-Verlag, Berlin-Heidelberg-New York.
184. Haderler, K.P. (1984) Population dynamics of killing parasites which reproduce in the host, *J. Math. Biol.*, **21**:5-55.
185. Haderler, K.P. (1989a) Pair formation in age-structured populations, *Acta Applicandae Mathematicae*, **14**:91-102.
186. Haderler, K.P. (1989b) Modeling AIDS in structured populations, *47th Session of the International Statistical Institute, Paris, August/September. Conf. Proc.*, **C1-2**:83-99.
187. Haderler, K.P. (1992) Periodic solutions of homogeneous equations, *J. Diff. Equations*, **95**:183-202.
188. Haderler, K.P. (1993), Pair formation with maturation period, *J. Math. Biol.*, **32**:1-15.
189. Haderler, K.P., and C. Castillo-Chavez (1995) A core group model for disease transmission, *Math Biosci.*, **128**:41-55.
190. Haderler, K.P. and J. Müller (1995) Vaccination in age structured populations I: The reproductive number. In *Models for Infectious Human Diseases: Their Structure and Relation to Data*, (V. Isham and G. Medley, Eds.), Cambridge University Press, Cambridge: pp. 90-101.
191. Haderler, K.P. and K. Nagoma (1990) Homogeneous models for sexually transmitted diseases, *Rocky Mountain J. Math.*, **20**:967-986.
192. Haderler, K.P. and P. van den Driessche (1997), Backward bifurcation in epidemic control, *Math. Biosci.* **146**: 15-35.
193. Haderler, K.P., R. Waldstätter, and A. Wörz-Busekros (1988) Models for pair-formation in bisexual populations, *J. Math. Biol.*, **26**:635-649.
194. Hanski, I.A. (1993) *Metapopulation Ecology*, Oxford University Press, Oxford.
195. Hanski, I.A. and M.E. Gilpin (1997) *Metapopulation Biology: Ecology, Genetics, and Evolution*, Academic Press, San Diego.
196. Hassell, M.P. (1975) Density dependence in single-species populations, *J. Animal. Ecol.*, **44**:283-295.
197. Hastings, A. (1983) Age-dependent predation is not a simple process, I. Continuous models, *Theor. Popul. Biol.*, **23**: 347-362.
198. Hayes, N.D. (1950) Roots of the transcendental equation associated with a certain differential-difference equation, *J. London Math. Soc.*, **25**:226-232.
199. Heesterbeek, J.A.P. (1992) R_0 , Thesis, CWI, Amsterdam.
200. Heesterbeek, J.A.P. and J.A.J Metz (1993) The saturating contact rate in marriage and epidemic models. *J. Math. Biol.*, **31**: 529-539.
201. Heesterbeek, J.A.P., and M.G. Roberts (1995) Threshold quantities for helminth infections, *J. Math. Biol.*, **33**: 425-434.
202. Heffernan, J.M., R.J. Smith? and L.M. Wahl (2005) Perspectives on the basic reproductive ratio, *J. Roy. Soc. Interface*, **2**: 281-293.
203. Hethcote, H.W. (1976) Qualitative analysis for communicable disease models, *Math. Biosci.*, **28**:335-356.
204. Hethcote, H.W. (1978) An immunization model for a heterogeneous population, *Theor. Pop. Biol.*, **14**:338-349.
205. Hethcote, H.W. (1989) Three basic epidemiological models. In *Applied Mathematical Ecology*, S. A. Levin, T.G. Hallam and L.J. Gross, eds., Biomathematics **18**, 119-144, Springer-Verlag, Berlin-Heidelberg-New York.
206. Hethcote, H.W. (1997) An age-structured model for pertussis transmission, *Math. Biosci.*, **145**: 89-136.
207. Hethcote, H.W. (2000) The mathematics of infectious diseases, *SIAM Review*, **42**: 599-653.

208. Hethcote, H.W. and S.A. Levin (1989) Periodicity in epidemic models. In : S.A. Levin, T.G. Hallam & L.J. Gross (eds) *Applied Mathematical Ecology*, Biomathematics **18**, Springer-Verlag, Berlin-Heidelberg-New York: pp. 193–211.
209. Hethcote, H.W., H.W. Stech and P. van den Driessche (1981) Periodicity and stability in epidemic models: a survey. In: S. Busenberg & K.L. Cooke (eds.) *Differential Equations and Applications in Ecology, Epidemics and Population Problems*, Academic Press, New York: 65–82.
210. Hethcote, H.W., and J.W. Van Ark. (1992) *Modeling HIV Transmission and AIDS in the United States*, Lect. Notes Biomath. **95**, Springer-Verlag, Berlin-Heidelberg-New York.
211. Hethcote, H.W., and J.A. Yorke (1984) *Gonorrhea Transmission Dynamics and Control*, Lect. Notes Biomath. **56**, Springer-Verlag, Berlin-Heidelberg-New York.
212. Hirsch, M. and S. Smale (1974) *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, Orlando, FL (1974).
213. Hodgkin, A.L. and A.F. Huxley (1952) A quantitative description of membrane current and its application to conduction and excitation in nerves, *J. Physiology*, **117**:500-544.
214. Hofbauer, J. (1981) General cooperation theorem for hypercycles, *Monatsheft Math.*, **91**: 223-240.
215. Holling, C.S. (1965) The functional response of predators to prey density and its role in mimicry and population regulation, *Mem. Entomol. Soc. Canada*, **45**: 1-60.
216. Holt, R.D. (1977) Predation, apparent competition, and the structure of prey communities, *Theor. Pop. Biol.*, **12**: 197-225.
217. Hoppensteadt, F.C. (1974) An age dependent epidemic model, *J. Franklin Inst.*, **297**: 325-333.
218. Hoppensteadt, F.C. (1975) *Mathematical Theories of Populations: Demographics, Genetics and Epidemics*, SIAM, Philadelphia.
219. Hsu Schmitz, S-F (1993) Some theories, estimation methods and applications of marriage functions in demography and epidemiology, *Ph. D. dissertation*, Cornell University.
220. Hsu Schmitz, S-F and C. Castillo-Chavez (1994) Parameter estimation in non-closed social networks related to the dynamics of sexually-transmitted diseases. In *Modeling the AIDS Epidemic: Planning, Policy, and Prediction*, E.H. Kaplan and M.L. Brandeau eds., 533-559, Raven Press, New York.
221. Hsu Schmitz, S-F and C. Castillo-Chavez, (1996) Completion of mixing matrices for non-closed social networks. In *Proceedings of the First World Congress of Nonlinear Analysts*, V. Lakshmikantham, ed., 3163-3173, de Gruyter, Berlin.
222. Huang, W., K. Cooke, and C. Castillo-Chavez (1990) Stability and bifurcation for a multiple group model for the dynamics of HIV/AIDS transmission, *SIAM J. of Applied Math.*, **52**(3):835-854.
223. Hudson, P.J., and A. P. Dobson (1995) Macroparasites: Observed patterns in naturally fluctuating animal populations. In: *Ecology of Infectious Diseases in Natural Populations* (B. T. Grenfell and A. P. Dobson, eds), 144-176, Cambridge University Press.
224. Huffaker, C.B. (1958) Experimental studies on predation: Dispersion factors and predator-prey oscillations, *Hilgardia* **27**:343-383.
225. Hurewicz, W. (1958) *Lectures in Differential Equations*, MIT Press, Cambridge.
226. Huang, W., K.L. Cooke, and C. Castillo-Chavez (1992) Stability and bifurcation for a multiple group model for the dynamics of HIV/AIDS transmission, *SIAM J. App. Math.*, **52**:835-854.
227. Hutchinson, G.E. (1948) Circular casual systems in ecology, *Ann. N.Y. Acad. Sci.*, **50**:221-246.
228. Hutson, V. and G.T. Vickers (1983) A criterion for permanent coexistence of species with an application to a two-prey one-predator system, *Math. Biosci.*, **63**:253-269.
229. Hyman, J.M. and Jia Li (2000) An intuitive formulation for the reproductive number for the spread of diseases in heterogeneous populations, *Math. Biosci.*, **167**: 65–86.
230. Hyman, J.M., Jia Li, and E. A. Stanley (1994) Threshold conditions for the spread of the HIV infection in age-structured populations of homosexual men, *J. Theor. Biol.*, **166**: 9–31.

231. Hyman, J.M., J. Li & E. A. Stanley (1999) The differential infectivity and staged progression models for the transmission of HIV, *Math. Biosci.*, **155**: 77–109.
232. Hyman, J.M. and E. A. Stanley (1988) Using mathematical models to understand the AIDS epidemic, *Math. Biosci.*, **90**: 415–474.
233. Iannelli, M. (1995) *Mathematical Theory of Age-Structured Population Dynamics*, Appl. Mathematical Monographs, C.N.R..
234. Isham, V. (1989) Estimation of the incidence of HIV infection, *Phil. Trans. Roy. Soc. Lond. B*, **325**: 113–121.
235. Ivlev, V.S. (1961) *Experimental Ecology of the Feeding of Fishes*, Yale University Press, New Haven.
236. (1983) Jack, J.J.B., D. Noble and R.W. Tsien (1983) *Electric Current Flow in Excitable Cells*, Oxford, UK.
237. Jacquez, J.A., C. P. Simon, and J. Koopman (1991) The reproductive number in deterministic models of contagious diseases, *Comm. Theor. Biol.*, **2**: 159–209.
238. Jewell, N.P., K. Dietz, and V.T. Farewell. (1991) *AIDS Epidemiology: Methodological Issues*, Birkhäuser, Boston-Basel-Berlin.
239. Kaplan, D. and L. Glass (1995) *Understanding Nonlinear Dynamics*, Springer-Verlag, Berlin-Heidelberg-New York.
240. Kendall, D.G. (1949) Stochastic processes and population growth, *J. Roy. Stat. Soc., Ser. B* **2**:230–264.
241. Kermack, W.O. and A.G. McKendrick (1927) A contribution to the mathematical theory of epidemics, *Proc. Royal Soc. London*, **115**:700–721.
242. Kermack, W.O. and A.G. McKendrick (1932) Contributions to the mathematical theory of epidemics, part. II, *Proc. Roy. Soc. London*, **138**:55–83.
243. Kermack, W.O. and A.G. McKendrick (1933) Contributions to the mathematical theory of epidemics, part. III, *Proc. Roy. Soc. London*, **141**:94–112.
244. Keyfitz, N. (1949) The mathematics of sex and marriage, *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, **4**:89–108.
245. (2000) Keymer, J.E., P.A. Marquet, J.X. Velasco-Hernandez, and S.A. Levin, Extinction thresholds and metapopulation persistence in dynamic landscapes, *Amer. Naturalist*, **156**: 478–494.
246. Kierstad, H. and L.B. Slobodkin (1953) The size of water masses containing plankton blooms, *J. Mar. Res.*, **12**:141–147.
247. Kingsland, S.E. (1985) *Modeling Nature: Episodes in the History of Population Ecology*, University of Chicago Press, Chicago.
248. Kolmogorov, A.N. (1936) Sulla teoria di Volterra della lotta per l'esistenza, *Giorn. Inst. Ital. Attuari*, **7**:74–80.
249. Kostova, T. and Jia Li (1996) Oscillations and stability due to juvenile versus adult competition, *Int. J. Comput. Math. Applic.*, **32**: 57–70.
250. Kostova, T., Jia Li, and M. Friedman (1999) Two models for competition between age classes, *Math. Biosci.*, **157**: 65–89.
251. Kurtz, T.G. (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes, *J. Appl. Probab.*, **7**:49–58.
252. Kurtz, T.G. (1971) Limit theorems for sequences of jump Markov processes approximating differential equations, *J. Appl. Prob.*, **8**:344–356.
253. Kretzschmar, M. (1989) A renewal equation with birth-death process as a model for parasitic infections, *J. Math. Biol.*, **27**:191–221.
254. Kretzschmar, M. (1993) Comparison of an infinite dimensional model for parasitic diseases with a related 2-dimensional system, *J. Math. Anal. Appl.*, **173**:235–260.
255. Kribs-Zaleta, C.M. and J.X. Velasco-Hernandez (2000), A simple vaccination model with multiple endemic states, *Math Biosci.* **164**: 183–201.
256. Lagakos, S.W., L. M. Barraj, and V. de Gruttola (1988) Nonparametric analysis of truncated survival data, with applications to AIDS, *Biometrika*, **75**: 515–523.
257. Lajmanovich, A. and J.A. Yorke (1976) A deterministic model for gonorrhea in a nonhomogeneous population, *Math. Biosci.*, **28**:221–236.

258. Lee, E.B. and L. Markus (1967) *Foundations of Optimal Control Theory*, Wiley, New York.
259. Leslie, P.H. (1945) On the use of matrices in certain population mathematics, *Biometrika*: **33**,183-212.
260. Leslie, P.H. (1948) Some further notes on the use of matrices in population mathematics, *Biometrika*, **35**:213-245.
261. Levin, J.J. and D.F. Shea (1972) On the asymptotic behaviour of the bounded solutions of some integral equations I, II, III, *J. Math. Anal. & Appl.*, **37**: 42-82, 288-326, 537-575.
262. Levin, S.A. (1970) Community equilibria and stability, and an extension of the competitive exclusion principle, *Am. Naturalist*, **104**:413-423.
263. Levin, S.A. (1983a) Co-evolution. In *Population Biology*, H.I. Freedman and C. Strobeck, eds., Lect. Notes Biomath. **52**, 328-334, Springer-Verlag, Berlin-Heidelberg-New York.
264. Levin, S.A. (1983b) Some approaches to the modeling of co-evolutionary interactions. In *Co-evolution*, M. Nitecki, ed., 21-66, University of Chicago Press, Chicago.
265. Levin, S.A. and R. M. May (1976) A note on difference-delay equations, *Theor. Pop. Biol.*, **9**:178-187.
266. Levin, S.A. and D. Pimentel (1981) Selection of intermediate rates of increase in parasite-host systems, *Amer. Naturalist*, **117**:308-315.
267. Levin, S.A., T.M. Powell, and J.H. Steele (eds.) (1993) *Patch Dynamics*, Lecture Notes in Biomathematics **96**, Springer Verlag.
268. Levins, R. (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control, *Bull. Entom. Soc. Amer.*, **15**: 237-240.
269. Lewis, E.G. (1942) On the generation and growth of a population, *Sankhyā*, **6**:93-96.
270. Li, Jia and T. G. Hallam (1988) Survival in continuous structured population models, *J. Math. Biol.*, **26**: 421-433.
271. Li, T.Y. and J.A. Yorke (1975) Period three implies chaos, *Amer. Math. Monthly*, **82**:985-992.
272. Lima, S. and M. Torres (1997) Models for the Transmission Dynamics of Gonorrhea in a Homosexually-Active Population, *Department of Biometrics, Cornell University, Technical Report Series*, BU-1426-M.
273. Londen, S-O. (1974) On the asymptotic behavior of the bounded solutions of a nonlinear Volterra equation, *S.I.A.M. J. Math. Anal.*, **5**:849-875.
274. Lorenz, E.N. (1963) Deterministic non-periodic flow, *J. Atmos. Science*, **20**:130-141.
275. Lotka, A.J. (1922) The stability of the normal age distribution, *Proc. Nat. Acad. Sci.*, **8**:339-345.
276. Lotka, A.J. (1925) *Elements of Physical Biology*, Williams & Wilkins, Baltimore; Reissued as *Elements of Mathematical Biology*, Dover, New York (1956).
277. Lotka, A.J. and F.R. Sharpe (1923) Contributions to the analysis of malaria epidemiology, *Am. J. Hygiene*, **3**, Supplement: 1-21.
278. Lubkin, S. and C. Castillo-Chavez (1996) A pair formation approach to modeling inheritance of social traits. In *Proceedings of the First World Congress of Nonlinear Analysts, Tampa, Florida*, V. Lakshmikantham, ed., 3227-3234, de Gruyter, Berlin.
279. Ludwig, D., D.D. Jones, and C.S. Holling (1978) Qualitative analysis of insect outbreak systems: The spruce budworm and forest, *J. Animal Ecol.*, **47**:315-332.
280. Luo, X. and C. Castillo-Chavez, (1993) Limit behavior of pair-formation for a large dissolution rate, *J. Math. Systems, Estimation, and Control*, **3**:247-264.
281. McFarland, D. (1972) Comparison of alternative marriage models. In *Population Dynamics*, T.N.E. Greville, ed., 89-106, Academic Press, New York-London.
282. McKendrick, A.G. (1926) Applications of mathematics to medical problems, *Proc. Edinburgh Math. Soc.*, **40**:98-130.
283. MacLulich, D.A. (1937) *Numbers of the Varying Hare*, Univ. of Toronto Studies, Biological Series, No. **43**, University of Toronto Press, Toronto.
284. McNeill, W.H. (1976) *Plagues and Peoples*, Doubleday, New York.
285. McNeill, W.H. (1992) *The Global Condition*, Princeton University Press, Princeton, N.J.
286. Mackey, M.C. and L. Glass (1977) Oscillation and chaos in physiological control systems, *Science*, **197**:287-289.

287. Malthus, T.R. (1798) *An Essay on the Principle of Population*, 1a Ed, J Johnson in St Paul's Churchyard, London.
288. Markus, L. (1956) Asymptotically autonomous differential systems. In: S. Lefschetz (ed) *Contributions to the Theory of Nonlinear Oscillations III*. Annals of Mathematics Studies , Princeton University Press, Princeton, N.J. **36**: 17–29.
289. May, R.M. (1974) *Stability and Complexity in Model Ecosystems*, Princeton University Press, Princeton, N.J.
290. May, R.M. (1976) Simple mathematical models with very complicated dynamics, *Nature*, **261**:459-467.
291. May R. (1981) *Theoretical ecology, principles and applications*, Second Edition, Sinauer Associates, Massachusetts.
292. May, R.M. (1986) Population dynamics of microparasitic infections. In *Mathematical Ecology; An Introduction*, T.G. Hallam and S.A. Levin, eds., Biomathematics **17**, 405-442, Springer-Verlag, Berlin-Heidelberg-New York.
293. May, R.M., and R. M. Anderson (1978) Regulation and stability of host-parasite population interactions. II. Destabilizing processes, *J. Animal Ecology*, **47**:249-267.
294. May, R.M. and R.M. Anderson (1983) Epidemiology and genetics in the co-evolution of parasites and hosts, *Phil. Trans. Royal Soc. London, Series B*, **219**:281-313.
295. May, R.M. and W.J. Leonard (1975) Nonlinear aspects of competition between species, *SIAM J. App. Math.*, **29**:243-253.
296. Maynard Smith, J. (1974) *Models in Ecology*, Cambridge University Press, Cambridge.
297. Medley, G.F., R. M. Anderson, D. R. Cox, and L. Billard (1987) Incubation period of AIDS in patients infected via blood transfusion, *Nature* **328**: 719-721.
298. Mena-Lorca, J. and H.W. Hethcote (1992) Dynamic models of infectious diseases as regulators of population size. *J. Math. Biol.*, **30**: 693–716.
299. Meyers, L.A. (2007) Contact network epidemiology: Bond percolation applied to infectious disease prediction and control, *Bull. Am. Math. Soc.*, **44**: 63–86.
300. Meyers, L.A., M.E.J. Newman and B. Pourbohloul(2006) Predicting epidemics on directed contact networks, *J. Theor. Biol.*, **240**: 400–418.
301. Meyers, L.A., B. Pourbohloul, M.E.J. Newman, D.M. Skowronski, and R.C. Brunham (2005), Network theory and SARS: predicting outbreak diversity, *J. Theor. Biol.*, **232**: 71–81.
302. Metz, J.A.J and O. Diekmann (1986) *The Dynamics of Physiologically Structured Populations*, Lect. Notes Biomath., **68**, Springer-Verlag, Berlin-Heidelberg-New York.
303. Miller, J.C. (2010) A note on a paper by Erik Volz: SIR dynamics in random networks. *J. Math. Biol.* DOI 10.1007/s00285-010-0337-9.
304. Miller, J.C. and E. Volz (2011) Simple rules govern epidemic dynamics in complex networks, to appear.
305. Miller, W. Jr. (1977) *Symmetry and Separation of Variables*, Encyclopedia of Mathematics and Its Applications **4**, Addison-Wesley, Reading MA.
306. Milton, J.G. and J. Bélair (1990) Chaos, noise and extinction in models of population growth, *Theor. Pop. Biol.*, **37**:273-290.
307. Mollison D.(ed.) (1995) *Epidemic Models: their structure and relation to data*, Cambridge University Press.
308. Monod, J. (1950) La technique de culture continue: Théorie et applications, *Ann. Inst. Pasteur*, **79**:390-410.
309. (2010) Mubayi A, G. Chowell, C. Castillo Chavez, C. Kribs-Zaleta, N. Siddiqui, N. Kumar, and P. Das, Transmission dynamics and underreporting of Kala azar in the Indian State of Bihar. *J Theor Biol.*, **262**: 177-185.
310. Murray, J.D. (1989) *Mathematical Biology*, Biomathematics **19**, Springer-Verlag, Berlin-Heidelberg-New York.
311. Murray, H.W, J.D. Berman, C.R. Davies, et al. (2005) Advances in leishmaniasis, *Lancet*, **366**: 1561-77.
312. Nagumo, J.S., S. Arimoto, and S. Yoshizawa (1962) An active pulse transmission line simulating nerve axon, *Proc. Inst. Radio Engineers*, **50**:2061-2071.

313. Communicable Diseases Australia, National Notifiable Diseases Surveillance System. <http://www9.health.gov.au/cda/Source/CDA-index.cfm>.
314. Neuhauser, C. (2001) Mathematical challenges in spatial ecology, *Notices Amer. Math. Soc.*, **48**: 1304–1314.
315. M.E.J. Newman (2002) The spread of epidemic disease on networks, *Phys. Rev. E*, **66**, 016128.
316. Newman, M.E.J. (2003) The structure and function of complex networks. *SIAM Review*, **45**: 167–256.
317. Newman, M.E.J., S.H. Strogatz & D.J. Watts (2001) Random graphs with arbitrary degree distributions and their applications, *Phys. Rev. E*, **64**, 026118.
318. Nicholson, A.J. (1954) An outline of the dynamics of animal populations, *Australian. J. Zoology*, **3**:9-65.
319. Nicholson, A.J. and V. A. Bailey (1935) The balance of animal populations, *Proc. Zool. Soc. London*, **3**:551-598.
320. Nisbet, R.M. and W.S.C. Gurney (1982) *Modeling Fluctuating Populations*, Wiley-Interscience, Chichester.
321. Nisbet, R.M. and W.S.C. Gurney (1983) The systematic formulation of population models for insects with dynamically varying instar duration, *Theor. Pop. Biol.*, **23**:114-135.
322. Nisbet, R.M. and W.S.C. Gurney (1986) The formulation of age structure models. In *Mathematical Ecology; An Introduction*, T.G. Hallam and S.A. Levin, eds., Biomathematics **17**, 95-115, Springer-Verlag, Berlin-Heidelberg-New York.
323. Nisbet, R.M. and L. Onyiah (1994) Population dynamic consequences of competition within and between age classes. *J. Math. Bio.*, **32**: 329–344.
324. Nishiura H, C. Castillo-Chavez, M. Safan, and G. Chowell (2009) Transmission potential of the new influenza A(H1N1) virus and its age-specificity in Japan. *Euro Surveill* **14**: pii = 19227.
325. Nishiura H, G. Chowell, and C. Castillo-Chavez (2011) Did modeling overestimate the transmission potential of pandemic (H1N1-2009)? Sample size estimation for post-epidemic seroepidemiological studies. *PLoS ONE* **6**: e17908. doi:10.1371/journal.pone.0017908
326. Nishiura H, G. Chowell, M. Safan, and C. Castillo-Chavez (2010) Pros and cons of estimating the reproduction number from early epidemic growth rate of influenza A (H1N1) 2009. *Theor Biol Med Model* **7**: 1.
327. Nold, A. (1980) Heterogeneity in diseases-transmission modeling, *Math. Biosci.*, **52**:227-240.
328. Novick, A. and L. Szilard (1950) Experiments with the chemostat on spontaneous mutation of bacteria, *Proc. Nat. Acad. Sci. USA*, **36**:708-719.
329. Oaten, A. and W.W. Murdoch (1975) Switching, functional response and stability in predator-prey systems, *Amer. Naturalist*, **109**:299-318.
330. Odum, E.P. (1953) *Fundamentals of Ecology* Saunders, Philadelphia.
331. Ortiz, A. R., H. T. Banks, C. Castillo-Chavez, G. Chowell, and X. Wang (2011) A discrete events delay differential system model for transmission of vancomycin-resistant enterococcus (VRE) in hospitals, *J. Inv. Ill-Posed Problems* **18**, to appear.
332. Ortiz, A. R., H. T. Banks, C. Castillo-Chavez, G. Chowell, and X. Wang (2011b) A deterministic methodology for estimation of parameters in dynamic Markov chain models, *J. Biol. Systems*, (in press).
333. Oster, G. (1978) The dynamics of nonlinear models with age structure. In *Studies in Mathematical Biology*, S.A. Levin, ed., Vol **16**, 411-438, Math. Assoc. of America.
334. Paine, R.T. (1966) Food web complexity and species diversity, *Amer. Naturalist*, **100**:65-75.
335. Parlett, P. (1972) Can there be a marriage function? In *Population Dynamics*, T.N.E. Greville, ed., 107-135, Academic Press, New York-London.
336. Park, T. (1948) Experimental studies of interspecies competition: I. Competition between populations of the flour beetles, *tribolium confusum* Duvall and *tribolium castaneum* Herbst, *Ecological Monographs*, **18**:267-307.
337. Pearl, R. and L.J. Reed (1920) On the rate of growth of the population of the United States since 1790 and its mathematical representation, *Proc. Nat. Acad. Sci.*, **6**:275-288.

338. Poincaré, H. (1881) Sur les courbes définies par des équation différentielles, *J. Math. Pures Appl.*, Vol. **1**(1881):375-427, Vol. **8**(1882):251-286, Vol. **11**(1885):162-244, Vol. **12**(1886):151-217.
339. Pollard, J.H. (1973) *Mathematical Models for the Growth of Human Populations*, Cambridge University Press, Cambridge.
340. Pontryagin, L.S., V.S. Boltyanskii, R.V. Gamkrelidze, and E.F. Mischenko (1962) *The Mathematical Theory of Optimal Processes*, Wiley-Interscience, New York. (translated from the Russian).
341. Raggett, G.F. (1982) Modeling the Eyam plague, *IMA Journal*, **18**:221-226.
342. Real, L.A. and J. H. Brown (1991) *Foundations of Ecology: Classic Papers with Commentaries*, University of Chicago Press, Chicago.
343. Renshaw, E. (1991) *Modeling Biological Populations in Space and Time*, Cambridge University Press, Cambridge.
344. Ricker, W.E. (1954) Stock and recruitment, *J. Fisheries Res. Board Canada*, **11**:559-623.
345. Ricker, W.E. (1958) *Handbook of Computations for Biological Statistics of Fish Populations*, Bull. Fisheries Res. Board Canada, 119.
346. Riley, S., C. Fraser, C.A. Donnelly, A.C. Ghani, L.J. Abu-Raddad, A.J. Hedley, G.M. Leung, L-M Ho, T-H Lam, T.Q. Thach, P. Chau, K-P Chan, S-V Lo, P-Y Leung, T. Tsang, W. Ho, K-H Lee, E.M.C. Lau, N.M. Ferguson, & R.M. Anderson (2003) Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions, *Science*, **300**: 1961-1966.
347. Rivera-Salguero, B. (2000) Scramble versus contest competition in a two patch system. *Department of Biometrics, Cornell University, Technical Report Series*, BU-1515-M.
348. Roche, P.W., V. Krause, M. Bartlett, D. Coleman, H. Cook, M. Counahan, C. Davis, L. Del Fabbro, C. Geile, R. Gilmore, R. Kampen, and M. Young (2004) Invasive pneumococcal disease in Australia, 2003, *Communicable Diseases Intelligence* **28**: 441-454.
349. Roche, P.W., V. L. Krause, M. Bartlett, D. Coleman, H. Cook, C. Davis, J. E. Fielding, R. Holland, C. Giele, R. Gilmour, and R. Kampen (2006) Invasive pneumococcal disease in Australia, 2004, *Communicable Diseases Intelligence*, **30**: 80-92.
350. Roche, P.W., V. Krause, and H. Cook (2007) Invasive pneumococcal disease in Australia, 2005, *Communicable Diseases Intelligence* **31**:86-100.
351. Rosenzweig, M.L. (1969) Why the prey curve has a hump, *Amer. Naturalist*, **103**:81-87.
352. Rosenzweig, M.L. (1971) Paradox of enrichment: Destabilization of exploitation ecosystems in ecological time, *Science*, **171**:385-387.
353. Rosenzweig, M.L. and R.H. MacArthur (1963) Graphical representation and stability conditions of predator-prey interactions, *Amer. Naturalist*, **97**:209-223.
354. Ross, R. (1911) *The Prevention of Malaria*, 2nd ed., (with Addendum), John Murray, London.
355. Saha, P. and S. Strogatz (1995) The birth of period 3, *Math. Mag.*, **68**:42-47.
356. Saints, K. (1987) Discrete and continuous models of age-structured population dynamics, *Senior Research Report*, Harvey Mudd College, Claremont, Cal.
357. Samuelson, P.A. (1941) Conditions that a root of a polynomial be less than unity in absolute value, *Ann. Math. Stat.*, **12**:360-364.
358. Sánchez B.N., P.A. Gonzalez, R.A. Saenz (2000) The influence of dispersal between two patches on the dynamics of a disease, *Department of Biometrics, Cornell University, Technical Report Series*, BU-1531-M.
359. Sánchez, D.A. (1978) Linear age-dependent population growth with harvesting, *Bull. Math. Biol.*, **40**:377-385.
360. Sánchez, D.A. (1968) *Ordinary Differential Equations and Stability Theory: An Introduction*, W. H. Freeman, San Francisco; reprinted by Dover, New York (1979).
361. Sandefur, J. (1990) *Discrete Dynamical Systems: Theory and Applications*, Oxford University Press, Oxford.
362. Schoener, T.W. (1973) Population growth regulated by intraspecific competition for energy or time: Some simple representations, *Theor. Pop. Biol.*, **4**:56-84.

363. Sharpe, F.R. and A.J. Lotka (1911) A problem in age distribution, *Phil. Magazine*, **21**:435-438.
364. Shim, E., H.T. Banks, and C. Castillo-Chavez (2006) Seasonality of rotavirus infection with its vaccination, *Mathematical Studies on Human Disease Dynamics* (A. Gumel, C. Castillo-Chavez, D.F. Clemence, and R. E. Mickens, eds.), Am. Math. Soc. **410**: 327-348.
365. Shulgin, B., L. Stone, and Z. Agur (1998) Pulse vaccination strategy in the SIR epidemic model, *Bull. Math. Biol.*, **60**:1123-1148.
366. Skellam, J.G. (1951) Random dispersal in theoretical populations, *Biometrika*, **38**:196-218.
367. Smith, F.E. (1963) Population dynamics in *Daphnia magna* and a new model for population growth, *Ecology*, **44**:651-663.
368. Smith, H.L. (1988) Systems of ordinary differential equations which generate an order preserving flow: A survey of results, *SIAM Rev.*, **30**:87-113.
369. Smith, H.L. and P. Waltman (1996) *The Theory of the Chemostat: Dynamics of Microbial Competition*, Cambridge Studies in Mathematical Biology **13**, Cambridge University Press, Cambridge.
370. Soberon, J.M. and C.M. Del Rio (1981) The dynamics of a plant-pollinator interaction, *J. Theor. Biol.*, **91**:363-378.
371. Soper, H.E. (1929) Interpretation of periodicity in disease prevalence, *J. Roy. Stat. Soc., Series B*, **92**:34-73.
372. Stone, L., B. Shulgin, Z. Agur (2000) Theoretical examination of the pulse vaccination in the SIR epidemic model, *Math. and Computer Modeling*, **31**:207-215.
373. Strogatz, S.H., (1994) *Nonlinear Dynamics and Chaos*, Addison-Wesley, Reading, Mass.
374. Strogatz, S.H. (2001) Exploring complex networks, *Nature*, **410**: 268-276.
375. Suazo, E., S.K. Suslov and J.M. Vega-Guzmán (2011a) The Riccati equation and a diffusion-type equation, *New York J. Math.*, **17a**: 225-244.
376. Suazo, E., S.K. Suslov and J.M. Vega-Guzmán (2011b) The Riccati system and a diffusion-type equation, in preparation.
377. Sud, A., N. Varma, R.K. Marwaha, F.M. Patel, A. Trehan, S. Singh, and S. Varma (2004) Visceral leishmaniasis in a non-endemic area of India, *Trop. Doctor*, **34**: 247-249.
378. Sutton, K.L., H. T. Banks, and C. Castillo-Chávez. (2008) Estimation of invasive pneumococcal disease dynamic parameters and the impact of conjugate vaccination in Australia, *Math. Biosc. & Eng.*, **5**: 175-204.
379. Sutton, K. L., H. T. Banks and C Castillo-Chavez (2010a) Using inverse problem methods with surveillance data in pneumococcal vaccination, *Math. & Computer Modelling* **51**:369-388.
380. Sutton, K, H.T. Banks and C Castillo-Chavez (2010b) Public vaccination policy using an age-structure model of pneumococcal infection dynamics, *J. Biol. Dyn.* **4**:169-195.
381. Thieme, H.R. (1994) Asymptotically autonomous differential equations in the plane, *Rocky Mountain J. Math.*, **24**:351-380.
382. Thieme, H.R. (2003) *Mathematics in Population Biology*, Princeton University Press, Princeton.
383. Thieme, H.R. and C. Castillo-Chavez (1989) On the role of variable infectivity in the dynamics of the human immunodeficiency virus. In *Mathematical and statistical approaches to AIDS epidemiology*, C. Castillo-Chavez, ed., Lect. Notes Biomath. **83**, 200-217, Springer-Verlag, Berlin-Heidelberg-New York.
384. Thieme, H.R. and C. Castillo-Chavez (1993) How may infection-age dependent infectivity affect the dynamics of HIV/AIDS?, *SIAM J. Appl. Math.*, **53**:1447-1479.
385. Tschumy, W.O. (1982) Competition between juveniles and adults in age-structured populations, *Theor. Popul. Biol.*, **21**: 255-268.
386. Tudor, D.W. (1985) An age-dependent epidemic model with application to measles, *Math. Biosci.*, **73**: 131-147.
387. U. S. Bureau of the Census (1975) *Historical statistics of the United States: colonial times to 1970*, Washington, D. C. Government Printing Office.
388. U.S. Bureau of the Census (1980) *Statistical Abstracts of the United States*, 101st edition.

389. U.S. Bureau of the Census (1991) *Statistical Abstracts of the United States*, 111th edition.
390. U.S. Bureau of the Census (1999) *Statistical Abstracts of the United States*, 119th edition.
391. van den Driessche, P. and J. Watmough (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosc.*, **180**: 29–48.
392. Velazquez, J.P. (2000) SIS nonlinear discrete time models with two competing strains. *Department of Biometrics, Cornell University, Technical Report Series*, BU-1514-M.
393. Verhulst, P.F. (1838) Notice sur la loi que la population suit dans son accroissement, *Corr. Math. et Phys.*, **10**:113–121.
394. Verhulst, P.F. (1845) Recherches mathématiques sur la loi d'accroissement de la population, *Mém Acad Roy, Brussels*, **18**:1–38.
395. Verhulst, P.F. (1847) Recherches mathématiques sur la loi d'accroissement de la population, *Mém Acad Roy, Brussels*, **20**:1–32.
396. Volterra, V. (1926) Variazioni e fluttuazioni del numero d'individui in specie animali conviventi, *Mem. Acad. Sci. Lincei*, **2**:31–13.
397. Volterra, V. (1931) *Leçons sur la Théorie Mathématique de la Lutte pour la Vie*, Gauthier-Villars, Paris.
398. Volz, E. (2008) SIR dynamics in random networks with heterogeneous connectivity, *J. Math. Biol.*, **56**: 293–310.
399. Waldstätter, R. (1989) Pair formation in sexually transmitted diseases. In *Mathematical and Statistical Approaches to AIDS Epidemiology*, C. Castillo-Chavez, ed., Lect. Notes Biomath. **83**, 260–274, Springer-Verlag, Berlin-Heidelberg-New York.
400. Waldstätter, R. (1990) Models for pair formation with applications to demography and epidemiology, Ph. D. thesis, University of Tübingen.
401. Waltman, P. (1986) *A Second Course in Ordinary Differential Equations*, Academic Press, Orlando, Fla.
402. Waltman, P. (1974) *Deterministic Threshold Models in the Theory of Epidemics*, Lect. Notes in Biomath., **1**, Springer-Verlag, Berlin-New York.
403. Watson, G.N. (1944) *A Treatise on the Theory of Bessel Functions*, Second Edition, Cambridge University Press, Cambridge.
404. Webb, G.F. (1985) *Theory of Nonlinear Age-Dependent Population Dynamics*, Marcel Dekker, New York.
405. Wilkie, A.D. (1989) Population projections for AIDS using an actuarial model, *Phil. Trans. Roy. Soc. Lond. B* **325**: 99–112.
406. Yodzis, P. (1989) *Introduction to Theoretical Ecology*, Harper and Row, New York.
407. Zeng, D., H. Chen, C. Castillo-Chavez, W. B. Lober and M. Thurmond (2011) *Infectious Diseases Bioinformatics and Biosurveillance*, Integrated Series in Information Systems **27**, Springer.
408. Zerpa O, M. Ulrich, R. Borges, et al. (2003) Epidemiological aspects of human and canine visceral leishmaniasis in Venezuela, *Pan American Journal of Public Health*, **13**: 239–245.

Index

\mathcal{R}_0 , *see* basic reproductive number

absorption, 24, 31
action potential, 119, 120, 219
activator, 328
adjoint variable, 255, 262, 263
age at infection, 421, 422
age distribution
 stable, 270–273, 278, 279, 287
 persistent, 278
age group, 289
age of infection, 351, 384
age progression rate, 289
age structure, 288
age-structured population, 79
agriculture, 31, 37, 432
AIDS, 345–347, 377, 411
Allee effect, 22, 55
allele, 332
analytic expressions, 17
Antarctic fin whale, 74
antibiotic, xiv
Antonine plagues, 347
aperiodic, *see* periodic
Assyria, 346
asymptotic stability, 20, *see* stability
asymptotic velocity, 322
attack rate, 354
autocatalytic reaction, 332
autonomous, 18
Aztecs, 347, 424

backward bifurcation, 441
bacteria, xviii, 346, 350
bacteria agents, 411
balance equation, 302, 330
bang-bang, *see* control

 control, 238
basic reproduction number, 353, 368, 412, 418
basic reproductive number, xv, 5
Bayes's theorem, 40
Bendixson theorem, 152
Bernardelli population wave, 272
Bernouilli, 349
Beverton-Holt model, 69
bifurcation
 backward, 441
 period-doubling, 88
 Hopf, 431
bifurcation curve, 44
biomass, 3
bionomic equilibrium, 244
birth, 432
birth control, xiii
birth cycle, 68
birth modulus, 274
birth rate, 8, 134, 289
 per capita, xiv, 4
births, 349, 411, 413
bistable, 45
Black Death, *see* bubonic plague
blood cells, 116
blood sample, 356
blood stream, xv
blood–brain barrier, 144
boundary condition, 303
branching process, 351, 362, 373
bubonic plague, xiii, 347, 356, 432, 434
Burgers equation, 339

cable equation, 333, 334
cannibalism, 80, 81
canonical form, 156
capacitance, 334

- capital asset, 261, 263
- carriers, 414
- carrying capacity, xiv, xv, 9, 413
- catastrophe, 27, 224
- catastrophic outbreak, 203
- Cauchy initial value problem, 336, 340
- census, 7, 10, 81, 448
- center, 140, 167, 188
- Chagas, 346
- chaos, 63
- chaotic behavior, 61, 65
- characteristic equation, 140
- characteristic equation, 72, 96, 336, 379, 419
- chemostat, 183
- chicken pox, 346, 422
- cholera, 345
- closed season, 240
- cluster, 346, 367
- coalescence, 33, 201, 225
- cobwebbing method, 53
- coexistence, xviii, 130, 218, 224
- cohort, 273, 352
- collapse, 55
- colonization, 294
- commensalism, 193
- communicable disease, 345, 413, 442
- community matrix, 129, 206, 207, 232
- compartment
 - infected, 395
 - uninfected, 394
- compartmental model, 349, 442
- compensation, 21
- competing disease strains, 445
- competition
 - contest, 65
 - intraspecies, 65
 - scramble, 65, 98
 - unqualified, 168
- competitive exclusion, xviii, 165, 170
- complex dynamics, xvii
- compounding, 247
- computer algebra system, *see* Maple, *see* Mathematica
- configuration model, 370
- connecting orbit, 235, 325
- connection matrix, 297
- constraint, 229
- contact, 367, 368, 412
- contact intervention, 368
- contact rate, 416
- contest competition, 98
- control
 - bang-bang, 238, 249
 - optimal, 238
- control group, 348
- control reproduction number, 379, 392
- control, optimal, 238
- convolution, 276
- Cortez, *see* Aztecs
- critical depensation, 21
- critical harvest rate, 28, 237
- critical transmissibility, 368
- culture, 126
 - batch, 126
 - continuous, 126
- cumulative distribution, 38
- current shadow price, 263
- cusp, 45
- data, 419, 431, 434
- death modulus, 274
- death rate, *see* mortality rate
 - per capita, 4
- debilitating, 345, 433
- degree, 362, 367
- degree distribution, 362, 367
- delay, xvii, 19
- delay logistic equation, 94
- delta function, 317
- demand curve, 244
- demographic, 346, 394, 400
- demographic equation, 446
- demographic equilibrium, 87, 446
- demographic process, xv
- dendritic branch, 333, 340
- density dependence, 289
- Density-dependent, 413, 416
- depensation model, 21
- destabilization, 101, 323
- determinant, 140
- deterministic model, *see* model
- diarrhea, 345
- difference equation
 - differential, 105
 - linear, 72
 - linear homogeneous, 49
 - nonlinear, 51
- differential equation
 - autonomous, 21
 - linear homogeneous, 19
 - nonautonomous, 16
- differential infectivity, 377
- differential–difference equation, 349
- differential–difference equations, 118
- diffusion, 24, 91, 293, 442
- diffusion coefficient, 303, 331
- diffusion equation, 301
- diffusive instability, 327

- diffusivity, 303, 331
- dimensionless, xv, 36, 112
- direction field, 20
- discount rate, 247
- discrete, viii
- discrete generations, 112
- disease
 - fatal, 425, 432
- disease evolution, 346
- disease mortality, 432
- disease outbreak, 367
- disease transmission, 301, 348, 405
- dispersal, 448
- distributed
 - delay, xvii, 91
- divergence, 330
- divorce, 160
- doubling time, 6
- drug use, 346
- Dulac criterion, 172

- Ebola virus, 345
- economic models, 78, 244
- economic optimization, 37
- edge, 368
- effort, 29
- eigenvalue, 140, 295, 306
 - complex, 272
 - dominant, 273
 - double, 140
- eigenvector, 156
- elimination, 349, 434
- emigration, 50
- endangered species, xv
- endemic, 133
- energy reserve, 202
- English boarding school, 360
- environmental activists, 37
- environmental landscape, xiii
- epidemic, v, xv, 346
- epidemiology, viii
- equilibrium, 19
 - age distribution, 279, 283, 285, 290
 - bionomic, 244
 - endemic, 133
 - eutrophic, 32
 - isolated, 136
 - monotone, 57
 - oligotrophia, 32
 - oscillatory, 57
 - stable, 150
- equilibrium analysis, 56, 413
- eradication, 420
- escapement, 68

- eutrophic equilibrium, 32
- eutrophication, 31
- excess
 - degree, 363
- excitable cell, 333
- exponential distribution, 40, 369
- exponential growth, xvi
- exposed period, 373
- extinction, xv
- extinction rate, 294
- Eyam, 356

- family
 - one-parameter, 5
- fecundity, 268
- Feigenbaum constant, 64
- fertility, 68, 272
- Fibonacci sequence, 52
- Ficks law, 302, 330
- final size relation, 354, 374, 392
- finite time horizon, 248
- fishery, 11, 21, 33, 68, 242, 244, *see* harvesting
 - collapse, 33
 - open access, 242
- FitzHugh–Nagumo system, 219
- fleas, 357
- flour beetles, 80, 289
- flow, 350, 404, 411
- fluid convection, 209
- flux, 302, 330
- focus, 140
- foliage, 43, 202
- foraging rate, 333
- forcing
 - periodic, 84
- forest, 199
- forestry, 31
- Fourier integral, 314, 315
- Fourier series, 307, 314
 - full, 313
- Fourier sine integral, 318, 320
- Fourier sine series, 307, 310, 311
- fox rabies, 360
- functional equation, 286
- functional response
 - predator, 43, 175, 176
- functional response predator, 236
- fundamental solution, 317, 338, 339

- Garki project, 349
- gender, xviii, 160, 346
- gene, 60, 279
- generating function, 367
- generations

- nonoverlapping, xvii
- overlapping, xvii
- genesis model, 279
- genetic, 346
- genetic disease, 346
- global behavior, 172
- globally asymptotically stable, 150
- gonorrhea, 13, 41, 346, 400, 411
- granulocytes, 116
- graph, 297
 - bi-directional, 297
- graphic solution, vii, 53, 58
- grazing, 91, 94
- Great Plague, 347, 356
- Green's theorem, 253
- growth curve, 22
- growth rate, xiv
 - delayed, 93
 - intrinsic, 11
 - per capita, 8
- haddock, 69
- half-life, 6
- halibut fishery, 11
- Hamiltonian, 255
- Han empire, 347
- harvesting
 - constant effort, 27, 28
 - constant rate, 27
 - instantaneous, 240
 - proportional, 28
- hazard of infection, 370
- health care, xiv, 405, 432
- heat equation, 340, 341
- heat kernel, 336, 338
- helminth agents, 411
- herbivore, 130, 221
- herd immunity, 359, 420
- heterogeneity, xviii, 372, 442
- heterosexual, 346, 404, 442
- HIV/AIDS, 346, 377
- home patch, 299, 300
- homoclinic, 233
- homosexual, 13, 41
- Hopf
 - bifurcation, 431
- host, 3, 466
- Hudson's Bay Company, 179
- hunting, 22, 240, 361
- hyperbola, 141, 224
- hysteresis, 32
- immature members, 70, 109
- immigration, 46, 50, 189, 349, 394
- immune response, 356
- immunity
 - herd, 359, 420
 - temporary, 349, 350, 427, 428, 445
- immunization, 350, 359, 420
- Incas, 347
- infected, 346, 367
- infected edge, 367
- infection
 - secondary, 42
- infectiousness, 348
- infective, 13
- infective period distribution, 441
- infectivity variable, 374
- inflection point, 18
- influenza, 345, 376, 437
- inhibitor, 328
- initial age distribution, 275
- initial boundary value problem, 303
- initial condition, 5
- initial data, 190
- initial value problem, 5
- inoculation, 349, 422
- instability, xviii
- integral equation, 276, 284, 349
- integro-differential equation, 91
- interaction, 368
 - facultative, 192
 - multispecies, xviii
 - obligatory, 192
 - terms, 130
- interepidemic period, 422
- internal patch dynamics, 298, 299
- interpolation, 448
- intra-specific competition, 289
- intracellular current, 333
- intracranial tumor, 144
- intraspecific competition, 289
- intrinsic growth rate, *see* growth rate
- invading species, 212
- invariant set, 148
- Invasive pneumococcal disease, 457
- ion channel, 218
- island, 296
- isocline, 170
 - predator, 176
 - prey, 176
- isolation, 350, 426
- iterate, 60, 63
- Jordan curve theorem, 149
- Jury criterion, 75, 78
- juveniles, 289, 458

- Kermack-McKendrick model, 350
- Kolmogorov forward equation, 341
- Kolmogorov model, 190
- Kolmogorov theorem, 191
- lags, *see* delay
- lake
 - eutrophic, xvii
 - eutrophication, xvii
 - irreversible, 35
 - oligotrophies, 31
 - reversible, 34
- Laplace equation, 331
- Laplace transform, 106, 277, 287
- Laplacian, 331
- larva, 80, 112
- least squares, 7
- Leishmaniasis, 453
- Leslie matrix, 268
- Leslie model, 190
- life expectancy, 421
- life span
 - average, 5, 345, 423
 - budworm, 43
 - generation, xv
 - tree, 43
- limit set, 148
- limiting system, 134
- limnological mechanisms, 31
- limpet, 145
- linearization, 19, 56
- logistic (differential) equation, xvi
- logistic difference equation, 58
- Lorenz equations, 208
- Lotka-Volterra equations, 123
- Lotka-Sharpe equation, 279
- lynx, Canadian, 179
- M-matrix, 397
- macroparasite, 492
- major epidemic, 362, 368
- malaria, xv, 345, 442
- Malthus, xiii, *see also* population model, 5
- management model, 380
- Maple, vii, 36
- marginal value, 263
- marriage function, 89
- mass action incidence, 351, 388
- Mathematica, vii, 36, *see* Maple, 92
- mating function, *see also* marriage function, 89
- Matlab, vii
- matrix, 417, 418
- maximum sustainable yield (MSY), 243
- McKendrick equation, 274
- mean field approximation, 332
- mean transmissibility, 367
- measles, 345, 347
- measurement, 43
- membrane battery, 333
- metapopulation, 293, 297
- metered model, 68
- microorganisms, 4, 345
- migration rate, *see* emigration
- minimum velocity, 325
- minor outbreak, 365, 368
- mixing
 - homogeneous, xviii, 14, 361, 442
- mobility, 299, 301
- model, vii, *see also* population model
 - Rosenzweig-MacArthur, 175
 - Beverton-Holt, 69
 - compartmental, 349, 442
 - depensation, 21
 - deterministic, 358
 - differential infectivity, 377
 - economic, 78, 244
 - genesis, 279
 - Kermack-McKendrick, 350, 354
 - Kolmogorov, 190
 - Leslie, 190
 - Leslie matrix, 268
 - metered, 68
 - multispecies, 126
 - neuron, 218
 - Nicholson's blowflies, 112
 - Nicholson-Bailey, 75
 - Ricker, 69
 - stochastic, 465
- model fitting, 348
- modeling assumptions, xv, *see* specific model
- modulus, birth/death, 274
- monotone
 - convergence, 100
 - decrease, 20, 182, 270, 355
- mortality, 345, 418
- mortality function, *see* modulus, death
 - modulus, 71
- mortality rate, xvi, 401
- mortality-harvesting curve, 109
- mosquitoes, 349, 401, 404, 442
- motion of particles, 302
- multispecies model, 126
- mutated strain, 400
- mutual inhibition, 156
- mutualism, xviii, 192
- natural death, 349, 419, 427
- neighbor, 364, 371

- network, 367, 368
- neural network, 60, 333
- neuron, 333
- neuron model, 218
- newborn, 68, 86, 422, 426, 435
 - vaccination, 135
- next generation matrix, 394, 396
- Nicholson's blowflies model, 112
- Nicholson-Bailey model, 75
- node, 139, 325, 370, 372
- nonlinear, 323
- North Atlantic plaice, 69
- nullcline, 176
- numerical simulation, 112, 205, 237, 380
- nutrients, 10, 126

- objective function, 238, 460
- observation, xiii
- occupancy model, 294
- occupied edge, 367, 368
- Ohms law, 333, 334
- optimal, *see* control
 - equilibrium population, 248
- optimal control, 238
- optimization, 251, 438
- orbit, 123
- Ordinary Least Squares, 434, 436
- Ornstein - Uhlenbeck process, 341
- oscillation, 65, 80, 173, 174, 180, 185, 205, 245, 272
- outbreak, 202, 346, 351, 465
- overfishing, 246
- overharvesting, 111
- overpopulation, xiii

- Paramecium, 185, 186
 - aurelia, 185, 188
 - caudatum, 185
- parameter estimation, 348, 434, 460
- parasite, xv, 75, 199, 453
 - helminth agents, 411
- partial differential equation, 288, 293, 301, 302
 - linear, 306
- partition, 288
- partner, 14, 160
- passenger pigeon, 22
- patches, xviii, 293, 294, 323, 447
- pathogen, 4, 419
- pattern in space, 324
- per capita, xiii, *see* asobirth, death, growth rates, xiii
 - birth rate, xiv, 4
 - death rate, 4
 - growth rate, 8
- period-doubling, 63, 100, 238
- period-doubling bifurcation, 88
- periodic orbits, 125, 155, 240
- periodic solution, 61, 148, 150, 221, 232
- periodic wave train, 332
- perpetual outbreak, 203
- Perron-Frobenius theorem, 271
- persistence, 212
- persistent
 - age distribution, 278
- perturbation, 3, 20, 142, 173, 195, 212
- pest, xv
- petri dish, 6
- phase plane, 123
- phase portrait, 125
- phosphorus, 31
- pigeon, *see* passenger
- plague, 346, 411, *see also* bubonic plague
- Poincaré-Bendixson theorem, 149, 172, 178
- pollinator, 191
- pollutants, 91
- Pontryagin maximum principle, 238
- population
 - decay, xv
 - density, xvi, 3, 61, 418
 - dynamics, xiii, 465
 - explosion, xv
 - model
 - exponential, xiv, *see also* Malthus
 - logistic, xiv
- population control, 431
- positive feedback, 328
- positive semiorbit, 148
- predation, 8, 43, 199, 333
- predator functional response, 176
- predator numerical response, 175
- predator-prey, 123, 229, 294, 329, 433
- predator-induced coexistence, 214
- predator-induced exclusion, 216
- predator-prey, xviii, 173
- prediction, 5
- probability, 38, 299, 367, 368
- probability density, 39
- probability distribution, 38
- protozoa, *see* Paramecium
- public health, 347, 348, 404, 415, 435
- pulse vaccination, 443

- qualitative
 - analysis, 17, 92, 424
 - behavior, xvii, 8, 135
 - dynamics, 16, 134
 - prediction, 5, 65, 193
 - quality of life, xiv

- quantitative prediction, 348
- quarantine, 348, 354, 377, 415, 426
- random contact, 351
- random motion, 303, 330
- rat, 357
- rate of transfer, 349, 427
- reaction–diffusion equation, 321, 329
- recolonization, 294
- recovery, 350, 395, 404, 411, 413, 445
- recruitment, 16, *see also* birth cycle
 - cycle, 67
 - delayed, 70, 98
 - function, 65
- recursive calculation, 52
- recycling, 31
 - forest, 205
 - phosphorus, 32
 - rate, 31
- refuge, 45, 189
- reinfection, 119, 346, 350, 411
- removed, 350, 411, 415
- renewal condition, 274
- renewal equation, 276
 - nonlinear, 284
- reproduction matrix, 71
- reproductive
 - curve, 68
 - life, xv
 - number, *see* basic reproductive
- resource limitations, xiii, 9, *see also*
 - competition, harvesting, 229
- resources, 289
- respiratory infection, 345, 457
- response delay, 91
- Riccati equation, 335
- Ricker model, 69
- robust, 17, 51
- Rosenzweig–MacArthur model, 175
- Routh–Hurwitz
 - conditions, 208
- Routh–Hurwitz conditions, 106, 419
- rubella, 345, 422
- rumor, 346, 414
- runoff, 31
- Saccharomyces exiguus*, 188
- saddle point, 139, 141
- salmon, xvii, 69
- sandhill crane, 29
- SARS, 345, 435
- schistosomiasis, 345
- second order, 52
 - difference equation, 76
 - partial differential equations, 304
- secondary infection, 348, 412
- sedimentation, 31
- selection, 406, 466
- self-limiting, 125, 177
- self-regulating, 130
- semiorbit, 148
 - bounded, 148
 - positive, 148
- Sennacherib, 346
- separation of variables, 8, 68, 124, 304, 413
- separatrix, 149
- SEIQR model, 378
- sequence, 49
- serological study, 356
- sexually transmitted disease, 389, 411
- shadow price, 263
- shark, 123
- Sharpe, 279
- sigmoid function, 31, 452
- singular control, 257, 258
- SIR model, 351, 407, 411, 419
- SIS model, 411, 425, 445, 447
- SITR model, 375
- sleeping sickness, 345
- smallpox, 347, 420, 434
- smoking, 158
- snowshoe hare, 179
- social interaction, 407
- Spaniards, 347
- spatial location, 293
- spiral point, 141
- spiral wave, 332
- spruce budworm, xvii, 25, 199, 204, 206
- stability, 17
 - absolute, 100
 - asymptotic, 23, 58
 - exponential, 114
 - neutral, 143
 - oscillatory, 114
- stable
 - age distribution, 270–273, 278, 279, 287
- stable age distribution, *see* age distribution
- stable pattern, 324, 332
- standard incidence, 388
- standardized solution, 337
- stasis, xv
- stochastic, 3, 341, 358, 426
- stock recruit, 69
- stress, 202
- superposition principle, 336
- surveillance data, 460
- survival, 168
 - equilibrium, 65, 81

- probability, 80
- survivor species, 213, 217
- survivorship, 69
- susceptibility, 348, 367, 431
- sustainable economic rent, 244
- switch point, 238
- switching function, 257

- Taylor approximation, viii, *see* Taylor's theorem
- Taylor's theorem, 19, 76, 95, 129, 141
- terminal condition, 251
- threshold, xv, 296, 348, 418
- tide pool, 145
- time scale of
 - births and deaths, 352
 - disease, 352
 - observation, xv
- tipping point, xv
- trajectories, 123
- trajectory, 156
- transformation property, 339
- transient, 311
- transition, 26, 195, 369, 441
- transmissibility, 367, 368
- transmission, 367, 368
- transmission intervention, 368
- transmission rate, 14, 440
- transportation network, 293
- travel between patches, 295, 296, 299
- treatment model, 354, 375
- tube, 302
 - infinite, 303
 - semi-infinite, 303
- tuberculosis, 345
- tumor, *see* intracranial
- typhus, 345

- uncertainty, 35, 438

- unchangeable, 159
- underreporting, 360, 453
- uniform persistence, 212
- uninhabitable patch, 296
- uniqueness, 12
 - of solutions, 18
- unstable, 21, 58, 433

- vaccination, xiv, 134, 348, 415
 - pulse, 443
- variable maturation, 107
- vector, 346, 404, 442, 455
- vector field, 150
- vector transmission, 404, 442
- vector-matrix notation, 23, 71, 74, 136, 207
- vegetation, 26, 130, 221
- Verhulst, xvi
 - difference equation, 51
- vertex, 362–365, 367, 368
- vertical transmission, 442
- viral agent, 345
- virus, 420, 458
- Volterra, *see also* Lotka-Volterra
- Von Foerster equation, 274
- vortex, 139

- wave traveling, 325
- West Nile virus, 346, 404
- WinPP, vii, viii, 92, 93
- worker, 159
 - inactive, 159
 - master, 159
 - positive, 159
 - reluctant, 159

- XPP, vii, viii, 92

- Yield, 29
- yield effort curve, 29