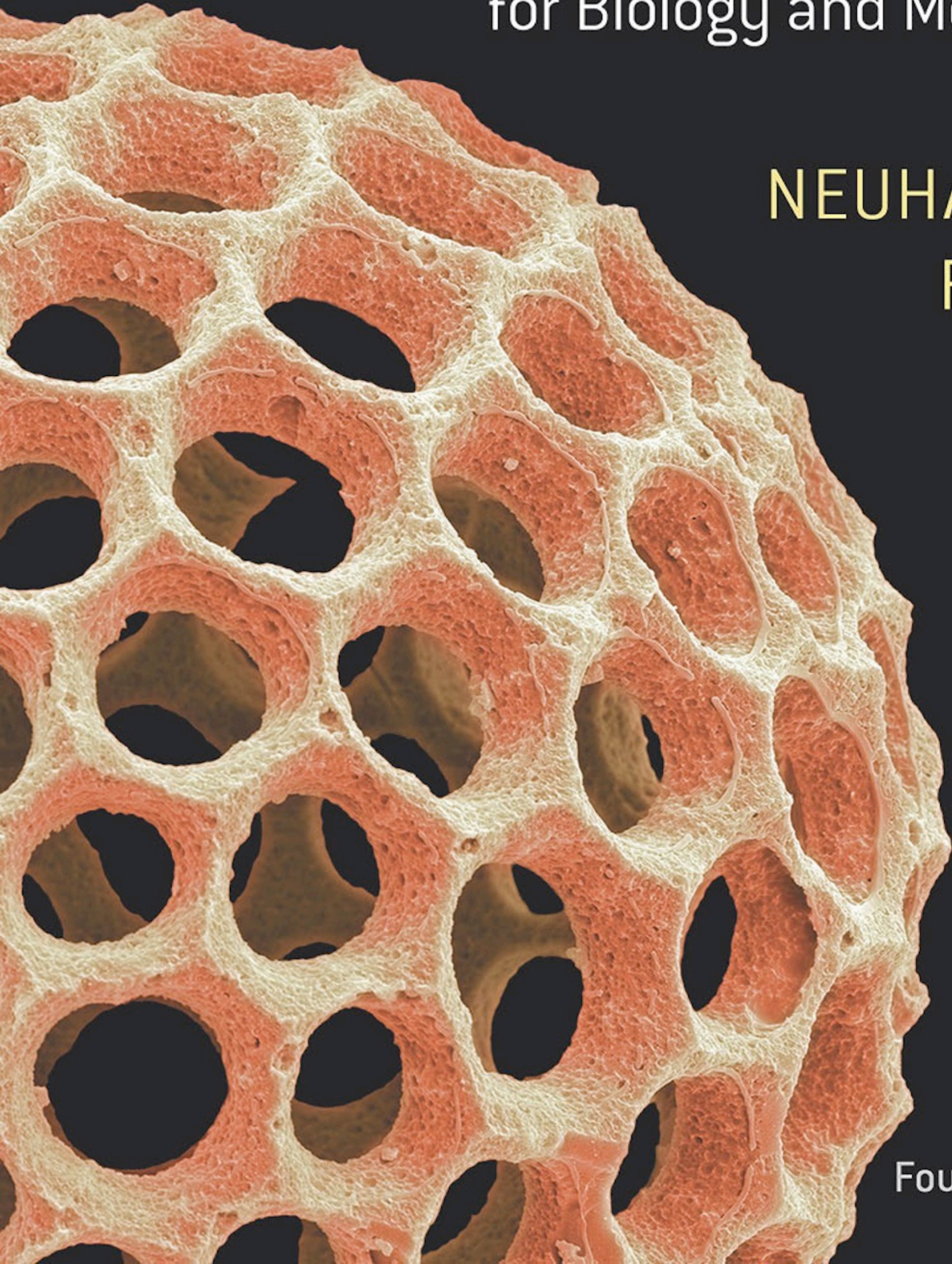
CALCULUS

for Biology and Medicine

NEUHAUSER
ROPER



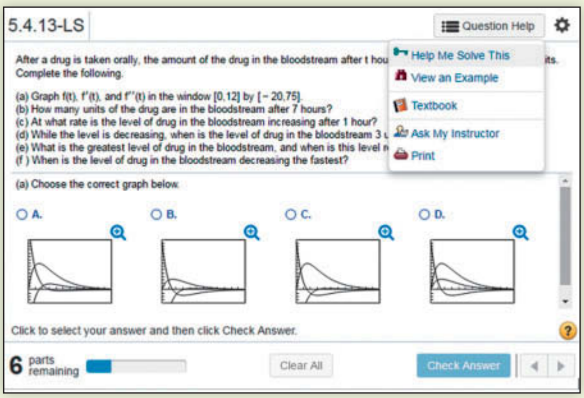
Fourth Edition

MyLab Math for *Calculus for Biology and Medicine*, 4e

(access code required)

Pearson MyLab

Used by over 3 million students a year, MyLab™ Math is the world's leading online program for teaching and learning mathematics. MyLab Math delivers assessment, tutorials, and multimedia resources that provide engaging and personalized experiences for each student, so learning can happen in any environment. For the first time, instructors and students using *Calculus for Biology and Medicine* can access text-specific online homework and resources in and out of the classroom.

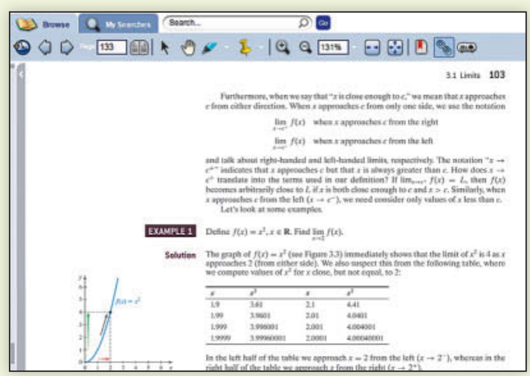


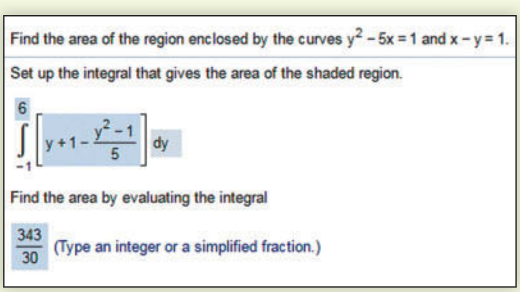
Exercises with Immediate Feedback

Homework and practice exercises for this text regenerate algorithmically to give students unlimited opportunity for practice and mastery. MyLab Math provides helpful feedback when students enter incorrect answers and includes the optional learning aids Help Me Solve This, View an Example, and/or the eText.

Complete eText

Students and instructors get unlimited access to the eText within any MyLab Math course using that edition of the textbook. The Pearson eText app allows existing subscribers to access their titles on an iPad or Android tablet for either online or offline viewing.





Questions that Deepen Understanding

MyLab Math includes a variety of question types designed to help students succeed in the course. In **Setup & Solve** questions, students show how they set up a problem as well as the solution, better mirroring what is required on tests. **Additional Conceptual Questions** were written by faculty at Cornell University and focus on deeper, theoretical understanding of the key concepts in calculus.

To learn more, visit pearson.com/mylab/math.

CALCULUS

for Biology and Medicine

Fourth Edition

CLAUDIA NEUHAUSER
University of Minnesota

MARCUS L. ROPER
University of California — Los Angeles

Pearson

Director, Portfolio Management: *Deirdre Lynch*
Executive Editor: *Jeff Weidenaar*
Editorial Assistant: *Jennifer Snyder*
Content Producer: *Patty Bergin*
Managing Producer: *Karen Wernholm*
Producer, Production & Digital Studio, Mathematics & Statistics: *Marielle Guiney*
Manager, Courseware QA: *Mary Durnwald*
Manager, Content Development, Math: *Kristina Evans*
Product Marketing Manager: *Emily Ockay*
Field Marketing Manager: *Evan St. Cyr*
Marketing Assistant: *Erin Rush*
Senior Author Support/Technology Specialist: *Joe Vetere*
Text Design: *Tamara Newnam*
Composition: *iEnergizer Aptara, Inc.*
Illustrations: *Network Graphics*
Cover Design: *Tamara Newnam*
Cover and Chapter opening Image: *Steve Gschmeissner/Science Photo Library/AGE Fotostock*

Library of Congress Cataloging-in-Publication Data

Names: Neuhauser, Claudia, 1962- author. | Roper, Marcus L., author.
Title: Calculus for biology and medicine.
Description: Fourth edition / Claudia Neuhauser, University of Minnesota, Marcus L. Roper, University of California—Los Angeles. | Boston : Pearson, [2018] | Includes bibliographical references and index.
Identifiers: LCCN 2017036101 | ISBN 9780134070049 (hardcover)
Subjects: LCSH: Biomathematics–Textbooks. | Medicine–Mathematics–Textbooks.
Classification: LCC QH323.5 .N46 2018 | DDC 570.1/51–dc23 LC record available at https://lccn.loc.gov/2017036101

1 17

Pearson

ISBN 13: 978-0-13-407004-9
ISBN 10: 0-13-407004-6

Contents

Throughout this Table of Contents we use an asterisk (*) for topics that are not directly used in the later sections of the text. That is, you can study the entire book without studying these topics. Some of these topics go deeper into rigorous definitions of limits and bounds (for example, 3.6 or 10.2). Others explore extensions of the major mathematical ideas, like stability in recurrence equations (in 5.7) or models using systems of recurrence equations (in 9.3.3 and 10.9). A third class of asterisked topics provides more modeling examples using the mathematical tools developed in the text (for example, 2.3.4, 5.9, 9.4, and 11.5). Instructors may decide for themselves which of these topics to cover.

Preface

The goal of *Calculus for Biology and Medicine* has remained constant from its inception:

> ***To show students how calculus is used to analyze phenomena in nature without compromising the rigor of the presentation of calculus principles.***

The result of this goal is a calculus text that has plentiful life and health sciences applications and that provides students with the knowledge and skills necessary to analyze and interpret mathematical models of a diverse array of phenomena in the living world. Since this text is written for college freshmen, the examples were chosen so that no formal training in biology is needed.

The rigor of the text prepares students well for more advanced courses in mathematics and statistics. Our hope is that students will find calculus concepts easier to understand and more interesting if they are related to their major and career aspirations. While the table of contents resembles that of a traditional calculus text, the content does not: Abstract calculus concepts are introduced in a biological context, and students learn how to transfer and apply these concepts to biological situations.

New to This Edition

- **Modeling** – The 4th Edition places much more emphasis on modeling biological situations. Students are instructed in the processes of modeling real-world situations and given many opportunities to practice these techniques in problems.
- **Applications** – The applications in the text have been greatly expanded in number. New applications include population genetics, pharmacology, and the evolution of microbial cooperation. Many of these applications are adapted from published studies and other current sources.
- **Technology** – The 4th Edition now includes clear student instructions on using spreadsheets to numerically solve equations, visualize data, and model biological processes. This material is clearly labeled so that instructors who prefer not to use it can easily omit it, or assign technology sections as optional readings to their students.
- **Approach** – The level of rigor of the text has been maintained, but we have made adjustments to how some topics are introduced to make the presentation accessible to students, better bridging the gap between what students already know and what they are attempting to learn. We have also streamlined or made optional material that is not useful for life sciences students (formal discussion of limits, continuity in multivariate functions, etc.). This material is maintained so that instructors may continue to teach it, and students who may transfer out of life sciences calculus courses into physical sciences and engineering calculus will still find the material that is covered in physical calculus; but the main current of the book is through topics that are directly needed for life sciences.
- **Writing** – Every attempt has been made in the new edition to use language that will enable students to better use the text as an independent learning resource. In some sections this required lengthening explanations that were overly terse in order to make them more accessible; in others, topics are introduced informally using examples directly taken from life sciences to motivate the more formal mathematical material that was the strength of the previous editions.
- **Prerequisites** – We added a Precalculus Skills Diagnostic Test at the beginning of the text to help students gauge whether review of precalculus topics is needed. Answers to the quiz are provided in the back of the book along with tips on what to review in Chapter 1 if refreshers are needed.
- **Design** – The book has been redesigned in full color to help students better use it and to help motivate students as they put in the hard work to learn the mathematics.

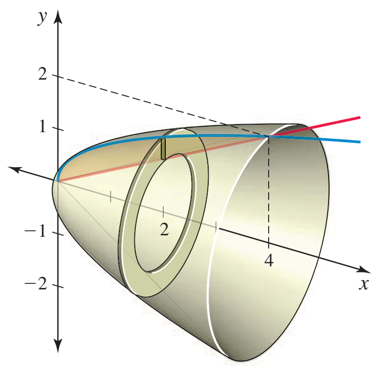


Figure 6.45 The solid of rotation for Example 14 can be made up of washer-like elements.

- **Figures** – Many figures were revised to take advantage of the new full-color design. Most notably, the 3-dimensional figures were re-rendered using the latest software. See the figure at the left for an example.
- **Biology Notes** – New "Bio Info" notes provide optional background information to support the narrative and exercises.
- **"Help Text" within Examples** – We added text (in blue type) within examples to explain the mathematical principle(s) applied in the steps of the solution. This text helps students understand the solution and emphasizes that each step in a mathematical argument is carefully justified.
- **Topic Coverage** – Based on feedback from reviewers, some new topics have been added to the text. The most significant among these are the following:
  - Many new mathematical models, including models for microbial cooperation, evolution, and epidemiology, and multicompartment models in pharmacology.
  - Expanded applications for optimization methods.
  - Tools for fitting models to real data.
  - Expanded discussion of methods for visualizing multivariate functions.

MyLab™ Math Online Homework – Last, but not least, the text now has online homework within MyLab Math. The MyLab Math course contains hundreds of algorithmically generated exercises that provide students with instant feedback, optional learning aids for many exercises, and the complete eBook. See below for additional features of MyLab Math for this text.

Features of the Text

A distinguishing feature of this text is the biological examples and exercises, which are notably *real* (many from published studies or other current sources), *relevant*, and *varied*. The References section at the back of this text contains an exhaustive list of the sources we used.

Examples and Explanations Each topic is inspired by biological examples. These motivating introductions are followed by a thorough discussion outside the life science context to enable students to become familiar with both the meaning and the mechanics of the mathematical topic. Finally, biological examples are presented to teach students how to apply the material in a life science context. Examples in the text are completely worked out, and the steps in the solutions are explained in blue text to the right of each step.

Exercises Calculus cannot be learned by watching someone do it. Because of this, *Calculus for Biology and Medicine* provides students with skill-based exercises as well as word problems. Word problems are an integral part of teaching calculus in a life science context. The word problems contained in the text are up-to-date and are adapted from either standard biology texts or original research. The exercises and word problems are at the end of each section and are organized by subsection to help students refer to specific subsections of content while completing homework. This also aids instructors in assigning homework problems.

Technology *Calculus for Biology and Medicine* assumes the availability of graphing calculators. This allows students to develop a much better visual understanding of the concepts in calculus. Beyond this, no special software is required.

Reflections and Outlook

Like many schools now, both UCLA and University of Minnesota offer life science students their own calculus track. Other universities are increasingly adopting separate calculus tracks to deal with the different needs of life sciences majors and physical science/engineering majors. There are many ways to design curricula for these courses, and faculty venturing into the recommendations offered by reports on new needs for life sciences education (such as *Bio2010: Transforming Undergraduate Education for Future Research Biologists* from the National Research Council and the National Academies, or *Scientific Foundations for Future Physicians* from the Association of the American Medical Colleges and the Howard Hughes Medical Institute) may be overwhelmed by the amount of quantitative training that is now expected for life science students and by how it goes far beyond what students can be prepared for with a single year of calculus. In this fourth edition of the textbook we have focused on retaining the strengths of the third edition, including giving students access to the rigorous foundations of mathematical ideas that will enable them to take further classes in math that are increasingly necessary for quantitative minded biologists. However, we have rewritten much of the material with an eye to eliminating barriers to study (e.g., by avoiding using expressions with multiple unknown constants in them).

We are also very mindful of the future needs of students to handle the large data streams created by new innovations in omics, personal medicine, and remote sensing. Much of the math underlying these new areas is outside what can be covered in this book, but Chapters 9 and 12 lay foundations for students who will go on to study bioinformatics. Additionally, we have brought data (and data fitting) increasingly into the book, especially in support of the new mathematical modeling topics we have introduced. Study of algorithms is supported by explicit directions on using spreadsheets to implement the algorithms. Any spreadsheet software can be used, but we have found Google Sheets to be especially effective in the classroom, since it allows spreadsheets to be simultaneously shared and edited across dozens of computers.

Chapter Summary

Chapter 1 This chapter reviews precalculus tools, including functions and methods for graphing data. Many students will have studied this material in their precalculus classes, so summaries are kept brief. Section 1.1 includes a diagnostic test that students can take (either by itself, or in conjunction with MyLab Math) to review their knowledge of these topics. The basic tools from algebra and trigonometry are summarized in Section 1.2. Section 1.3 then describes the functions that students need to be familiar with for this book, including exponential and logarithmic functions. Section 1.4 focuses on graphing, including log-log and semi-log plots and translating verbal descriptions of biological phenomena into graphs.

Chapter 2 This chapter covers recurrence equations (or discrete time models) and sequences. Importantly, we use this chapter to introduce $\sum$-notation for summing series. This notation is used throughout the text. We also use this chapter to introduce mathematical modeling, including the assumptions that are built into models, parsing verbal descriptions, and comparing models against data. Our examples here are drawn from population growth and physiological modeling of how drugs pass through a patient's body.

Chapter 3 Limits and continuity are key concepts for understanding the conceptual parts of calculus. Visual intuition is emphasized before the theory is discussed. We show how the bisection method can be used as a practical tool for solving equations. The formal definition of limits is given at the end of the chapter in an optional section.

Chapter 4 We start with an intuitive and visual description of the derivative before giving a formal definition. Then, before we go into the mechanics of differentiation, we describe interpretations of the derivative in different contexts (including chemical reactions), building students' intuition further. Differentiation rules are discussed and broken into readily digestible chunks to give students time to acquaint themselves with them. Error propagation and differential equations are the main applications.

Chapter 5 This chapter presents biological and more traditional applications of differentiation. We maintain the 3rd Edition's approach that derives results on functional extrema rigorously from the Mean Value Theorem. But we also explain to skeptical students why calculus-based tools for analyzing functions and drawing their graphs are still relevant when computers allow functions to be so readily plotted. We have also enlarged the number of applications for optimization, including models from physiology and population genetics. We also added a new section on differential equation-based models (again focusing on population growth and the passage of drugs through the body), so that students encounter these vital applications before they meet integration. Finally, we introduce antiderivatives, in anticipation of studying integration in Chapter 6. Analysis of recurrence equations is covered in an optional section.

Chapter 6 Integration is motivated geometrically. We also describe the definition of the integral via Riemann sums in a way that has been greatly simplified from the 3rd Edition. In particular, students can study this material without needing to know $\sum$-notation and without the full formalism of partitions. In our experience, this makes this difficult topic much easier for students without an overall compromise on the level of rigor in their understanding. The fundamental theorem of calculus and its consequences are discussed in depth. We discuss applications for integration, but have reduced the amount of required material in this chapter to focus only on applications that are directly relevant to life sciences, such as calculating the mean of a function and its cumulative change.

Chapter 7 This chapter contains integration techniques, focusing on the techniques that are immediately necessary for solving differential equations, including integration by parts, by substitution, and a tailored introduction to the method of partial fractions. Material on Taylor polynomials and on using tables of integrals (a technique now largely made redundant by computer algebra packages) is covered in optional sections at the end of the chapter.

Chapter 8 This chapter provides an introduction to differential equations, covering separable equations and linear first order equations. The treatment is not complete, but it will equip students with both analytical and graphical skills to analyze differential equations. The chapter showcases and interprets mathematical models from many areas of biology, including microbial cooperation, ecology, and epidemiology. Addition of integrating factors allows us to discuss two-compartment models, which are widely used for studying the movement of drugs through the body.

Chapter 9 Linear models, and the matrix methods needed to solve them, are central to modern methods in bioinformatics. The material in this chapter introduces students to the basic concepts needed to study multivariate functions in Chapters 10 and 11. However, although the treatment of eigenvalues and eigenvectors emphasizes their importance to models of change (both recurrence equations and systems of differential equations), matrix math is introduced in a way designed to provide students with a firm platform for further study in linear algebra for bioinformatics applications.

Chapter 10 This is an introduction to multidimensional calculus. Since students often struggle with the transition from univariate functions to multivariate functions, we have expanded the introductory material to build up student intuition more gradually, with more examples (including practical examples like heat index) and larger discussion of how functions can be visualized. The main mathematical topics are partial derivatives and linearization of vector-valued functions. We cover at length finding extrema of functions (including under constraints). Although this topic is not needed for Chapter 11, optimization has many uses (and we highlight its application to least squares estimation of fitting parameters), and we find this section worthy of class time. The final sections provide optional material on systems of recurrence equations and on partial differential equation models.

Chapter 11 Both graphical and analytical tools are developed to enable students to analyze systems of differential equations. The material is divided into linear and nonlinear systems. Understanding the stability of linear systems in terms of vector fields, eigenvectors, and eigenvalues helps students to master the more difficult analysis of nonlinear systems. Theory is explained before applications are given. Extensive examples (with accompanying problems) showcase applications of these tools in ecology, epidemiology, and physiology.

Chapter 12 This chapter introduces some fundamental probabilistic and statistical tools, taking students from counting (i.e., combinatorial) approaches to probability, through important distributions that arise when modeling stochastic processes. Throughout students are introduced to fundamental ideas for working with data: estimating probability distributions from histograms, fitting linear models, and calculating and interpreting summary statistics.

How to Use This Book

By design this book contains more material than can be covered in one year. The intent is to allow for schools and instructors to have more flexibility in the choice of material covered. Topics labeled with an asterisk (*) in the Table of Contents may be omitted at the instructor's discretion. They include sections going more rigorously into the definitions of limits and continuity as well as many of the modeling and applications sections.

The book's content can be arranged to support any length of course, from one quarter to three semesters. Chapter 1 is precalculus material. Students should have been exposed to this material before starting their first course in calculus. This said, we find it highly useful for students to self-study this material before starting the class, which they can do more easily using the new Precalculus Skills Diagnostic Test. Additionally, we often cover in class the material in Section 1.4 (in particular on how to graph data using logarithmic and semi-logarithmic axes). Sections 2.1 and 2.2 give students a minimal introduction to sequences, series, and $\sum$-notation. However, we strongly recommend Section 2.3 as an introduction to deriving, solving, and interpreting mathematical models, before students meet models again in the calculus context.

Chapters 3 and 4 must be covered in that order before any of the other sections are covered. In addition to Chapters 3–4, the following sections can be chosen:

One semester—integration emphasis 5.1–5.6, 5.10, 6.1–6.3 (without 6.3.4 and 6.3.5)

One semester—differential equation emphasis 5.1–5.6, 5.9–5.10, 6.1, 6.2, 8.2, 8.3 (without solving any of the differential equations)

One semester—probability emphasis Chapter 3 (except 3.6), Chapter 4 (without 4.11), 5.1–5.4, 5.10, 6.1, 6.2, 7.1, 7.2.1, 12.1–12.5 (without 12.5.5), 12.6 (if time permits)

Two quarters 5.1–5.6, 5.8, 5.10, 6.1–6.2, 6.3.1 and 6.3.2, Chapter 7, Chapter 8

Two semesters or three quarters 5.1–5.6, 5.10, 6.1–6.2, 6.3.1 and 6.3.2, Chapters 7, 8, and 9 (without 9.2.4 or 9.4), 10.1, 10.3, 10.4, 11.1–11.4

Four quarters or three semesters All sections that are not labeled optional (with *); optional sections should be chosen as time permits

MyLab Math Online Course (access code required)

Used by over 3 million students a year, MyLab Math is the world's leading online program for teaching and learning mathematics. MyLab Math delivers assessment, tutorials, and multimedia resources that provide engaging and personalized experiences for each student, so learning can happen in any environment. For the first time, instructors teaching with *Calculus for Biology and Medicine* can assign text-specific online homework and other resources to students outside of the classroom.

To learn more about how MyLab Math combines proven learning applications with powerful assessment, visit **pearson.com/mylab/math** or contact your Pearson representative.

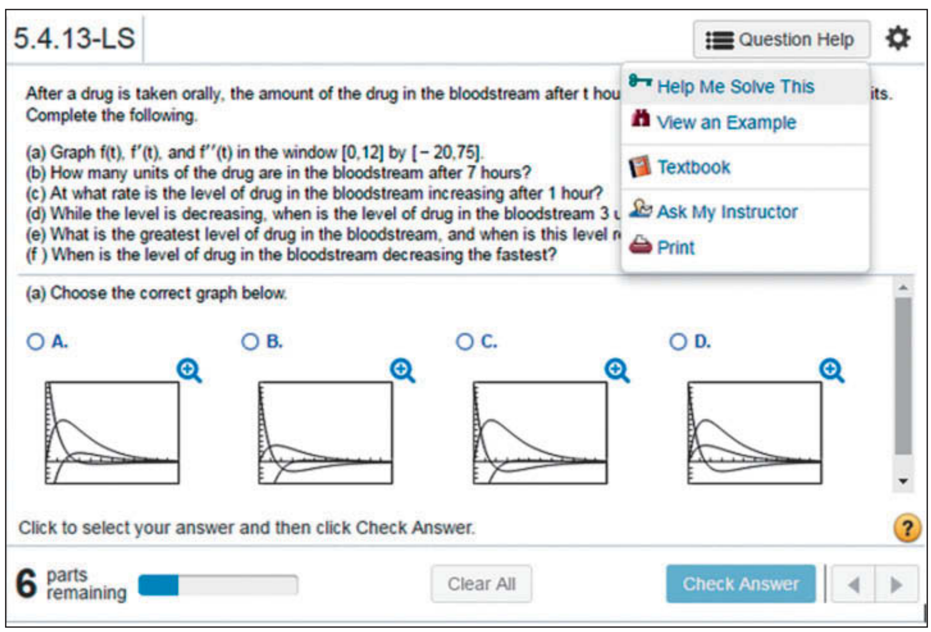
Preparedness

One of the biggest challenges in calculus courses is making sure students are adequately prepared with the prerequisite skills needed to successfully complete their course work. MyLab Math supports students with precalculus content and just-in-time remediation. Instructors can create quizzes to assess necessary prerequisite skills, then automatically assign personalized remediation for any gaps in skills that are identified.

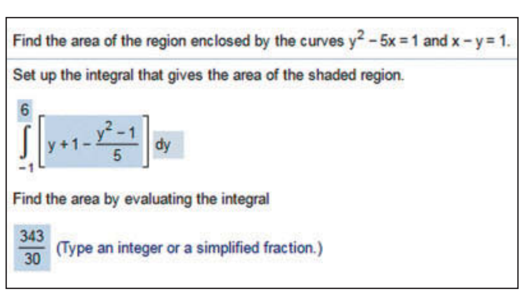
Building Understanding

MyLab Math's online homework offers students immediate feedback and tutorial assistance that helps them build understanding of key concepts.

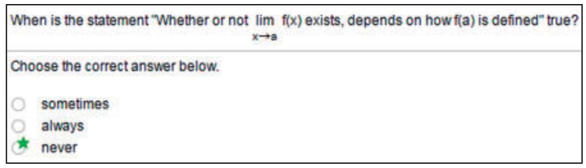
- **Exercises with immediate feedback**—the assignable exercises for this text regenerate algorithmically to give students unlimited opportunity for practice and mastery. MyLab Math provides helpful feedback when students enter incorrect answers and includes optional learning aids including Help Me Solve This, View an Example, and an eText.



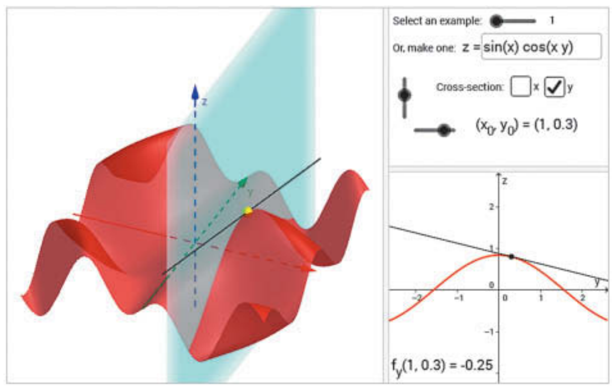
- **Setup and Solve Exercises** ask students to first describe how they will set up and approach the problem. This reinforces students' conceptual understanding of the process they are applying and promotes long-term retention of the skill.



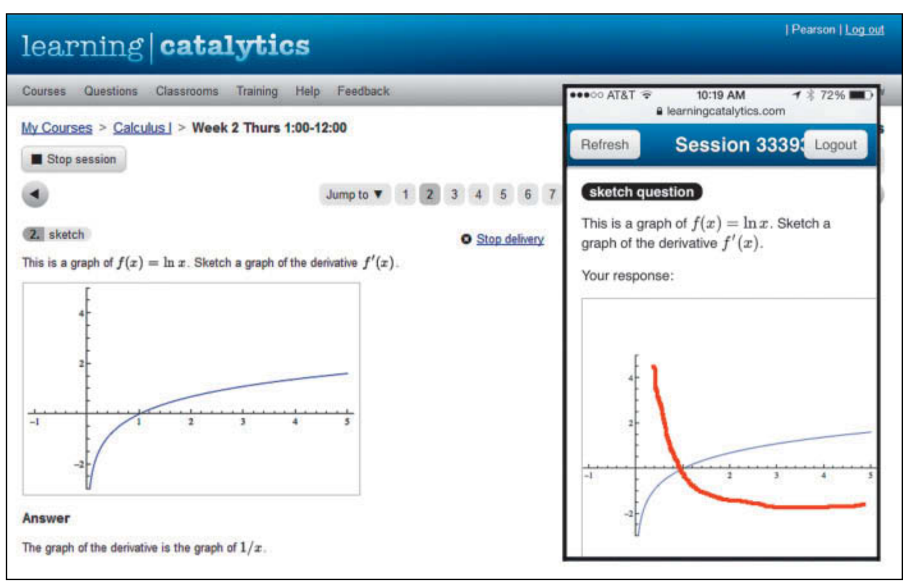
- **Additional Conceptual Questions** focus on deeper, theoretical understanding of the key concepts in calculus. These questions were written by faculty at Cornell University under a National Science Foundation grant and are also assignable through Learning Catalytics.



- **Interactive Figures** have been added to support teaching and learning. The figures are designed to be used in lecture as well as by students independently. They are editable using the freely available GeoGebra software.



- **Learning Catalytics™** is a student response tool that uses students' smartphones, tablets, or laptops to engage them in more interactive tasks and thinking during lecture. Learning Catalytics fosters student engagement and peer-to-peer learning with real-time analytics. Learning Catalytics is available to all MyLab Math users.

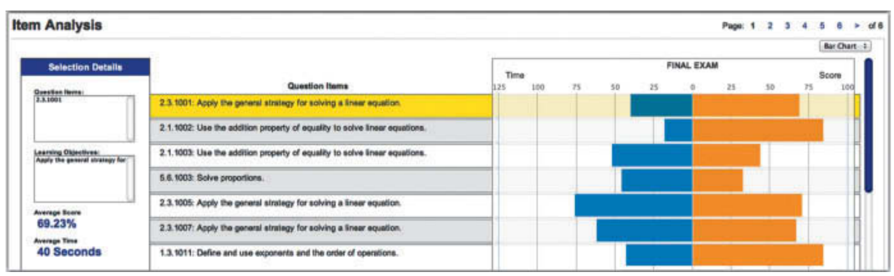


- **Complete eText** is available to students through their MyLab Math courses for the lifetime of the edition, giving students unlimited access to the eText within any course using that edition of the textbook.

- **Mathematica** manual and projects, **Maple** manual and projects, and **TI Graphing Calculator** manual utilize the most current versions of Maple and Mathematica, as well as the TI-84 Plus and TI-89. Each provides detailed guidance for integrating the software package or graphing calculator throughout the course, including syntax and commands.
- **Accessibility** and achievement go hand in hand. MyLab Math is compatible with the JAWS screen reader, and enables multiple-choice and free-response problem types to be read and interacted with via keyboard controls and math notation input. MyLab Math also works with screen enlargers, including ZoomText, MAGic, and SuperNova. More information is available at **www.pearson.com/mylab/math/accessibility**.

Instructor Support

- **Comprehensive gradebook** with enhanced reporting functionality allows you to efficiently manage your course. The gradebook meets all FERPA requirements.
  - Reporting Dashboard provides insight to view, analyze, and report learning outcomes. Student performance data is presented at the class, section, and program levels in an accessible, visual manner so you'll have the information you need to keep your students on track.
  - Item Analysis tracks class-wide understanding of particular exercises so you can refine your class lectures or adjust the course/department syllabus. Just-in-time teaching has never been easier!



- **Training and Support** – MyLab Math comes from an experienced partner with educational expertise and an eye on the future. Whether you are just getting started with MyLab Math, or have a question along the way, we're here to help you learn about our technologies and how to incorporate them into your course. To learn more about how MyLab Math helps students succeed, visit **www.pearson.com/mylab/math/support** or contact your Pearson representative.

Supplements

Instructor's Solutions Manual (download only) Provides fully worked-out solutions to every textbook exercise, including the Chapter Review problems. Available for download online within MyLab Math.

Student's Solutions Manual Provides fully worked-out solutions to the odd-numbered exercises in the sections and Chapter Reviews. Available in print (ISBN-13: 978-013-412269-4) or downloadable within MyLab Math.

Acknowledgments

This book would not have been possible without the help of numerous people. The faculty members below contributed their expertise to this project. We are most grateful to them.

Nandini Bhattacharya, *University of California Santa Cruz* *

Adena Calden, *University of Massachusetts Amherst* *

Youn-Sha Chan, *University of Houston Downtown*
Alberto Corso, *University of Kentucky* *
Hao Gao, *University of California Los Angeles* *
Guillermo Goldsztein, *Georgia Institute of Technology* *
Yvette Hester, *Texas A&M University* *
Peter Howard, *Texas A&M University*
Yang Kuang, *Arizona State University* *
Glenn Lahodny, Jr., *Texas A&M University* *
Glenn Ledder, *University of Nebraska Lincoln* *
Petr Lisonek, *Simon Fraser University*
Lawrence Marx, *University of California Davis*
Edward Migliore, *University of California Santa Cruz* *
Douglas Norton, *Villanova University* *
Heather Ramsey, *Texas A&M University* *
Patrick Shipman, *Colorado State University* *
Jeong-Mi Yoon, *University of Houston Downtown*

* Reviewers who contributed to the preparation of this edition

Thanks to students we have taught at the University of Minnesota, the University of California Los Angeles, and the University of California Davis for their constructive criticism and enthusiasm. We owe a special thanks to George Lobell, former mathematics editor at Prentice Hall, who made this book possible in the first place, and to those who helped shepherd this 4th Edition: Jeff Weidenaar, Patty Bergin, Jenn Snyder, Marielle Guiney, Kristina Evans, Emily Ockay, and other staff at Pearson; Kelly Ricci at Aptara; Ron Weickart at Network Graphics; Roger Lipsett (answers and solutions); Paul Anagnostopoulos (typist); and Paul Lorczak and Alberto Corso (accuracy checking).

Thanks also to the National Science Foundation, which has supported new approaches to teaching calculus for life sciences at UCLA under grant number DMS-1351860.

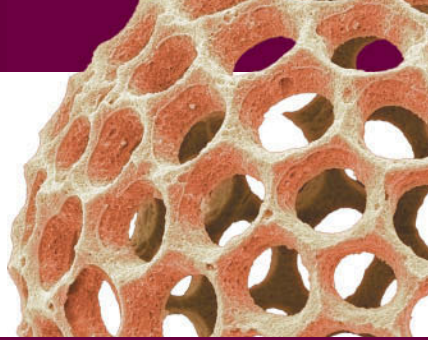
Marcus would like to express his love and gratitude to his wife, Dr. Mechel Henry, MD, for her support during this book-writing journey, and to his children, Eliot and Charlotte. You have given me so many wonderful reasons for missing deadlines on this book!

Claudia Neuhauser
neuha001@umn.edu
University of Minnesota

Marcus Roper
mroper@math.ucla.edu
University of California Los Angeles

CHAPTER

1



Preview and Review

The chapter begins with a diagnostic test on precalculus skills. Sections 1.2 and 1.3 serve as a review of algebra, trigonometry, and precalculus, material needed to master the topics covered in this book. Section 1.4 reviews graphing functions and introduces the important concept of plotting data or functions on transformed axes to determine how two variables are related. Section 1.4 also includes a subsection on visualizing verbal descriptions of biological phenomena.

A Brief Overview of Calculus

Calculus has two main ingredients: differentiation and integration. Differentiation allows us to calculate how quickly a function is changing (for example, the rate at which a population of organisms is growing). Integration allows us to calculate the area under curves. Although knowing the area under a curve may not seem very important, we will learn in Chapter 6 that integration is the opposite of differentiation (that is, integrating the rate of change of a function gets us back to the original function). Often the first step to finding a function is to find an equation for its derivative; many phenomena in biology can be modeled using **differential equations**, equations that govern the derivative or rate of change of a function. For example, in Chapter 5 we will learn how to find the rate of change of the number of cells growing in a flask. Integration enables to solve these differential equations and find a function to calculate the number of organisms in the flask at any given time. In addition to developing the theory of differential and integral calculus, you will be introduced to many examples of differential equations to describe biological phenomena including population growth, the speed of chemical reactions, the firing of neurons, and the spread of invasive species into new habitats.

The use of quantitative reasoning is becoming increasingly more important in biology—for instance, in modeling interactions among species in a community, describing the activities of neurons, explaining genetic diversity in populations, and predicting the impact of global warming on vegetation. Today, calculus (Chapters 2–11) and probability and statistics (Chapter 12) are among the most important quantitative tools of a biologist.

Section 1.1 Precalculus Skills Diagnostic Test

In this chapter we review the following important precalculus topics:

1. Algebra

2. Trigonometry

3. Visualizing functions and data

4. Translating word descriptions of biological phenomena into sketches.

We recommend that if you are self-studying this chapter you start by taking a diagnostic test, which will show you which areas you may need to review before moving on to the calculus material. The answers to these questions are in the back of this book; the answer key also tells you which subsection to review if you are unable to answer a question.

1. Write the equation of a straight line
(a) with slope 1 and y-intercept at $y = 5$.
(b) with slope 2 and passing through the point $(x, y) = (2, 3)$.
(c) passing through the points $(x, y) = (2, 1)$ and $(x, y) = (4, 7)$.

2. To convert between the temperature measured in degrees Fahrenheit (°F) and degrees Celsius (°C) we use the following formula:

$$y = \frac{5}{9}(x - 32),$$

where x is the temperature given in degrees Fahrenheit and y is the temperature given in degrees Celsius.
(a) If the temperature in Los Angeles in February is 80°F, what is the temperature in degrees Celsius?
(b) If the temperature in Rochester in February is −10°C, what is the temperature in degrees Fahrenheit?
(c) Is there any temperature that reads the same in both degrees Fahrenheit and degrees Celsius?

3. Describe in words the set of points (x, y) satisfying the equation:

$$(x + 1)^2 + (y - 5)^2 = 9.$$

4. (a) Convert the angle $\theta = \frac{\pi}{7}$ from radians to degrees.
(b) Find all solutions of the equation $\sin x = -\frac{\sqrt{3}}{2}$.
(c) Show that $1 + \tan^2\theta = \sec^2\theta$.
(d) Find all solutions for $x \in [0, \pi]$ of the equation $\cos 3x = \frac{1}{\sqrt{2}}$.

5. (a) Simplify the following expressions:
(i) $\dfrac{2^3}{2^{\frac{2}{3}}}$ **(ii)** $2^{\frac{2}{3}} \times 4$
(b) If $4^x = \frac{1}{2}$, find x.
(c) Evaluate $\log_{10} 10000$.
(d) Simplify $\log_{10} 3x + \log_{10} 5x$.
(e) Solve for x: $\ln(x^2) + \ln x = 2$.
(f) If $\ln x = 3$, calculate $\log_{10} x$.

6. (a) Find the (complex) roots of the quadratic equation $x^2 + x + 1 = 0$, simplifying your answer as much as possible.
(b) Evaluate $(1 + i) \times (2 - i)$ and simplify your answer.
(c) Show that if $z = a + ib$ for $a, b \in \mathbf{R}$ then $z + \bar{z}$ is real.

7. (a) Determine the ranges of the following functions:
(i) $f(x) = x^2, x \in [-1, 1]$. **(ii)** $f(x) = x^2, x \in \mathbf{R}$.
(b) If $f(x) = \sqrt{x}$ and $g(x) = (x + 1)^2$ find:
(i) $f(3)$. **(ii)** $(f \circ g)(x)$. **(iii)** $(g \circ f)(4)$.
(c) If $f(x) = |x|$ show that $(f \circ f)(x) = f(x)$.

8. (a) For large values of x, which of the following polynomials will return the largest value?
(i) $p_1(x) = x$ **(ii)** $p_2(x) = \frac{1}{2}x^2$ **(iii)** $p_3(x) = \frac{1}{3}x^3$
(b) The steady flow of fluid in a pipe with circular cross-section obeys ***Poiseuille's equation***. The fastest flow occurs at the center of the pipe and there is no flow where the fluid touches the walls. If r is the distance from the center of the pipe and a is the radius of the pipe, then the velocity u varies with r according to

$$u(r) = u_0\left(1 - \frac{r^2}{a^2}\right),$$

where u_0 is the maximum velocity in the pipe (see Figure 1.1).
(i) What is the degree of $u(r)$ as a polynomial in r?
(ii) What is the domain of this function?
(iii) What is the range of this function?
(iv) $u(r)$ decreases with distance from the center line. At what distance does $u(r)$ decrease to $\frac{1}{2}$ of its maximum value?

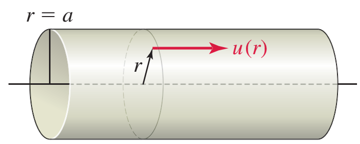


Figure 1.1 Diagram of the geometry of a circular pipe for Question 8. At distance r from the center of the pipe the flow is $u(r)$.

9. Metabolism of a particular drug in the body is described by ***Michaelis-Menten*** kinetics. If the concentration is c, the rate of metabolization (r, or the amount removed from the blood in one hour) is given by the formula

$$r(c) = \frac{c}{c + 10}.$$

(a) For which of the following concentrations is the rate of metabolization largest?
(i) $c = 1$ **(ii)** $c = 2$ **(iii)** or $c = 3$
(b) Suppose $c = 5$ initially, then the patient takes a pill and more drug is absorbed; the concentration of c doubles to 10. Does the rate of metabolization
(i) double? **(ii)** more than double? **(iii)** less than double?
(c) If $c = 5$ initially, how much would the concentration have to increase to get the rate of metabolization to double?
(d) If $c = 10$ initially, is there any concentration increase that would get the rate of metabolization to double?

10. Species-area curves are used to predict how the amount of diversity (number of species) in a particular habitat decreases if the habitat shrinks. A typical species-area relationship between N, the number of species, and A, the area of the habitat, is:

$$N = kA^z$$

where k and z are positive constants. Assume $z = \frac{1}{5}$ to answer the following questions.
(a) If the initial area is A_0, calculate how much the habitat area must shrink for the number of species to be halved.
(b) How much must it shrink to reduce the number of species to one-third of its starting value?
(c) Conversely, how much extra habitat needs to be added to increase the number of species to twice its starting value?

11. The size N of a population of cells can be modeled by an exponential law:

$$N = N_0 e^{rt}$$

where t is the time elapsed since the population growth began and N_0 and r are positive constants.
(a) Calculate N_0 and r given the following population-size data. There were initially 1000 cells in the population (i.e., $N(0) = 1000$). At $t = 2$ there are 1000 more cells in the population (i.e., $N(2) = 2000$).
(b) Using the formula and the parameters from part (a), calculate how much time must elapse before 1000 more cells are added to the population (i.e., for what t does $N(t) = 3000$)?
(c) How much time must elapse before the population doubles from its size at $t = 2$ (i.e., for what t does $N(t) = 4000$)?

12. Calculate the inverses of the following functions (i.e., find $f^{-1}(x)$ such that $(f^{-1} \circ f)(x) = x$).
(a) $f(x) = x^2 + 1$ for $x \geq 0$.
(b) $f(x) = 2\ln(x + 1)$ for $x > -1$. **(c)** $f(x) = x^5$.

13. (a) Combine the terms in the following expressions into a single logarithm:

(i) $\ln(x) + \ln(x^2 + 1)$ **(ii)** $\log(x^{1/3}) - \log((x+1)^{1/3})$
(iii) $2 + \log_2(x)$

(b) Use the change of base formula to turn the logarithms below into natural logs:

(i) $\log_2 7$ **(ii)** $\log 6$ **(iii)** $\log_x 2$

14. (a) Give the period and amplitude of the following trigonometric functions:

(i) $f(x) = -2\sin x$ **(ii)** $f(x) = 2\cos 3x$
(iii) $f(x) = 3\cos\frac{\pi x}{2}$

(b) Find the range and maximum domain of:

(i) $f(x) = \tan x$ **(ii)** $f(x) = \cos x$

(c) Explain how the curve $y = 3\cos 2x$ is related to the curve $y = \cos x$.

15. The function $f(x)$ is drawn in Figure 1.2. Sketch the graphs of:
(a) $f(x + 2)$ **(b)** $f(x) + 1$ **(c)** $-f(x)$ **(d)** $f(-x)$
(e) $f(\frac{x}{2})$

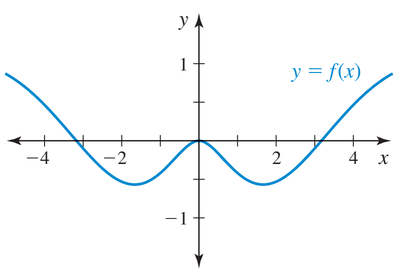


Figure 1.2 Plot of the function $f(x)$ for Question 15.

16. The typical weights (in kilograms) of five popular dog breeds are shown on the logarithmic number line shown in Figure 1.3.
(a) What is the typical mass of the following breeds:

(i) Chihuahua **(ii)** Labrador retriever **(iii)** St. Bernard

(b) How much heavier is an adult Dalmatian than a puppy?
(c) A typical house cat weighs 5 kg. Copy the number line axes and draw a point on the number line to represent the house cat.

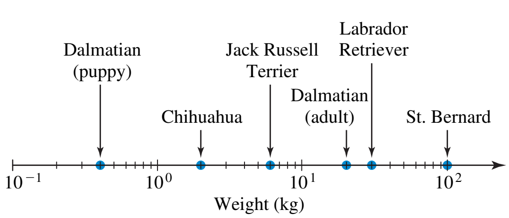


Figure 1.3 Weights of popular dog breeds in kilograms, plotted on a logarithmic number line for Question 16.

17. Each of the graphs in Figures 1.4 through 1.7 shows how one quantity (plotted on the y-axis) varies with a second quantity (plotted on the x-axis). In each case state whether the graph shows

1. a power law dependence $y = kx^a$ for some constants k and a. In this case, give the value of a.

or

2. an exponential dependence $y = ka^x$ for some constants k and a. In this case, give the value of a.
(a) Dependence of number of languages spoken (D) on the area (A) of the globe that is sampled.
(b) Frequency of earthquakes in Southern California (N) with shaking magnitude larger than m.
(c) Size of a bacterial population growing in a microfluidic chamber (N) as a function of time (t).
(d) Number of HIV viruses in 1 ml of blood (N) against time (t) for a patient receiving HIV treatment.

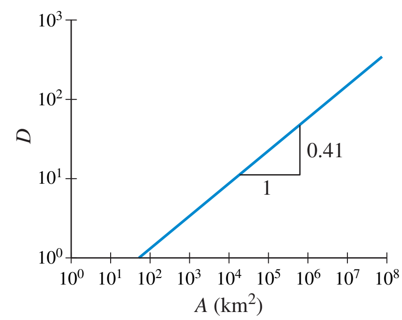


Figure 1.4 Question 17(a). Number of languages spoken, D as a function of area, A. Adapted from ***Gomes et al. (1999)***.

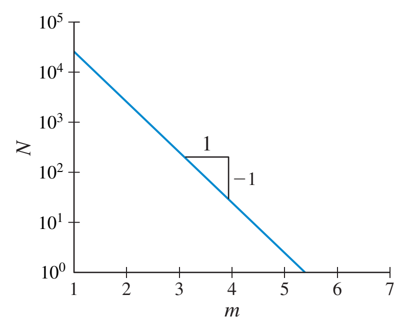


Figure 1.5 Question 17(b). Number of earthquakes (N) with magnitude larger than m, in one year as a function of m. Graph adapted from ***Rundle et al. (2003)***.

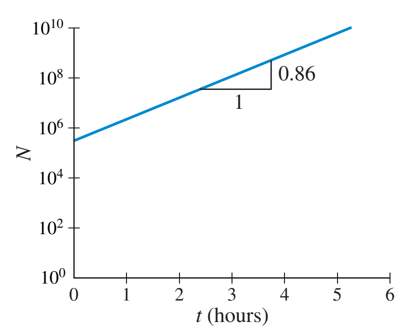


Figure 1.6 Question 17(c). Bacterial population size (N) as a function of time (t). Data adapted from ***Balaban et al. (2004)***.

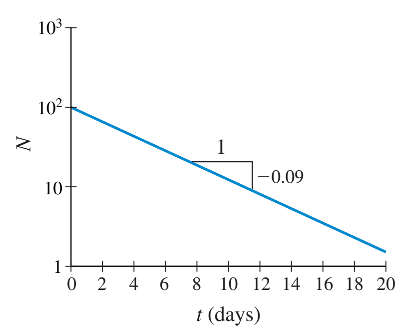


Figure 1.7 Question 17(d). Number of HIV viruses (N) as a function of time (t). Data adapted from ***Ho et al. (2004)***.

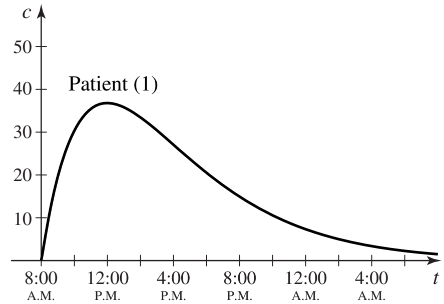


Figure 1.8 Question 18. Concentration of an ADHD drug (c) in a patient's blood, as a function of time (t).

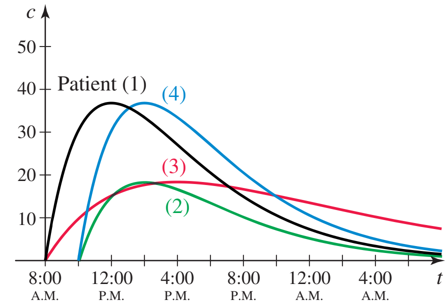


Figure 1.9 Question 18. Identify the curve corresponding to each patient.

18. You are studying how medication that is used to treat Attention Deficit Hyperactivity Disorder (ADHD) is metabolized. A healthy patient (patient 1) takes a dose of the drug at 8 A.M. and you measure the concentration in their blood plasma at hourly intervals. You obtain the data shown in Figure 1.8.

Three other patients also receive the drug, but follow slightly different dose regimens. Your task is to identify which curve in Figure 1.9 describes which patient based on the word descriptions below.

(a) Patient A receives the same dose as patient 1 but at 10:00 rather than 8:00 A.M.

(b) Patient B receives an extended release form of the drug at 8:00 A.M., which takes longer to enter the bloodstream.

(c) At 10:00 A.M. Patient C receives half the dose that is given to patient A.

1.2 Preliminaries

This section reviews some of the concepts and techniques from algebra and trigonometry that are frequently used in calculus. The problems at the end of the section will help you reacquaint yourself with this material.

1.2.1 The Real Numbers

Figure 1.10 The real-number line.

The **real numbers** can most easily be visualized on the **real-number line** (see Figure 1.10), on which numbers are ordered so that if $a < b$, then a is to the left of b. Sets (collections) of real numbers are typically denoted by the capital letters A, B, C, etc. To describe the set A, we write

$$A = \{x : \text{condition}\}$$

where "condition" tells us which numbers are in the set A. The most important sets in calculus are **intervals**. We use the following notations: If $a < b$, then

$$\text{the } \textbf{open} \text{ interval } (a, b) = \{x : a < x < b\}$$

and

$$\text{the } \textbf{closed} \text{ interval } [a, b] = \{x : a \le x \le b\}$$

We also use **half-open** intervals:

$$[a, b) = \{x : a \le x < b\} \qquad \text{and} \qquad (a, b] = \{x : a < x \le b\}$$

Unbounded intervals are sets of the form $\{x : x > a\}$. Here are the possible cases:

$$[a, \infty) = \{x : x \ge a\}$$
$$(-\infty, a] = \{x : x \le a\}$$
$$(a, \infty) = \{x : x > a\}$$
$$(-\infty, a) = \{x : x < a\}$$

The symbols "∞" and "$-\infty$" mean "plus infinity" and "minus infinity," respectively. These symbols are *not* real numbers, but are used merely for notational convenience. The real-number line, denoted by $\mathbf{R}$, does not have endpoints, and we can write $\mathbf{R}$ in the following equivalent forms:

$$\mathbf{R} = \{x : -\infty < x < \infty\} = (-\infty, \infty)$$

The location of the number 0 on the real-number line is called the **origin**, and we can measure the distance of the number x to the origin. For instance, -5 is 5 units to the left of the origin. A convenient notation for measuring distances from the origin on the real-number line is the absolute value of a real number.

Definition The **absolute value** of a real number a, denoted by $|a|$, is

$$|a| = \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a < 0 \end{cases}$$

For example, $|-7| = -(-7) = 7$. We can use absolute values to find the distance between any two numbers x_1 and x_2 as follows:

$$\text{distance between } x_1 \text{ and } x_2 = |x_1 - x_2|$$

Note that $|x_1 - x_2| = |x_2 - x_1|$. So the distance between x_1 and x_2 is, as you would expect, the same as the distance between x_2 and x_1. To find the distance between -2 and 4, we compute $|-2-4| = |-6| = 6$, or $|4-(-2)| = |4+2| = 6$.

We will frequently need to solve equations containing absolute values, for which the following property is useful:

Property of Absolute Value Equations Let $b \geq 0$. Then $|a| = b$ is equivalent to $a = \pm b$.

EXAMPLE 1 Solve $|x-4| = 2$.

Solution Applying the Property of Absolute Value Equations, we obtain: $x - 4 = \pm 2$, (i.e., $x = 4 \pm 2$). So either $x = 4 - 2 = 2$ or $x = 4 + 2 = 6$. The solutions, illustrated graphically in Figure 1.11, are therefore $x = 6$ and $x = 2$. The points of intersection of $y = |x-4|$ and $y = 2$ are at $x = 6$ and $x = 2$. Solving $|x-4| = 2$ can also be interpreted as finding the two numbers that have distance 2 from 4. •

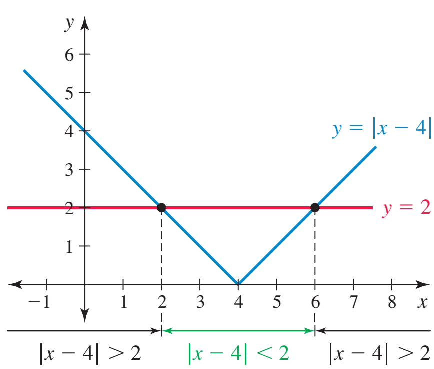


Figure 1.11 The graph of $y = |x-4|$ and $y = 2$. The points of intersection are at $x = 6$ and $x = 2$.

When there are absolute value signs on both sides of the equation there can be $\pm$ on both sides of the equation as illustrated in the next example.

EXAMPLE 2 Solve $|\frac{3}{2}x - 1| = |\frac{1}{2}x + 1|$.

Solution Applying the Property of Absolute Value Equations rule to both sides of the equation we can replace the absolute value signs by $\pm$ on both sides:

$$\pm\left(\frac{3}{2}x - 1\right) = \pm\left(\frac{1}{2}x + 1\right)$$

Now the $\pm$ signs are independent of each other. We can choose $+$ on the left side and $-$ on the right side. So it seems that there are four possibilities to consider: $(+,+)$,

$(+, -)$, $(-, +)$, $(-, -)$, where $(+, -)$ denotes choosing $+$ sign on the right side and $-$ sign on the left side. However, some of the possibilities give equivalent answers; for example:

$$\text{Taking} \quad (+,+) \quad \text{gives:} \qquad \left(\frac{3}{2}x - 1\right) = \left(\frac{1}{2}x + 1\right) \tag{1.1}$$

while

$$\text{Taking} \quad (-,-) \quad \text{gives:} \qquad -\left(\frac{3}{2}x - 1\right) = -\left(\frac{1}{2}x + 1\right). \tag{1.2}$$

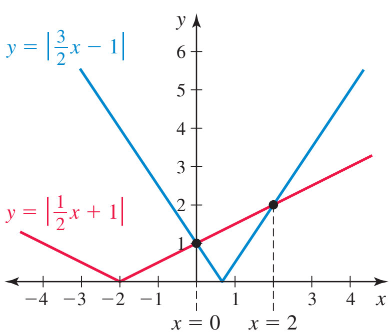


Figure 1.12 The graphs of $y = |\frac{3}{2}x - 1|$ and $y = |\frac{1}{2}x + 1|$. The points of intersection are at $x = 0$ and $x = 2$.

If we multiply both sides by the term -1 we can turn (1.2) into (1.1), so the two equations are equivalent.

With this in mind we need only consider only the choices of sign: $(+, +)$ and $(+, -)$:

$$\text{Taking } (+,+) \text{ gives } \left(\frac{3}{2}x - 1\right) = \left(\frac{1}{2}x + 1\right) \qquad (-,-) \text{ gives the same equation.}$$

$$x = 2. \qquad \text{Subtract } \tfrac{1}{2}x \text{ from both sides, add 1 to both sides.}$$

$$\text{Taking } (+,-) \text{ gives } \left(\frac{3}{2}x - 1\right) = -\left(\frac{1}{2}x + 1\right) \qquad (-,+) \text{ gives the same equation.}$$

$$2x = 0 \qquad \text{Add } \tfrac{1}{2}x \text{ to both sides, add 1 to both sides.}$$

$$x = 0.$$

A graphical solution of this example is shown in Figure 1.12. •

Returning to Example 1, where we found the two points whose distance from 4 was equal to 2, we can also try to find those points whose distance from 4 is less than (or greater than) 2. This amounts to solving inequalities with absolute values. Looking back at Figure 1.11, we see that the set of x-values whose distance from 4 is less than 2 (i.e., $|x - 4| < 2$) is the interval $(2, 6)$. Similarly, the set of x-values whose distance from 4 is greater than 2 (i.e., $|x - 4| > 2$) is the union of the two intervals $(-\infty, 2)$ and $(6, \infty)$, or $(-\infty, 2) \cup (6, \infty)$.

To solve absolute-value inequalities, the following two properties are useful:

> **Using the Absolute Value in Inequalities** Let $b > 0$. Then
>
> **1.** $|a| < b$ is equivalent to $-b < a < b$.
>
> **2.** $|a| > b$ is equivalent to $a > b$ or $a < -b$.

EXAMPLE 3

(a) Solve $|2x - 5| < 3$. **(b)** Solve $|4 - 3x| \geq 2$.

Solution

(a) We rewrite $|2x - 5| < 3$ as

$$-3 < 2x - 5 < 3$$

$$2 < 2x < 8 \qquad \text{Add 5 to all three parts}$$

$$1 < x < 4 \qquad \text{Divide all three parts by 2}$$

The solution is therefore the set $\{x : 1 < x < 4\}$. In interval notation, the solution can be written as the open interval $(1, 4)$.

(b) To solve $|4 - 3x| \geq 2$, we go through the following steps:

Either:	$4 - 3x \geq 2$	or	$4 - 3x \leq -2$	When both sides of inequality are divided or multiplied by a negative number, the inequality must be reversed.
	$-3x \geq -2$		$-3x \leq -6$	
	$x \leq \frac{2}{3}$		$x \geq 2$	

The solution is the set $\{x : x \leq \frac{2}{3} \text{ or } x \geq 2\}$, or, in interval notation, $(-\infty, \frac{2}{3}] \cup [2, \infty)$. •

1.2.2 Lines in the Plane

We will frequently encounter situations in which the relationship between quantities can be described by a **linear equation**. For example, the maximum bite strength of spotted hyenas increases with age of the hyena. Specifically, if x is the age of the hyena in months and y is the maximum bite force (in newtons) that the hyena is capable of exerting, then as shown by ***Binder and Van Valkenburgh (2000)***:

$$y = 166.0 + 12.7x \tag{1.3}$$

Equation (1.3) is an example of a **linear equation**, and we say that x and y satisfy a linear equation.

The graph of a linear equation is a straight line. The equation of the straight line can be written using any of three different forms:

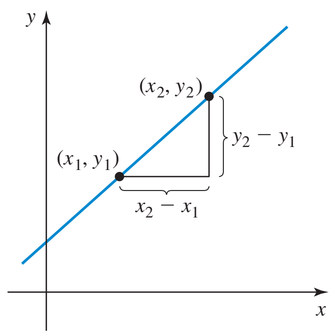


Figure 1.13 The slope of a straight line.

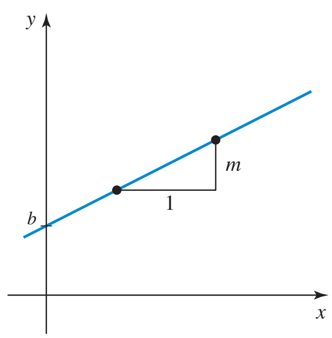


Figure 1.14 The slope x-intercept form of the equation for a straight line requires the slope, m, and y-intercept, b. The triangle shows that for every unit the line travels in the x-direction it goes up m units in the y-direction.

1. The **standard** form of a linear equation is given by
$$Ax + By + C = 0$$
where A, B, and C are constants, A and B are not both equal to 0, and x and y are the two variables.
2. If the two points (x_1, y_1) and (x_2, y_2) lie on a straight line, then the **slope** of the line is
$$m = \frac{y_2 - y_1}{x_2 - x_1}$$
(See Figure 1.13.) Two points (or one point and the slope) are sufficient to determine the equation of a straight line.
If you are given one point and the slope, provided m is finite you can use the **point–slope** form of a straight line to write its equation, given by
$$y - y_0 = m(x - x_0)$$
where m is the slope and (x_0, y_0) is a point on the line. If you are given two points, first compute the slope and then use one of the points and the slope to find the equation of the straight line in point–slope form.
3. Lastly, the **slope–intercept** form is:
$$y = mx + b$$
where m is the slope and b is the y-intercept, which is the point of intersection of the line with the y-axis; the y-intercept has coordinates $(0, b)$. (See Figure 1.14.)

Definition **Forms of Linear Equations**

$$\begin{array}{rl} Ax + By + C = 0 & \text{(Standard Form)} \\ y - y_0 = m(x - x_0) & \text{(Point–Slope Form)} \\ y = mx + b & \text{(Slope–Intercept Form)} \end{array}$$

EXAMPLE 4 Determine, in slope–intercept form, the equation of the line passing through $(-2, 1)$ and $(3, -\frac{1}{2})$.

Solution The slope of the line is

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{-\frac{1}{2} - 1}{3 - (-2)} = \frac{-\frac{3}{2}}{5} = -\frac{3}{10}$$

Using the point–slope form:

$$y - 1 = -\frac{3}{10}(x - (-2)) \qquad m = -\tfrac{3}{10},\ (x_0, y_0) = (-2, 1)$$

or, in slope–intercept form,

$$y = -\frac{3}{10}x + \frac{2}{5}$$

We could have used the other point, $(3, -\frac{1}{2})$, and obtained the same result.

Two special cases occur if either $m = 0$ or m is undefined (see Figure 1.15).

Equations for Horizontal and Vertical Lines:

$$\begin{array}{ll} y = k & \text{horizontal line (slope 0)} \\ x = h & \text{vertical line (slope undefined)} \end{array}$$

In the next example, we show how to determine the slope and the y-intercept of a given straight line.

EXAMPLE 5 Determine the slope and the y-intercept of the line $3y - 2x + 9 = 0$.

Solution Rearrange into slope-intercept form by putting y and x on opposite sides of the equation:

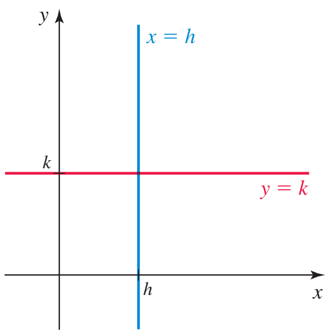


Figure 1.15 The horizontal line $y = k$ and the vertical line $x = h$.

$$3y = 2x - 9.$$

$$y = \frac{2x}{3} - 3 \qquad \text{Put in form } y = mx + b \text{, so divide by 3.}$$

We can now read off the slope $m = \frac{2}{3}$ and the y-intercept $b = -3$. •

When two quantities x and y are linearly related so that

$$y = mx$$

we say that y is **proportional** to x, with m denoting the **constant of proportionality**, and we write

$$y \propto x$$

The symbol $\propto$ is read "is proportional to." If we write Equation (1.3) in the form

$$y - 166.0 = 12.7x$$

then $y - 166.0$ is proportional to the hyena's age, x, with constant of proportionality 12.7, and we can write

$$y - 166.0 \propto x$$

There are two more properties of straight lines we wish to mention. When two lines l_1 and l_2 in the plane either (a) have no points in common or (b) are identical, they are called **parallel**, denoted by $l_1 \parallel l_2$. The following criterion is useful in deciding whether two lines are parallel: Two lines l_1 and l_2 are parallel ($l_1 \parallel l_2$) if and only if their slopes are identical. For two nonvertical lines l_1 and l_2 with slopes m_1 and m_2, respectively, the criterion becomes

$$l_1 \parallel l_2 \qquad \text{if and only if} \qquad m_1 = m_2$$

Parallel lines are either identical or do not intersect anywhere.

Two lines l_1 and l_2 are called **perpendicular** ($l_1 \perp l_2$) if they intersect at an angle of 90°. The following criterion is useful for deciding whether two lines are perpendicular: Two nonvertical lines are perpendicular if and only if their slopes are negative reciprocals. That is, if l_1 and l_2 are nonvertical lines with slopes m_1 and m_2, then

$$l_1 \perp l_2 \qquad \text{if and only if} \qquad m_1 m_2 = -1$$

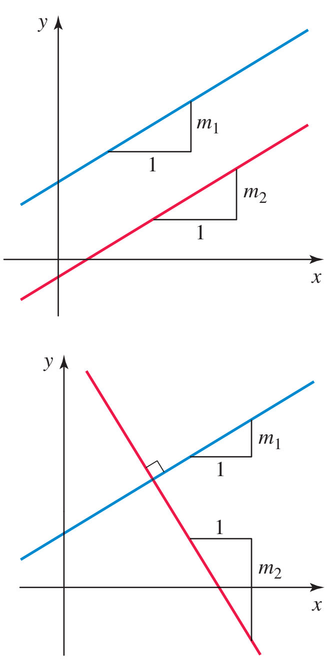


Figure 1.16 Two lines with slopes m_1 and m_2 are parallel if and only if $m_1 = m_2$, and perpendicular if and only if $m_1 m_2 = -1$.

We will prove this result in Problem 54 at the end of this section. Figure 1.16 summarizes the slope conditions for parallel and perpendicular lines.

1.2.3 Equation of the Circle

A **circle** is the set of all points at a given distance, called the **radius**, from a given point, called the **center**. If r is the radius of the circle and (x_0, y_0) is its center (see Figure 1.17), then, using the Pythagorean theorem, we find that for all points (x, y) lying on the circle

$$r^2 = (x - x_0)^2 + (y - y_0)^2$$

If $r = 1$ and $(x_0, y_0) = (0, 0)$, the circle is called the **unit circle**.

EXAMPLE 6 Find the equation of the circle with center (2, 3) and passing through (5, 7).

Solution Using the Pythagorean theorem, we can compute the distance in the plane between (2, 3) and (5, 7):

$$\sqrt{(5-2)^2 + (7-3)^2} = \sqrt{9+16} = 5$$

Thus, this circle has radius 5 and center (2, 3), and its equation is

$$25 = (x-2)^2 + (y-3)^2$$

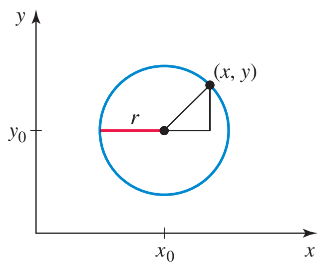


Figure 1.17 Circle with radius r centered at (x_0, y_0).

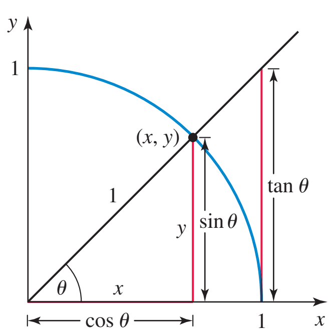


Figure 1.18 The trigonometric functions on a unit circle.

1.2.4 Trigonometry

We will need a few results from trigonometry. Recall that angles are measured in either degrees or radians and that a complete revolution on a unit circle (Figure 1.18) corresponds to 360°, or 2π. For reasons that will become clear, the radians are preferred over degrees in calculus. To convert between degree and radian measure, we use the formula

$$\boxed{\frac{\theta \text{ measured in degrees}}{360^\circ} = \frac{\theta \text{ measured in radians}}{2\pi}}$$

For instance, to convert 23° into radian measure we compute

$$\theta = 23^\circ \frac{2\pi}{360^\circ} = 0.401$$

To convert $\frac{\pi}{6}$ into degrees we compute

$$\theta = \frac{\pi}{6}\frac{360^\circ}{2\pi} = 30^\circ$$

There are four trigonometric functions that you should be familiar with: **sine**, **cosine**, **tangent**, and **secant**; the other two, cotangent and cosecant, are rarely used. The six are defined on a unit circle (see Figure 1.18) and are abbreviated as sin, cos, tan, sec, cot, and csc, respectively. Recall that a positive angle is measured counterclockwise from the positive x-axis, whereas a negative angle is measured clockwise. If (x, y) is a point on the circle, and θ is the angle between (x, y) and the x-axis, then:

Definition **Trigonometric Functions**

$$\sin\theta = \frac{y}{1} = y \qquad \csc\theta = \frac{1}{\sin\theta} = \frac{1}{y}$$

$$\cos\theta = \frac{x}{1} = x \qquad \sec\theta = \frac{1}{\cos\theta} = \frac{1}{x}$$

$$\tan\theta = \frac{y}{x} \qquad \cot\theta = \frac{1}{\tan\theta} = \frac{x}{y}$$

There are a number of frequently used trigonometric identities. First, since $\tan\theta = y/x$ with $y = \sin\theta$ and $x = \cos\theta$, it follows that

$$\boxed{\tan\theta = \frac{\sin\theta}{\cos\theta}}$$

Now, applying the Pythagorean theorem to the triangle in Figure 1.18 and using the notation $\sin^2\theta = (\sin\theta)^2$, we find that

$$\boxed{\sin^2\theta + \cos^2\theta = 1}$$

Next, if we divide the preceding identity by $\cos^2\theta$, we obtain

$$\frac{\sin^2\theta}{\cos^2\theta} + 1 = \frac{1}{\cos^2\theta}$$

Using $\tan\theta = \sin\theta/\cos\theta$ and $\sec\theta = 1/\cos\theta$, we can write this as

$$\boxed{\tan^2\theta + 1 = \sec^2\theta}$$

In the next example, we solve a trigonometric equation.

EXAMPLE 7 Solve: $2\sin\theta\cos\theta = \cos\theta$ on $[0, 2\pi)$.

Solution Bring $\cos\theta$ to the left side and factor $\cos\theta$ to obtain

$$\cos\theta(2\sin\theta - 1) = 0$$ Do not cancel $\cos\theta$ because we might have $\cos\theta = 0$.

That is,

$$\cos\theta = 0 \qquad \text{or} \qquad 2\sin\theta - 1 = 0$$

Solving $\cos\theta = 0$, we find that $\theta = \frac{\pi}{2}$ or $\theta = \frac{3\pi}{2}$
Solving $2\sin\theta - 1 = 0$, we get $\sin\theta = \frac{1}{2}$, which yields $\theta = \frac{\pi}{6}$ or $\theta = \frac{5\pi}{6}$
The solution set is therefore $\{\frac{\pi}{6}, \frac{\pi}{2}, \frac{5\pi}{6}, \frac{3\pi}{2}\}$.

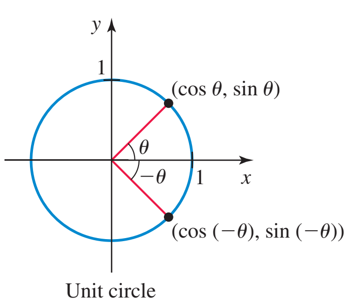


Figure 1.19 Using the unit circle to define trigonometric identities.

Figure 1.19 yields the following two identities when we compare the two angles θ and $-\theta$ (a positive angle is measured counterclockwise from the positive x-axis, whereas a negative angle is measured clockwise):

$$\boxed{\sin(-\theta) = -\sin\theta \qquad \text{and} \qquad \cos(-\theta) = \cos\theta}$$

Some exact trigonometric values are collected in Table 1-1. Rewriting Table 1-1 will make it easier to re-create the table in case you forget the exact values. Using $\tan\theta = \sin\theta/\cos\theta$, you immediately get the values for $\tan\theta$.

TABLE 1-1 Some Exact Trigonometric Values

Angle θ	**0 (0°)**	$\frac{\pi}{6}$ **(30°)**	$\frac{\pi}{4}$ **(45°)**	$\frac{\pi}{3}$ **(60°)**	$\frac{\pi}{2}$ **(90°)**
sin θ	$\frac{1}{2}\sqrt{0} = 0$	$\frac{1}{2}\sqrt{1} = \frac{1}{2}$	$\frac{1}{2}\sqrt{2} = \frac{1}{\sqrt{2}}$	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{4} = 1$
cos θ	$\frac{1}{2}\sqrt{4} = 1$	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{2} = \frac{1}{\sqrt{2}}$	$\frac{1}{2}\sqrt{1} = \frac{1}{2}$	$\frac{1}{2}\sqrt{0} = 0$

1.2.5 Exponentials and Logarithms

Exponentials and logarithms are particularly important in biological contexts.

An exponential is an expression of the form

$$a^r$$

where a is called the **base** and r the **exponent**. To ensure a^r is real, we will assume a is positive or that r is an integer or a rational number of the form p/q where p is an integer and q is a non-zero integer. We summarize some of the properties of an exponential as follows:

Properties of Exponentials

$$a^r a^s = a^{r+s} \qquad (ab)^r = a^r b^r$$

$$\frac{a^r}{a^s} = a^{r-s} \qquad \left(\frac{a}{b}\right)^r = \frac{a^r}{b^r}$$

$$a^{-r} = \frac{1}{a^r} \qquad (a^r)^s = a^{rs}$$

EXAMPLE 8 Evaluate the following exponential expressions:

(a) $3^2 3^{5/2}$ **(b)** $\frac{2^{-4}2^3}{2^2}$ **(c)** $\frac{a^k a^{3k}}{a^{5k}}$

Solution

(a) $3^2 3^{5/2} = 3^{2+5/2} = 3^{9/2}$

(b) $\frac{2^{-4}2^3}{2^2} = \frac{2^{-1}}{2^2} = 2^{-1-2} = 2^{-3} = \frac{1}{2^3} = \frac{1}{8}$

(c) $\frac{a^k a^{3k}}{a^{5k}} = a^{k+3k-5k} = a^{-k} = \frac{1}{a^k}$ •

Logarithms allow us to solve equations with unknown exponents.

$$x = \log_a y \quad \text{is equivalent to} \quad y = a^x$$

In other words, a logarithm is an exponent. The expression $\log_a y$ is the exponent on the base a that yields the number y. Logarithms are defined only for $y > 0$ (where the base is assumed to be positive and different from 1).

For example, if we are given the equation

$$2^x = 8$$

we can see that $x = 3$ is a solution of this equation so

$$x = \log_2 8 = 3$$ Work from the definition with $a \to 2$, $y \to 8$.

EXAMPLE 9 Find the real number x that satisfies

(a) $\log_3 x = -2$ **(b)** $\log_{1/2} 8 = x$

Solution

(a) We write the equation in the equivalent form

$$x = 3^{-2}$$ Work from the definition with $a \to 3$, $x \to -2$, $y \to x$.

$$x = \frac{1}{3^2} = \frac{1}{9}$$ Simplify the exponential.

(b) We write the equation in the equivalent form

$$\begin{aligned}\left(\frac{1}{2}\right)^x &= 8 && a \to \tfrac{1}{2}, x \to 8, y \to x\\ 2^{-x} &= 8\\ 2^{-x} &= 2^3 && 2^3 = 8\\ 2^x &= 2^{-3}\end{aligned}$$

Setting the exponents equal to each other, we find that $x = -3$. Note that, in order to compare exponents, the bases must be the same. •

Some important properties of logarithms are as follows:

Properties of Logarithms

$$\log_a(xy) = \log_a x + \log_a y$$

$$\log_a\left(\frac{x}{y}\right) = \log_a x - \log_a y$$

$$\log_a x^r = r \log_a x$$

The most important logarithm is the **natural logarithm**, which has the number e as its base. The number e is an irrational number whose value is approximately 2.7182818. The natural logarithm is written $\ln x$; that is, $\log_e x = \ln x$.

In some math textbooks $\ln x$ is written as $\log x$. We will follow the convention that $\log x = \log_{10} x$.

EXAMPLE 10 Assume that x and y are positive, and simplify the following expressions:

(a) $\log_3(9x^2)$ **(b)** $\log_5 \frac{x^2+3}{5x}$ **(c)** $-\ln \frac{1}{2}$ **(d)** $\ln \frac{3x^2}{\sqrt{y}}$

Solution

(a) $\log_3(9x^2) = \log_3 9 + \log_3 x^2 = 2 + 2\log_3 x$

(b) $\log_5 \frac{x^2+3}{5x} = \log_5(x^2+3) - \log_5 5 - \log_5 x$

$= \log_5(x^2+3) - 1 - \log_5 x$ $\log_5(x^2+3)$ cannot be simplified any further.

(c) $-\ln \frac{1}{2} = \ln(\frac{1}{2})^{-1} = \ln 2$

(d) $\ln \frac{3x^2}{\sqrt{y}} = \ln 3 + \ln x^2 - \ln \sqrt{y}$

$= \ln 3 + 2\ln x - \frac{1}{2}\ln y$ $\ln \sqrt{y} = \ln y^{1/2} = \frac{1}{2}\ln y$ •

In algebra, you learned how to solve equations of the form $e^{2x} = 3$ or $\ln(x+1) = 5$. We will need to do this frequently. The key to solving such equations are the two identities

$$\log_a a^x = x \quad \text{and} \quad a^{\log_a x} = x$$

The next example illustrates how to use these identities.

EXAMPLE 11 Solve for x.

(a) $e^{2x} = 3$ **(b)** $\ln(x+1) = 5$ **(c)** $5^{2x-1} = 2^x$

Solution

(a) To solve $e^{2x} = 3$ for x, we take logarithms to base e on both sides:

$$\begin{aligned}\ln e^{2x} &= \ln 3\\ 2x &= \ln 3 && \ln e^{2x} = 2x\\ x &= \frac{1}{2}\ln 3\end{aligned}$$

(b) To solve $\ln(x+1) = 5$, we write the equation in exponential form:

$$\begin{aligned} e^{\ln(x+1)} &= e^5 \\ x + 1 &= e^5 \\ x &= e^5 - 1 \end{aligned}$$

(c) To solve $5^{2x-1} = 2^x$ for x, we observe that the bases on the left hand side is 5, and on the right hand side 2. We therefore cannot compare the exponents directly. Instead, we take logarithms on both sides. Any positive base (different from 1) for the logarithm would work, and we choose base e, since it is the most commonly used base in calculus. Doing so yields

$$\begin{aligned} \ln 5^{2x-1} &= \ln 2^x \\ (2x-1)\ln 5 &= x \ln 2 && \text{Simplify.} \\ 2x \ln 5 - x \ln 2 &= \ln 5 && \text{Solve for } x. \\ x(2\ln 5 - \ln 2) &= \ln 5 && \text{Factor.} \\ x &= \frac{\ln 5}{2\ln 5 - \ln 2} && \text{Divide by } 2\ln 5 - \ln 2. \end{aligned}$$

1.2.6 Complex Numbers and Quadratic Equations

The square of a real number is always nonnegative. However, there are situations in which we need to take a square root of a negative number. Since the resulting square root cannot be a real number, we introduce a new symbol, which we denote by i, that will allow us to deal with this case. We set

$$i^2 = -1$$

The symbol i is called the **imaginary unit**. Thus, instead of writing $\sqrt{-17}$, for instance, we can now write $\sqrt{-17} = \sqrt{-1}\,\sqrt{17} = i\sqrt{17}$.

The symbol i allows us to introduce a new number system, the set of **complex numbers**:

A **complex number** is a number of the form

$$z = a + bi$$

where a and b are real numbers. The real number a is the **real part** of $a + bi$, and the real number b is the **imaginary part**.

For instance, $-3 + 7i$ has real part -3 and imaginary part 7, and $2 - 5i$ has real part 2 and imaginary part -5. Since $a + 0i = a$, it follows that the set of real numbers is a subset of the set of complex numbers. Complex numbers of the form bi are called **purely imaginary numbers**.

Two complex numbers are equal if their respective real and imaginary parts are equal; that is,

$$a + bi = c + di \qquad \text{if and only if} \qquad a = c \quad \text{and} \quad b = d$$

To add two complex numbers, we use the following rule:

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

This rule says that real and imaginary parts are added separately. To calculate the product of two complex numbers, we proceed as follows:

$$\begin{aligned}(a+bi)(c+di) &= ac + adi + bci + bdi^2 \\ &= ac + (ad+bc)i - bd \\ &= (ac-bd) + (ad+bc)i\end{aligned}$$

Note that we used $i^2 = -1$ in the penultimate step. There is no need to memorize the product of two complex numbers, since we can always compute it by the distributive law.

EXAMPLE 12 Simplify the expressions:

(a) $(2+3i) - (5-6i)$ **(b)** $(5-3i)(1+2i)$

Solution **(a)** $(2+3i) - (5-6i) = 2 + 3i - 5 + 6i = -3 + 9i$,

(b) $(5-3i)(1+2i) = 5 + 10i - 3i - 6i^2 = 5 + 7i - (6)(-1) = 11 + 7i$. •

> If $z = a + bi$ is a complex number, its **conjugate**, denoted by $\overline{z}$, is defined as
>
> $$\overline{z} = a - bi$$

For complex numbers z and w, it can be shown (see Problems 113–115) that

> **Properties of Complex Numbers**
>
> $$\overline{(\overline{z})} = z$$
>
> $$\overline{z+w} = \overline{z} + \overline{w}$$
>
> $$\overline{zw} = \overline{z}\,\overline{w}$$

Furthermore, if we multiply a complex number by its conjugate, we find that

$$\begin{aligned}z\overline{z} &= (a+bi)(a-bi) = a^2 - abi + abi - b^2i^2 \\ &= a^2 + b^2\end{aligned}$$

That is,

> If $z = a + bi$, then $z\overline{z} = a^2 + b^2$

EXAMPLE 13 Let $z = 3 + 2i$.

(a) Find $\overline{z}$. **(b)** Compute $z\overline{z}$.

Solution **(a)** $\overline{z} = 3 - 2i$.

(b) $z\overline{z} = (3+2i)(3-2i) = 9 - 4i^2 = 9 + 4 = 13$. •

We encounter complex numbers primarily when we solve quadratic equations. Recall that, to solve

$$ax^2 + bx + c = 0$$

for $a \neq 0$, we use the quadratic formula

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where $x_{1,2}$ refers to the two solutions x_1 (with the "+" sign) and x_2 (with the "−" sign). We will assume that a, b, and c are all real.

EXAMPLE 14 Solve $x^2 + 4x + 5 = 0$.

Solution

$$x_{1,2} = \frac{-4 \pm \sqrt{4^2 - (4)(1)(5)}}{(2)(1)} \quad \text{Use the quadratic formula.}$$

$$= \frac{-4 \pm \sqrt{16 - 20}}{2} = \frac{-4 \pm \sqrt{-4}}{2} \quad \text{Simplify.}$$

If we allowed solutions only in the real-number system, we would conclude that $x^2 + 4x + 5 = 0$ has no solutions. But if we allow solutions in the complex number system, we find that

$$x_{1,2} = \frac{-4 \pm \sqrt{4i^2}}{2} = \frac{-4 \pm 2i}{2} = \frac{2(-2 \pm i)}{2} = -2 \pm i$$

That is, $x_1 = -2 + i$ and $x_2 = -2 - i$. •

The term $b^2 - 4ac$ under the square root sign in the quadratic formula is called the **discriminant**. If the discriminant is nonnegative, the two solutions of the corresponding quadratic equation are real. (When the discriminant is equal to 0, the two solutions are identical.) If the discriminant is negative, the two solutions are complex conjugates of each other.

EXAMPLE 15 Without solving $2x^2 - 3x + 7 = 0$, what can you say about the solution?

Solution We compute the discriminant

$$b^2 - 4ac = (-3)^2 - (4)(2)(7) = 9 - 56 = -47 < 0 \quad \text{The discriminant is negative.}$$

Since the discriminant is negative, the equation $2x^2 - 3x + 7 = 0$ has two complex solutions, which are conjugates of each other. •

Section 1.2 Problems

1.2.1

1. Find the two numbers that have distance 4 from -1 by **(a)** measuring the distances on the real-number line and **(b)** solving an appropriate equation involving an absolute value.

2. Find all pairwise distances between the numbers -5, 2, and 7 by **(a)** measuring the distances on the real-number line and **(b)** computing the distances by using absolute values.

3. Solve the following equations:

(a) $|2x + 4| = 6$ **(b)** $|x - 3| = 2$

(c) $|2x - 3| = 5$ **(d)** $|1 - 5x| = 6$

4. Solve the following equations:

(a) $|2x + 4| = |5x - 2|$ **(b)** $|1 + 2u| = |5 - u|$

(c) $|4 + \frac{t}{2}| = |\frac{3}{2}t - 2|$ **(d)** $|2s - 6| = |3 - s|$

5. Solve the following inequalities:

(a) $|5x - 2| \le 4$ **(b)** $|3 - 4x| > 8$

(c) $|7x + 4| \ge 3$ **(d)** $|3 + 2x| < 7$

6. Solve the following inequalities:

(a) $|2x + 3| < 6$ **(b)** $|3 - 4x| \ge 2$

(c) $|x + 5| \le 1$ **(d)** $|7 - 2x| < 0$

1.2.2

In Problems 7–42, determine the equation of the line that satisfies the stated requirements. Put the equation in standard form.

7. The line passing through (3, 2) with slope -2

8. The line passing through $(2, -1)$ with slope $\frac{1}{4}$

9. The line passing through $(0, -2)$ with slope -3

10. The line passing through $(-3, 5)$ with slope 1/2

11. The line passing through $(-2, -3)$ and (1, 4)

12. The line passing through $(-1, 4)$ and $(1, -4)$

13. The line passing through (0, 3) and (2, 1)

14. The line passing through $(1, -1)$ and (4, 5)

15. The horizontal line through $(4, \frac{1}{4})$

16. The horizontal line through $(0, -1)$

17. The vertical line through $(-2, 0)$

18. The vertical line through $(2, -3)$

19. The line with slope 3 and y-intercept (0, 2)

20. The line with slope -1 and y-intercept (0, 5)

21. The line with slope 1/2 and y-intercept (0, 2)

22. The line with slope $-1/3$ and y-intercept $(0, 1/3)$

23. The line with slope -2 and x-intercept $(1, 0)$

24. The line with slope 1 and x-intercept $(-2, 0)$

25. The line with slope $-1/2$ and x-intercept $(-1/2, 0)$

26. The line with slope $1/5$ and x-intercept $(-1/2, 0)$

27. The line passing through $(2, -3)$ and parallel to $x+2y-4=0$

28. The line passing through $(1, 2)$ and parallel to $x-2y+4=0$

29. The line passing through $(-1, -1)$ and parallel to the line passing through $(0, 2)$ and $(3, 0)$

30. The line passing through $(2, -1)$ and parallel to the line passing through $(0, -4)$ and $(2, 1)$

31. The line passing through $(1, 4)$ and perpendicular to $2y-5x+7=0$

32. The line passing through $(1, -1)$ and perpendicular to $x-2y+3=0$

33. The line passing through $(5, -1)$ and perpendicular to the line passing through $(-2, 1)$ and $(1, -2)$

34. The line passing through $(4, -1)$ and perpendicular to the line passing through $(-2, 0)$ and $(1, 1)$

35. The line passing through $(1, 3)$ and parallel to the horizontal line passing through $(3, -1)$

36. The line passing through $(1, 5)$ and parallel to the horizontal line passing through $(2, 1)$

37. The line passing through $(-2, 3)$ and parallel to the vertical line passing through $(2, 1)$

38. The line passing through $(3, 1)$ and parallel to the vertical line passing through $(-1, -2)$

39. The line passing through $(1, -3)$ and perpendicular to the horizontal line passing through $(-1, -1)$

40. The line passing through $(1, 3)$ and perpendicular to the horizontal line passing through $(3, 2)$

41. The line passing through $(7, 3)$ and perpendicular to the vertical line passing through $(-1, -7)$

42. The line passing through $(-2, 5)$ and perpendicular to the vertical line passing through $(-1, 4)$

43. To convert a length measured in feet to a length measured in centimeters, we use the facts that a length measured in feet is proportional to a length measured in centimeters and that 1 ft corresponds to 30.5 cm. If x denotes the length measured in ft and y denotes the length measured in cm, then $y = 30.5\,x$

(a) Use the relationship to convert the following measurements into centimeters:

(i) 6 ft **(ii)** 4 in **(iii)** 1 ft, 7 in **(iv)** 20.5 in

(b) Rearrange the formula to show how the length x, measured in feet, can be obtained from the length y, measured in cm.

(c) Use the relationship to convert the following measurements into ft:

(i) 195 cm **(ii)** 12 cm **(iii)** 48 cm

44. (a) To convert the weight of an object from kilograms (kg) to pounds (lb), you use the facts that a weight measured in kilograms is proportional to a weight measured in pounds and that 1 kg corresponds to 2.20 lb. Find an equation that relates weight measured in kilograms to weight measured in pounds.

(b) Use your answer in (a) to convert the following measurements:

(i) 63 lb **(ii)** 5 lb **(iii)** 2.5 kg **(iv)** 76 kg

45. Assume that the distance a car travels is proportional to the time it takes to cover the distance. Find an equation that relates distance and time if it takes the car 15 min to travel 10 mi. What is the constant of proportionality if distance is measured in miles and time is measured in hours?

46. Assume that the number of seeds a plant produces is proportional to its aboveground biomass. Find an equation that relates number of seeds and aboveground biomass if a plant that weighs 213 g has 13 seeds.

47. Experimental study plots are often squares of length 1 m. If 1 ft corresponds to 0.305 m, compute the area of a square plot of length 1 m in ft^2.

48. Large areas are often measured in hectares (ha) or in acres. If 1 ha = 10,000 m^2 and 1 acre = 4046.86 m^2, how many acres is 1 hectare?

49. To convert the volume of a liquid measured in ounces to a volume measured in liters, we use the fact that 1 liter equals 33.81 ounces. Denote by x the volume measured in ounces and by y the volume measured in liters. Assume a linear relationship between these two units of measurements.

(a) Find the equation relating x and y.

(b) A typical soda can contains 12 ounces of liquid. How many liters is this?

50. To convert a distance measured in miles to a distance measured in kilometers, we use the fact that 1 mile equals 1.609 kilometers. Denote by x the distance measured in miles and by y the distance measured in kilometers. Assume a linear relationship between these two units of measurements.

(a) Find an equation relating x and y.

(b) The distance between Los Angeles and Las Vegas is 434 km. How many miles is this?

51. In the United States, measurements in recipes are usually given in cups. In Europe, measurements are usually given in grams. One cup of flour weighs 120 g.

(a) A U.S. recipe requires two and a half cups of flour. How many grams is this?

(b) A British recipe calls for 225 g of flour. How many cups is this?

(c) Write a formula to convert flour measurements in grams to measurements in cups.

52. (a) To measure temperature, three scales are commonly used: Fahrenheit, Celsius, and Kelvin. These scales are linearly related. The Celsius scale is devised so that 0°C is the freezing point of water and 100°C is the boiling point of water. If you are more familiar with the Fahrenheit scale, then you know that water freezes at 32°F and boils at 212°F. Find a linear equation that relates temperature measured in degrees Celsius and temperature measured in degrees Fahrenheit.

(b) The normal body temperature in humans ranges from 97.6°F to 99.6°F. Convert this temperature range into degrees Celsius.

(c) Is there any temperature that reads the same in Celsius and Fahrenheit?

53. **Measuring Brain Activity** fMRI is a method for inferring brain activity by measuring changes in blood flow to different parts of the brain (blood flow changes can be measured noninvasively using an MRI scanner). The technique works because of a correlation between blood flow rate (which can be measured using the scanner) and brain activity. ***Logothetis et al. (2001)*** showed in experiments on macaque monkeys that blood flow (y) is linearly related to brain activity (x). Both x and y are measured on scales from 0 to 1.

(a) Here are two data points from ***Logothetis et al. (2001)***:

x	y
0.16	0.52
1.0	1.0

Find a formula for y in terms of x.

(b) Find the blood flow rate (y) corresponding to each of the following brain activity measurements.

(i) $x = 0.5$ **(ii)** $x = 0.9$ **(iii)** $x = 0$

(c) It is most useful to have a formula for brain activity (x) in terms of blood flow (y), since blood flow can be measured. Derive this formula from your answer to part (a).

54. Use the following steps to show that if two nonvertical lines l_1 and l_2 with slopes m_1 and m_2, respectively, are perpendicular, then $m_1 m_2 = -1$: Assume that $m_1 < 0$ and $m_2 > 0$.

(a) Use a graph to show that if θ_1 and θ_2 are the respective angles of inclination of the lines l_1 and l_2, then $\theta_1 = \theta_2 + \frac{\pi}{2}$. (The angle of inclination of a line is the angle $\theta \in [0, \pi)$ between the line and the positively directed x-axis.)

(b) Use the fact that $\tan(\pi - x) = -\tan x$ to show that $m_1 = \tan\theta_1$ and $m_2 = \tan\theta_2$.

(c) Use the fact that $\tan(\frac{\pi}{2} - x) = \cot x$ and $\cot(-x) = -\cot x$ to show that $m_1 = -\cot\theta_2$.

(d) From the latter equation, deduce the truth of the claim set forth at the beginning of this problem.

1.2.3

55. Find the equation of a circle with center $(1, -2)$ and radius 2.

56. Find the equation of a circle with center $(2, 3)$ and radius 4.

57. **(a)** Find the equation of a circle with center $(2, 5)$ and radius 4.

(b) Where does the circle intersect the y-axis?

(c) Does the circle intersect the x-axis? Explain.

58. **(a)** Find all possible radii of a circle centered at $(2, -5)$ so that the circle intersects only one axis.

(b) Find all possible radii of a circle centered at $(2, -5)$ so that the circle intersects both axes.

59. Find the center and the radius of the circle given by the equation $(x+2)^2 + y^2 = 25$.

60. Find the center and the radius of the circle given by the equation $(x+1)^2 + (y-3)^2 = 9$.

61. Find the center and the radius of the circle given by the equation $0 = x^2 + y^2 + 6x + 2y - 12$. (To do this, you must complete the squares.)

62. Find the center and the radius of the circle given by the equation $x^2+y^2+2x-4y+1 = 0$. (To do this, you must complete the squares.)

1.2.4

63. **(a)** Convert $65°$ to radian measure.

(b) Convert $\frac{11\pi}{12}$ to degree measure.

64. **(a)** Convert $-15°$ to radian measure.

(b) Convert $\frac{7}{4}\pi$ to degree measure.

65. Evaluate the following expressions without using a calculator:

(a) $\sin(-\frac{\pi}{4})$ **(b)** $\cos(\frac{5\pi}{6})$ **(c)** $\tan(\frac{\pi}{6})$

66. Evaluate the following expressions without using a calculator:

(a) $\sin(\frac{5\pi}{4})$ **(b)** $\cos(-\frac{11\pi}{6})$ **(c)** $\tan(\frac{\pi}{3})$

67. **(a)** Find the values of $\alpha \in [0, 2\pi)$ that satisfy $\cos\alpha = -\frac{1}{2}\sqrt{3}$.

(b) Find the values of $\alpha \in [0, 2\pi)$ that satisfy $\tan\alpha = \frac{1}{\sqrt{3}}$.

68. **(a)** Find the values of $\alpha \in [0, 2\pi)$ that satisfy $\cos\alpha = -\frac{1}{2}\sqrt{2}$.

(b) Find the values of $\alpha \in [0, 2\pi)$ that satisfy $\sec\alpha = \sqrt{2}$.

69. Show that the identity $1 + \tan^2\theta = \sec^2\theta$ follows from $\sin^2\theta + \cos^2\theta = 1$.

70. Show that the identity $1 + \cot^2\theta = \csc^2\theta$ follows from $\sin^2\theta + \cos^2\theta = 1$.

71. Solve $\cos^2\theta - 2 = 2\sin\theta$ on $[0, 2\pi)$.

72. Solve $\sec^2 x = \sqrt{3}\tan x + 1$ on $[0, \pi)$.

1.2.5

73. Evaluate the following exponential expressions:

(a) $2^4 8^{-2/3}$ **(b)** $\dfrac{3^3 3^{-1/2}}{3^{1/2}}$ **(c)** $\dfrac{5^k (25)^{k-1}}{5^{2-k}}$

74. Evaluate the following exponential expressions:

(a) $(2^4 2^{-3/2})^2$ **(b)** $\left(\dfrac{6^{5/2}6^{2/3}}{6^{1/3}}\right)^3$ **(c)** $\left(\dfrac{3^{-2k+3}}{3^{4+k}}\right)^3$

75. Which real number x satisfies

(a) $\log_4 x = -2$? **(b)** $\log_{1/3} x = -3$? **(c)** $\log_{10} x = -2$?

76. Which real number x satisfies

(a) $\log_2 x = -3$? **(b)** $\log_{1/4} x = -\frac{1}{2}$? **(c)** $\log_3 x = 0$?

77. Which real number x satisfies

(a) $\log_{1/2} 32 = x$? **(b)** $\log_{1/3} 81 = x$? **(c)** $\log_{10} 0.001 = x$?

78. Which real number x satisfies

(a) $\log_3 81 = x$? **(b)** $\log_5 \frac{1}{25} = x$? **(c)** $\log_{10} 1000 = x$?

79. Simplify the following expressions:

(a) $-\ln\frac{1}{3}$ **(b)** $\log_4(x^2 - 4)$ **(c)** $\log_2 4^{3x-1}$

80. Simplify the following expressions:

(a) $-\log_3 \frac{1}{4}$ **(b)** $\log\left(\frac{x^3-x}{x-1}\right)$ **(c)** $\ln\left(e^{x-2}\right)$

81. Solve for x.

(a) $e^{3x-1} = 2$ **(b)** $e^{-2x} = 10$ **(c)** $e^{x^2-1} = 10$

82. Solve for x.

(a) $5^x = 625$ **(b)** $4^{4x} = 256$ **(c)** $10^{2x} = 0.0001$

83. Solve for x.

(a) $\ln(x-3) = 5$ **(b)** $\ln(x+2) + \ln(x-2) = 1$

(c) $\log_3 x^2 - \log_3 2x = 2$

84. Solve for x.
(a) $\log_3(2x - 1) = 2$ **(b)** $\ln(2 - 3x) = 0$
(c) $\log(x) - \log(x + 1) = \log(\frac{2}{3})$

1.2.6

In Problems 85–92, simplify each expression and write it in the standard form a + bi.

85. $(3 - 2i) - (-5 + 2i)$ **86.** $(6 + i) - 4i$
87. $(4 - 2i) + (9 + 4i)$ **88.** $(6 - 4i) + (2 + 5i)$
89. $4(5 + 3i)$ **90.** $(2 - 3i)(3 + 2i)$
91. $(6 - i)(6 + i)$ **92.** $(-4 - 3i)(4 + 3i)$

In Problems 93–98, let $z = 1 + 2i$, $u = 2 - 3i$, $v = 1 - 5i$, and $w = 1 + i$. Compute the following expressions:

93. $\overline{z}$ **94.** $z + u$ **95.** $\overline{z + v}$
96. $\overline{v - w}$ **97.** $\overline{vw}$ **98.** $\overline{uz}$

In Problems 99–104, solve each quadratic equation in the complex number system.

99. $2x^2 - 3x + 2 = 0$ **100.** $x^2 + x + 1 = 0$
101. $-x^2 + x + 2 = 0$ **102.** $x^2 + 2x + 3 = 0$
103. $x^2 + x + 6 = 0$ **104.** $-2x^2 + 4x - 3 = 0$

In Problems 105–110, first determine whether the solutions of each quadratic equation are real or complex without solving the equation. Then solve the equation.

105. $3x^2 - 4x - 7 = 0$ **106.** $x^2 - x - 1 = 0$
107. $3x^2 - x - 4 = 0$ **108.** $4x^2 - x + 1 = 0$
109. $3x^2 - 5x + 6 = 0$ **110.** $-3x^2 - x - 4 = 0$

111. If $z = a + bi$, find $z + \overline{z}$ and $z - \overline{z}$.

112. If $z = a + bi$, find $\overline{z}$. Use your answer to compute $\overline{(\overline{z})}$, and compare your answer with z.

113. Show $\overline{(\overline{z})} = z$.

114. Show $\overline{z + w} = \overline{z} + \overline{w}$.

115. Show $\overline{zw} = \overline{z}\,\overline{w}$.

1.3 Elementary Functions

1.3.1 What Is a Function?

Scientists often study relationships between quantities, such as how enzyme activity depends on temperature or how the length of a fish is related to its age. To describe such relationships mathematically, the concept of a *function* is useful.

> **Definition** A **function** f is a rule that assigns each element x in the set A exactly one element y in the set B. The element y is called the **image** (or **value**) of x under f and is denoted by $f(x)$ (read "f of x"). The set A is called the **domain** of f, the set B is called the **codomain** of f, and the set $f(A) = \{y : y = f(x)$ for some $x \in A\}$ is called the **range** of f.

To define a function, we use the notation

$$\begin{aligned} f : A &\to B \\ x &\mapsto f(x) \end{aligned}$$

where A and B are subsets of the set of real numbers. Frequently, we simply write $y = f(x)$ and call x the **independent** variable and y the **dependent** variable. We can illustrate functions graphically in the x–y plane. In Figure 1.20, we see the graph of $y = f(x)$, with domain A, codomain B, and range $f(A)$.

There are many ways that the function $f(x)$ can be specified. If the set A has only finitely many elements, we can make a list of the values of $f(x)$, with one value for each $x \in A$. If A contains infinitely many points, $f(x)$ could be given by a graph as in Figure 1.20, or it could be expressed algebraically, such as $f(x) = x^2$. The codomain, B, and range, $f(A)$, play slightly different roles. For each $x \in A$, the image $f(x)$ must lie in the codomain B. But not every point in the codomain must be the image of some $x \in A$. Only for points in the range, say $y \in f(A)$, does there exist some $x \in A$ with $f(x) = y$. And this x need not be unique; for each $y \in f(A)$ there may be several elements $x \in A$ for which $f(x) = y$. For instance, let

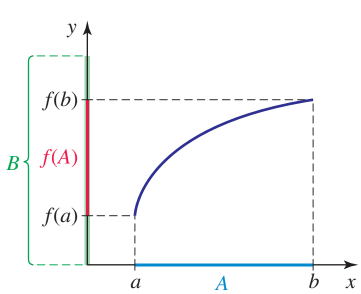


Figure 1.20 A function $f(x)$ with domain A, codomain B, and range $f(A)$.

$$\begin{aligned} f : \mathbf{R} &\to \mathbf{R} \\ x &\mapsto x^2 \end{aligned}$$

The codomain of f is $\mathbf{R}$, but the range of f is only $[0, \infty)$ because the square of a real number is nonnegative; that is, $f(\mathbf{R}) = [0, \infty) \neq \mathbf{R}$. Given any $y \in f(\mathbf{R}) = [0, \infty)$, we can find some $x \in \mathbf{R}$ with $f(x) = y$; take $x = \sqrt{y}$. If $y \neq 0$, then there are two points in $\mathbf{R}$ whose image is y, namely, $+\sqrt{y}$ and $-\sqrt{y}$. The domain of a function need not be the largest possible set on which we can define the function, as $\mathbf{R}$ is in the preceding example. For instance, we could have defined f on a smaller set, such as $[0, 1]$, calling the new function g, given by

$$\begin{aligned} g: [0,1] &\to \mathbf{R} \\ x &\mapsto x^2 \end{aligned}$$

Although the same rule is used for f and g, the two functions are not the same, because their respective domains are different.

> Two functions f and g are **equal** if and only if
>
> **1.** f and g are defined on the same domain, and
>
> **2.** $f(x) = g(x)$ for all x in the domain.

EXAMPLE 1 Let

$$\begin{aligned} f_1 : [0,1] &\to \mathbf{R}, \\ x &\mapsto x^2 \end{aligned} \qquad \begin{aligned} f_2 : [0,1] &\to \mathbf{R} \\ x &\mapsto \sqrt{x^4} \end{aligned} \qquad \text{and} \qquad \begin{aligned} f_3 : \mathbf{R} &\to \mathbf{R} \\ x &\mapsto x^2 \end{aligned}$$

Determine which of these functions are equal.

Solution Because f_1 and f_2 are defined on the same domain and $f_1(x) = f_2(x) = x^2$ for all $x \in [0, 1]$, it follows that f_1 and f_2 are equal.

Neither f_1 nor f_2 is equal to f_3, because the domain of f_3 is different from the domains of f_1 and f_2. •

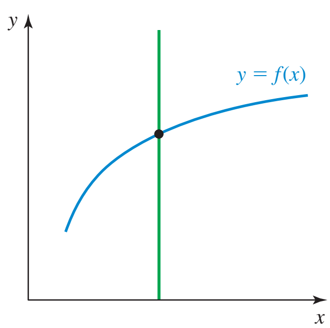


Figure 1.21 The vertical line test shows that the graph of $y = f(x)$ is a function.

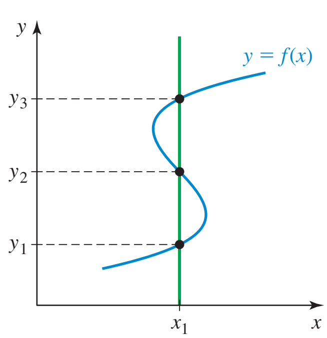


Figure 1.22 The vertical line test shows that the graph of $y = f(x)$ is not a function. The point x_1 is assigned three different images y_1, y_2, y_3.

The choices of domains for the functions that we have thus far considered may look somewhat arbitrary (and they are arbitrary in the examples we have seen so far). In applications, however, there is often a natural choice of domain. For instance, if we look at a certain plant response (such as total biomass or the ratio of above to below biomass) as a function of nitrogen concentration in the soil, then, given that nitrogen concentration cannot be negative, the domain for this function could be the set of nonnegative real numbers. As another example, suppose we define a function that depends on the fraction of a population infected with a certain virus; then a natural choice for the domain of this function would be the interval $[0, 1]$ because a fraction of a population must be a number between 0 and 1.

In our definition of a function, we stated that a function is a rule that assigns, to each element $x \in A$, *exactly* one element $y \in B$. When we graph $y = f(x)$ in the x–y plane, there is a simple test to decide whether or not $f(x)$ is a function:

> **Vertical Line Test** If each vertical line intersects the graph of $y = f(x)$ at most once, then $f(x)$ is a function.

Figure 1.21 shows the graph of a function: Each vertical line intersects the graph of $y = f(x)$ at most once. The graph of $y = f(x)$ in Figure 1.22 is not a function, since there are x-values that are assigned to more than one y-value, as illustrated by the vertical line that intersects the graph more than once.

Sometimes functions show certain symmetries. For example, in Figure 1.23, $f(x) = x$ is symmetric about the origin; that is, $f(x) = -f(-x)$. In Figure 1.24, $g(x) = x^2$ is symmetric about the y-axis; that is, $g(x) = g(-x)$. In the first case, we say that f is *odd*;

in the second case, that g is *even*. To check whether a function is even or odd, we use the following definition:

> **Symmetries of Functions** A function $f : A \to B$ is called
>
> **1. even** if $f(x) = f(-x)$ for all $x \in A$, and
>
> **2. odd** if $f(x) = -f(-x)$ for all $x \in A$.

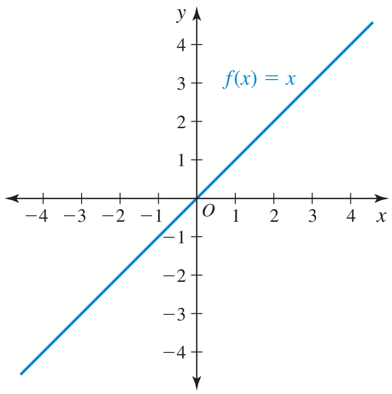


Figure 1.23 The graph of $y = x$ is symmetric about the origin.

Using this criterion, we can show that $f(x) = x$, $x \in \mathbf{R}$, is an odd function:

$$-f(-x) = -(-x) = x = f(x) \qquad \text{for all } x \in \mathbf{R}$$

Likewise, to show that $g(x) = x^2$, $x \in \mathbf{R}$, is an even function, we compute

$$g(-x) = (-x)^2 = x^2 = g(x) \qquad \text{for all } x \in \mathbf{R}$$

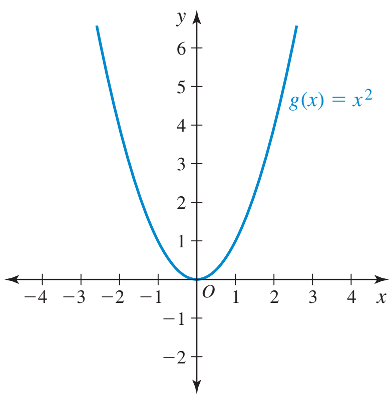


Figure 1.24 The graph of $y = x^2$ is symmetric about the y-axis.

We will now look at the case where one quantity is given as a function of another quantity that, in turn, can be written as a function of yet another quantity. To illustrate this situation, suppose we are interested in the abundance of a predator, which depends on the abundance of a herbivore, which, in turn, depends on the abundance of plant biomass. If we denote the plant biomass by x and the herbivore biomass by u, then x and u are related via a function g, namely, $u = g(x)$. Likewise, if we denote the predator biomass by y, then u and y are related via a function f, namely, $y = f(u)$. We can express the predator biomass as a function of the plant biomass by substituting $g(x)$ for u. That is, we find $y = f[g(x)]$. Functions that are defined in such a way are called composite functions.

> **Definition** If $g : A \to B$ and $f : C \to D$ are functions, the **composite function** $f \circ g$ (also called the **composition** of f and g) is a function defined by:
>
> $$(f \circ g)(x) = f[g(x)]$$
>
> for each $x \in A$ with $g(x) \in C$.

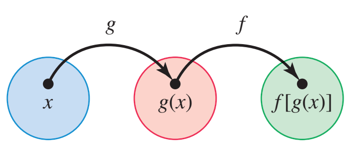


Figure 1.25 The composition of functions.

The composition of functions is illustrated in Figure 1.25. We call g the inner function and f the outer function. When defining any function we give its domain. To calculate $f(u)$, u must lie in the domain of f. So to calculate $f(g(x))$, $g(x)$ must be in the domain of f. This is not necessarily true for all $x \in A$, as Example 3, below, makes clear. Hence we must take for the domain of $f \circ g$ only the points $x \in A$ for which $g(x)$ is in the domain of f.

EXAMPLE 2 If $f(x) = \sqrt{x}$, $x \geq 0$, and $g(x) = x^2 + 1$, $x \in \mathbf{R}$, find

(a) $(f \circ g)(x)$ and **(b)** $(g \circ f)(x)$.

Solution To define the function we need both an algebraic rule for calculating its values, and the domain and codomain.

(a) To find $(f \circ g)(x)$, we set $f(u) = \sqrt{u}$ and $g(x) = x^2 + 1$. Then

$$y = f(u) = f[g(x)] = f(x^2 + 1) = \sqrt{x^2 + 1}$$

To determine the domain of $f \circ g$, we observe that the domain of the inner function g is $\mathbf{R}$ and its range is $[1, \infty)$. Since the range of g is contained in the domain of the outer function f (because $[1, \infty) \subset [0, \infty)$), the domain of $f \circ g$ is $\mathbf{R}$.

(b) To find $(g \circ f)(x)$, we set $g(u) = u^2 + 1$ and $f(x) = \sqrt{x}$. Then

$$y = g(u) = g[f(x)] = g(\sqrt{x}) = (\sqrt{x})^2 + 1 = x + 1$$

To determine the domain of $g \circ f$, we observe that the domain of the inner function f is $[0, \infty)$ and its range is $[0, \infty)$. The range of f is contained in the domain of the outer function g (because $[0, \infty) \subset \mathbf{R}$), so the domain of $g \circ f$ is $[0, \infty)$.

In the last example, you should observe that $f \circ g$ is different from $g \circ f$, which implies that the order in which you compose functions is important. The notation $f \circ g$ means that you apply g first and then f. In addition, you should pay attention to the domains of composite functions. In the next example, the domain is harder to find.

EXAMPLE 3 If $f(x) = \sqrt{x-2}, x \geq 2$, and $g(x) = \sqrt{x}, x \geq 0$, find $(f \circ g)(x)$ together with its domain.

Solution We compute

$$(f \circ g)(x) = f[g(x)] = f(\sqrt{x}) = \sqrt{\sqrt{x} - 2}$$

This part was not difficult. However, finding the domain of $f \circ g$ is more complicated. The domain of the inner function g is the interval $[0, \infty)$; and, the range of g is also the interval $[0, \infty)$. The domain of f is only $[2, \infty)$, which means that the range of g is *not* contained in the domain of f. We therefore need to restrict the domain of g to ensure that its range is contained in the domain of f. We can choose only values of x such that $g(x) \in [2, \infty)$. Since $g(x) = \sqrt{x}$, we need to restrict x to $[4, \infty)$. Thus, for every $x \in [4, \infty)$, $g(x) \in [2, \infty)$, which is the domain of f. Therefore,

$$(f \circ g)(x) = \sqrt{\sqrt{x} - 2}, \quad x \geq 4$$

See Figure 1.26. •

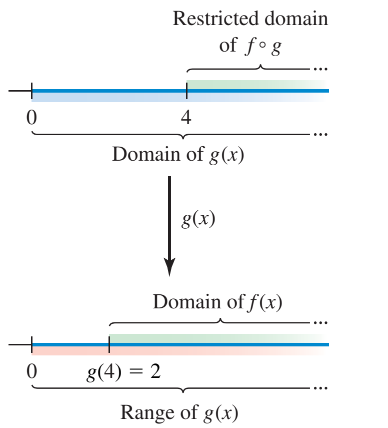


Figure 1.26 Finding the domain of a composite function: The domain of $g(x)$ must be restricted in Example 3.

In the subsections that follow, we introduce the basic functions that are used throughout the remainder of this book.

1.3.2 Polynomial Functions

Polynomial functions are the simplest elementary functions.

> **Definition** A **polynomial** function is a function of the form
>
> $$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$
>
> where n is a nonnegative integer and $a_0, a_1, \ldots, a_n$ are (real-valued) constants with $a_n \neq 0$. The coefficient a_n is called the **leading coefficient**, and n is called the **degree** of the polynomial function. The largest possible domain of f is $\mathbf{R}$.

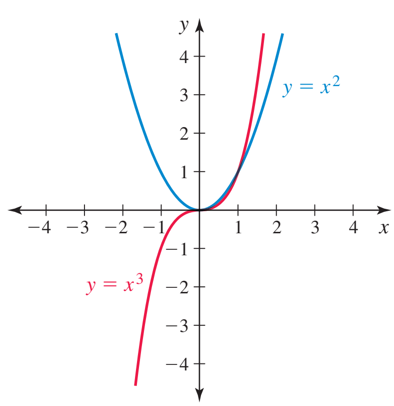


Figure 1.27 The graphs of $y = x^n$ for $n = 2$ and $n = 3$.

We have already encountered polynomials, namely, the constant function $f(x) = c$, the linear function $f(x) = mx + b$, and the quadratic function $f(x) = ax^2$. The constant, nonzero function has degree 0, the linear function has degree 1, and the quadratic function has degree 2. Other examples are $f(x) = 4x^3 - 3x + 1$, $x \in \mathbf{R}$, which is a polynomial of degree 3, and $f(x) = 2 - x^7$, $x \in \mathbf{R}$, which is a polynomial of degree 7. In Figure 1.27, we display $y = x^n$ for $n = 2$ and 3. Looking at the figure, we see that $y = x^n$ is an even function when $n = 2$ and an odd function when $n = 3$. This property holds in general: $y = x^n$ is an even function when n is even and an odd function when n is odd. We can show this algebraically by using the criterion in Section 1.3.1. (See Problem 28 at the end of this section.)

Polynomials arise naturally in many situations. We present two examples.

EXAMPLE 4 Bio Info • Diversity Index One way to measure the diversity of a population of organisms is to calculate its Gini-Simpson diversity index, H. In its simplest incarnation consider a population of yeast cells; each yeast cell is one of two types (call them "red" and "green"). The diversity index of the population is the probability that if two cells are picked at random, they are different colors (i.e., one is red and the other is green).

If p is the proportion of cells of red-type, we will show in Chapter 12 that the diversity index H can be calculated from p using the formula:

$$H(p) = 2p(1 - p).$$

$H(p)$ is a polynomial in p. What is its degree? What is the domain and range of the function $H(p)$?

Solution Multiplying out we see that $H(p) = 2p - 2p^2$, so the degree of H is 2. Since p represents the proportion of red cells, it must be between 0 and 1, so the domain of H is $[0, 1]$. To calculate the range we can use a graphical calculator to plot $H(p)$ for $p \in [0, 1]$ (see Figure 1.28).

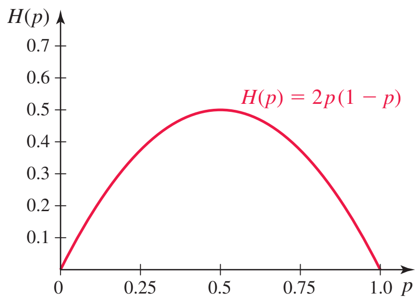


Figure 1.28 Gini-Simpson diversity index has domain [0, 1] and range [0, 1/2].

We see from the plot that $H(p)$ takes the smallest value, 0, when $p = 0$ or 1. (When $p = 0$, all cells are green-type, so there is zero chance that two randomly chosen cells will be of different types.)

The maximum value of p occurs when $p = \frac{1}{2}$ (i.e., $0 \leq H(p) \leq H\left(\frac{1}{2}\right) = \frac{1}{2}$). So the range of H is $[0, \frac{1}{2}]$. In Chapter 5 we will learn how to find the maximum values of functions like $H(p)$ without needing to plot them. •

EXAMPLE 5 **A Chemical Reaction** Consider the reaction rate of the chemical reaction

$$\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{AB}$$

in which the molecular reactants A and B form the molecular product AB. The rate at which this reaction proceeds depends on how often A and B molecules collide. The **law of mass action** states that the rate at which this reaction proceeds is proportional to the product of the respective concentrations of the reactants. Here, *concentration* means the number of molecules per fixed volume. If we denote the reaction rate by R and the concentration of A and B by [A] and [B], respectively, then the law of mass action says that

$$R \propto [\mathrm{A}] \cdot [\mathrm{B}]$$

Introducing the proportionality factor k, we obtain

$$R = k[\mathrm{A}] \cdot [\mathrm{B}]$$

Note that $k > 0$, because [A], [B], and R are positive. We assume now that the reaction occurs in a closed vessel; that is, we add specific amounts of A and B to the vessel at the beginning of the reaction and then let the reaction proceed without further additions. Assume there is no product, AB, present initially, and the initial concentrations of A and B are 2 and 5 respectively. When the concentration of AB reaches x, calculate the rate at which the product is being produced, $R(x)$. Show that $R(x)$ is a polynomial in x. Find the degree and the domain and range of the function $R(x)$.

Solution First note that we are not given the rate constant k. So we should expect k to show up in our answer as an unknown constant. Now $R(x) = k[\mathrm{A}][\mathrm{B}]$, so we need [A] and [B]. As the reaction proceeds [A] and [B] are used up. It takes one unit of A (and one of B) to make a unit of AB. Thus if there are x units of AB, x units of A and x of B must

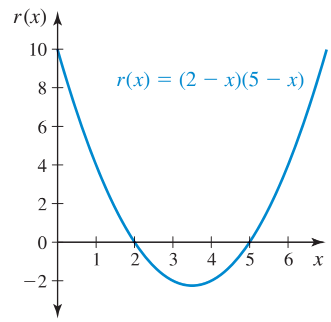


Figure 1.29 Reaction rate $r(x)$ as a function of the amount, x, of AB that has been produced.

have been used up. Thus [A] $= 2 - x$ (starting concentration minus the concentrate used up in the reaction) and [B] $= 5 - x$, so

$$\begin{aligned} R(x) &= k[\mathrm{A}][\mathrm{B}] \\ &= k(2-x)(5-x) \\ &= k(10 - 7x + x^2). \qquad \text{Multiply out.} \end{aligned}$$

So $R(x)$ is a polynomial of degree 2 in x. Although we don't know k, it appears only as an overall constant multiplying every term of the polynomial. So if we can calculate $r(x) = (2-x)(5-x)$, we can multiply it by k to get $R(x)$. $r(x)$ has no unknown constants, so it can be plotted using a graphical calculator (see Figure 1.29).

We can see from the plot that $r(x) = 0$ (so $R(x) = 0$) if $x = 2$ or $x = 5$. But what is the true domain of the function? Initially we are given that $x = 0$ (no AB is present). As the reaction proceeds, x increases, so $R(x)$ decreases. When x reaches 2, $R(x) = 0$; this is because [A] $= 0$ (i.e., all A has been used up). At this point the reaction will stop (indeed if $x \geq 2$, then [A] < 0, which is not possible). So our domain is $[0, 2]$. From the plot we can see that over this domain $r(x)$ ranges from 10 to 0. So $0 \leq R(x) \leq 10k$, that is, the range is $R([0, 2]) = [0, 10k]$. •

1.3.3 Rational Functions

Rational functions are built from polynomial functions.

> **Definition** A **rational** function is the quotient of two polynomial functions $p(x)$ and $q(x)$:
>
> $$f(x) = \frac{p(x)}{q(x)} \quad \text{for } q(x) \neq 0$$

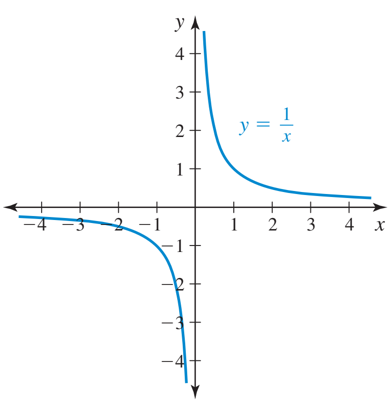


Figure 1.30 The graph of $y = \frac{1}{x}$ for $x \neq 0$.

Since division by 0 is not allowed, we must exclude those values of x for which $q(x) = 0$. Here are a couple of examples of rational functions, together with their largest possible domains:

$$y = \frac{1}{x}, \quad x \neq 0$$

$$y = \frac{x^2 + 2x - 1}{x - 3}, \quad x \neq 3$$

An important example of a rational function is the hyperbola, together with its largest possible domain:

$$y = \frac{1}{x}, \quad x \neq 0$$

The graph of y is shown in Figure 1.30.

Bio Info • Rational functions are often useful for building mathematical descriptions of population growth rates or of the rates of certain reactions. Throughout this text, we will encounter populations whose sizes change with time. The change in population size is described by the **growth rate**. Roughly speaking, the growth rate tells you how much a population changes during a small time interval. (The growth rate is analogous to the velocity of a car. Velocity is also a rate; it tells you how much the position changes in a small time interval. We will give a precise definition of rates in Section 4.1.) The **per capita** growth rate is the growth rate divided by the population size. The per capita growth rate is also called the **specific** growth rate. The next example introduces a function that is frequently used to describe growth rates.

EXAMPLE 6 **Monod Growth Function** There is a function that is frequently used to describe the per capita growth rate of a population of organisms when the rate depends on the concentration of some nutrient and becomes saturated for large enough nutrient concentrations. If we denote the concentration of the nutrient by N, then the per capita growth rate $r(N)$ is given by the *Monod growth function*

$$r(N) = \frac{aN}{k+N}, \quad N \geq 0$$

where a and k are positive constants. The domain of this function is $N \geq 0$. Assuming $k = 1$, find its range and describe in words the effect of changing N upon $r(N)$.

Solution $r(N)$ contains the unknown constant, a, and again we expect this feature in our answer. However, just like k in Example 5, a shows up as an overall constant of multiplication, so we can get the properties of $r(N)$ from the function $s(N) = \frac{N}{N+1}$. (We can get $r(N)$ by evaluating $s(N)$ and multiplying by a.) $s(N)$ has no unknown constants, so it can be plotted using a graphical calculator (see Figure 1.31).

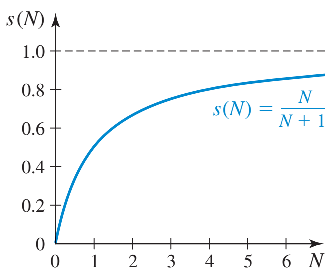


Figure 1.31 Per capita growth rate ($s(N)$) with the unknown constant a removed, as a function of population size (N).

The rate of growth increases with N, but at a slower and slower rate for larger values of N. Doubling N has a much bigger effect when N is small than when N is already large. Ultimately $s(N)$ tends to, but does not reach a maximum value of 1, meaning that $r(N)$ tends to a maximum value a. The range of the function $r(N)$ is therefore $[0, a)$. •

1.3.4 Power Functions

Definition A **power** function is of the form

$$f(x) = x^r$$

where r is a real number.

Examples of power functions, with their largest possible domains, are

$$y = x^{1/3}, \quad x \in \mathbf{R}$$

$$y = x^{5/2}, \quad x \geq 0$$

$$y = x^{1/2}, \quad x \geq 0$$

$$y = x^{-1/2}, \quad x > 0$$

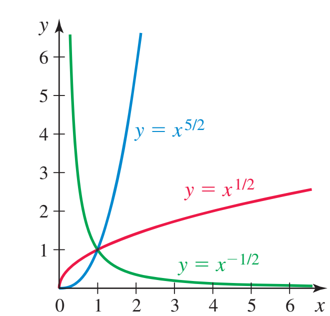


Figure 1.32 Some power functions with rational exponents.

Polynomials of the form $y = x^n, n = 1, 2, \ldots$, are a special case of power functions. Since power functions may involve even roots, as in $y = x^{3/2} = (\sqrt{x})^3$, we frequently need to restrict their domain.

Figure 1.32 compares the power functions $y = x^{5/2}$, $y = x^{1/2}$, and $y = x^{-1/2}$ for $x > 0$. Pay close attention to how the exponent determines the ranking according to size for x between 0 and 1 and for $x > 1$. We find that $x^{5/2} < x^{1/2} < x^{-1/2}$ for $0 < x < 1$, but $x^{5/2} > x^{1/2} > x^{-1/2}$ for $x > 1$.

EXAMPLE 7 **Scaling and Allometry** Power functions are frequently found in scaling relations between biological variables (e.g., organ sizes). These are relations of the form

$$y \propto x^r$$

where r is a nonzero real number. That is, y is proportional to some power of x. Recall that we can write this relationship as an equation if we introduce the proportionality factor k:

$$y = kx^r$$

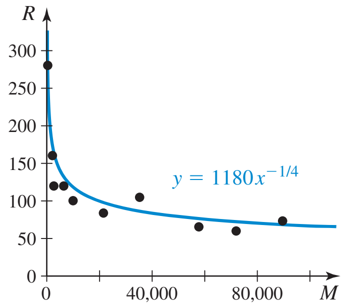


Figure 1.33 Resting heart rate, R, as a function of body mass, M, for mammals. Data adapted from ***Savage et al. 2004.***

For example, the resting heart rates, R, of mammals at rest (in beats/min) were shown by ***Savage et al. (2004)*** to decrease with the mass of the mammal, M, in grams according to the following relationship:

$$R = 1180M^{-1/4}.$$

This relationship, along with data from real animals, is shown in Figure 1.33.

This relationship is to be interpreted in a statistical sense; it is obtained by fitting a curve to data points. Heart rates actually vary between animals of the same species depending on the condition of the animal, and the data points for different species are typically scattered around the fitted curve. In Chapter 12 we will learn methods for fitting models to data. •

The next example relates the volume and the surface area of a cube. This relationship is not to be understood in a statistical sense, because it is an exact relationship resulting from geometric considerations.

EXAMPLE 8 Suppose that we wish to know the relationship between the surface area S and the volume V of a sphere. The scaling relations of each of these quantities with the diameter L of the sphere are as follows:

$$S \propto L^2 \quad \text{or} \quad S = k_1 L^2$$
$$V \propto L^3 \quad \text{or} \quad V = k_2 L^3$$

Here, k_1 and k_2 denote the constants of proportionality. (We label them with different subscripts to indicate that they might be different.) To express S in terms of V, we must first solve L in terms of V and then substitute L in the equation for S. Because $L = (V/k_2)^{1/3}$, it follows that

$$S = k_1 \left[\left(\frac{V}{k_2} \right)^{1/3} \right]^2 = \frac{k_1}{k_2^{2/3}} V^{2/3}$$

Introducing the constant of proportionality $k = k_1/k_2^{2/3}$, we find that

$$S = kV^{2/3}, \qquad \text{or simply} \qquad S \propto V^{2/3}$$

In words, the surface area of a sphere scales with the volume in proportion to $V^{2/3}$. We can now ask, for instance, by what factor the surface area increases when we double the volume. When we double the volume, we find that the resulting surface area, denoted by S', is

$$S' = k(2V)^{2/3} = 2^{2/3} \underbrace{kV^{2/3}}_{S}$$

That is, the surface area increases by a factor of $2^{2/3} \approx 1.587$ if we double the volume of the sphere. This scaling has implications on heat retention in animals: A larger body has a relatively smaller surface area and will retain more heat. •

1.3.5 Exponential Functions

In our study of exponential functions, let's first look at an example that illustrates where they occur.

EXAMPLE 9 Bio Info · Exponential Growth Bacteria reproduce asexually by cellular fission, in which the parent cell splits into two daughter cells after duplication of the genetic material. This division may happen as often as every 20 minutes; under ideal conditions, every cell in a population of bacteria will split in that time, doubling the size of the population.

Let us measure time such that one unit of time corresponds to the doubling time of the bacteria. If we denote the size of the population at time t by $N(t)$, then the function

$$N(t) = 2^t, \quad t \geq 0$$

has the property of doubling its value every time t is increased by 1, since:

$$N(t+1) = 2^{t+1} = 2 \cdot 2^t = 2N(t) \tag{1.4}$$

The function $N(t) = 2^t$, $t \geq 0$, is an exponential function because the variable t is in the exponent. We call the number 2 the base of the exponential function.

We find that when $t = 0$, $N(0) = 1$; that is, there is just one individual in the population at time $t = 0$. If, at time $t = 0$, 40 individuals were present in the population, we would write $N(0) = 40$ and

$$N(t) = 40 \cdot 2^t, \quad t \geq 0$$

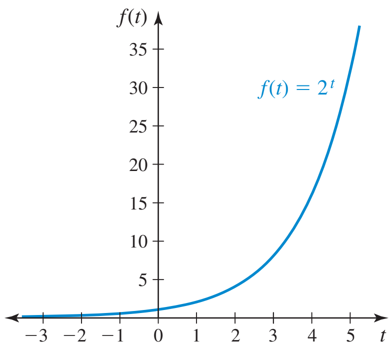


Figure 1.34 The function $f(t) = 2^t$, $t \in \mathbf{R}$.

You can verify that this formula for $N(t)$ also satisfies $N(t+1) = 2N(t)$.

It is often desirable not to specify the initial number of individuals in the equation describing $N(t)$. This approach has the advantage that the equation for $N(t)$ then describes a more general situation, in the sense that we can use the same equation for different initial population sizes. We often denote the population size at time 0 by N_0 (read "N sub 0") instead of $N(0)$. The equation for $N(t)$ is then

$$N(t) = N_0 2^t, \quad t \geq 0$$

We can verify that $N(0) = N_0 2^0 = N_0$ and that $N(t+1) = N_0 2^{t+1} = 2(N_0 2^t) = 2N(t)$.

The function $f(t) = 2^t$ can be defined for all $t \in \mathbf{R}$; its graph is shown in Figure 1.34.

Here is the definition of an exponential function:

Definition The function f is an **exponential** function with base a if

$$f(x) \propto a^x$$

$\propto$ means "proportional to," i.e., $f(x) = ka^x$. for some constant k, see Section 1.2.2

where a is a positive constant other than 1. The largest possible domain of f is $\mathbf{R}$.

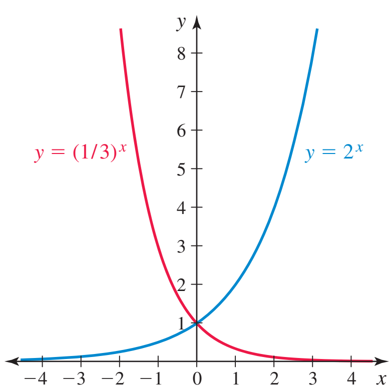


Figure 1.35 Exponential growth and exponential decay, for two functions $f(x) = a^x$ with $a > 1$ and $0 < a < 1$.

When $a = 1$, $f(x) = 1$ is simply the constant function, which we don't count as an exponential function.

The basic shape of the exponential function $f(x) = a^x$ depends on the base a; two examples are shown in Figure 1.35. As x increases, the graph of $f(x) = 2^x$ shows a rapid increase, whereas the graph of $f(x) = (1/3)^x$ shows a rapid decrease toward 0. We find the rapid increase whenever $a > 1$ and the rapid decrease whenever $0 < a < 1$. Therefore, we say that we have exponential *growth* when $a > 1$ and exponential *decay* when $0 < a < 1$.

Recall that $a^0 = 1$ and $a^{1/k} = \sqrt[k]{a}$, where k is a positive integer. In Subsection 1.2.5, we summarized the properties of exponentials. The most important base in calculus is base-e, where $e = 2.718\ldots$, which we encountered in Subsection 1.2.5. The number e is called the **natural exponential base**. The exponential function with base e is alternatively written as $\exp(x)$. That is,

$$\exp(x) = e^x$$

The advantage of this alternative form can be seen when we try to write something like $e^{x^2/\sqrt{x^3+1}}$: $\exp(x^2/\sqrt{x^3+1})$ is easier to read. More generally, if $g(x)$ is a function in x, then we can write, equivalently, either

$$\exp[g(x)] \qquad \text{or} \qquad e^{g(x)}$$

The next two examples provide an important application of exponential functions.

EXAMPLE 10

Bio Info • Radioactive Decay Radioactive isotopes such as carbon 14 are used to determine the absolute age of fossils or minerals, establishing an absolute chronology of the geological time scale. This technique was discovered in the early years of the 20th century and is based on the property of certain atoms to transform spontaneously by giving off protons, neutrons, or electrons. The phenomenon, called *radioactive decay*, occurs at a constant rate that is independent of environmental conditions. The method was used, for instance, to trace the successive emergence of the Hawaiian islands, from the oldest, Kauai, to the youngest, Hawaii (which is about 100,000 years old).

Carbon 14 is formed high in the atmosphere. It is radioactive and decays into nitrogen (N^{14}). There is an equilibrium between atmospheric carbon 12 (C^{12}) and carbon 14 (C^{14})—a ratio that has been relatively constant over a fairly long period. When plants capture carbon dioxide (CO_2) molecules from the atmosphere and build them into a product (such as cellulose), the initial ratio of C^{14} to C^{12} is the same as that in the atmosphere. Once the plants die, however, their uptake of CO_2 ceases, and the radioactive decay of C^{14} causes the ratio of C^{14} to C^{12} to decline. Because the law of radioactive decay is known, the change in ratio provides an accurate measure of the time since the plants' death.

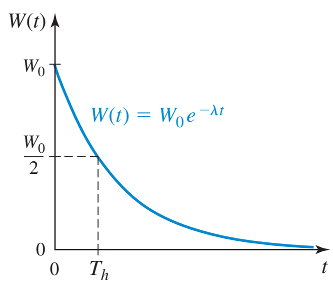


Figure 1.36 The function $W(t) = W_0 e^{-\lambda t}$.

According to the radioactive decay law, if the amount of C^{14} at time t is denoted by $W(t)$, with $W(0) = W_0$, then

$$W(t) = W_0 e^{-\lambda t}, \quad t \geq 0$$

where $\lambda > 0$ (λ is the lowercase Greek letter lambda) denotes the **decay rate**. The function $W(t) = W_0 e^{-\lambda t}$ is another example of an exponential function. Its graph is shown in Figure 1.36. $W(t)$ is an exponential function according to the definition because $W(t) \propto e^{-\lambda t} = (e^{-\lambda})^t$ i.e., $a = e^{-\lambda}$.

Frequently, the decay rate is expressed in terms of the **half-life** of the material, which is the length of time that it takes for half of the isotope to decay. If we denote this time by T_h, then (see Figure 1.36)

$$W(T_h) = \frac{1}{2} W_0 = W_0 e^{-\lambda T_h}$$

from which we obtain

$$\frac{1}{2} = e^{-\lambda T_h}$$

$$2 = e^{\lambda T_h} \qquad \text{Divide 1 by both sides}$$

Recall from algebra (or Subsection 1.2.5) that, to solve for the exponent λT_h, we must take logarithms on both sides. Since the exponent has base e, we use natural logarithms and find that

$$\ln 2 = \lambda T_h$$

Solving for T_h or λ yields

Half-Life Formulas

$$T_h = \frac{\ln 2}{\lambda} \qquad \text{or} \qquad \lambda = \frac{\ln 2}{T_h}$$

It is known that the half-life of C^{14} is 5730 years. Hence,

$$\lambda = \frac{\ln 2}{5730 \text{ years}}$$

Note that the unit "years" appears in the denominator. It is important to carry the units along. When we compute λt in the exponent of $e^{-\lambda t}$, we need to measure t in units of years in order for the units to cancel properly. For example, suppose $t = 2000$ years; then

$$\lambda t = \frac{\ln 2}{5730 \text{ years}} 2000 \text{ years} = \frac{(\ln 2)(2000)}{5730} \approx 0.2419$$

and we see that "years" appears in both the numerator and the denominator and thus can be canceled.

An application of the C^{14} dating method is given in the next example.

EXAMPLE 11 Suppose that, on the basis of their C^{12} content, samples of wood found in an archeological excavation site contain about 23% as much C^{14} as does living plant material. Determine when the wood was cut.

Solution The ratio of the current amount of C^{14} to the amount in living plant material is expressed as

$$0.23 = \frac{W(t)}{W(0)} = e^{-\lambda t}$$

Taking logarithms (base e) on both sides, we obtain

$$\ln(0.23) = -\lambda t$$

With $\lambda = \ln 2/(5730 \text{ years})$ from Example 10,

$$t = \frac{-1}{\lambda} \ln 0.23 = \frac{-5730 \text{ years}}{\ln 2} \ln 0.23 = 12{,}150 \text{ years}$$

Use a calculator to evaluate the natural logarithms.

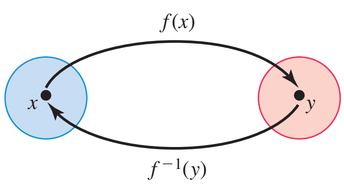


Figure 1.37 The function $y = f(x)$ and its inverse.

1.3.6 Inverse Functions

Before we can introduce logarithmic functions, we must understand the concept of inverse functions. Roughly speaking, the inverse of a function f reverses the effect of f. That is, if $y = f(x)$, that is, f maps x into y, then the inverse function, denoted by f^{-1} (read "f inverse"), takes y and maps it back into x. (See Figure 1.37.) Not every function has an inverse: Because an inverse function is a function itself, we require that every value y in the range of f be mapped into exactly one value x. In other words, for a function to have an inverse, it must be that whenever $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$ or, equivalently, that if $f(x_1) = f(x_2)$ then $x_1 = x_2$. (For f to be a function, each x in the domain must be mapped to exactly one point in the codomain. For an inverse to exist, each point in the codomain must be the image of at most one point in the domain.)

Functions that have the property "$x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$" [or, equivalently, "$f(x_1) = f(x_2)$ implies $x_1 = x_2$"] are called **one to one**.

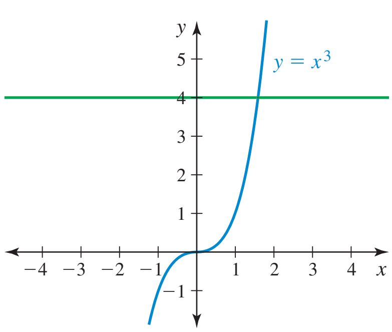


Figure 1.38 Horizontal line test successful.

> **Horizontal Line Test** If you know what the graph of a particular function looks like over its domain, then it is easy to determine whether or not the function is one to one: If no horizontal line intersects the graph of the function f more than once, then f is one to one.

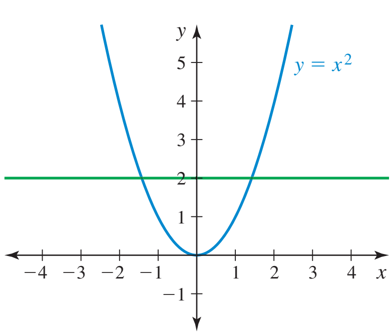


Figure 1.39 Horizontal line test unsuccessful.

We illustrate this test in Figures 1.38 and 1.39.

Now consider $y = x^3$ and $y = x^2$, for $x \in \mathbf{R}$. The function $y = x^3$, $x \in \mathbf{R}$, has an inverse function, because $x_1^3 \neq x_2^3$ whenever $x_1 \neq x_2$. (See Figure 1.38.) The function $y = x^2$, $x \in \mathbf{R}$, does *not* have an inverse function, because $x_1 \neq x_2$ does not imply $x_1^2 \neq x_2^2$ (or, equivalently, $x_1^2 = x_2^2$ does not imply $x_1 = x_2$; see Figure 1.39). The equation $x_1^2 = x_2^2$ implies that $x_1 = \pm x_2$. For instance, both -2 and 2 are mapped into 4, and we find that $f(-2) = f(2)$ but $-2 \neq 2$. To invert this function, we would have to map 4 into -2 and 2, but then it would no longer be a function by our definition. By restricting the domain of $y = x^2$ to, say, $x \geq 0$, we can define an inverse function of $y = x^2, x \geq 0$.

Here is the formal definition of an inverse function:

Definition Let $f : A \to B$ be a one-to-one function with range $f(A)$. The **inverse** function f^{-1} has domain $f(A)$ and range A and is defined by

$$f^{-1}(y) = x \quad \text{if and only if} \quad y = f(x)$$

for all $y \in f(A)$.

EXAMPLE 12 Show that the function $f(x) = x^3 + 1, x \geq 0$ is one to one, and find its inverse.

Solution First, note that $f(x)$ is one to one. To see this quickly, graph the function and apply the horizontal line test. (See Figure 1.40.) Be aware, though, that unless you know what the graph looks like over its entire domain, the graphical approach can be misleading. To demonstrate it algebraically, start with $f(x_1) = f(x_2)$ and show that this implies $x_1 = x_2$:

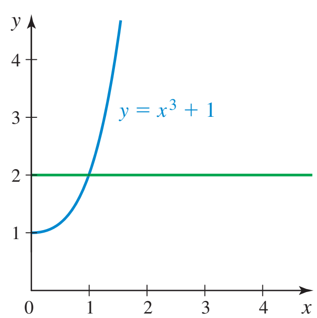


Figure 1.40 The graph of $f(x) = x^3 + 1$ in Example 12. The horizontal line test is successful.

$$\begin{aligned} f(x_1) &= f(x_2) \\ x_1^3 + 1 &= x_2^3 + 1 \\ x_1^3 &= x_2^3 \\ x_1 &= x_2 \quad \text{Take the cube root on both sides} \end{aligned}$$

Which tells us that $f(x)$ has an inverse. Now we will find f^{-1}.

To find an inverse function, we follow three steps:

1. Write $y = f(x)$:
$$y = x^3 + 1$$
2. Solve for x in terms of y:
$$\begin{aligned} x^3 &= y - 1 \\ x &= \sqrt[3]{y - 1} \end{aligned}$$
The range of f is $[1, \infty)$, and this range becomes the domain of f^{-1}, so we obtain
$$f^{-1}(y) = \sqrt[3]{y - 1}, \quad y \geq 1$$
Typically, we write functions in terms of x. To do this, we need to interchange x and y in $x = f^{-1}(y)$. This is the third step:
3. Interchange x and y:
$$y = f^{-1}(x) = \sqrt[3]{x - 1}, \ x \geq 1$$

•

Geometric Relationship Between $f(x)$ and $f^{-1}(x)$. Given a point x on the x-axis, the graph of $y = f(x)$ enables us to read the value of $f(x)$ from the height of the graph (as in Figure 1.41). Conversely, if we have a point y on the y-axis, and we want $f^{-1}(y)$, we can read across from y to the curve $y = f(x)$. The distance to the curve gives us the value of $x = f^{-1}(y)$, as Figure 1.41 shows.

What if we want our curve to represent f^{-1} directly? To get $y = f^{-1}(x)$ we can just interchange x and y, just like Step 3 of Example 12 (see Figure 1.42).

Now the x-axis is vertical and the y-axis is horizontal. Otherwise the graph is identical. To make the x-axis horizontal, we need to reflect the whole graph about the line $y = x$ (see Figure 1.43).

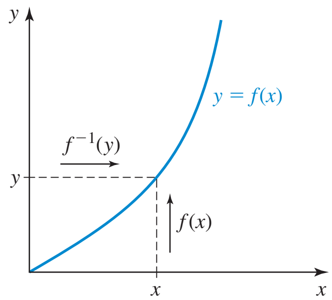


Figure 1.41 The graph of $y = f(x)$ allows $f(x)$ to be calculated for each x on the x-axis. It also allows $f^{-1}(y)$ to be calculated for each y on the y-axis.

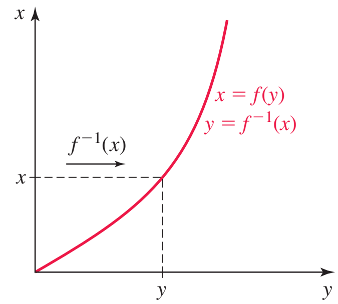


Figure 1.42 To plot $f^{-1}(x)$ interchange the labels on the x- and y-axes from Figure 1.41.

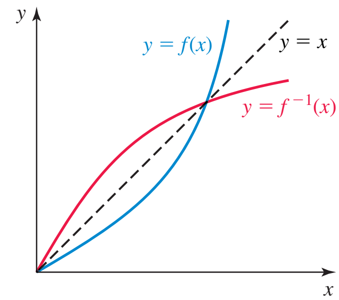


Figure 1.43 The graph of $y = f^{-1}(x)$ is obtained by reflecting the curve $y = f(x)$ about the line $y = x$.

In general the graphs of $y = f(x)$ and $y = f^{-1}(x)$ are related geometrically by reflection about $y = x$.

Caution! f^{-1} is not the reciprocal of f (i.e., $1/f$). This difference is further explained in Problem 74 at the end of this section.

As mentioned in the beginning of this subsection, the inverse of a function f reverses the effect of f. If we first apply f to x and then f^{-1} to $f(x)$, we obtain the original value x. Likewise, if we first apply f^{-1} to x and then f to $f^{-1}(x)$, we obtain the original value x. That is, if $f : A \to B$ has an inverse f^{-1}, then

$$f^{-1}[f(x)] = x \qquad \text{for all } x \in A$$
$$f[f^{-1}(x)] = x \qquad \text{for all } x \in f(A)$$

1.3.7 Logarithmic Functions

Recall from algebra (or Subsection 1.2.5) that, to solve the equation

$$e^x = 3$$

for x, you must take logarithms on both sides:

$$x = \ln 3$$

In other words, applying the natural logarithm undoes the operation of raising e to the x power. Thus, the natural logarithm is the inverse of the exponential function, and conversely, the exponential function is the inverse of the logarithmic function.

We will now define the inverse of the exponential function $f(x) = a^x$, $x \in \mathbf{R}$. The base a can be any positive number, except 1.

Definition The inverse of $f(x) = a^x$ is called the **logarithm to base *a*** and is written $f^{-1}(x) = \log_a x$.

The maximum domain of $f(x) = a^x$ is the set of all real numbers, and its range is the set of all positive numbers. Since the range of f is the domain of f^{-1}, we find that the maximum domain of $f^{-1}(x) = \log_a x$ is the set of positive numbers.

Because $y = \log_a x$ is the inverse function of $y = a^x$, we can find the graph of $y = \log_a x$ by reflecting the graph of $y = a^x$ about the line $y = x$. Recall that the graph of $y = a^x$ had two basic shapes, depending on whether $0 < a < 1$ or $a > 1$. (See Figure 1.35.) Figure 1.44 illustrates the graphs of $y = a^x$ and $y = \log_a x$ when $a > 1$, and Figure 1.45 illustrates the graphs of $y = a^x$ and $y = \log_a x$ when $0 < a < 1$.

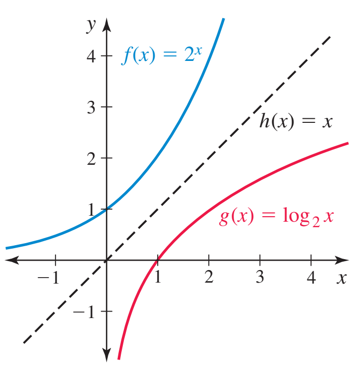


Figure 1.44 The graph of $y = a^x$ and the graph of $y = \log_a x$ for $a = 2$.

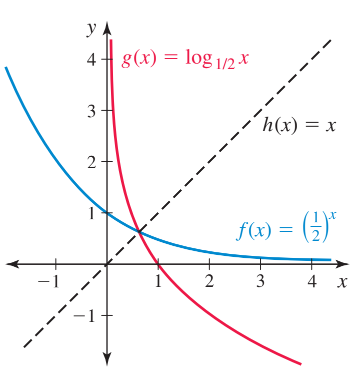


Figure 1.45 The graph of $y = a^x$ and the graph of $y = \log_a x$ for $a = \frac{1}{2}$.

Relationship Between Exponential and Logarithmic Functions

1. $\log_a a^x = x$ for $x \in \mathbf{R}$; i.e., if $y = \log_a x$ then $x = a^y$
2. $a^{\log_a x} = x$ for $x > 0$; i.e., if $y = a^x$ then $x = \log_a y$

It is important to remember that the logarithm is defined only for positive numbers; that is, $y = \log_a x$ is defined only for $x > 0$. The logarithm satisfies the following properties:

Properties of Logarithms

$$\log_a(st) = \log_a s + \log_a t$$

$$\log_a\left(\frac{s}{t}\right) = \log_a s - \log_a t$$

$$\log_a s^r = r\log_a s$$

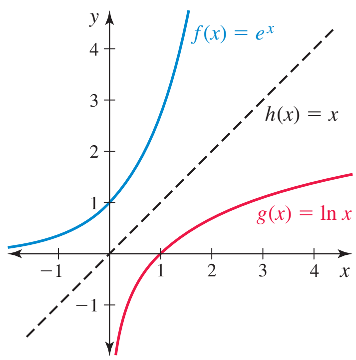


Figure 1.46 The graphs of $y = e^x$ and $y = \ln x$.

The inverse of the exponential function with the natural base e is denoted by $\ln x$ and is called the **natural logarithm** of x. The graphs of $y = e^x$ and $y = \ln x$ are shown in Figure 1.46. Note that both e^x and $\ln x$ are increasing functions. However, whereas e^x climbs very quickly for large values of x, $\ln x$ increases very slowly for large values of x. As with all inverse functions, each can be obtained as the reflection of the other about the line $y = x$.

The logarithm to base 10 is frequently written as $\log x$ (i.e., the base of 10 in $\log_{10} x$ is omitted).

EXAMPLE 13 Simplify the following expressions:

(a) $\log_2[8(x-2)]$ **(b)** $\log_3 9^x$ **(c)** $\ln e^{3x^2+1}$

Solution **(a)** We simplify as follows:

$$\log_2[8(x-2)] = \log_2 8 + \log_2(x-2) = 3 + \log_2(x-2)$$

No further simplification is possible.

(b) Simplifying yields

$$\log_3 9^x = x\log_3 9 = x\log_3 3^2 = 2x$$

$\log_3 9 = 2$ because $\log_3 9$ denotes the exponent to which we must raise 3 in order to get 9.

(c) We use the fact that $\ln x$ and e^x are inverse functions and find that

$$\ln e^{3x^2+1} = 3x^2 + 1$$

Any exponential function with base a can be written as an exponential function with base e. Likewise, any logarithm to base a can be written in terms of the natural logarithm.

Change of Base Formulas

$$a^x = \exp[x\ln a] = e^{x\ln a}$$

$$\log_a x = \frac{\ln x}{\ln a}$$

First, let's express a^x in base e. Notice that $a = e^{\ln a}$, so $a^x = (e^{\ln a})^x = e^{x\ln a} = \exp(x\ln a)$.

Second, let's express $\log x$ in base e (i.e., in terms of natural logarithms). Because $y = \log_a x$ means $a^y = x$, we have

$$x = \exp(y \ln a)$$

$$\ln x = y \ln a \qquad \text{Take log base } e \text{ of both sides.}$$

$$\frac{\ln x}{\ln a} = y = \log_a x.$$

EXAMPLE 14 Write the following expressions in terms of base e:

(a) 2^x **(b)** 10^{x^2+1} **(c)** $\log_3 x$ **(d)** $\log_2(3x - 1)$

Solution

(a) $2^x = \exp(x \ln 2) = e^{x \ln 2}$

(b) $10^{x^2+1} = \exp(\ln 10^{x^2+1}) = \exp\left[(x^2 + 1) \ln 10\right] = e^{(x^2+1)\ln 10}$

(c) $\log_3 x = \dfrac{\ln x}{\ln 3}$

(d) $\log_2(3x - 1) = \dfrac{\ln(3x - 1)}{\ln 2}$ •

EXAMPLE 15 A colony of bacteria is growing exponentially. The number of cells at time t is given by

$$N(t) = N_0 e^{\lambda t}$$

where N_0 is the initial number of cells in the colony (see 1.3.5).

(a) What is the doubling time (i.e., time taken for the number of cells to reach $2N_0$)?

(b) If $N_0 = 1000$ ($= 10^3$ cells) and the mean doubling time is 1 hour, when will the population reach $N = 10^9$ cells?

Solution

(a) After the doubling time t_d elapses, we have

$$N(t_d) = 2N_0$$

$$N_0 e^{\lambda t_d} = 2N_0$$

$$e^{\lambda t_d} = 2$$

$$\lambda t_d = \ln 2 \qquad \text{Take the natural log of both sides.}$$

$$t_d = \frac{\ln 2}{\lambda}.$$

We can rewrite the size of the population as:

$$N(t) = N_0 e^{(\ln 2)t/t_d} = N_0 2^{t/t_d}.$$

From this expression we can see that $N(t + t_d) = 2N(t)$, that is, the number of cells doubles once from $t = 0$ to $t = t_d$, again from $t = t_d$ to $t = 2t_d$, and again from $t = 2t_d$ to $t = 3t_d$, and so on.

(b) We are given that $t_d = 1$ hour, so we have

$$\lambda = \frac{\ln 2}{t_d} \qquad \text{Rearrange } t_d = \frac{\ln 2}{\lambda}$$

$$= \frac{\ln 2}{1 \text{ hr}}$$

$$\approx 0.693/\text{ hr} \qquad \text{Just as in Example 10, we need to keep the units of } \lambda.$$

Now we want to calculate t for which $N(t) = 10^9$.

$$
\begin{aligned}
N_0 e^{\lambda t} &= 10^9 \\
e^{\lambda t} &= 10^6 && \text{Divide both sides by } N_0 = 1000. \\
\lambda t &= \ln 10^6 = 6 \ln 10 && \text{Take ln of both sides and simplify the right-hand side.} \\
t &= \frac{6 \ln 10}{\lambda} = \frac{6 \ln 10}{0.693\ \text{hr}^{-1}} = 19.9\ \text{hr}.
\end{aligned}
$$

1.3.8 Trigonometric Functions

The trigonometric functions are examples of periodic functions.

Definition A function $f(x)$ is **periodic** if there is a positive constant a such that

$$f(x + a) = f(x)$$

for all x in the domain of f. If a is the smallest number with this property, we call it the **period** of $f(x)$.

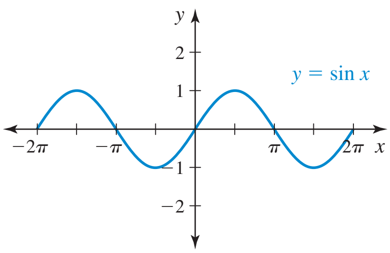


Figure 1.47 The graph of $y = \sin x$.

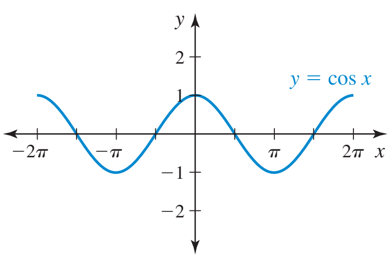


Figure 1.48 The graph of $y = \cos x$.

We begin with the sine and cosine functions. In Subsection 1.2.4, we recalled the definition of sine and cosine on a unit circle. There, $\sin\theta$ and $\cos\theta$ represented trigonometric functions of angles, and θ was measured in degrees or radians. Now we define the trigonometric functions as functions of *real numbers*. For instance, we define $f(x) = \sin x$ for $x \in \mathbf{R}$. The value of $\sin x$ is then, *by definition*, the sine of an angle of x radians (and similarly for all the other trigonometric functions).

The graphs of the sine and cosine functions are shown in Figures 1.47 and 1.48, respectively.

The sine function, $y = \sin x$, is defined for all $x \in \mathbf{R}$. Its range is $-1 \le y \le 1$. Likewise, the cosine function, $y = \cos x$, is defined for all $x \in \mathbf{R}$ with range $-1 \le y \le 1$. Both functions are periodic with period 2π. That is, $\sin(x + 2\pi) = \sin x$ and $\cos(x + 2\pi) = \cos x$. [We also have $\sin(x + 4\pi) = \sin x$, $\sin(x + 6\pi) = \sin x, \ldots$, and $\cos(x + 4\pi) = \cos x$, $\cos(x + 6\pi) = \cos x, \ldots$, but, by convention, we use the smallest possible value to specify the period.] We see from Figures 1.47 and 1.48 that the graph of the cosine function can be obtained by shifting the graph of the sine function a distance of $\pi/2$ units to the left. (We will discuss horizontal shifts of graphs in more detail in Section 1.4.)

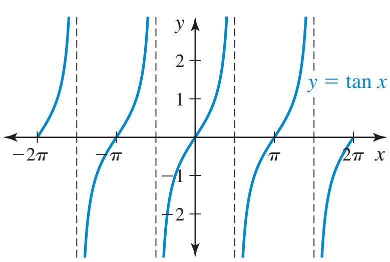


Figure 1.49 The graph of $y = \tan x$.

To define the tangent function, $y = \tan x$, recall that

$$\tan x = \frac{\sin x}{\cos x} \qquad \text{tan } x \text{ is not defined if } \cos x = 0.$$

Because $\cos x = 0$ for values of x that are odd integer multiples of $\pi/2$, the domain of $\tan x$ consists of all real numbers with the exception of odd integer multiples of $\pi/2$. The range of $y = \tan x$ is $-\infty < y < \infty$. The graph of $y = \tan x$ is shown in Figure 1.49, from which we see that $\tan x$ is periodic with period π.

The graphs of the remaining three trigonometric functions are shown in Figures 1.50–1.52. Recall that $\sec x = \frac{1}{\cos x}$, $\csc x = \frac{1}{\sin x}$, and $\cot x = \frac{1}{\tan x}$. It follows that the

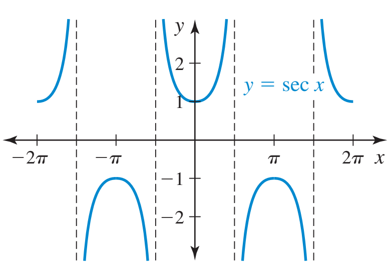


Figure 1.50 The graph of $y = \sec x$.

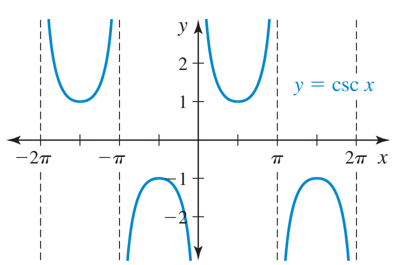


Figure 1.51 The graph of $y = \csc x$.

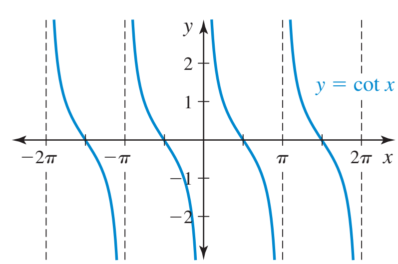


Figure 1.52 The graph of $y = \cot x$.

domain of the secant function $y = \sec x$ consists of all real numbers with the exception of odd integer multiples of $\pi/2$; the range is $|y| \geq 1$. The domain of the cosecant function $y = \csc x$ consists of all real numbers with the exception of integer multiples of π; the range is $|y| \geq 1$. The domain of the cotangent function $y = \cot x$ consists of all real numbers with the exception of integer multiples of π; the range is $-\infty < y < \infty$.

Since the sine and cosine functions are of particular importance, we now describe them in more detail. Consider the function

$$f(x) = a\sin(kx) \qquad \text{for } x \in \mathbf{R}$$

where a is a real number and $k \neq 0$. Now, $f(x)$ takes on values between $-a$ and a. We call $|a|$ the **amplitude**. The function $f(x)$ is periodic. To find the period p of $f(x)$, we set

$$|k|p = 2\pi \qquad \text{or} \qquad p = \frac{2\pi}{|k|}$$

Because the cosine function can be obtained from the sine function by a horizontal shift, we can define the amplitude and period analogously for the cosine function. That is, $f(x) = a\cos(kx)$ has amplitude $|a|$ and period $p = 2\pi/|k|$.

EXAMPLE 16 Compare $f(x) = 3\sin\left(\frac{\pi}{4}x\right)$ and $g(x) = \sin x$.

Solution The amplitude of $f(x)$ is 3, whereas the amplitude of $g(x)$ is 1. The period p of $f(x)$ satisfies $\frac{\pi}{4}p = 2\pi$ or $p = 8$, whereas the period of $g(x)$ is 2π. Graphs of $f(x)$ and $g(x)$ are shown in Figure 1.53. •

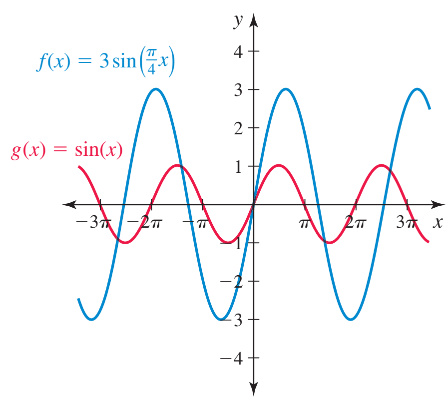


Figure 1.53 The graphs of $y = 3\sin(\frac{\pi}{4}x)$ and $g(x) = \sin x$ in Example 16.

Section 1.3 Problems

1.3.1

In Problems 1–4, state the range for the given functions. Graph each function.

1. $f(x) = x^2, x \in \mathbf{R}$

2. $f(x) = x^2, x \in [0, 2]$

3. $f(x) = x^2, -2 < x \leq 0$

4. $f(x) = x^2, -\frac{1}{2} < x < \frac{3}{2}$

5. (a) Show that, for $x \neq 1$,

$$\frac{x^2-1}{x+1} = x - 1$$

(b) Are the functions $f(x)$ and $g(x)$ equal?

$$f(x) = \frac{x^2-1}{x+1}, \quad x \neq -1$$

$$g(x) = x - 1, \quad x \in \mathbf{R}$$

6. (a) Show that

$$2|x-2| = \begin{cases} 2(x-2) & \text{for } x \geq 2 \\ 2(2-x) & \text{for } x \leq 2 \end{cases}$$

(b) Are the functions $f(x)$ and $g(x)$ equal?

$$f(x) = \begin{cases} 4-2x & \text{for } 0 \leq x \leq 2 \\ 2x-4 & \text{for } 2 \leq x \leq 3 \end{cases}$$

$$g(x) = 2|x-2|, \ x \in [0, 3]$$

In Problems 7–12, sketch the graph of each function and decide in each case whether the function is (i) even, (ii) odd, or (iii) does not show any obvious symmetry. Then use the criteria in Subsection 1.3.1 to check your answers.

7. $f(x) = 3x$

8. $f(x) = 3x^2$

9. $f(x) = |3x|$

10. $f(x) = 2x - 1$

11. $f(x) = -|x|$ **12.** $f(x) = 3x^3$

13. Suppose that $f(x) = x^2, x \in \mathbf{R}$ and $g(x) = 3 + x, x \in \mathbf{R}$.

(a) Show that $(f \circ g)(x) = (3 + x)^2, x \in \mathbf{R}$.

(b) Show that $(g \circ f)(x) = 3 + x^2, x \in \mathbf{R}$.

14. Suppose that $f(x) = \sqrt{x}, x \geq 0$, and $g(x) = 1 - 2x, x \in \mathbf{R}$.

(a) Show that $f \circ g(x) = \sqrt{1-2x}, x \leq \frac{1}{2}$.

(b) Show that $g \circ f(x) = 1 - 2\sqrt{x}, x \geq 0$.

15. Suppose that $f(x) = 1 - x, x \in \mathbf{R}$, and $g(x) = \sqrt{x}, x \geq 0$.

(a) Find $(f \circ g)(x)$ together with its domain.

(b) Find $(g \circ f)(x)$ together with its domain.

16. Suppose that $f(x) = \frac{1}{x+1}, x \neq -1$, and $g(x) = 2x^2, x \in \mathbf{R}$.

(a) Find $(f \circ g)(x)$. **(b)** Find $(g \circ f)(x)$.

In both (a) and (b), find the domain.

17. Suppose that $f(x) = \frac{1}{x}, x \neq 0$, and $g(x) = \sqrt{x}, x \geq 0$.

(a) Find $(f \circ g)(x)$ together with its domain.

(b) Find $(g \circ f)(x)$ together with its domain.

18. Suppose that $f(x) = x^4, x \geq 3$, and $g(x) = \sqrt{x+1}, x \geq 3$. Find $(f \circ g)(x)$ together with its domain.

19. Suppose that $f(x) = x^2, x \geq 0$, and $g(x) = \sqrt{x}, x \geq 0$. Typically, $f \circ g \neq g \circ f$, but this is an example in which the order of composition does not matter. Show that $f \circ g = g \circ f$.

20. Suppose that $f(x) = x^4, x \in \mathbf{R}$. For each of the following functions $g(x)$, determine whether $(f \circ g)(x) = (g \circ f)(x)$ or not.

(a) $g(x) = x + 1, x \in \mathbf{R}$. **(b)** $g(x) = \sqrt{x}, x \in \mathbf{R}$.

(c) $g(x) = \frac{1}{x}, x > 0$. **(d)** $g(x) = -x, x \in \mathbf{R}$.

(e) $g(x) = |x|, x \in \mathbf{R}$.

1.3.2

T **21.** Use a graphing calculator to graph $f(x) = x^2, x \geq 0$, and $g(x) = x^3, x \geq 0$, together. For which values of x is $f(x) > g(x)$, and for which is $f(x) < g(x)$?

T **22.** Use a graphing calculator to graph $f(x) = x^3, x \geq 0$, and $g(x) = x^5, x \geq 0$, together. When is $f(x) > g(x)$, and when is $f(x) < g(x)$?

T **23.** Graph $y = x^n, x \geq 0$, for $n = 1, 2, 3$, and 4 in one coordinate system. Where do the curves intersect?

T **24. (a)** Graph $f(x) = x, x \geq 0$, and $g(x) = x^2, x \geq 0$, together, in one coordinate system.

(b) For which values of x is $f(x) \geq g(x)$, and for which values of x is $f(x) \leq g(x)$?

T **25. (a)** Graph $f(x) = x^2$ and $g(x) = x^3$ for $x \geq 0$, together, in one coordinate system.

(b) Show algebraically that $x \geq x^2$ for $0 \leq x \leq 1$.

(c) Show algebraically that $x \leq x^2$ for $x \geq 1$.

26. Show algebraically that if $n \geq m, x^n \leq x^m$, for $0 \leq x \leq 1$, and $x^n \geq x^m$, for $x \geq 1$.

27. (a) Show that $y = x^4, x \in \mathbf{R}$, is an even function.

(b) Show that $y = x^3, x \in \mathbf{R}$, is an odd function.

28. Show that

(a) $y = x^n, x \in \mathbf{R}$, is an even function when n is an even integer.

(b) $y = x^n, x \in \mathbf{R}$, is an odd function when n is an odd integer.

29. Species Diversity In Example 4 we discussed the Gini-Simpson diversity index, which we defined to be the probability that two individuals randomly chosen from a population of two cell types are genetically different. We gave a formula for calculating Gini-Simpson diversity index, H:

$$H(p) = 2p(1-p) \quad p \in [0, 1].$$

Here p is the proportion of red individuals in the population. Conversely, we could ask what is the probability that the two individuals are genetically identical. Call this probability $I(p)$. It is given by

$$I(p) = 2p^2 - 2p + 1.$$

(a) The function $I(p)$ is known as the Simpson index. Explain why the domain of I is $p \in [0, 1]$.

(b) Use a graphing calculator to plot $I(p)$.

(c) What is the range $I([0, 1])$?

30. Chemical Reactions In Example 5 we considered the chemical reaction

$$A + B \to AB.$$

Suppose that the rate of the reaction is k, as in the example, but that the starting concentrations of the two reactants are $[A] = 3, [B] = 1$.

(a) Explain why the rate of the reaction is then $R(x) = k(3 - x)(1 - x)$, where x is the concentration of the product AB.

(b) What is the domain for $R(x)$? And what is its range?

(c) Suppose instead that the starting concentrations of A and B were $[A] = 6, [B] = 3$. Write the reaction rate $R(x)$ as a function of the concentration, x, of AB. Find the domain and range of $R(x)$.

31. Suppose that a beetle walks up a tree along a straight line at a constant speed of 1 meter per hour. What distance will the beetle have covered after 1 hour, 2 hours, and 3 hours? Write an equation that expresses the distance (in meters) as a function of the time (in hours), and show that this function is a polynomial of degree 1.

32. Suppose that a fungal disease originates in the middle of an orchard, initially affecting only one tree. The disease spreads out radially at a constant speed of 3 feet per day. What area will be affected after 2 days, 4 days, and 8 days? Write an equation that expresses the affected area as a function of time, measured in days, and show that this function is a polynomial of degree 2.

1.3.3

In Problems 33–36, for each function, find the largest possible domain and determine the range.

T **33.** $f(x) = \dfrac{1}{1-x}$ T **34.** $f(x) = \dfrac{2x+1}{(x-2)(x+3)}$

T **35.** $f(x) = \dfrac{x-2}{x^2-9}$ T **36.** $f(x) = \dfrac{1}{1-x^2}$

T **37.** Compare $y = \frac{1}{x}$ and $y = \frac{1}{x^2}$ for $x > 0$ by graphing the two functions. Where do the curves intersect? Which function is greater for small values of x? for large values of x?

38. Let n and m be two positive integers with $m \leq n$. Answer the following questions about $y = x^{-n}$ and $y = x^{-m}$ for $x > 0$: Where do the curves intersect? Which function is greater for small values of x? for large values of x?

39. Let

$$f(x) = \frac{x}{x+1}, \quad x > -1$$

(a) Use a graphing calculator to graph $f(x)$.

(b) On the basis of the graph in (a), determine the range of $f(x)$.

(c) For which values of x is $f(x) = 3/4$?

(d) On the basis of the graph in (a), determine how many solutions $f(x) = a$ has, where a is in the range of $f(x)$.

40. Let

$$f(x) = \frac{2}{3+x}, \quad x > -3$$

(a) Use a graphing calculator to graph $f(x)$.

(b) Find the range of $f(x)$.

(c) For which values of x is $f(x) = 1$?

(d) Based on the graph in (a), explain in words why, for any value a in the range of $f(x)$, you can find exactly one value $x \geq 0$ such that $f(x) = a$. Determine x for general a by solving $f(x) = a$.

41. Let

$$f(x) = \frac{2x+1}{1+x}, \quad x \geq 0$$

(a) Use a graphing calculator to graph $f(x)$.

(b) Find the range of $f(x)$.

(c) For which values of x is $f(x) = 5/4$?

(d) On the basis of the graph in (a), explain in words why, for any value a in the range of $f(x)$, you can find exactly one value $x \geq 0$ such that $f(x) = a$. Determine x by solving $f(x) = a$.

In Problems 42–44, we discuss the Monod growth function, which was introduced in Example 6 of this section.

42. In Example 6 we met the Monod growth function. The most general form of this function has two constants in it:

$$r(N) = \frac{aN}{k+N}. \qquad a \text{ and } k \text{ are positive constants.}$$

(Compare with Example 6, where we took $k = 1$.) In this question we will consider how the function changes if the constants a or b are changed.

(a) Graph $r(N)$ for (i) $a = 5$ and $k = 1$, (ii) $a = 5$ and $k = 3$, and (iii) $a = 8$ and $k = 1$. Place all three graphs in one coordinate system.

(b) By comparing the graphs of (a)(i) and (a)(iii), describe in words what happens when you change a.

(c) By comparing the graphs of (a)(i) and (a)(ii), describe in words what happens when you change k.

43. The Monod growth function $r(N)$ describes growth as a function of nutrient concentration N. Assume that

$$r(N) = \frac{5N}{1+N}, \quad N \geq 0$$

Find the percentage increase when the nutrient concentration is doubled from $N = 0.1$ to $N = 0.2$. Compare this result with what you find when you double the nutrient concentration from $N = 20$ to $N = 40$. This is an example of *diminishing return*.

44. In Example 6 we met the Monod growth function. The most general form of this function has two constants in it:

$$r(N) = \frac{aN}{k+N}. \qquad a \text{ and } k \text{ are positive constants.}$$

(Compare with Example 6 where we took $k = 1$.) In this question we will consider how, given some experimental data, we can determine values for a and k to fit the Monod growth function to the data.

First, we measure growth rate for three values of N:

N	$r(N)$
0	0
2	1.5
4	2

We want to find the values of a and k that would fit the Monod growth function to this data. Write out the equations for $r(0)$, $r(2)$, and $r(4)$:

$$r(0) : 0 = 0 \tag{1.5}$$

$$r(2) : \frac{2a}{k+2} = 1.5 \tag{1.6}$$

$$r(4) : \frac{4a}{k+4} = 2. \tag{1.7}$$

Equation (1.5) is automatically satisfied. We need to pick values of a and k that satisfy (1.6) and (1.7). To do this, we need to eliminate one variable so that we have one equation in one unknown.

(a) To eliminate a, divide (1.6) into (1.7) (i.e., divide the left-hand side of (1.7) by the left-hand side of (1.6) and the right-hand side of (1.7) by the right-hand side of (1.6)):

$$\frac{2(k+2)}{k+4} = \frac{2}{1.5} = \frac{4}{3}.$$

(i) Solve this equation for k.

(ii) Substitute your value for k back into (1.6) and solve for a.

(iii) What if you instead substitute your value for k from (i) into (1.7), and solve for a? Do you get a different answer?

(b) Suppose in a different experiment you measured the following data:

N	$r(N)$
0	0
1	1
3	2.25

Calculate values for a and k to fit the Monod growth function to this data.

(c) Suppose in a different experiment you measured the following data:

N	$r(N)$
0	0.5
1	1
3	1.5

Are there any values for a and k that would fit the Monod growth function to this data?

45. Let

$$f(x) = \frac{x^2}{4+x^2}, \quad x \geq 0$$

(a) Use a graphing calculator to graph $f(x)$.

(b) On the basis of your graph in (a), find the range of $f(x)$.

(c) What happens to $f(x)$ as x gets larger?

46. We sometimes talk about functions increasing at accelerating (or decelerating) rates. To clarify what we mean by these terms, we will consider some specific examples.

(a) First, the Monod growth function:

$$r(N) = \frac{N}{2+N} \quad N \geq 0.$$

(i) Calculate $r(0.1)$ and $r(0.2)$. How much does $r(N)$ change when N is increased by 0.1, from $N = 0.1$?

(ii) Now calculate $r(2)$ and $r(2.1)$. Show that r increases less from when N is initially 2 and is increased by 0.1 than when N is initially 0.1 and is increased by 0.1.

(iii) Calculate $r(4)$ and $r(4.1)$. Show that r increases less when N is initially 4 than it does when N is 2 initially.

r is increasing with N, but the effect of increasing N by the same amount is smaller for $N = 4$ than $N = 2$, and smaller for $N = 2$ than $N = 0.1$. So we say that $r(N)$ is increasing at a *decelerating rate*.

(b) Now consider the growth function

$$s(N) = \frac{N^2}{4+N^2} \quad N \geq 0$$

This is an example of a *sigmoidal* function.

(i) Plot $s(N)$ and $r(N)$ on the same axes. Show that $s(N)$ and $r(N)$ have the same range.

(ii) Compare the increases in $s(N)$ when N is increased by 0.1, from

0 to 0.1, 2 to 2.1, and 4 to 4.1.

Since s increases more from 2 to 2.1 than from 0 to 0.1, we say that s increases at an accelerating rate between $N = 0$ and $N = 2$.

(iii) Show that the increase of $s(N)$ is *decelerating* for large N.

1.3.4

In Problems 47–50, use a graphing calculator to sketch the graphs of the functions.

T **47.** $y = x^{3/2}, x \geq 0$

T **48.** $y = x^{1/3}, x \geq 0$

T **49.** $y = x^{-1/3}, x > 0$

T **50.** $y = 2x^{-7/8}, x > 0$

T **51. (a)** Graph $y = x^{-1/2}, x > 0$, and $y = x^{1/2}, x \geq 0$, together, in one coordinate system.

(b) Show algebraically that $x^{-1/2} \geq x^{1/2}$ for $0 < x \leq 1$.

(c) Show algebraically that $x^{-1/2} \leq x^{1/2}$ for $x \geq 1$.

T **52. (a)** Graph $y = x^{5/2}, x \geq 0$, and $y = x^{1/2}, x \geq 0$, together, in one coordinate system.

(b) Show algebraically that $x^{5/2} \leq x^{1/2}$ for $0 \leq x \leq 1$. (*Hint*: Show that $x^{1/2}/x^{-1/2} = x \leq 1$ for $0 < x \leq 1$.)

(c) Show algebraically that $x^{5/2} \geq x^{1/2}$ for $x \geq 1$.

We introduced power laws using the relationship between mammal body mass (M) measured in grams and resting heart rate (R).* Savage et al. (2004) *showed that the two are related by a power law:* $R(M) = 1180M^{-1/4}$. *This relationship is to be interpreted in a statistical sense; it does not exactly predict the resting heart rate, as the following question shows.

53. (a) Professor R.'s dog Molly weights 25.4 kg. Predict Molly's heart rate from the formula for $R(M)$.

(b) Professor R. weighs 78 kg. Predict his resting heart rate from the formula for $R(M)$.

(c) In reality Molly's resting heart rate is about 40 beats per minute, and Professor R.'s heart rate is about 60 beats per minute. Calculate the relative error from using the heart beat formula. The formula for relative error is:

$$\frac{|\text{real heart beat rate} - \text{predicted heart beat rate}|}{\text{predicted heart beat rate}}.$$

In each of the questions 54–56, use a measured power law to make predictions.

54. Sabretooth Tiger Bite Force Estimating the mass of extinct animals is very difficult, since estimates have to be made based on fossil skeletons and without knowledge of the muscles and other tissues that make up much of an animal's mass. ***van Valkenburgh (1990)*** measured body mass and skull length in living carnivores and found that, if y is body mass measured in kg and x is skull length measured in mm,

$$y = kx^{3.13}.$$

(a) A gray wolf (*Canis lupus*) weighs 45 kg, and has a skull length of 275 mm. Calculate the coefficient of proportionality k in the power law.

(b) Estimate the weight of the sabretooth tiger (*Smilodon populator*). *S. populator* fossils have skull lengths of 350 mm.

(c) How strong was *Smilodon's* bite? ***Hartstone-Rose, Perry, and Morrow (2012)*** studied the correlation between bite force and body mass. They found that, if bite force (measured in newtons, or N) is f and body mass (measured in kg) is y, then

$$f = cy^{0.96}$$

for some constant c.

(i) The tiger has a bite force of 7980 N, and a mass of 200 kg. Calculate the constant c.

(ii) Estimate the bite force of *Smilodon populator* using your body mass estimate from part (b).

T **55. Language Diversity** ***Gomes et al. (1999)*** found a power law relationship between land area and language diversity. Specifically, if A is the area of a region measured in km^2 and D is the number of different languages spoken by the people occupying that region, then

$$D = 0.2A^{0.41}. \qquad (1.8)$$

(a) Estimate from this formula the number of languages spoken in the United States (area: 9.857×10^6 km^2).
[In reality, there are 337 languages within the United States, but your answer from (a) agrees well with the 176 indigenous languages spoken within the United States.]

(b) Equation (1.8) is a statistical relationship rather than a strict mathematical formula. To see this, make a plot showing both equation (1.8) and the following data points on the same axes.

Country	Area (km^2)	# Languages
Cameroon	475,400	230
Belgium	30,500	3
China	9,597,000	129
India	3,287,600	122
Mexico	1,943,900	60
Kenya	569,100	68
United Arab Emirates	83,600	7

(i) How many countries lie near the line given by the formula?

(ii) Are there any countries that do not lie on the line given by the formula?

(iii) Use equation (1.8) to predict the number of languages in the UCLA campus (the area of the UCLA campus is 1.7 km^2). Does your answer make sense?

56. Cat Tongue Motion ***Reis et al. (2010)*** investigated how the frequency of tongue motion (f, number of licks per second) for a cat drinking water depends on the cat's mass (M, measured in kg). They found that

$$f = cM^{-1/6} \tag{1.9}$$

that is, frequency of licks is proportional to (body mass)$^{-1/6}$.

(a) Given the following data

Species	Mass (kg)	Frequency (s^{-1})
Domestic cat	4.8	3.55
Tiger	175.0	1.70

calculate the constant c in (1.9).

(b) Calculate the lapping frequency for Professor R.'s slightly tubby pet cat, Texas Ranger (mass, $M = 6.4$ kg).

(c) Calculate the lapping frequency for a cheetah (mass, $M = 40$ kg).

(d) Estimate from equation (1.9) the mass of a cat whose lapping frequency is less than 1 s^{-1}.

1.3.5

57. Assume that a population size at time t is $N(t)$ and that $N(t) = 2^t, t \geq 0$.

(a) Find the population size for $t = 0, 1, 2, 3$, and 4.

(b) Graph $N(t)$ for $t \geq 0$.

58. Assume that a population size at time t is $N(t)$ and that $N(t) = 20 \cdot 2^t, t \geq 0$.

(a) Find the population size at time $t = 0$.

(b) Show that $N(t) = 20e^{t \ln 2}, t \geq 0$.

(c) How long will it take until the population size reaches 1000? [*Hint*: Find t so that $N(t) = 1000$.]

59. The half-life of C^{14} is 5730 years. If a sample of C^{14} has a mass of 20 micrograms at time $t = 0$, how much is left after 2000 years?

60. The half-life of C^{14} is 5730 years. If a sample of C^{14} has a mass of 40 micrograms at time 0, how long will it take until **(a)** 10 grams and **(b)** 5 grams are left?

61. After 7 days, a particular radioactive substance decays to half of its original amount. Find the decay rate of this substance.

62. After 4 days, a particular radioactive substance decays to 30% of its original amount. Find the half-life of this substance.

63. Polonium 210 (Po^{210}) has a half-life of 140 days.

(a) If a sample of Po^{210} has a mass of 100 micrograms, find a formula for the mass after t days.

(b) How long would it take this sample to decay to 10% of its original amount?

(c) Sketch the graph of the amount of mass left after t days.

64. The half-life of C^{14} is 5730 years. Suppose that wood found at an archeological excavation site contains about 35% as much C^{14} (in relation to C^{12}) as does living plant material. Determine when the wood was cut.

65. The half-life of C^{14} is 5730 years. Suppose that wood found at an archeological excavation site is 10,000 years old. How much C^{14} (based on C^{12} content) does the wood contain relative to living plant material?

66. The age of rocks of volcanic origin can be estimated with isotopes of argon 40 (Ar^{40}) and potassium 40 (K^{40}). K^{40} decays into Ar^{40} over time. If a mineral that contains potassium is buried under the right circumstances, argon forms and is trapped. Since argon is driven off when the mineral is heated to very high temperatures, rocks of volcanic origin do not contain argon when they are formed. The amount of argon found in such rocks can therefore be used to determine the age of the rock. Assume that a sample of volcanic rock contains 0.00047% K^{40}. The sample also contains 0.000079% Ar^{40}. How old is the rock? (The decay rate of K^{40} to Ar^{40} is 5.335×10^{-10}/yr.)

67. A growing population contains $N(t)$ individuals $N(t)$ at time t is modeled by the equation

$$N(t) = N_0 e^{rt}$$

where N_0 denotes the population size at time 0. The constant r is called the **intrinsic rate of growth**.

(a) Plot $N(t)$ as a function of t if $N_0 = 100$ and $r = 2$. Compare your graph against the graph of $N(t)$ when $N_0 = 100$ and $r = 3$. Which population grows faster?

(b) You are given the following data for the size of the population.

t	$N(t)$
0	100
2	300

(i) Calculate the parameters N_0 and r to make the mathematical model fit the data.

(ii) When will the population size first reach 1000 individuals?

(iii) When will the population size first reach 10,000 individuals?

68. Earthquake Prediction ***Rundle et al. (2003)*** showed that earthquakes in Southern California obey an exponential distribution—that is, if $N(m)$ is the number of earthquakes in a given year whose magnitude exceeds m, then

$$N(m) = c \cdot 10^{-m}$$

where c is a positive constant.

(a) Suppose in a given year there are 10 earthquakes of magnitude 5 or above.

(i) Calculate the constant c.

(ii) How many earthquakes will have magnitudes exceeding 2? (2 is the threshold at which earthquakes can be felt by most people.)

(iii) How many earthquakes will have magnitude exceeding 6? (6 is the threshold for an earthquake to be regarded as strong.)

1.3.6

69. Which of the following functions is one to one (use the horizontal line test)?

(a) $f(x) = x^2, x \geq 0$ **(b)** $f(x) = x^2, x \in \mathbf{R}$

(c) $f(x) = \sqrt{x}, x \geq 0$ **(d)** $f(x) = \ln x, x > 0$

(e) $f(x) = \frac{1}{x^2}, x \neq 0$ **(f)** $f(x) = \frac{1}{x^2}, x > 0$

70. (a) Show that $f(x) = x^3 - 1$, $x \in \mathbf{R}$, is one to one, and find its inverse together with its domain.

(b) Graph $f(x)$ and $f^{-1}(x)$ in one coordinate system, together with the line $y = x$, and convince yourself that the graph of $f^{-1}(x)$ can be obtained by reflecting the graph of $f(x)$ about the line $y = x$.

71. (a) Show that $f(x) = x^2 + 1$, $x \geq 0$, is one to one, and find its inverse together with its domain.

(b) Graph $f(x)$ and $f^{-1}(x)$ in one coordinate system, together with the line $y = x$, and convince yourself that the graph of $f^{-1}(x)$ can be obtained by reflecting the graph of $f(x)$ about the line $y = x$.

72. (a) Show that $f(x) = x^2 - x$, $x \geq \frac{1}{2}$, is one to one, and find its inverse together with its domain.

(b) Graph $f(x)$ and $f^{-1}(x)$ in one coordinate system, together with the line $y = x$, and convince yourself that the graph of $f^{-1}(x)$ can be obtained by reflecting the graph of $f(x)$ about the line $y = x$.

73. (a) Show that $f(x) = 1/x^3$, $x > 0$, is one to one, and find its inverse together with its domain.

(b) Graph $f(x)$ and $f^{-1}(x)$ in one coordinate system, together with the line $y = x$, and convince yourself that the graph of $f^{-1}(x)$ can be obtained by reflecting the graph of $f(x)$ about the line $y = x$.

74. The reciprocal of a function $f(x)$ can be written as either $1/f(x)$ or $[f(x)]^{-1}$. The point of this problem is to make clear that a reciprocal of a function has nothing to do with the inverse of a function. As an example, let $f(x) = x + 1$, $x \in \mathbf{R}$. Find both $[f(x)]^{-1}$ and $f^{-1}(x)$, and compare the two functions. Graph all three functions together.

1.3.7

T **75.** Find the inverse of $f(x) = 3^x$, $x \in \mathbf{R}$, together with its domain, and graph both functions in the same coordinate system.

T **76.** Find the inverse of $f(x) = 4^x$, $x \in \mathbf{R}$, together with its domain, and graph both functions in the same coordinate system.

T **77.** Find the inverse of $f(x) = (\frac{1}{4})^x$, $x \in \mathbf{R}$, together with its domain, and graph both functions in the same coordinate system.

T **78.** Find the inverse of $f(x) = (\frac{1}{3})^x$, $x \in \mathbf{R}$, together with its domain, and graph both functions in the same coordinate system.

79. Simplify the following expressions:

(a) $2^{5\log_2 x}$ **(b)** $4^{3\log_4 x}$ **(c)** $5^{5\log_{1/5} x}$

(d) $4^{3\log_2 x}$ **(e)** $2^{3\log_{1/2} x}$ **(f)** $8^{\log_{1/2} x}$

80. Simplify the following expressions:

(a) $\log_4 16^x$ **(b)** $\log_2 16^x$ **(c)** $\log_3 27^x$

(d) $\log_{1/2} 4^x$ **(e)** $\log_{1/2} 8^{-x}$ **(f)** $\log_3 9^{-x}$

81. Simplify the following expressions:

(a) $\ln x^2 + \ln x^{-1}$ **(b)** $\ln x^4 - \frac{1}{3}\ln x^{-2}$

(c) $\ln(x^2 - 1) - \ln(x + 1)$ **(d)** $\frac{1}{2}\ln x^{-1} + \ln x^{-3}$

82. Simplify the following expressions:

(a) $e^{3\ln x}$ **(b)** $e^{-\ln(x^2+1)}$

(c) $e^{-2\ln(1/x)}$ **(d)** $e^{-2\ln x}$

83. Write the following expressions in terms of base e, and simplify:

(a) 3^x **(b)** 4^{x^2-1} **(c)** 2^{-x-1} **(d)** 3^{-4x+1}

84. Write the following expressions in terms of base e:

(a) $\log_2(x^2 - 1)$ **(b)** $\log_3(5x + 1)$

(c) $\log(x + 2)$ **(d)** $\log_2(2x^2 - 1)$

85. Show that the function $y = (1/2)^x$ can be written in the form $y = e^{-\mu x}$, where μ is a positive constant. Determine μ.

86. Show that if $a > 1$, then the function $y = a^x$ can be written in the form $y = e^{\mu x}$, where μ is a positive constant. Write μ in terms of a.

87. Earthquakes Strength The Richter magnitude scale is used to measure the strength of earthquakes. The magnitude m of an earthquake is calculated from the amplitude of shaking, A (measured in μm, where $1\ \mu\text{m} = 10^{-6}$ m), measured by a seismometer, and from the distance of the seismometer to the epicenter of the earthquake, D (measured in km), using the following formula.

$$m = \log A - 2.48 + 2.76 \log D.$$

(a) A seismometer distance 100 km from the earthquake epicenter measures shaking with an amplitude of 100 μm. Calculate m.

(b) The smallest amplitude of shaking that most people can feel is 1 mm (1 mm $= 10^3\mu$m). Calculate the smallest magnitude of earthquake a person might feel if they were 10 km away from the earthquake epicenter.

(c) An earthquake is measured to have magnitude $m = 7.2$. Calculate the amplitude of shaking if

(i) $D = 10$ km from the epicenter.

(ii) $D = 100$ km from the epicenter.

(d) Measured at the same distance from the epicenter, an increase of 1 in the Richter magnitude of an earthquake (e.g., from $m = 3$ to $m = 4$) corresponds to what factor increase in the amplitude of shaking?

88. Species Diversity In Example 4 we introduced the Gini-Simpson index for measuring the diversity of a region containing two different types of individuals. Another index of diversity that is very commonly used is the Shannon diversity index. For a region that contains two different types of organisms, and in which a proportion p of organisms are of type 1, and a proportion $1 - p$ are of type 2, the Shannon diversity index is given by the formula:

$$H(p) = -p \ln p - (1 - p)\ln(1 - p) \tag{1.10}$$

(a) What is the domain of the function H?

(b) Show by plotting H against p that the maximum value for the Shannon diversity index occurs when $p = 1/2$, that is, when both types of organisms are equally abundant in the population.

(c) Show that H has the following symmetry: $H(1 - p) = H(p)$. Explain why this implies that if we swap the labels on the two types of organisms (so that type 1 is now called type 2, and type 2 is now called type 1), then the Shannon diversity index will not change.

(d) If type 1 goes extinct (so that $p = 0$), what happens to the Shannon diversity index?

(i) First, explain why we cannot directly evaluate the formula for H if $p = 0$.

(ii) Then consider what happens if $p \neq 0$ but is very small, say $p = 0.001$. Then you should be able to use (1.10) to evaluate $H(p)$.

(iii) Make a table of values of $H(p)$ if $p = 0.001$, 10^{-4}, 10^{-5}, 10^{-6}. Show that H gets closer to 0 as p gets smaller. Because of this observation we define $H(0) = 0$ and $H(1) = 1$. We will return to this idea when we discuss continuity in Chapter 3.

(e) If we use (1.10) to evaluate H when $0 < p < 1$ and define $H(0) = H(1) = 0$, what is the range of the function $H(p)$?

1.3.8

In Problems 89–94, for each given pair of functions, use a graphing calculator to compare the functions. Describe what you see.

89. $y = \sin x$ and $y = 2\sin x$

90. $y = \cos x$ and $y = \cos 2x$

91. $y = \cos x$ and $y = 2\cos x$

92. $y = \cos x$ and $y = \cos(x+1)$

93. $y = \tan x$ and $y = 2\tan x$

94. $y = \tan x$ and $y = \tan(x+1)$

95. Find the amplitude and the period of $f(x)$:

$$f(x) = 2\sin\left(\frac{x}{2}\right), \quad x \in \mathbf{R}$$

96. Find the amplitude and the period of $f(x)$:

$$f(x) = 3\cos 4x, \quad x \in \mathbf{R}$$

97. Find the amplitude and the period of $f(x)$:

$$f(x) = -4\sin(\pi x), \quad x \in \mathbf{R}$$

98. Find the amplitude and the period of $f(x)$:

$$f(x) = -\frac{3}{2}\sin\left(\frac{\pi}{3}x\right), \quad x \in \mathbf{R}$$

99. The mean monthly precipitation ($p(t)$, measured in inches/month), in Los Angeles varies over the course of year according to Figure 1.54, where t is the time elapsed in months from the beginning of this year ($0 \le t \le 12$). Write the formula for the function $p(t)$.

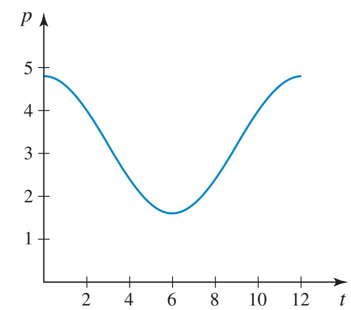


Figure 1.54 Monthly precipitation in Los Angeles. Simplified from real data.

100. The growth rate of the bread mold *Neuropora crassa* varies over the course of a day, because the fungus responds to changes in light. Your measure the growth rate for growth rate $G(t)$ in mm/hr as a function of time t in hours (see Figure 1.55).

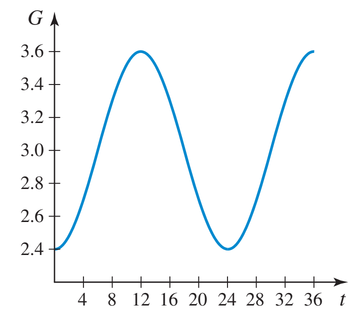


Figure 1.55 Growth rate of *Neurospora crassa* as a function of time. Graph is simplified from real data.

(a) What is the period of the function $G(t)$?

(b) Write a formula for $G(t)$ as a function of time t.

101. Use the fact that $\sec x = \frac{1}{\cos x}$ to explain why the maximum domain of $y = \sec x$ consists of all real numbers except odd integer multiples of $\pi/2$.

102. Use the fact that $\cot x = \frac{1}{\tan x}$ to explain why the maximum domain of $y = \csc x$ consists of all real numbers except integer multiples of π.

1.4 Graphing

In the preceding section, we introduced the functions most important to our study. You must be able to graph the following functions without a calculator: $y = c$, x, x^2, x^3, $1/x$, e^x, $\ln x$, $\sin x$, $\cos x$, $\sec x$, and $\tan x$. This will help you to sketch functions quickly and to come up with an analytical description of a function based on a graph. In this section, you will learn how to obtain new functions from these basic functions and how to graph them. In addition, we will introduce important transformations that are often used to analyze data.

1.4.1 Graphing and Basic Transformations of Functions

In this subsection, we will recall some basic transformations rules: vertical and horizontal translations, reflections about $x = 0$ and $y = 0$, and stretching and compressing.

The graph of

$$y = f(x) + a$$

is a **vertical translation** of the graph of $y = f(x)$. If $a > 0$, the graph of $y = f(x)$ is shifted up a units; if $a < 0$, the graph of $y = f(x)$ is shifted down $|a|$ units.

This definition is illustrated in Figure 1.56, where we display $y = x^2$, $y = x^2 + 2$, and $y = x^2 - 2$.

> The graph of
> $$y = f(x - c)$$
> is a **horizontal translation** of the graph of $y = f(x)$. If $c > 0$, the graph of $y = f(x)$ is shifted c units to the right; if $c < 0$, the graph of $y = f(x)$ is shifted $|c|$ units to the left.

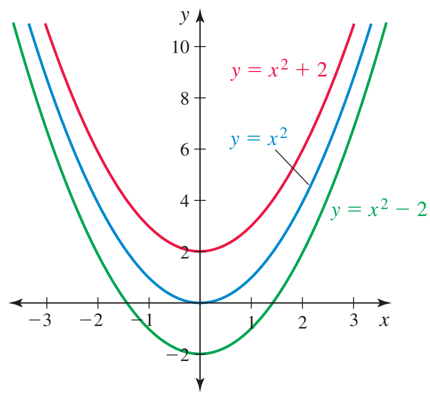


Figure 1.56 The graphs of $y = x^2 + 2$ and $y = x^2 - 2$ are obtained from $y = x^2$ by translation in the y-direction.

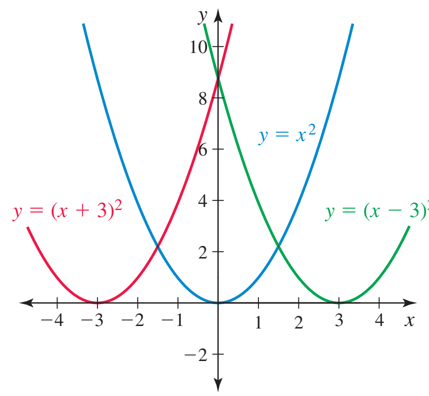


Figure 1.57 The graphs of $y = (x - 3)^2$ and $y = (x + 3)^2$ are obtained from $y = x^2$ by translation in the x-direction.

This definition is illustrated in Figure 1.57, where we display $y = x^2$, $y = (x - 3)^2$, and $y = (x + 3)^2$. ($y = (x + 3)^2$ is shifted to the *left*, since $y = (x - (-3))^2$ so $c = -3 < 0$.)

> The graph of $y = -f(x)$ is a reflection of the graph of $y = f(x)$ about the x-axis. The graph of $y = f(-x)$ is a reflection of $y = f(x)$ about the y-axis.

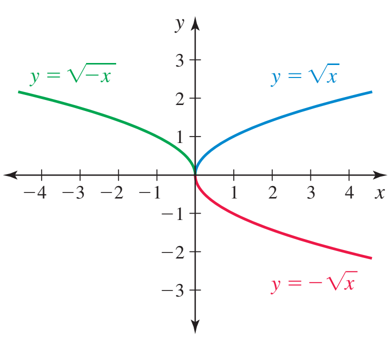


Figure 1.58 Reflections about the x-axis and the y-axis.

Reflections about the x-axis ($y = 0$) and the y-axis ($x = 0$) are illustrated in Figure 1.58. We graph $y = \sqrt{x}$; its reflection about the x-axis, $y = -\sqrt{x}$; and its reflection about the y-axis, $y = \sqrt{-x}$.

> Multiplying a function by a factor between 0 and 1 compresses the graph of the function in the y-direction; multiplying a function by a factor greater than 1 stretches the graph of the function in the y-direction.

These operations are illustrated in Figure 1.59.

We illustrate the preceding transformations in the next two examples.

EXAMPLE 1 Explain how the graph of

$$y = 2\sin\left(x - \frac{\pi}{4}\right) \quad \text{for } x \in \mathbf{R}$$

can be obtained from the graph of $y = \sin x$, $x \in \mathbf{R}$.

Solution We transform $y = \sin x$ in two steps, illustrated in Figure 1.60. First we shift $y = \sin x$ to the right $\frac{\pi}{4}$ units. This yields $y = \sin(x - \frac{\pi}{4})$. Then we multiply $y = \sin(x - \frac{\pi}{4})$ by 2. This corresponds to stretching $y = \sin(x - \frac{\pi}{4})$ by the factor 2. •

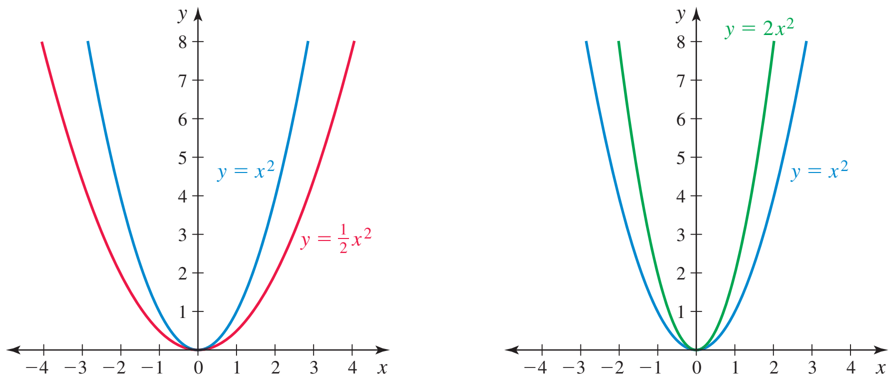


Figure 1.59 The graph of $y = \frac{1}{2}x^2$ can be obtained from $y = x^2$ by compressing in the y-direction. The graph of $y = 2x^2$ can be obtained from $y = x^2$ by stretching in the y-direction.

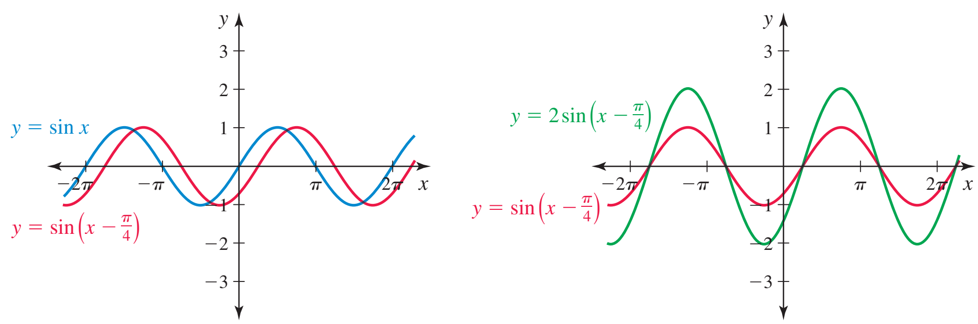


Figure 1.60 Using a sequence of transformations to turn $y = \sin x$ into $y = 2\sin(x - \frac{\pi}{4})$.

EXAMPLE 2 Explain how the graph of $y = -\sqrt{x-3} - 1$, $x \geq 3$, can be obtained from the graph of $y = \sqrt{x}$, $x \geq 0$.

Solution We transform $y = \sqrt{x}$ in three steps, illustrated in Figure 1.61. First we shift $y = \sqrt{x}$ three units to the right and obtain $y = \sqrt{x-3}$. Then we reflect $y = \sqrt{x-3}$ about the x-axis, which yields $y = -\sqrt{x-3}$. Finally, we shift $y = -\sqrt{x-3}$ down one unit. This is the graph of $y = -\sqrt{x-3} - 1$. •

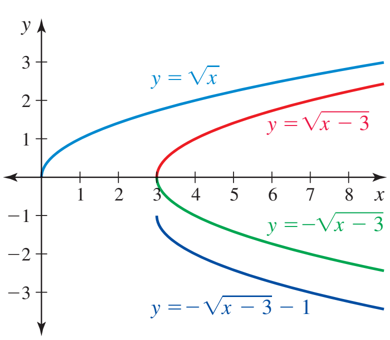


Figure 1.61 Using a sequence of transformations to turn $y = -\sqrt{x}$ into $y = -\sqrt{x-3} - 1$.

1.4.2 The Logarithmic Scale

Bio Info • Important Biological Length Scales We often encounter sizes that vary over a wide range. For instance, lengths in the metric system are measured in meters (m). (A meter is a bit longer than a yard: 1 meter is equal to 1.0936 yards.) A longer metric unit that is commonly used is a kilometer (km), which is 1000 m. Shorter commonly used metric units are a millimeter (mm), which is 1/1000 of a meter; a micrometer (μm), which is 1/1,000,000 (one-millionth) of a meter; and a nanometer (nm), which is 1/1,000,000,000 (one-billionth) of a meter.

Here are some examples of important biological lengths: A ribosome is about 20 nm ($= 2 \times 10^{-8}$ m), a poxvirus is about 400 nm (4×10^{-7} m), a bacterium is about 1 μm ($= 10^{-6}$ m), a tardigrade (or "water bear") is about 1.2 mm ($= 1.2 \times 10^{-3}$ m), an adult human is about 1.8 m, a blue whale is between 25 and 35 m. The tallest tree is Hyperion, a 116-m-tall coast redwood in Redwood National Park. The largest single organisms are colonies of clonal plants or fungi. "Pando," a colony of aspen in Utah, is 1.1 km wide, while a colony of honey mushrooms in Oregon is 3.8 km wide ***(DeWoody et al., 2008; Partes et al. 2003)***.

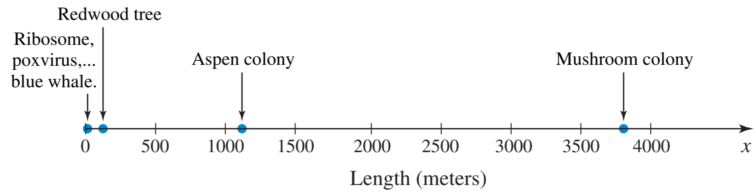


Figure 1.62 Important biological lengths represented on a number line. Note that many points are clustered together on the left.

Let's represent these different length scales on a single number line (Figure 1.62). We see that it is very hard to represent data covering multiple orders of magnitude on a single number line: Only the two or three largest lengths can be distinguished. We want a number line that allows data with a very large range to be distinguished. This can be accomplished by plotting on a logarithmic scale, in which the spacing between 0.1 and 1 is the same as the spacing between 1 and 10, or 10 and 100, or 100 and 1000, and so on (Figure 1.63).

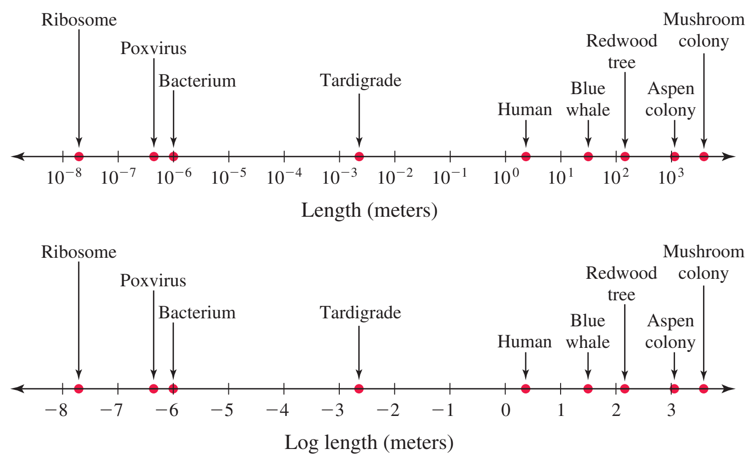


Figure 1.63

If rather than recording the length on the number axes, we record the log of the length, we see how the transformed scale is constructed by plotting the exponents of the real lengths on a linear scale.

Note that, although the exponents on the number line are negative, the numbers themselves are all positive. The origin corresponds to a length of 1, moving left means getting smaller and smaller (but never reaching 0), and moving right the data becomes larger and larger.

The lengths in the preceding examples differed by many factors of 10. Instead of saying that quantities differ by many factors of 10, we will say that they differ by many **orders of magnitude**: If two quantities differ by a factor of 10, they differ by one order of magnitude. (If they differ by a factor of 100, they differ by two orders of magnitude, and so on.) Orders of magnitude are approximate comparisons: A 1.8-m-tall human and a 25-m-long blue whale differ by about one order of magnitude.

The usefulness of the logarithmic scale is that it devotes equal space to each order of magnitude.

EXAMPLE 3 Display the numbers 0.003, 0.1, 0.5, 6, 200, and 4000 on a logarithmic scale.

Solution To display the numbers, we need to take logarithms first:

x	0.003	0.1	0.5	6	200	4000
$\log x$	−2.5229	−1	−0.3010	0.7782	2.3010	3.6021

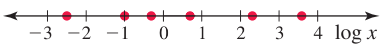


Figure 1.64 Example 3.

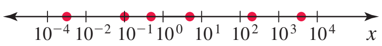


Figure 1.65 Example 3.

Since $\log 0.003 = -2.5229$, we find this number 2.5229 units to the left of 0 on the logarithmic scale. Similarly, since $\log 0.1 = -1$, this number is one unit to the left of 0, and $\log 200 = 2.3010$ is 2.3010 units to the right of 0 (Figure 1.64).

In the biological literature, x rather than $\log x$ is used to label logarithmic number lines. The locations of the numbers are the same; only the labeling changes. That is, 0.003 would be -2.5229 units to the left of the origin of the line (which is now at 10°, or 1). This way of drawing the line is shown in Figure 1.65. •

1.4.3 Transformations into Linear Functions

When you look through a biology textbook, you very likely find graphs like the ones in Figures 1.66 and 1.67. In either graph, you see a straight line (with data points scattered about it). In Figure 1.66, the vertical axis is logarithmically transformed and the horizontal axis is on a linear scale; in Figure 1.67, both axes are logarithmically transformed. Why do we display data like this, and what do these graphs mean?

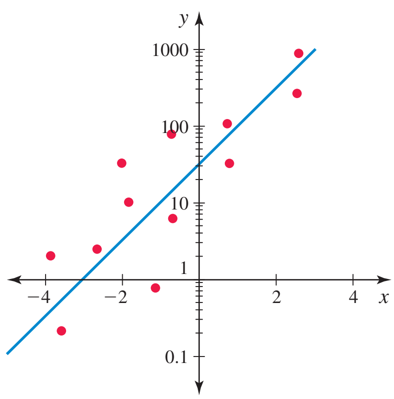


Figure 1.66 x and y data are linearly related when the y-axis is logarithmically transformed.

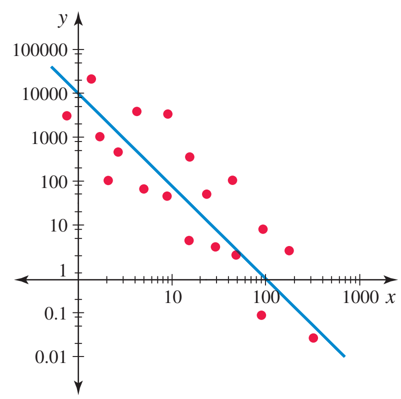


Figure 1.67 x and y data are linearly related when both x- and y-axis are logarithmically transformed.

The first question is quick to answer: Straight lines (or linear relationships) are easy to recognize visually. If transforming data results in data points lying along a straight line, we should do the transformation, because, as we will see, this will allow us to obtain a functional relationship between quantities. Now on to the second question: What do these graphs mean?

Exponential Functions. Let's look at Figure 1.66 and redraw just the straight line, using $\log y$ (instead of y) on the vertical axis and x on the horizontal axis (Figure 1.68). Set $Y = \log y$ and forget for a moment where the graph came from. We see a linear relationship between Y and x—a relationship of the form

$$Y = c + mx$$

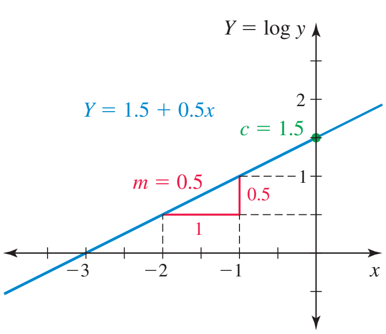


Figure 1.68 Figure 1.66 redrawn. Now the vertical axis is labeled $\log y$. We can read off the value of the slope (m) and y-intercept (c).

where c is the Y-intercept and m is the slope. From Figure 1.68 we measure: $c = 1.5$ and $m = 0.5$. That is, we have

$$Y = 1.5 + 0.5x$$

$$\log y = 1.5 + 0.5x \qquad Y = \log y$$

Exponentiating both sides, we find that

$$y = 10^{1.5+0.5x} = 10^{1.5}(10^{0.5})^x$$

Since $10^{1.5} \approx 31.62$ and $10^{0.5} \approx 3.162$, we can write the preceding equation as

$$y = (31.62)(3.162^x) \tag{1.11}$$

Since $y \propto a^x$ with $a = 3.162$, we can say that y is an exponential function of x.

A graph in which the vertical axis is on a logarithmic scale and the horizontal axis is on a linear scale is called a **log-linear plot** or a **semilog plot**.

If we display an exponential function of the form $y = ba^x$ on a semilog plot, a straight line results.

To see this, we take logarithms to base 10 on both sides of $y = ba^x$:

$$\log y = \log(ba^x) \tag{1.12}$$

Using the properties of logarithms, we simplify the right-hand side to

$$\log(ba^x) = \log b + \log a^x = \log b + x \log a$$

If we set $Y = \log y$, then (1.12) becomes

$$Y = \log b + (\log a)x \tag{1.13}$$

Comparing this equation with the general form of a linear function $Y = c + mx$, we see that the Y-intercept is $\log b$ and the slope is $\log a$. You do not need to memorize this statement, since you can always do the calculation that resulted in (1.13), but you should memorize the fact that an exponential function results in a straight line on a semilog plot. If $a > 1$, the slope of the line is positive; if $0 < a < 1$, the slope of the line is negative.

EXAMPLE 4 Graph $y = 2.5 \cdot 3^x$, $x \in \mathbf{R}$ on a semilog plot.

Solution We take logarithms first:

$$\begin{aligned} \log y &= \log(2.5 \cdot 3^x) \\ &= \log 2.5 + x \log 3 = 0.3979 + 0.4771x \end{aligned}$$

Use a calculator to evaluate the logarithms

The graph is shown in Figure 1.69. Note that the origin of the coordinate system is where $x = 0$ and $y = 1$ (or $\log y = 0$). The labeling on the vertical axis is for y, and we see that the labels are multiples of 10. The y intercept is at $y = 2.5$; to locate this point we use the fact that $\log 2.5 = 0.3979$ so 2.5 is 0.3979 units above the origin on the logarithmic scale. To complete the graph we need also the slope of the line: $m = 0.4771$. The line with these features is shown in Figure 1.69. •

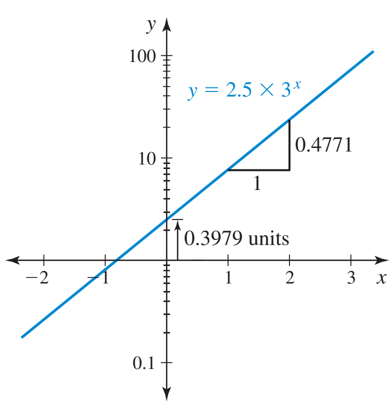


Figure 1.69 The graph of $y = 2.5 \times 3^x$ on a semilog plot.

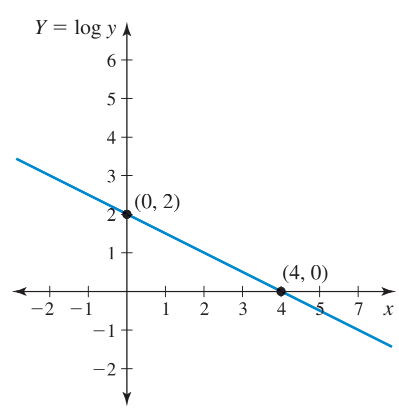


Figure 1.70 The graph for Example 5. The line goes through the points (0, 2) and (4, 0).

EXAMPLE 5 The variables x and y are related by as shown in Figure 1.70. Find y as a function of x.

Solution Figure 1.70 shows a semilog plot. We set $Y = \log y$. Then, in an x–Y graph, the Y-intercept is 2, and, using the two points (0, 2) and (4, 0), we find that the slope of the line is $m = (0-2)/(4-0) = -0.5$. Hence, we can use the slope intercept equation at a straight line:

$$Y = c + mx = 2 - 0.5x$$
$$\log y = 2 - 0.5x \qquad Y = \log y.$$
$$y = 10^{2-0.5x} = 10^2(10^{-0.5})^x = (100)(0.3162)^x \qquad \text{Take exponentials of both sides.}$$

Power Functions. Let's look back at Figure 1.67. There, both axes are logarithmically transformed. We redraw just the straight line, using $Y = \log y$ on the vertical axis and $X = \log x$ on the horizontal axis (Figure 1.71).

We measure the slope and y-intercept of the line. From Figure 1.71 we find that

$$c = 4 \qquad \text{and} \qquad m = -2 \qquad Y = c + mX \text{ where } m \text{ is slope and } c \text{ is } Y\text{-intercept}$$

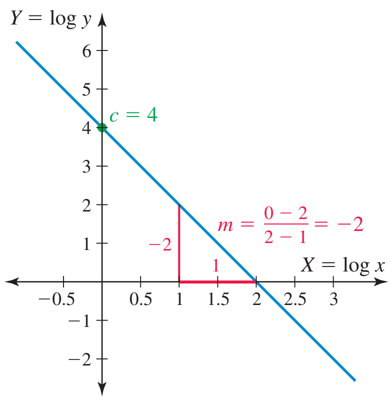


Figure 1.71 Figure 1.67 redrawn. Now the vertical axis is labeled $\log y$ and the horizontal axis is labeled $\log x$.

With $Y = \log y$ and $X = \log x$, the linear equation becomes

$$\log y = 4 - 2\log x$$

Exponentiating both sides, we get

$$y = 10^{4-2\log x} = 10^4(10^{\log x^{-2}}) = 10^4 x^{-2} \qquad 10^{\log a} = a$$

The function $y = 10^4 x^{-2}$ is a power function.

A graph in which both the vertical and the horizontal axis are logarithmically scaled is called a **log-log plot** or **double-log plot**.

> If a power function $y = bx^r$ is plotted in a double-log plot, a straight line results. The slope of the straight line is r, and the intercept is $\log b$.

To see this, we take logarithms to base 10 on both sides of $y = bx^r$:

$$\log y = \log(bx^r) \tag{1.14}$$

Using the properties of logarithms on the right-hand side of (1.14), we get

$$\log(bx^r) = \log b + \log x^r = \log b + r\log x$$

If we set $Y = \log y$ and $X = \log x$, then (1.14) becomes

$$Y = \log b + rX$$

Comparing this equation with the general form of a linear function, $Y = c + mX$, we see that the Y-intercept is $\log b$ and the slope is r. If $r > 0$, the slope is positive. If $r < 0$, the slope is negative.

EXAMPLE 6 Graph $y = 100x^{-2/3}$, $x > 0$, on a double-log plot.

Solution We take logarithms first:

$$\log y = \log(100x^{-2/3}) = \log 100 - \frac{2}{3}\log x = 2 - \frac{2}{3}\log x. \qquad \log(x^{-2/3}) = -\tfrac{2}{3}\log x \quad \log 100 = 2$$

If we set $Y = \log y$ and $X = \log x$. Then:

$$Y = 2 - \frac{2}{3}X$$

This is the equation of a straight line with X-intercept 3 and Y-intercept 2 (and thus slope $-2/3$). We graph this function in Figure 1.72, where we have X and Y on the two axes. If we use x and y on the two axes (Figure 1.73), the labels change: The y-intercept is now 100 (corresponding to $\log 100 = 2$) and the x-intercept is 1000 (corresponding to $\log 1000 = 3$). Note that the origin in Figure 1.72 is $X = 0$ and $Y = 0$; the origin in Figure 1.73 is $x = 1$ and $y = 1$. •

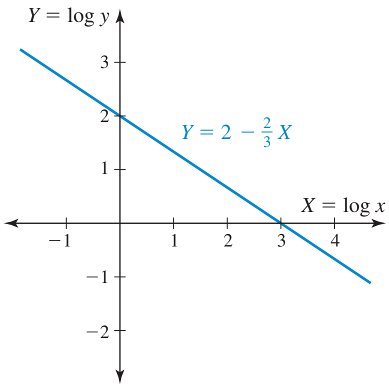


Figure 1.72 The graph $y = 100x^{-2/3}$ on a double-log plot where the axes are labeled $X = \log x$ and $Y = \log y$.

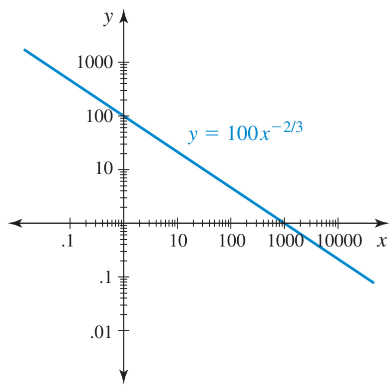


Figure 1.73 The graph $y = 100x^{-2/3}$ on a double-log plot where the axes are labeled x and y.

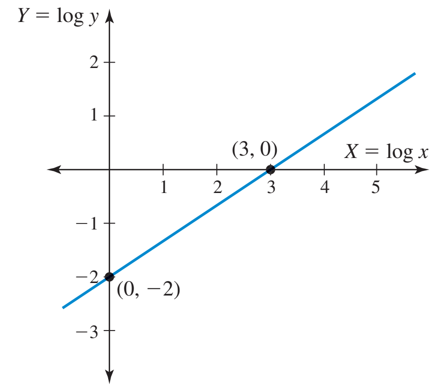


Figure 1.74 The graph of the function for Example 7.

EXAMPLE 7 Find the functional relationship between x and y on the basis of the graph in Figure 1.74.

Solution If $Y = \log y$ and $X = \log x$, then, in an X–Y graph, the Y-intercept is -2 and, using the two points $(0, -2)$ and $(3, 0)$, we calculate the slope as

$$m = \frac{0 - (-2)}{3 - 0} = \frac{2}{3}$$

Hence the equation is

$$Y = -2 + \frac{2}{3}X$$

With $Y = \log y$ and $X = \log x$, we find that

$$\log y = -2 + \frac{2}{3}\log x$$

and, after exponentiating both sides of this equation, we get

$$y = 10^{-2+\frac{2}{3}\log x} = 10^{-2}10^{\log x^{2/3}} = (0.01)x^{2/3} \qquad \frac{2}{3}\log x = \log x^{2/3}$$

Thus, the functional relationship between x and y is a power function of the form

$$y = (0.01)x^{2/3}$$ •

Applications

EXAMPLE 8 **Self-Thinning of Plants** When growing plants at sufficiently high initial densities, we often observe as the plants grow the number of plants decreases. This property is called *self-thinning*.

When the average weight of a plant is plotted on a log-log plot as a function of the density of survivors, we frequently find that the data lie along a straight line with slope $-3/2$. Assume that, for a particular plant, such a relationship holds for plant densities between 10^2 and 10^4 plants per square meter and that, at a density of 100 plants per square meter, the average weight per plant is about 10 grams. Find the functional relationship between dry weight and plant density, and graph this function on a log-log plot.

Solution Since the relationship between density (x) and dry weight (y) follows a straight line with slope $-3/2$ on a log-log plot, we set

$$\log y = C - \frac{3}{2} \log x \quad \text{for } 10^2 \leq x \leq 10^4$$

where C is a constant. To find C, we use the fact that when $x = 100$, $y = 10$. Therefore,

$$\log 10 = C - \frac{3}{2} \log 100$$

$$\Rightarrow 1 = C - \frac{3}{2} \times 2 = C - 3 \text{ which implies that} \qquad C = 4$$

Hence,

$$\log y = 4 - \frac{3}{2} \log x$$

Exponentiating both sides, we find that

$$y = 10^4 x^{-3/2} \quad \text{for } 10^2 \leq x \leq 10^4$$

The graph of this function on a log-log scale is shown in Figure 1.75. •

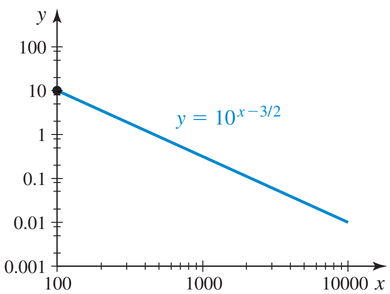


Figure 1.75 The graph of the function for Example 8 showing plant weight (y) against density (x). The line has slope $-3/2$ and goes through the point $(x, y) = (100, 10)$.

EXAMPLE 9 Polonium 210 (Po^{210}) is a radioactive material. To determine the half-life of Po^{210} experimentally, we measure the amount of radioactive material left after time t for various values of t. When we plot the data on a semilog plot, we find that we can fit a straight line to the curve. The slope of the straight line is -0.0022/day. Find the half-life of Po^{210}.

Solution Radioactive decay follows the equation

$$W(t) = W(0)e^{-\lambda t} \quad \text{for } t \geq 0$$

where $W(t)$ is the amount of radioactive material left after time t. If we take logs, we obtain:

$$\log W(t) = \log W(0) - \lambda t \log e$$

Logarithms are all base 10 for a semilog plot.

If we plot $W(t)$ as a function of t on a semilog plot, we obtain a straight line with slope $-\lambda \log e$. Matching this slope with the number given in the example, we obtain

$$\lambda \log e = 0.0022/\text{day}$$

Solving for λ yields

$$\lambda = \frac{1}{\log e} 0.0022/\text{day}$$

To find the half-life T_h, we use the formula (see Subsection 1.3.5)

$$T_h = \frac{\ln 2}{\lambda} = \frac{\ln 2}{0.0022} (\log e) \text{ days}$$
$$\approx 136.8 \text{ days}$$

•

EXAMPLE 10 Light intensity in lakes decreases with depth. Denote by $I(z)$ the light intensity at depth z, with $z = 0$ representing the surface. Then the percentage surface radiation at that reaches depth z, denoted by PSR(z), is computed as

$$\text{PSR}(z) = 100 \frac{I(z)}{I(0)}$$

PSR(0) = 100

When we graph the percentage surface radiation as a function of depth on a semilog plot, a straight line results. An example of such a curve is given in Figure 1.76, where the coordinate system is rotated clockwise by 90° so that the depth axis points downward. Derive an equation for $I(z)$ on the basis of the graph.

Solution We see that the dependent variable, $100I(z)/I(0)$, is logarithmically transformed, whereas the independent variable, z, is on a linear scale. The graph is a straight line. We thus find that

$$\log \text{PSR} = \log\left(100\frac{I(z)}{I(0)}\right) = c + mz$$

where c is the intercept on the percentage surface radiation axis and m is the slope. We see that

$$c = \log 100$$

and, using the points (0, 100) and (30, 1), we get

$$m = \frac{\log 100 - \log 1}{0 - 30} = -\frac{2}{30} = -\frac{1}{15}$$

Hence,

$$\log\left(100\frac{I(z)}{I(0)}\right) = \log 100 - \frac{1}{15}z$$

The left-hand side simplifies to $\log 100 + \log \frac{I(z)}{I(0)}$. After canceling $\log 100$ on both sides and exponentiating both sides, we find that

$$\frac{I(z)}{I(0)} = 10^{-(1/15)z} = \exp[\ln 10^{-(1/15)z}]$$

Thus

$$I(z) = I(0)e^{-(\frac{1}{15}\ln 10)z}$$

The number $\frac{1}{15}\ln 10$ is called the *vertical attenuation coefficient*. The magnitude of this number tells us how quickly light is absorbed in a lake. •

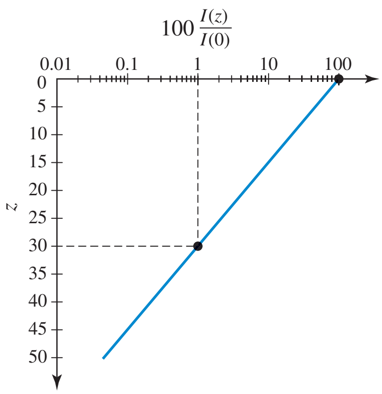


Figure 1.76 The graph of percentage surface radiation as a function of depth.

1.4.4 From a Verbal Description to a Graph

Being able to sketch a graph on the basis of a verbal explanation of some phenomenon is an extremely useful skill since a graph can summarize a complex situation that can be more easily communicated and remembered. Let's look at an example.

EXAMPLE 11 **Ecosystem Diversity** According to ***Rosenzweig and Abramsky (1993):***

> The relationship of primary productivity and species diversity is not simple. But within such regions, and perhaps even larger ones, a pattern is emerging: as productivity rises, first diversity increases, then it declines.

Bio Info • The primary productivity of an ecosystem is the total rate at which all plants in the ecosystem convert light into chemical energy. Species diversity is a measure of the number of different species present and in what proportions (see Example 4 in Section 1.3 for one method of quantifying diversity)

If we wanted to translate this verbal description into a graph, we would first determine the independent and the dependent variable. Here, we consider species diversity as a function of primary productivity; hence, primary productivity is the independent variable and species diversity is the dependent variable. We will therefore use a coordinate system whose horizontal axis denotes primary productivity and whose vertical axis denotes species diversity. Since both primary productivity and species diversity are nonnegative, we need to draw only the first quadrant (Figure 1.77).

Going back to the quote, we see that as productivity increases, diversity first increases, then decreases. The graph in Figure 1.78 illustrates this behavior.

Note that the exact shape of the curve cannot be inferred from the quote and will depend on the system studied. The graphs in Figure 1.79 all have features that agree with the description by Rosenzweig and Abramsky. The shape of the curves in Figure 1.79 is quantitatively different from the graph in Figure 1.78, but both have the same qualitative features of an initial increase followed by a decrease.

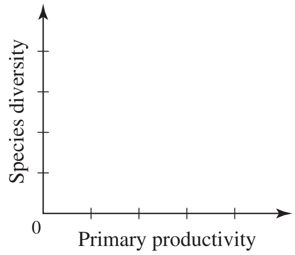


Figure 1.77 The coordinate system for species diversity as a function of primary productivity can be restricted to the first quadrant.

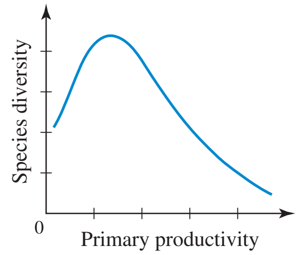


Figure 1.78 The graph of species diversity as a function of primary productivity is hump shaped.

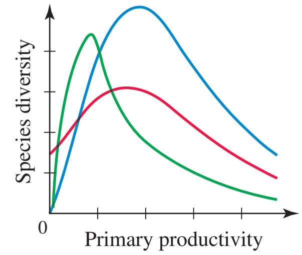


Figure 1.79 Different graphs capturing the same qualitative relationship as Figure 1.78.

As another example, we will look at the *functional response* of a predator to its prey density. •

EXAMPLE 12

Predator Response Curves The functional response of a predator to its prey density relates the number of prey consumed per predator (the dependent variable) to the prey density (the independent variable). ***Holling (1959)*** introduced three basic types of response. Type 1 describes a response in which the number of prey eaten per predator as a function of prey density rises linearly to a plateau. The type 2 functional response increases at a decelerating rate and eventually levels off. The type 3 functional response is S shaped, or sigmoidal, and also eventually levels off.

Now let's translate these three ways into graphs. All graphs will be plotted with prey density on the horizontal axis and the number of prey eaten per predator on the vertical axis. Since both prey density and number of prey eaten per predator are nonnegative variables, we need to draw only the first quadrant.

Even though this was not mentioned, we will assume that when the prey density is equal to zero, the number of prey eaten per predator will also be zero. This means that the three functional response curves all go through the origin.

The type 1 functional response first increases linearly (i.e., results in a straight line) and then reaches a plateau (stays constant) (See Figure 1.80.)

The type 2 functional response is described as a function that increases at a decelerating rate. This means that the function will increase less quickly as prey density increases (Figure 1.81). In contrast to the type 1 functional response, the type 2 response will continue to increase and approach, but not actually reach, the plateau at a finite value of prey density.

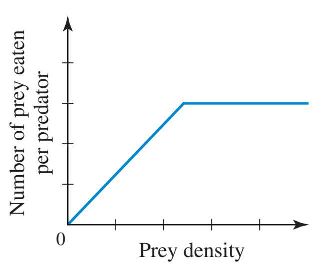


Figure 1.80 The type 1 functional response.

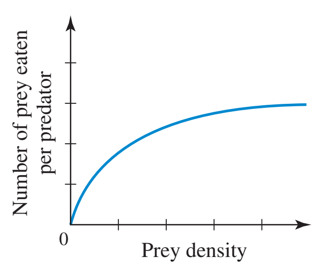


Figure 1.81 The type 2 functional response.

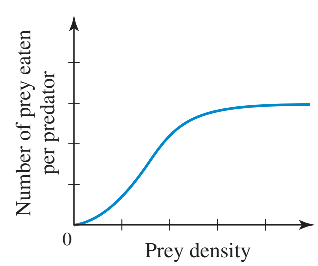


Figure 1.82 The type 3 functional response.

The type 3 functional response is described as sigmoidal. Sigmoidal curves are characterized by an initial accelerating increase followed by an increase at a decelerating rate (Figure 1.82). Similar to the type 2 functional response, the type 3 functional response approaches a plateau as prey density increases, and will not reach the plateau at a finite value of prey density. •

For each example discussed so far, the functional relationship depended on just one variable, such as the number of prey eaten per predator as a function of prey density. Often, however, a response depends on more than one independent variable. The next example presents a response that depends on two independent variables, and shows how to draw a graph of this more complex relationship.

EXAMPLE 13 **Seed Germination Success** The successful germination of seeds depends on both temperature and humidity. When the humidity is too low, seeds tend not to germinate at all, regardless of the temperature. Germination success is highest for intermediate values of temperature. Finally, seeds tend to germinate better when humidity levels are higher.

One way to translate this information into a graph is to graph germination success as a function of temperature for different levels of humidity. (In Chapter 10 we will meet other methods for graphing functions with more than one dependent variable.) If we measure temperature in Fahrenheit or Celsius, we can restrict the graphs to the first quadrant (Figure 1.83), since the temperature needs to be well above freezing for germination to occur (the temperature at which freezing starts is 0°C, or 32°F). Germination success will be between 0 and 100%. To sketch the graphs, it is better not to label the axes beyond what we provided in Figure 1.83, because we do not know the exact numerical response.

There is enough information to provide three graphs: one for low humidity, one for intermediate humidity, and one for high humidity. We will graph them all in one coordinate system, so that it is easier to compare the different responses. The graph for low humidity is a horizontal line where germination success is 0%. For intermediate and high humidity, the graphs are hump shaped, since germination success is highest for intermediate values of temperature. The graph for high humidity is above the graph for intermediate humidity, because seeds tend to germinate better when humidity levels are higher (Figure 1.84). •

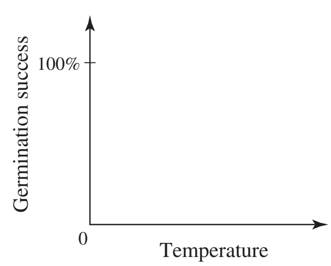


Figure 1.83 The coordinate system for germination success as a function of temperature can be restricted to the first quadrant. Germination success will be between 0 and 100% temperature will always be above 0.

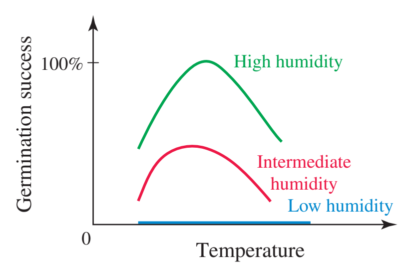


Figure 1.84 Germination success as a function of temperature for three humidity levels (low, intermediate, high).

Section 1.4 Problems

1.4.1

In Problems 1–22, sketch the graph of each function. Do not use a graphing calculator. (Assume the largest possible domain.)

1. $y = x^2 + 1$

2. $y = -(x-2)^2 + 1$

3. $y = x^3 - 2$

4. $y = (x+1)^3$

5. $y = -2x^2 - 3$

6. $y = -(2-x)^2 + 2$

7. $y = 3 + 1/x$

8. $y = \dfrac{x+1}{x}$

9. $y = \dfrac{x}{x+1}$

10. $y = 1 + \dfrac{1}{(x+2)^2}$

11. $y = \exp(x-2)$

12. $y = -\exp(x)$

13. $y = e^{-(x+3)}$

14. $y = 2e^{x-1}$

15. $y = \ln(x+1)$

16. $y = \ln(x-3)$

17. $y = -\ln(x-1) + 1$

18. $y = -\ln(3-x)$

19. $y = 2\sin(x + \pi/4)$

20. $y = 2\sin(-x)$

21. $y = \cos(\pi x)$

22. $y = -2\cos(\pi x/4)$

23. Explain how the following functions can be obtained from $y = x^2$ by basic transformations:

(a) $y = x^2 - 2$ **(b)** $y = (x-1)^2 + 1$ **(c)** $y = -2(x+2)^2$

24. Explain how the following functions can be obtained from $y = x^3$ by basic transformations:

(a) $y = x^3 - 1$ **(b)** $y = -x^3 - 1$ **(c)** $y = -3(x-1)^3$

25. Explain how the following functions can be obtained from $y = 1/x$ by basic transformations:

(a) $y = 1 - \frac{1}{x}$ **(b)** $y = -\frac{1}{x-1}$ **(c)** $y = \frac{x}{x+1}$

26. Explain how the following functions can be obtained from $y = 1/x^2$ by basic transformations:

(a) $y = \frac{1}{x^2} + 1$ **(b)** $y = -\frac{1}{(x+1)^2}$ **(c)** $y = -\frac{1}{x^2} - 2$

27. Explain how the following functions can be obtained from $y = e^x$ by basic transformations:

(a) $y = e^x + 3$ **(b)** $y = e^{-x}$ **(c)** $y = 2e^{x-2} + 3$

28. Explain how the following functions can be obtained from $y = e^x$ by basic transformations:

(a) $y = e^{-x} - 1$ **(b)** $y = -e^x + 1$ **(c)** $y = -e^{x-3} - 2$

29. Explain how the following functions can be obtained from $y = \ln x$ by basic transformations:

(a) $y = \ln(x-1)$ **(b)** $y = -\ln x + 1$ **(c)** $y = \ln(x+3) - 1$

30. Explain how the following functions can be obtained from $y = \ln x$ by basic transformations:

(a) $y = \ln(1-x)$ **(b)** $y = \ln(2+x) - 1$

(c) $y = -\ln(2-x) + 1$

31. Explain how the following functions can be obtained from $y = \sin x$ by basic transformations:

(a) $\sin(\pi x)$ **(b)** $\sin\left(x + \frac{\pi}{4}\right)$

(c) $-2\sin(\pi x + 1)$

32. Explain how the following functions can be obtained from $y = \cos x$ by basic transformations:

(a) $y = 1 + 2\cos x$ **(b)** $y = -\cos\left(x + \frac{\pi}{4}\right)$

(c) $y = -\cos\left(\frac{\pi}{2} - x\right)$

1.4.2

33. Find the following numbers on a number line that is on a logarithmic scale (base 10): 0.003, 0.03, 3, 5, 30, 50, 1000, 3000, and 30000.

34. Find the following numbers on a number line that is on a logarithmic scale (base 10): 0.03, 0.7, 1, 2, 5, 10, 17, 100, 150, and 2000.

35. (a) Find the following numbers on a number line that is on a logarithmic scale (base 10): 10^2, 10^{-3}, 10^{-4}, 10^{-7}, and 10^{-10}.

(b) Can you find 0 on a number line that is on a logarithmic scale?

(c) Can you find negative numbers on a number line that is on a logarithmic scale?

36. (a) Find the following numbers on a number line that is on a logarithmic scale (base 10):

(i) $10^{-3}, 2 \times 10^{-3}, 3 \times 10^{-3}$ **(ii)** $10^{-1}, 2 \times 10^{-1}, 3 \times 10^{-1}$

(iii) $10^2, 2 \times 10^2, 3 \times 10^2$

(b) From your answers to (a), how many units (on a logarithmic scale) is **(i)** 2×10^{-3} from 10^{-3} **(ii)** 2×10^{-1} from 10^{-1} and **(iii)** 2×10^2 from 10^2?

(c) From your answers to (a), how many units (on a logarithmic scale) is **(i)** 3×10^{-3} from 10^{-3} **(ii)** 3×10^{-1} from 10^{-1} and **(iii)** 3×10^2 from 10^2?

In Problems 37–42, insert an appropriate number in the blank space.

37. The longest known species of worm is the earthworm *Microchaetus rappi* of South Africa; in 1937, a 6.7-m-long specimen was collected from the Transvaal. The shortest worm is *Chaetogaster annandalei*, which measures less than 0.51 mm in length. *M. rappi* is ____ order(s) of magnitude longer than *C. annandalei*.

38. Both the La Plata river dolphin (*Pontoporia blainvillei*) and the sperm whale (*Physeter macrocephalus*) belong to the suborder Odontoceti (individuals that have teeth). A La Plata river dolphin weighs between 30 and 50 kg, whereas a sperm whale weighs between 35,000 and 40,000 kg. A sperm whale is ____ order(s) of magnitude heavier than a La Plata river dolphin.

39. The Etruscan shrew (*Suncus etruscus*) is by weight the smallest mammal in the world, weighing 1.8 g on average. By comparison, the largest mammal in the world, the blue whale (*Balaenoptera musculus*), may weigh as much as 190 tonnes (190×10^3 kg). A blue whale is ____ order(s) of magnitude heavier than an Etruscan shrew.

40. Compare a square with side length 1 m against a square with side length 100 m. The area of the larger square is ____ order(s) of magnitude larger than the area of the smaller square.

41. The diameter of a typical bacterium is about 0.5 to 1 μm. An exception is the bacterium *Epulopiscium fishelsoni*, which is about 600 μm long and 80 μm wide. The volume of *E. fishelsoni* is about ____ order(s) of magnitude larger than that of a typical bacterium. (*Hint*: Approximate the shape of a typical bacterium by a sphere and the shape of *E. fishelsoni* by a cylinder.)

42. The length of a typical bacterial cell is about one-tenth that of a small eukaryotic cell. Consequently, the cell volume of a bacterium is about ____ order(s) of magnitude smaller than that of a small eukaryotic cell. (*Hint*: Approximate the shapes of both types of cells by spheres.)

1.4.3

In Problems 43–46, when* $\log y$ *is graphed as a function of* x*, a straight line results. Graph straight lines, each given by two points, on a log-linear plot, and determine the functional relationship. (The original* x–y *coordinates are given.)

43. $(x_1, y_1) = (0, 5)$, $(x_2, y_2) = (3, 1)$

44. $(x_1, y_1) = (-1, 4)$, $(x_2, y_2) = (2, 8)$

45. $(x_1, y_1) = (-1, 1)$, $(x_2, y_2) = (1, 3)$

46. $(x_1, y_1) = (1, 4)$, $(x_2, y_2) = (4, 1)$

In Problems 47–54, use a logarithmic transformation to find a linear relationship between the given quantities and graph the resulting linear relationship on a log-linear plot.

47. $y = 3 \times 10^{-2x}$ **48.** $y = 10^{1.5x}$

49. $y = 2e^{-1.2x}$ **50.** $y = 1.5e^{2x}$

51. $y = 5 \times 2^{4x}$ **52.** $y = 3 \times 5^{-1.3x}$

53. $y = 4 \times 3^{2x}$ **54.** $y = 2^{x+1}$

In Problems 55–58, when* log *y is graphed as a function of* log *x, a straight line results. Graph straight lines, each given by two points, on a log-log plot, and determine the functional relationship. (The original x–y coordinates are given.)

55. $(x_1, y_1) = (1, 2), (x_2, y_2) = (5, 1)$

56. $(x_1, y_1) = (2, 7), (x_2, y_2) = (7, 2)$

57. $(x_1, y_1) = (4, 2), (x_2, y_2) = (8, 8)$

58. $(x_1, y_1) = (3, 4), (x_2, y_2) = (1, 3)$

In Problems 59–66, use a logarithmic transformation to find a linear relationship between the given quantities and graph the resulting linear relationship on a log-log plot.

59. $y = 2x^5$ **60.** $y = 3x^2$

61. $y = x^6$ **62.** $y = 2x^9$

63. $y = 2x^{-2}$ **64.** $y = 6x^{-1}$

65. $y = 4x^{-3}$ **66.** $y = 9x^{-3}$

In Problems 67–72, use a logarithmic transformation to find a linear relationship between the given quantities and determine whether a log-log or log-linear plot should be used to graph the resulting linear relationship.

67. $f(x) = 3x^{1.7}$ **68.** $g(s) = 9e^{-1.65}$

69. $N(t) = 130 \times 2^{1.2t}$ **70.** $I(u) = 4.8u^{-0.89}$

71. $R(t) = 3.6t^{1.2}$ **72.** $L(t) = 2^{-2.7y+1}$

73. The following table is based on a functional relationship between x and y that is either an exponential or a power function:

x	y
1	1.8
2	2.07
4	2.38
10	2.85
20	3.28

Use an appropriate logarithmic transformation and a graph to decide whether the table comes from a power function or an exponential function, and find the functional relationship between x and y.

74. The following table is based on a functional relationship between x and y that is either an exponential or a power function:

x	y
0.5	1.21
1	0.74
1.5	0.45
2	0.27
2.5	0.16

Use an appropriate logarithmic transformation and a graph to decide whether the table comes from a power function or an exponential function, and find the functional relationship between x and y.

75. The following table is based on a functional relationship between x and y that is either an exponential or a power function:

x	y
−1	0.398
−0.5	1.26
0	4
0.5	12.68
1	40.18

Use an appropriate logarithmic transformation and a graph to decide whether the table comes from a power function or an exponential function, and find the functional relationship between x and y.

76. The following table is based on a functional relationship between x and y that is either an exponential or a power function:

x	y
0	0
0.5	0.34
1	2.70
1.5	9.11
2	21.60

Use an appropriate logarithmic transformation and a graph to decide whether the table comes from a power function or an exponential function, and find the functional relationship between x and y.

77. The following table is based on a functional relationship between x and y that is either an exponential or a power function:

x	y
0.1	0.045
0.5	1.33
1	5.7
1.5	13.36
2	24.44

Use an appropriate logarithmic transformation and a graph to decide whether the table comes from a power function or an exponential function, and find the functional relationship between x and y.

78. The following table is based on a functional relationship between x and y that is either an exponential or a power function:

x	y
0.1	0.067
0.5	0.22
1	1.00
1.5	4.48
2	20.09

Use an appropriate logarithmic transformation and a graph to decide whether the table comes from a power function or an exponential function, and find the functional relationship between x and y.

So far, we have always used base 10 for a logarithmic transformation. The reason for this is that our number system is based on base 10 and it is therefore easy to logarithmically transform numbers of the form . . . , 0.01, 0.1, 1, 10, 100, 1000, . . . when we use base 10. In Problems 79–82, use the indicated base to logarithmically transform each exponential relationship so that a linear relationship results. Then use the indicated base to graph each relationship either in log or semilog transformed coordinates so that a straight line results.

79. $y = 2^x$; base 2 **80.** $y = 5^x$; base 5

81. $y = 2^{-x}$; base 2 **82.** $y = 3e^{-2x}$; base 3

83. Suppose that $N(t)$ denotes a population size at time t and satisfies the equation

$$N(t) = 2e^{3t} \quad \text{for } t \geq 0$$

(a) If you graph $N(t)$ as a function of t on a semilog plot, a straight line results. Explain why.

(b) Graph $N(t)$ as a function of t on a semilog plot, and determine the slope of the resulting straight line.

84. Suppose that you follow a population over time. When you plot your data on a semilog plot, a straight line with slope 0.03 results. Furthermore, assume that the population size at time 0 was 20. If $N(t)$ denotes the population size at time t, what function best describes the population size at time t?

85. Species–Area Curves Many studies have shown that the number of species on an island increases with the area of the island. Frequently, the functional relationship between the number of species, S, and the area, A, is approximated by $S = CA^z$, where z is a constant that depends on the particular species and habitat in the study. (Actual values of z range from about 0.2 to 0.35.) Suppose that the best fit to your data points on a log-log scale is a straight line. Is your model $S = CA^z$ an appropriate description of your data? If yes, how would you find z?

86. Michaelis–Menten Equation Enzymes serve as catalysts in many chemical reactions in living systems. The simplest such reactions transform a single substrate into a product with the help of an enzyme. The Michaelis–Menten equation describes the rate of such enzymatically controlled reactions. The equation, which gives the relationship between the initial rate of the reaction, v, and the concentration of the substrate, s, is

$$v(s) = \frac{v_{\max} s}{s + K}$$

where $v_{\max}$ is the maximum rate at which the product may be formed and K is called the Michaelis–Menten constant. Note that this equation has the same form as the Monod growth function. Given some data on the reaction rate v, for different substrate concentrations s, we would like to infer the parameters K and $v_{\max}$.

(a) The graph of v against s is nonlinear, so it is hard to determine K and $v_{\max}$ directly from a graph of the function $v(s)$. In the remaining parts of this question you will be guided to transform your plot into one in which the dependent variable depends linearly on the independent variable. First plot, using a graphing calculator, or by hand, $v(s)$ for the following values of K and $v_{\max}$:

$$(K, v_{\max}) = (1, 1), \quad (2, 1), \quad (1, 2)$$

(b) Show that the Michaelis–Menten equation can be written in the form

$$\frac{1}{v} = \frac{K}{v_{\max}} \frac{1}{s} + \frac{1}{v_{\max}}$$

This formula is known as the Lineweaver–Burk equation and shows that there is a linear relationship between $1/v$ and $1/s$.

(c) Sketch the graph of the Lineweaver–Burk equation. Use a coordinate system in which $1/s$ is on the horizontal axis and $1/v$ is on the vertical axis. Show that the resulting graph is a line that intersects the horizontal axis at $-1/K$ and the vertical axis at $1/v_{\max}$.

(d) Suppose you measured the following data for s and v.

s	v
1	0.5
5	1.5
10	2.0

Make a plot with $1/s$ on the horizontal axis and $1/v$ on the vertical axis, and estimate the parameters K and $v_{\max}$ for this data.

87. *(Continuation of Problem 86)* Estimating $v_{\max}$ and K from the Lineweaver–Burk graph as described in Problem 86 is not always satisfactory. A different transformation typically yields better estimates (see ***Dowd and Riggs, 1965***).

(a) Show that the Michaelis–Menten equation can be written as

$$\frac{v}{s} = \frac{v_{\max}}{K} - \frac{1}{K} v$$

and explain why this transformation results in a straight line when you graph v on the horizontal axis and $\frac{v}{s}$ on the vertical axis.

(b) Explain how you can estimate $v_{\max}$ and K from the graph.

(c) You measure the following data for s and v.

s	v
1	0.182
10	1.000
20	1.333

By making a plot with v on the horizontal axis and v/s on the vertical axis, estimate the parameters K and $v_{\max}$ from this data.

88. Oxygen Consumption In a case study in which the maximal rates of oxygen consumption (in ml/s) of nine species of wild African mammals were plotted against body mass (in kg) on a log-log plot, it was found that the data points fell on a straight line with slope approximately equal to 0.8 and vertical-axis intercept approximately equal to 0.105. Find an equation that relates maximal oxygen consumption and body mass. (Adapted from ***Reiss, 1989***).

89. Body Proportions In vertebrates, embryos and juveniles have large heads relative to their overall body size. As the animal grows older, proportions change; for instance, the ratio of skull length to body length diminishes. That this is the case not only for living vertebrates, but also for fossil vertebrates, is shown by the following example: (Adapted from ***Benton and Harper, 1997***).

Ichthyosaurs are a group of marine reptiles that appeared at the same time as the early dinosaurs They were fish shaped and comparable in size to dolphins. In a study of 20 fossil skeletons, the following allometric relationship between skull length, S (measured in cm) and backbone length, B (measured in cm) was found:

$$S = 1.162 B^{0.93}$$

(a) Choose suitable transformations of S and B so that the resulting relationship is linear. Plot the transformed relationship, and find the slope and the y-intercept.

(b) Explain why the scaling equation confirms that juveniles had relatively large heads. (*Hint*: Compute the ratio of S to B for a number of different values of B—say, 10 cm, 100 cm, 500 cm—and compare.)

90. Bite Strength The bite strength of finches increases as body mass increases. ***Van de May and Bout (2004)*** made the following measurements of bird mass (B, measured in g) and wind beat frequency (f, measured in Hz):

Species	Body Mass (B, Measured in g)	Bite Strength (S, Measured in N)
Java sparrow	30.4	9.6
Red-billed firefinch	6.9	1.2
Double barred finch	9.7	1.9

Assume that there is a power-law dependence of S upon B:

$$S = aB^c,$$

where a and c are some constants. By plotting log S against log B, estimate the parameters a and c.

91. Hummingbird Flight Hummingbird wing-beat frequency decreases as bird mass increases. ***Altshuler et al. (2010)*** made the following measurements of bird size (measured in mass, B, in g) and wind beat frequency (frequency, f, in Hz).

Species	Body Mass (B, Measured in g)	Wing Beat Frequency (f, Measured in Hz)
Giant hummingbird	22.025	14.99
Volcano hummingbird	2.708	43.31
Blue-mantled thornbill	6.000	29.27

Assume that there is a power-law dependence of f upon B:

$$f = bB^a,$$

for some constants a and b. By plotting $\log f$ against $\log B$, estimate the parameters a and b.

92. Animal Urination The urination speed of animals increases with the body mass of the mammal. ***Yang et al. (2014)*** made the following measurements of urination speed (u, measured in ml/s) against animal body mass (M, measured in kg):

Animal	M (kg)	u (ml/s)
Cow	635	450
Cat	3.5	2.40
Dog	18.14	17.07

It is believed that there is a power-law relationship between urination speed and body mass, namely

$$u = bM^a$$

for some constants a and b.

(a) By plotting log u against log M, estimate the parameters a and b.

(b) Estimate the urination speed u for a human adult (you can assume that $M = 80$ kg).

93. Drug Absorption After a patient takes the painkiller acetaminophen (often sold under the brand name Tylenol), the concentration of drug in their blood increases at first, as the painkiller is absorbed into the blood, and then starts to decrease as the drug is metabolized or removed by the liver. In one study, the concentration of drug (c, measured in μg/ml) was measured in a patient as a function of time (t, measured in hours since the drug was administered). The data in this example is taken from ***Rowling et al. (1977)***.

t	c
1	10.61
1.5	8.73
2	7.63
3	5.55
4	3.97
5	3.01
6	2.39

(a) You want to determine from the data whether the relationship between concentration and time follows a power law

$$c = at^b$$

for some set of constants a and b, or whether it instead follows an exponential law

$$c = kd^t$$

for some constants k and d.
Explain how you could plot the data with transformed horizontal and vertical axes to determine which mathematical model is correct.

(b) By plotting $\log c$ against $\log t$ in one graph, and $\log c$ against t in another, explain why the data supports the second model (exponential decay) better than it supports the first model.

(c) From your plot of $\log c$ against t, estimate the parameter d.

94. Light intensity in lakes decreases exponentially with depth. If $I(z)$ denotes the light intensity at depth z, with $z = 0$ representing the surface, then

$$I(z) = I(0)e^{-\alpha z}, \quad z \geq 0$$

where α is a positive constant called the *vertical attenuation coefficient*. Figure 1.85 shows the percentage surface radiation, defined as $100I(z)/I(0)$, as a function of depth in different lakes.

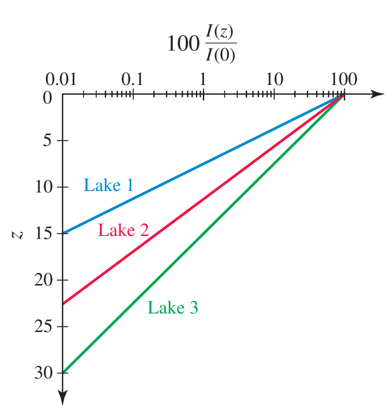


Figure 1.85 Light intensity as a function of depth for Problem 94.

(a) On the basis of the graph, estimate α for each lake.

(b) Reproduce a graph like the one in Figure 1.85 for Lake Constance (Germany) in May ($\alpha = 0.768\,\mathrm{m}^{-1}$) and December ($\alpha = 0.219\,\mathrm{m}^{-1}$) (data from ***Tilzer et al., 1982***).

(c) Explain why the graphs are straight lines.

95. Algal Growth in Lakes The absorption of light in a uniform water column follows an exponential law; that is, the intensity $I(z)$ at depth z is

$$I(z) = I(0)e^{-\alpha z}$$

where $I(0)$ is the intensity at the surface (i.e., when $z = 0$) and α is the *vertical attenuation coefficient*. (We assume here that α is constant. In reality, α depends on the wavelength of the light penetrating the surface.)

(a) Suppose that 10% of the light is absorbed in the uppermost meter. Find α. What are the units of α?

(b) What percentage of the remaining intensity at 1 m is absorbed in the second meter? What percentage of the remaining intensity at 2 m is absorbed in the third meter?

(c) What percentage of the initial intensity remains at 1 m, at 2 m, and at 3 m?

(d) Plot the light intensity as a percentage of the surface intensity on both a linear plot and a log-linear plot.

(e) Relate the slope of the curve on the log-linear plot to the attenuation coefficient α.

(f) The level at which 1% of the surface intensity remains is of biological significance. Approximately, it is the level where algal growth ceases. The zone above this level is called the *euphotic zone*. Express the depth of the euphotic zone as a function of α.

(g) Compare a very clear lake with a polluted river. Is the attenuation coefficient α for the clear lake greater or smaller than the attenuation coefficient α for the polluted river?

96. This problem relates to self-thinning of plants (Example 8). If we plot the logarithm of the over age weight of a plant, $\log w$, against the logarithm of the density of plants, $\log d$ (base 10), a straight line with slope $-3/2$ results. Find the equation that relates w and d, assuming that $w = 1$ g when $d = 10^3\,\mathrm{m}^{-2}$.

In Problems 97–102, find each functional relationship on the basis of the given graph.

97. Figure 1.86

98. Figure 1.87

99. Figure 1.88

100. Figure 1.89

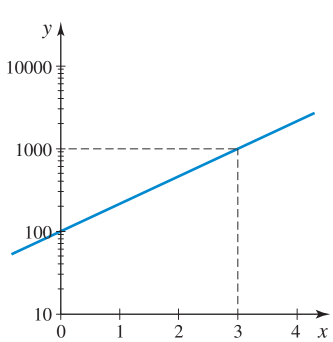


Figure 1.86 Graph for Problem 97.

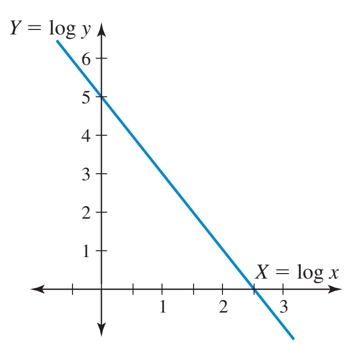


Figure 1.87 Graph for Problem 98.

101. Figure 1.90 (*Hint*: This relationship is different from the ones considered so far. The x-axis is logarithmically transformed, but the y-axis is linear.)

102. Figure 1.91 (*Hint*: This relationship is different from the ones considered so far. The x-axis is logarithmically transformed, but the y-axis is linear.)

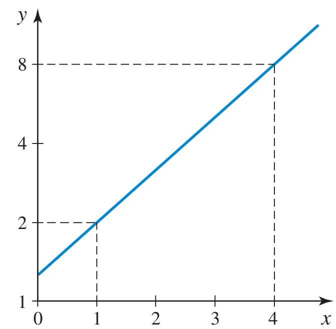


Figure 1.88 Graph for Problem 99.

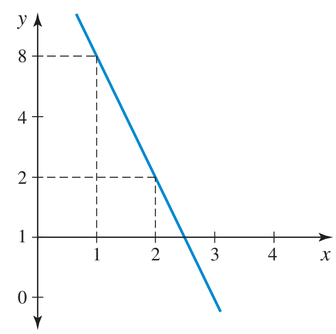


Figure 1.89 Graph for Problem 100.

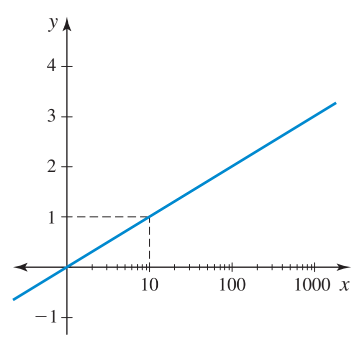


Figure 1.90 Graph for Problem 101.

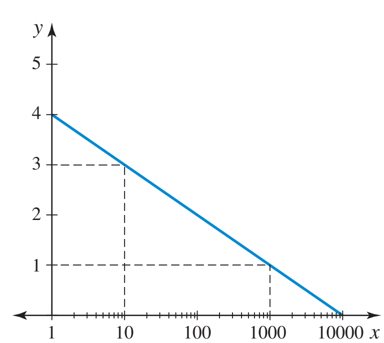


Figure 1.91 Graph for Problem 102.

103. The energy ΔG need to transport a mole of solute across a membrane from concentration c_1 to one of concentration c_2 follows the equation

$$\Delta G = 2.303RT\ \log \frac{c_2}{c_1}$$

where $R = 1.99$ kcal K^{-1} kmol^{-1} is the universal gas constant and T is temperature measured in kelvins (K). Plot ΔG as a function of the concentration ratio c_2/c_1 when $T = 298$ K (25°C). Use a coordinate system in which the vertical axis is on a linear scale and the horizontal axis is on a logarithmic scale.

104. Logistic Transformation Suppose that

$$f(x) = \frac{1}{1 + e^{-(b+mx)}} \tag{1.15}$$

where b and m are constants.
A function of the form (1.15) is called a **logistic** function. The logistic function was introduced by the Dutch mathematical biologist Verhulst around 1840 to describe the growth of populations with limited food resources.

(a) Show that

$$\ln \frac{f(x)}{1 - f(x)} = b + mx \tag{1.16}$$

This transformation is called the *logistic transformation*.

(b) Given some data, a table of values of x, and the function $f(x)$ at each value x, explain how to plot the data to produce a straight line, and to estimate the constant b and m from the model.

1.4.4

Problem 105 addresses the relationship between species diversity and primary productivity in an ecosystem (see Example 11 for definitions of these terms).

105. Not every study of species diversity as a function of productivity produces a hump-shaped curve. ***Owen (1988)*** studied rodent assemblages in Texas and found that the number of species was a decreasing function of productivity. Sketch a graph that would describe this situation.

106. Species diversity in a community may be controlled by disturbance frequency. The intermediate disturbance hypothesis states that species diversity is greatest at intermediate disturbance levels. Sketch a graph of species diversity as a function of disturbance frequency that illustrates this hypothesis.

107. ***Preston (1962)*** investigated the dependence of number of bird species on island area in the West Indian islands. He found that the number of bird species increased at a decelerating rate as island area increased. Sketch this relationship.

108. Phytoplankton converts carbon dioxide to organic compounds during photosynthesis. This process requires sunlight. It has been observed that the rate of photosynthesis is a function of light intensity: The rate of photosynthesis increases approximately linearly with light intensity at low intensities, saturates at intermediate levels, and decreases slightly at high intensities. Sketch a graph of the rate of photosynthesis as a function of light intensity.

109. Brown lemming densities in the tundra areas of North America and Eurasia show cyclic behavior: Over three to four years, lemming densities build up, and they typically crash the next year. Sketch a graph that describes this situation.

110. ***Yang et al. (2014)*** measured the time that animals of different sizes spend urinating. For animals larger than 1kg, the time spent urinating increases (slowly) with animal size. The smallest animal in their study was a cat (mass 5kg, duration of urination 18s) and the largest was an elephant (mass 5000 kg, duration of urination 29s). Make a sketch of time spent urinating as a function of animal size.

111. A study of ***Borcherts (1994)*** investigated the relationship between stem water storage and wood density in a number of tree species in Costa Rica. The study showed that water storage is inversely related to wood density; that is, higher wood density corresponds to lower water content. Sketch a graph of water content as a function of wood density that illustrates this situation.

112. Species richness can be a hump-shaped function of productivity. In the same coordinate system, sketch two hump-shaped graphs of species richness as a function of productivity, one in which the maximum occurs at low productivity and one in which the maximum occurs at high productivity.

113. The size distribution of zooplankton in a lake is typically a hump-shaped curve; that is, the number of zooplankton of a given size increases with size up to a critical size and then decreases with size for organisms larger than that critical size. ***Brooks and Dodson (1965)*** studied the effects of introducing a planktivorous fish in a lake. They found that the composition of zooplankton after the fish was introduced shifted to smaller individuals. In the same coordinate system, sketch the size distribution of zooplankton before and after the introduction of the planktivorous fish.

114. *Daphnia* is a genus of zooplankton that comprises a number of species. The body growth rate of *Daphnia* depends on food concentration. A minimum food concentration is required for growth: Below this level, the growth rate is negative; above, it is positive. In a study by ***Gliwicz (1990)***, it was found that growth rate is an increasing function of food concentration and that the minimum food concentration required for growth decreases with increasing size of the animal. Sketch two graphs in the same coordinate system, one for a large and one for a small *Daphnia* species, that illustrates this situation.

115. ***Grant (1982)*** investigated egg weight as a function of adult body weight among 10 species of Darwin's finches. He found that the relationship between the logarithm of the average egg size and the logarithm of the average body size is linear and that smaller species lay smaller eggs and larger species lay larger eggs. Graph this relationship.

116. ***Grant et al. (1985)*** investigated the relationship between mean wing length and mean weight among males of populations of six ground finch species. They found a positive and nearly linear relationship between these two quantities. Graph this relationship.

117. ***James et al. (2008)*** studied how nuclei interact in chimeric fungi (a chimera is an organism containing two or more different types of cells or nuclei). They wanted to study whether the two types of nuclei were cooperating or competing. They compared the growth rates of three different types of fungus on media containing different amounts of citric acid. First they grew two pure strains:

(Strain 1) A fungus containing only nuclei of type 1, which grew best in the absence of citric acid.

(Strain 2) A fungus containing only nuclei of type 2, which grew best at high citric acid concentration.

(a) Make a sketch of growth rate as a function of citric acid concentration. Your sketch should include two lines, one for strain 1 and one for strain 2.

(b) James et al. then compared the two pure strains with a chimeric fungus containing both types of nuclei. They considered two possible scenarios:

(Scenario 1) The two nuclei compete, and the fungus grows less well than both pure strains under all concentrations.

(Scenario 2) The two types of nuclei cooperate, and the fungus grows better than both pure strains at all citric acid concentrations.

Add two lines, representing scenarios 1 and 2 respectively, to your sketch from part (a).

118. In Example 13, we discussed germination success as a function of temperature for varying levels of humidity. We can also consider germination success as a function of humidity for various levels of temperature. Sketch the following graphs of germination success as a function of humidity: one for low temperature, one for intermediate temperature, and one for high temperature.

119. ***Boulinier et al. (2001)*** studied the dynamics of forest bird communities in forests of different sizes. They found that the extinction rate of area-sensitive species declined with forest size, whereas the mean extinction rate of non-area-sensitive species did not depend on forest size. In the same coordinate system, graph the mean extinction rate as a function of forest size for **(a)** an area-sensitive species and **(b)** a non-area-sensitive species.

Chapter 1 Review

Key Terms

Discuss the following definitions and concepts:

1. Real numbers
2. Intervals: open, closed, half-open
3. Absolute value
4. Proportional
5. Lines: standard form, point–slope form, slope–intercept form
6. Parallel and perpendicular lines
7. Circle: radius, center, equation of circle, unit circle
8. Angle: radians, degrees
9. Trigonometric identities
10. Complex numbers: real part, imaginary part
11. Function: domain, codomain, range, image
12. Symmetry of functions: even, odd
13. Composition of functions
14. Polynomial
15. Degree of a polynomial
16. Chemical reaction: law of mass action
17. Rational function
18. Growth rate/decay rate
19. Specific growth rate and per capita growth rate
20. Monod growth function
21. Power function
22. Scaling relations
23. Exponential function
24. Exponential growth/decay
25. Natural exponential base
26. Radioactive decay
27. Half-life
28. Inverse function, one to one
29. Logarithmic function
30. Relationship between exponential and logarithmic functions
31. Periodic function
32. Trigonometric function
33. Amplitude, period
34. Translation: horizontal, vertical
35. Logarithmic scale
36. Order of magnitude
37. Logarithmic transformation
38. Log-log plot
39. Semilog plot

Review Problems

1. Population Growth Suppose that the number of bacteria in a petri dish is given by

$$B(t) = 1000e^{0.1t}$$

where t is measured in hours.

(a) How many bacteria are present at $t = 0, 1, 2, 3,$ and 4?

(b) Find the time t when the number of bacteria reaches 100,000.

2. Population Decline Suppose that a pathogen is introduced into a population of bacteria at time 0. The number of bacteria then declines as

$$B(t) = 25{,}000e^{-1.5t}$$

where t is measured in hours.

(a) How many bacteria are left after 3 hours?

(b) How long will it be until only 1% of the initial number of bacteria are left?

3. Chemical Reaction The chemical reaction

$$C + D \to E$$

occurs at a rate k. The initial concentrations of C and D are $[C] = 1.5$, $[D] = 2.5$.

(a) Write a formula for the rate of the reaction $R(y)$ as a function of y, the concentration of C molecules. What is the domain and range of the function R?

(b) Write a formula for the rate of reaction $R(z)$ as a function of z, the concentration of D molecules. What is the domain and range of the function R?

4. The von Bertalanffy equation describes the rate of growth of certain fishes. The length of the fish as a function of time, t, since hatching is given by a formula

$$L(t) = L_\infty(1 - e^{kt})$$

where L_∞ and k are constants. Suppose you have the following data for $L(t)$ and you want to infer the constants k and L_∞:

t	$L(t)$
0	0
1	1
2	1.5

(a) Write out the equations that L_∞ and k need to satisfy, and reduce to a single equation for e^k. (*Hint*: By dividing your equations you can turn them into a rational function for $u = e^k$. You can simplify this function.)

(b) Solve your equations to obtain values of k and L_∞ that fit the mathematical model to your data.

5. Hypothetical Plants To compare logarithmic and exponential growth, we consider two hypothetical plants species A and species B, but that exhibit rather different growth rates. The height of species A (measured in feet) at age t (measured in years) is given by

$$L(t) = \ln(t) + 1, \quad t \geq 1$$

The height E (measured in feet) of species B at age t (measured in years), is given by

$$E(t) = e^{t-1}, \quad t \geq 1$$

(a) Show that the plants have the same height at time $t = 1$.

(b) Find the heights of each plant after 2, 10, 100 years.

(c) How long does it take for each plant to double its height (i.e., grow to 2 feet tall)?

(d) How long does it take to double height again (i.e., to reach 4 feet tall)? Compare with your answer to part (c).

(e) How long does it take each plant to reach 3 feet tall? Compare to your answer to part (c).

6. Population Growth In Chapter 3 of *The Origin of Species* (Darwin, 1859), Charles Darwin asserts that a "struggle for existence inevitably follows from the high rate at which all organic beings tend to increase. . . . Although some species may be now increasing, more or less rapidly, in numbers, all cannot do so, for the world would not hold them."

> There is no exception to the rule that every organic being naturally increases at so high a rate, that, if not destroyed, the earth would soon be covered by the progeny of a single pair. Even slow-breeding man has doubled in twenty-five years, and at this rate, in a few thousand years, there would literally not be standing room for his progeny.

Was Darwin correct? Let's calculate how human population density increases with time.

Starting with a single pair, compute the world's population after 1000 years and after 2000 years under Darwin's assumption that the world's population doubles every 25 years, and find the resulting population densities (number of people per square foot). To answer the last part, you need to know that the earth's diameter is about 7900 mi, the surface of a sphere is $4\pi r^2$, where r is the radius of a sphere, and the continents make up about 29% of the earth's surface. (Note that 1 mi = 5280 ft.)

7. Population Growth Assume that a population grows q% each year. How many years will it take the population to double in size? Give the functional relationship between the doubling time T and the annual percentage increase q. Produce a table that shows the doubling time T as a function of q for $q = 1, 2, \ldots, 10$, and graph T as a function of q. What happens to T as q gets closer to 0?

8. The Logistic Model for Population Growth For many organisms the rate of reproduction depends on the amount of space and resources available to each organism, which depends on how many other organisms these resources must be shared with. A model for such effects is to make the rate of reproduction (number of offspring produced in one year minus the number of deaths in one year) depend on the number of organisms, N. One model for this effect is the logistic equation,

$$R(N) = rN\left(1 - \frac{N}{K}\right),$$

where r and K are positive constants.

(a) Assume initially that $K = 100$. Draw a graph of $R(N)$ and find its domain and range.

(b) We want to understand the role played in the model by the parameters r and K. First set $K = 100$, and plot the function $R(N)$ for $r = 1, 2, 3$. Describe in words how the function changes. [*Hint*: Measure the maximum reproduction rate.]

(c) Now set $r = 1$, and plot the function for $K = 100$, 200, 300. K is often called the carrying capacity of the environment in which the organism lives. Explain why this term is used by looking at your graph.

9. Fish Yield (*Adapted from* ***Moss, 1980***) ***Oglesby (1977)*** investigated the relationship between annual fish yield (Y) and summer phytoplankton chlorophyll concentration (C). Fish yield was measured in grams dry weight per square meter per year, and the chlorophyll concentration was measured in micrograms per liter. Data from 19 lakes, mostly in the Northern Hemisphere, yielded the following relationship:

$$\log_{10} Y = 1.17 \log_{10} C - 1.92 \tag{1.17}$$

(a) Plot $\log_{10} Y$ as a function of $\log_{10} C$.

(b) Find the relationship between Y and C; that is, write Y as a function of C. Explain the advantage of the log–log transformation resulting in (1.17) versus writing Y as a function of C. [*Hint*: Try to plot Y as a function of C, and compare with your answer in (a).]

(c) Find the predicted yield (Y_p) as a function of the current yield (Y_c) if the current summer phytoplankton chlorophyll concentration were to double.

(d) By what percentage would the summer phytoplankton chlorophyll concentration need to increase to obtain a 10% increase in fish yield?

10. Radioactive Decay (*Adapted from* ***Moss, 1980***) To trace the history of a lake, a sample of mud from a core is taken and dated. One dating method uses radioactive isotopes. The C^{14} method is effective for sediments that are younger than 60,000 years. The $C^{14} : C^{12}$ ratio has been essentially constant in the atmosphere over a long time, and living organisms take up carbon in that ratio. Upon death, the uptake of carbon ceases and C^{14} decays, which changes the $C^{14} : C^{12}$ ratio according to

$$\left(\frac{C^{14}}{C^{12}}\right)_{\text{at time } t} = \left(\frac{C^{14}}{C^{12}}\right)_{\text{initial}} \times e^{-\lambda t}$$

where t is the time since death.

(a) If the $C^{14} : C^{12}$ ratio in the atmosphere is 10^{-12} and the half-life of C^{14} is 5730 years, find an expression for t, the age of the material being dated, as a function of the $C^{14} : C^{12}$ ratio in the material being dated.

(b) Use your answer in (a) to find the age of a mud sample from a core for which the $C^{14} : C^{12}$ ratio is 1.61×10^{-13}.

11. Fossil Coral Growth (*Adapted from* ***Futuyama, 1995****, and* ***Dott and Batten, 1976***) Corals deposit a single layer of lime each day. In addition, seasonal fluctuation in the thickness of the layers allows for grouping them into years. In modern corals, we can count 365 layers per year. J. Wells, a paleontologist, counted such growth layers on fossil corals. To his astonishment, he found that corals that lived about 380 million years ago had about 400 daily layers per year.

(a) Today, the earth rotates about its axis every 24 hours and revolves around the sun every $365\frac{1}{4}$ days. The speed of the earth's orbit around the sun does not change with time. Astronomers have determined that the earth's rotation has slowed down in recent centuries at the rate of about 2 seconds every 100,000 years. That is, 100,000 years ago, a day was 2 seconds shorter than today. Extrapolate the slowdown back to the Devonian, and determine the length of a day and the length of a year back when Wells's corals lived. (*Hint*: The number of hours per year remains constant. Why?)

(b) Find a linear equation that relates geologic time (in millions of years) to the number of hours per day at a given time.

(c) Algal stromatolites also show daily layers. A sample of some fossil stromatolites showed 400 to 420 daily layers per year. Use your answer in (b) to date the stromatolites.

12. HIV Concentration in Blood ***Ho et al. 1995*** studied how HIV-1 virus is cleared from the blood by using anti-retroviral drugs to halt viral reproduction. Here is some data from one patient in their study.

Time (days after drug is administered)	Concentration of Virus (# thousand copies per mL of blood)
1	78.9
4	52.8
8	17.6
11	9.5
15	5.1

(a) By plotting the data with suitably transformed axes, show that the number of viruses decreases exponentially with time.

(b) Estimate from your graph the time it takes for the number of viruses to be halved.

(c) Does the number of viruses ever reach 0?

T **13. Model for Aging** The probability that an individual lives beyond age t is called the *survivorship function* and is denoted by $S(t)$. The Weibull model is a popular model in reliability theory and in studies of biological aging. Its survivorship function is described by two parameters, λ and β, and is given by

$$S(t) = \exp[-(\lambda t)^{\beta}]$$

Mortality data from a *Drosophila melanogaster* population in Dr. Jim Curtsinger's lab at the University of Minnesota were collected and fitted to this model separately for males and females ***(Pletcher, 1998)***. The following parameter values were obtained (t was measured in days):

Sex	λ	β
Males	0.019	3.41
Females	0.022	3.24

(a) Use a graphing calculator to sketch the survivorship function for both the female and male populations.

(b) For each population, find the value of t for which the probability of living beyond that age is 1/2.

(c) If you had a male and a female of this species, which would you expect to live longer?

14. Circadian Rhthym The circadian clock of the bread mold *Neurospora crassa* is controlled by the gene *frq*. Figure 1.92 shows some idealized data for the level of *frq* expression over time t, measured in hours.

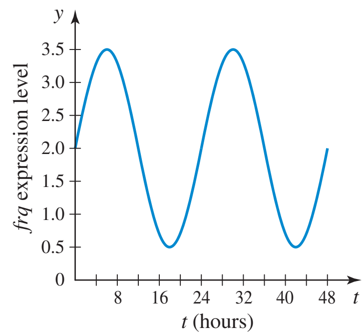


Figure 1.92 Expression of the *frq* gene over 48 hours.

(a) What is the period of the oscillations in *frq* expression levels?

(b) Write a formula for $E(t)$ as a function of t.

(c) frq^7 is mutant form of the fungus. The amplitude of frq^7 oscillations is the same as *frq*, but the period is increased to 29 hours. Draw the graph of frq^7 expression for $0 \le t \le 48$ hrs (assume the levels of *frq* and frq^7 are initially the same).

15. Chemical Reaction The rate of an enzymatic reaction is frequently described by the Michaelis–Menten equation

$$v = \frac{ax}{k+x}$$

where v is the rate of the reaction, x is the concentration of the substrate, a is the maximum reaction rate, and k is the substrate concentration at which the rate is half of the maximum rate. This curve describes how the reaction rate depends on the substrate concentration.

(a) Show that when $x = k$, the rate of the reaction is half the maximum rate.

(b) Show that an 81-fold change in substrate concentration is needed to change the rate from 10% to 90% of the maximum rate, regardless of the value of k.

16. Lake Acidification Atmospheric pollutants can cause acidification of lakes (by acid rain). This can be a serious problem for lake organisms; for instance, in fish the ability of hemoglobin to transport oxygen decreases with decreasing pH levels of the water. Experiments with the zooplankton *Daphnia magna* showed a negligible decline in survivorship at pH = 6, but a marked decline in survivorship at pH = 3.5, resulting in no survivors after just eight hours. Illustrate graphically the percentage survivorship as a function of time for pH = 6 and pH = 3.5.

17. Lake Chemistry The pH level of a lake controls the concentrations of harmless ammonium ions (NH_4^+) and toxic ammonia (NH_3) in the lake. For pH levels below 8, concentrations of NH_4^+ ions are little affected by changes in the pH value, but they decline over many orders of magnitude as pH levels increase beyond pH = 8. By contrast, NH_3 concentrations are negligible at low pH, increase over many orders of magnitude as the pH level increases, and reach a high plateau at about pH = 10 (after which levels of NH_3 are little affected by changes in pH levels). Illustrate the behavior of $[NH_4^+]$ and $[NH_3]$ graphically.

18. Development and Growth Egg development times of the zooplankton *Daphnia longispina* depend on temperature. It takes only about 3 days at 20°C, but almost 20 days at 5°C, for an egg to develop and hatch. When graphed on a log–log plot, egg development time (in days) as a function of temperature (in °C) is a straight line.

(a) Sketch a graph of egg development time as a function of temperature on a log-log plot.

(b) Use the data to find the function that relates egg development time and temperature for *D. longispina*.

(c) Use your answer in (b) to predict egg development time of *D. longispina* at 10°C.

(d) Suppose you measured egg development time in hours and temperature in Fahrenheit. Would you still find a straight line on a log-log plot?

19. Resource Consumption Organisms consume resources. The rate of resource consumption, denoted by v, depends on resource concentration, denoted by S. The *Blackman* model of resource consumption assumes a linear relationship between resource consumption rate and resource concentration: Below a threshold concentration (S_k), the consumption rate increases linearly with $S = 0$ when $v = 0$; when $S = S_k$, the consumption rate v reaches its maximum value v_{max}; for $S > S_k$, the resource consumption rate stays at the maximum value v_{max}. A function like this, with a sharp transition, cannot be described analytically by just one expression; it needs to be defined piecewise:

$$v = \begin{cases} g(S) & \text{for } 0 \le S < S_k \\ v_{max} & \text{for } S \ge S_k \end{cases}$$

Find $g(S)$, and graph the resource consumption rate v as a function of resource concentration S.

20. Light Intensity Light intensity in lakes decreases exponentially with depth. If $I(z)$ denotes the light intensity at depth z, with $z = 0$ representing the surface, then

$$I(z) = I(0)e^{-\alpha z}, \quad z \ge 0$$

where α is a positive constant called the *vertical attenuation coefficient*. This coefficient depends on the wavelength of the light and on the amount of dissolved matter and particles in the water. In the following, we assume that the water is pure:

(a) About 65% of red light (720 nm) is absorbed in the first meter. Find α.

(b) About 5% of blue light (475 nm) is absorbed in the first meter. Find α.

(c) Explain in words why a diver would not see red hues a few meters below the surface of a lake.

21. Population Growth. A colony of yeast cells grows exponentially with time; that is, the number of cells $N(t)$ increases with time t, measured in hours, according to the formula

$$N(t) = N_0 e^{rt}$$

Where N_0 and r are constants that you wish to determine.

(a) Suppose you measure the following data for the population size at $t = 0$ and $t = 2$:

t	$N(t)$
0	2000
1	4000

Calculate the values of N_0 and r that fit the mathematical model to this formula.

(b) When, according to your model, will the population size reach 10,000 cells?

(c) You realize that your cell counter is accurate only to 10%, meaning that if you count 2000 cells, the real population size is somewhere between 2000 − 10% = 1800 and 2000 + 10% = 2200.

Calculate the maximum and minimum values of r that are consistent with your measurements.

22. Population Growth Assume that the population size $N(t)$ at time $t \ge 0$ is given by

$$N(t) = N_0 e^{rt}$$

with $N_0 = N(0)$. The parameter r is called the *average annual growth rate*.

(a) Show that

$$r = \ln \frac{N(t+1)}{N(t)} \tag{1.18}$$

Formula (1.18) is used, for instance, by the U.S. Census Bureau to track world population growth.

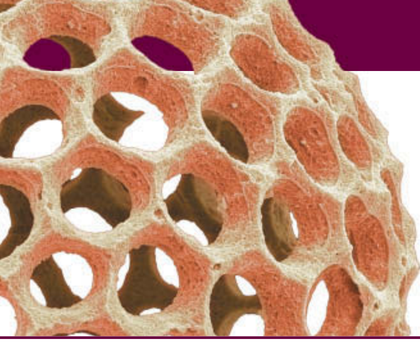
(b) Suppose a population doubles in size within a single year.

(i) What is the percent increase of the population during that year?

(ii) What is the average annual growth rate in percent during that year, according to (1.18)?

(c) Suppose the average annual growth rate of a population is 1.3%. How many years will it take the population to double in size?

(d) To calculate the doubling time of a growing population with a constant average annual growth rate, we divide the percent average annual growth rate into 70. Apply this "Rule of 70" to (c) and compare your answers. Derive the "Rule of 70."



CHAPTER

2

Discrete-Time Models, Sequences, and Difference Equations

In this chapter, we discuss models for populations that reproduce at discrete times and we develop some of the theory needed to analyze this type of model. The models give snapshots of the state of the population after 0, 1, 2, 3, . . . generations. Specifically, we will learn how to

- describe discrete-time models of population growth and decay—with tables and graphs, explicitly as functions of time, and recursively from one time step to the next;
- calculate terms in a sequence, a_n, from formulas for the nth term and recursively;
- use Σ notation to represent the sum of terms in a sequence a_n;
- calculate limits and fixed points of sequences;
- describe the relationship between fixed points and limits of a sequence;
- give examples of density-dependent population growth models;
- use recurrence equations to model population growth and the passage of medications through the body.

2.1 Exponential Growth and Decay

2.1.1 Modeling Population Growth in Discrete Time

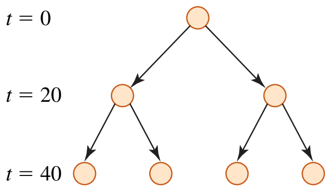


Figure 2.1 Bacteria split every 20 units of time.

Imagine that we observe bacteria that divide every 20 minutes and that, at the start of the experiment, there was one bacterium. How will the number of bacteria change over time? We call the time when we started the observation time 0. At time 0, there is one bacterium. After 20 minutes, the bacterium splits in two, so there are two bacteria at time 20. Twenty minutes later, each of the bacteria splits again, resulting in four bacteria at time 40, and so on (Figure 2.1).

We can produce a table that describes the growth of this population:

Time (min)	0	20	40	60	80	100	120
Population size	1	2	4	8	16	32	64

We can simplify the description of the growth of the bacterial population if we measure time in more convenient units. We say that one unit of time equals 20 minutes. Two units of time then corresponds to 40 minutes, three units of time to 60 minutes, and so on. We reproduce the table of population growth with these new units:

Time (20 min)	0	1	2	3	4	5	6
Population size	1	2	4	8	16	32	64

The new time units make it easier to write a general formula for the population size at time t. Denoting by $N(t)$ the population size at time t, where t is now measured in the new units (one unit is equal to 20 minutes), we guess from the second table that

$$N(t) = 2^t, \quad t = 0, 1, 2, \ldots \tag{2.1}$$

We encountered this function in Section 1.2 when we discussed exponential functions. There, the function was defined for all $t \geq 0$, whereas now, the function is defined only for nonnegative integer values; that is, for $t = 0, 1, 2, \ldots$. We call this kind of function discrete. Equation (2.1) allows us to determine the population size at any discrete time t directly, without first calculating the population sizes at all previous time steps. For instance, at time $t = 5$, we find that $N(5) = 2^5 = 32$, as shown in the second table, or, at time $t = 10$, $N(10) = 2^{10} = 1024$. The graph of $N(t) = 2^t$ is shown in Figure 2.2.

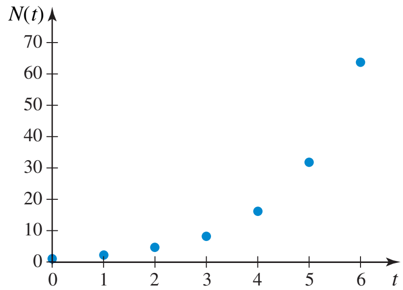


Figure 2.2 The size of a bacterial population that doubles in each unit of time is given by $N(t) = 2^t$.

The function $N(t) = 2^t$, $t = 0, 1, 2, \ldots$, is an exponential function, and we call the type of population growth that it represents **exponential growth**. The base 2 reflects the fact that the population size doubles every unit of time.

We will often write N_t instead of $N(t)$. The subscript notation is used only for discrete functions $N(t)$. So, instead of writing $N(t) = 2^t$, $t = 0, 1, 2, \ldots$, we can write $N_t = 2^t$, $t = 0, 1, 2, \ldots$.

So far, we assumed that $N(0) = N_0 = 1$. Let's see what N_t looks like if $N_0 = 100$. Regardless of N_0, the population size doubles every unit of time. We obtain the following table, where time is again measured in units of 20 minutes:

Time (20 min)	0	1	2	3	4	5	6
Population size	100	200	400	800	1600	3200	6400

We can guess the general form of N_t with $N_0 = 100$ from the table:

$$N_t = 100 \cdot 2^t, \quad t = 0, 1, 2, \ldots$$

We see that the initial population size $N_0 = 100$ appears as a multiplicative factor in front of the term 2^t. If we do not want to specify a numerical value for the population size N_0 at time 0, we can write

$$N_t = N_0 \cdot 2^t, \quad t = 0, 1, 2, \ldots$$

We already mentioned that the base 2 indicates that the population size doubles every unit of time. Replacing 2 by another number, we can describe other populations. For instance,

$$N_t = 3^t, \quad t = 0, 1, 2, \ldots$$

describes a population with $N_0 = 1$ and that triples in size every unit of time. The corresponding table is

Time	0	1	2	3	4
Population size	1	3	9	27	81

This model corresponds to a scenario where each bacterium divides into three in each unit of time.

Now that we have some experience with exponential growth in discrete time, we give the general formula:

Definition Exponential Population Growth

$$N_t = N_0 R^t, \quad t = 0, 1, 2, \ldots \tag{2.2}$$

The parameter R is a positive constant called the **growth constant**.

The constant N_0 is nonnegative and denotes the population size at time 0. The assumptions $R > 0$ and $N_0 \geq 0$ are made for biological reasons: Negative values for R or N_0 would result in negative population sizes.

EXAMPLE 1 Suppose a population of cells reproduces every 15 minutes and we measure its size every 30 minutes:

Time (min)	0	30	60	90	120	150	180
Population size	1	4	16	64	256	1024	4096

Write a formula for time $t = 0, 1, 2, \ldots$ when (a) one unit of time is 30 minutes, (b) one unit of time is 60 minutes, and (c) one unit of time is 15 minutes.

Solution **(a)** We see from the values listed in the table that when one unit of time is 30 minutes, the population quadruples every unit of time, with $N_0 = 1$. Thus,

$$N_t = 4^t, \quad t = 0, 1, 2, \ldots$$

(b) This time, we see from the values in the table that when one unit of time is equal to 60 minutes, the population grows by a factor of 16 each unit of time. Again, $N_0 = 1$. Hence,

$$N_s = 16^s, \quad s = 0, 1, 2, \ldots$$

We could also have arrived at this answer by noting that the time step in (b) is twice that of the time step in (a). In other words, when one unit of time elapses in (b), two units of time elapse in (a):

t	0	1	2	3	4	5	6
s	0		1		2		3

We find that $t = 2s$. If we substitute $2s$ for t in (a), we find that

$$N_t = 4^t \quad \text{yields} \quad N_s = 4^{2s} = 16^s$$

for $s = 0, 1, 2, \ldots$.

(c) When one unit of time is 15 minutes, and we use the variable $u = 0, 1, 2, \ldots$ to denote time, it follows that $t = u/2$ and

$$N_t = 4^t \quad \text{yields} \quad N_u = 4^{u/2} = 2^u$$

for $u = 0, 1, 2, \ldots$.

The function $N_t = N_0 R^t$, $t = 0, 1, 2, \ldots$, is an exponential function. We discussed exponential functions in the previous chapter. There, we said that a function $f(x)$ is exponential if $f(x) = ka^x$, where k and a are constants and $x \in \mathbf{R}$.

In Chapter 1, we learned how $f(x) = ka^x$, $x \in \mathbf{R}$, behaves for different values of a. We can analyze the behavior of $N_t = N_0 R^t$ in a similar manner (Figure 2.3). We see that when $R > 1$, the population size N_t increases indefinitely; when $R = 1$, the population size N_t stays the same for all $t = 0, 1, 2, \ldots$; and when $0 < R < 1$, the population size N_t declines and approaches 0 as t increases. $R > 1$ corresponds to an exponentially growing population. $R < 1$ corresponds to an exponentially decaying population. $R = 1$ is a population with constant size.

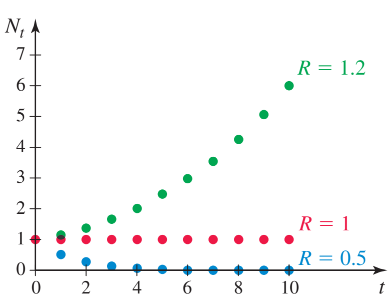


Figure 2.3 The graphs of $N_t = R^t$, $t = 0, 1, 2, \ldots, 10$, for three different values of R: $R = 0.5$, $R = 1$, and $R = 1.2$. When $R > 1$, the population grows with time. When $R < 1$, the population decays with time.

2.1.2 Recurrence Equations

When we constructed the tables for the bacterial population size with $R = 2$ at consecutive time steps, we doubled the population size from time step to time step. In other words, we computed the population size at time $t + 1$ on the basis of the population size at time t, using the equation

$$N_{t+1} = 2N_t \tag{2.3}$$

Equation (2.3) is a rule that is applied repeatedly to go from one time step to the next and is called a **recurrence equation** or a **recursion**. We say that the population size is defined **recursively** by Equation (2.3). Equation (2.3) alone is not enough to completely specify the sequence. We also need to know the *initial condition* on N_t, that is the value of N_0.

Suppose we want to use Equation (2.3) and the initial condition to calculate the population size at time $t = 4$, namely, N_4. We can apply Equation (2.3) to calculate N_1 from N_0 (i.e., set $t = 0$);

$$N_1 = 2 \cdot N_0.$$

We can then calculate N_2 from N_1 (i.e., set $t = 1$):

$$N_2 = 2 \cdot N_1 = 4N_0.$$

Similarly $N_3 = 2 \cdot N_2 = 8N_0$ (setting $t = 2$ in Equation (2.3)). Finally $N_4 = 2 \cdot N_3 = 16N_0$ (setting $t = 3$ in Equation (2.3)). Note that to calculate N_4 we first needed to calculate N_1, N_2, and N_3.

We thus have two equivalent ways to describe this population: For $t = 0, 1, 2, \ldots,$

$$N_t = 2^t N_0 \quad \text{is equivalent to} \quad N_{t+1} = 2N_t$$

More generally, if $N_t = N_0 R^t$ then N_t obeys a recurrence equation:

$$N_{t+1} = RN_t \tag{2.4}$$

with the initial condition being the value of N_0.

Applying (2.4) repeatedly, we obtain

$$\begin{aligned} N_1 &= RN_0 \\ N_2 &= RN_1 = R^2N_0 \\ N_3 &= RN_2 = R^3N_0 \\ N_4 &= RN_3 = R^4N_0 \\ &\vdots \\ N_t &= RN_{t-1} = R^tN_0 \end{aligned}$$

We say that $N_t = N_0R^t$ is a **solution** of the recurrence equation $N_{t+1} = RN_t$ with initial condition N_0 at time 0.

2.1.3 Visualizing Recurrence Equations

If N_t is defined by a recurrence equation and we do not know the formula for the solution of the recurrence equation, then it can be hard to deduce the behavior of N_t as t increases. In this section we discuss two kinds of plots that can be used to understand the function N_t even without solving the recurrence equation on a plot of N_{t+1} against N_t.

N_{t+1} Against N_t. We can plot N_{t+1} on the vertical axis against N_t on the horizontal axis.

Plotted in this way, exponential growth recurrence equation

$$N_{t+1} = RN_t \tag{2.5}$$

is a straight line through the origin with slope R (Figure 2.4). Since $N_t \geq 0$ for biological reasons, we restrict the graph to the first quadrant.

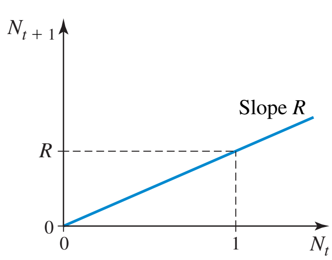


Figure 2.4 The exponential growth recursion $N_{t+1} = RN_t$ when $R > 0$.

What can we learn about the recurrence equation from this graph? Given any value of N_t, the height of the graph allows us to read out N_{t+1}. For example, if $R = 2$ then the graph has slope 2. Then in Figure 2.5 we use the graph to calculate N_{t+1} for each of $N_t = 3, 10,$ and 20. Notice that we did not need to know the initial condition N_0 to be able to plot this graph.

Reproductive Rate Against N_t. N_{t+1} is the number of organisms present at time $t + 1$ while N_t is the number of parents (organisms present at time t). So the number of organisms *added* between time t and time $t+1$ is $N_{t+1} - N_t$. It turns out to be most helpful to express this in terms of the number of organisms added per parent, namely as

$$\frac{N_{t+1} - N_t}{N_t} = \frac{N_{t+1}}{N_t} - 1.$$

We call this quantity the **reproductive rate**.

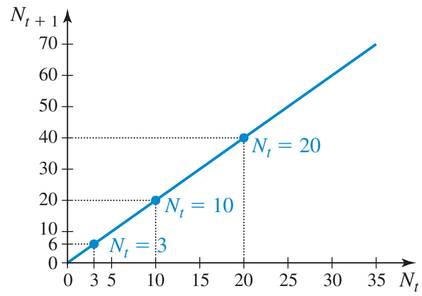


Figure 2.5 We can use the graph of N_{t+1} against N_t to calculate N_{t+1} for different values of N_t.

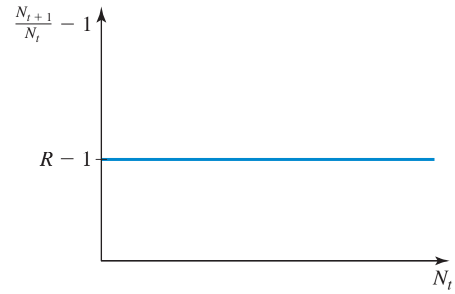


Figure 2.6 The graph of reproductive rate $\frac{N_{t+1}}{N_t} - 1$ as a function of N_t for a population with exponential growth.

To help interpret the reproductive rate, consider a population with exponential growth, $N_{t+1} = RN_t$; for this population the reproductive rate is a constant:

$$\frac{N_{t+1}}{N_t} - 1 = R - 1 \tag{2.6}$$

For each organism present in the population at time t, $R-1$ extra organisms are added to the population for time $t+1$. If $R > 1$, then the population size increases with each time step, whereas if $R < 1$, then the population size decreases each time step. In either case, a plot of $N_{t+1}/N_t - 1$ against N_t is just a flat horizontal line (Figure 2.6). If the reproductive rate doesn't change with N_t, we say that the reproduction rate is **density independent**.

This model eventually becomes biologically unrealistic, since any population will sooner or later experience food or habitat limitations that will limit its growth. We will discuss models that include such limitations in Section 2.3, but we show the need for models in which reproduction rates are density dependent by considering some real data for the growth of a population of yeast cells in 10 ml of growth medium in a laboratory (this data is based on the measurements of ***T. Carlson, 1913***).

First we plot the number of cells N_t as a function of time t, measured in hours, in Figure 2.7. We see from this plot that population growth is not exponential. Although the population grows initially, growth stops once the population size reaches around 6.6×10^6 cells. Real populations often have a **carrying capacity**, that is the organisms' environment can only sustain a limited number of organisms. As the size of the population approaches the carrying capacity, population growth slows and eventually stops. We can see this for the yeast cell population more clearly in a plot of reproductive rate against population size, as in Figure 2.8.

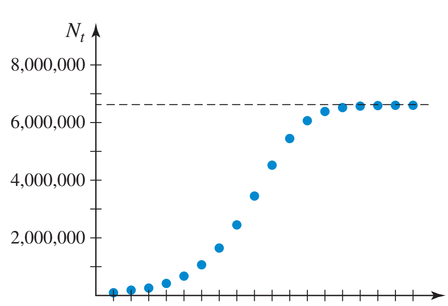


Figure 2.7 Number of cells, N_t, as a function of time t, for a population of yeast cells. The data in this plot is taken from ***T. Carlson (1913)***.

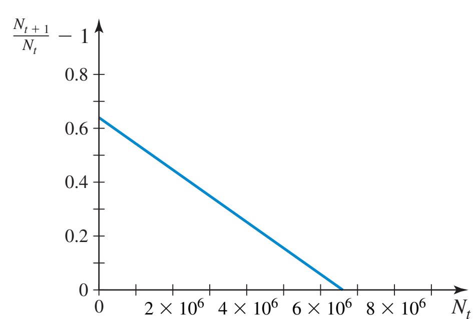


Figure 2.8 Reproductive rate for the same population shown in Figure 2.7.

We see from Figure 2.8 that reproductive rate $N_{t+1}/N_t - 1$ is **density dependent**. Specifically reproductive rate decreases with N_t, eventually reaching 0, meaning that the population stops growing when it reaches the carrying capacity of the test-tube.

Section 2.1 Problems

2.1.1

In Problems 1–4, produce a table for $t = 0, 1, 2, \ldots, 5$ and graph the function N_t.

1. $N_t = 3^t$

2. $N_t = 6 \cdot 2^t$

3. $N_t = \left(\frac{1}{3}\right)^t$

4. $N_t = 0.2(0.8)^t$

In Problems 5–7, give a formula for $N(t)$, $t = 0, 1, 2, \ldots,$ on the basis of the information provided.

5. $N_0 = 2$; population doubles every 20 minutes; one unit of time is 20 minutes

6. $N_0 = 4$; population doubles every 40 minutes; one unit of time is 40 minutes

7. $N_0 = 1$; population doubles every 40 minutes; one unit of time is 80 minutes

8. Suppose $N_t = 20 \cdot 4^t$, $t = 0, 1, 2, \ldots,$ and one unit of time corresponds to 3 hours. Determine the amount of time it takes the population to double in size.

9. Suppose $N_t = 100 \cdot 2^t$, $t = 0, 1, 2, \ldots,$ and one unit of time corresponds to 2 hours. Determine the amount of time it takes the population to triple in size.

10. Suppose you measure the following data for the size of a population of bacteria:

t	0	1	2	3
N_t	10	30	90	270

Write down a formula for the population size, N_t, as a function of time, t.

11. A strain of bacteria reproduces asexually every hour. That is, every hour, each bacterial cell splits into two cells. If, initially, there is one bacterium, find the number of bacterial cells after 1 hour, 2 hours, 3 hours, 4 hours, and 5 hours.

12. A strain of bacteria reproduces asexually every 30 minutes. That is, every 30 minutes, each bacterial cell splits into two cells. If, initially, there is one bacterium, find the number of bacterial cells after 1 hour, 2 hours, 3 hours, 4 hours, and 5 hours.

13. A strain of bacteria reproduces asexually every 42 minutes. That is, every 42 minutes, each bacterial cell splits into two cells. If, initially, there is 1 bacterium, how long will it take until there are 1024 bacteria?

14. A strain of bacteria reproduces asexually every 24 minutes. That is, every 24 minutes, each bacterial cell splits into two cells. If, initially, there is 1 bacterium, how long will it take until there are 512 bacteria?

15. A strain of bacteria reproduces asexually every 10 minutes. That is, every 10 minutes, each bacterial cell splits into two cells. If, initially, there are 5 bacteria, how long will it take until there are 320 bacteria?

16. A strain of bacteria reproduces asexually every 12 minutes. That is, every 12 minutes, each bacterial cell splits into two cells. If, initially, there are 10 bacteria, how long will it take until there are 160 bacteria?

17. Find the exponential growth equation for a population that doubles in size every unit of time and that has 40 individuals at time 0.

18. Find the exponential decay equation for a population that halves in size every unit of time and that has 1024 individuals at time 0.

19. Find the exponential growth equation for a population that has a reproductive rate of 4 and has 20 individuals at time 0.

20. Find the exponential growth equation for a population that triples in size every unit of time and that has 72 individuals at time 0.

21. Find the exponential growth equation for a population that quadruples in size every unit of time and that has five individuals at time 0.

22. Find the exponential growth equation for a population whose size increases by 50% in each unit of time and that has 32 individuals at time 0.

2.1.2

23. Find the recursion for a population that doubles in size every unit of time and that has 11 individuals at time 0.

24. Find the recursion for a population that triples in size every unit of time and that has 6 individuals at time 0.

25. Find the recursion for a population that quadruples in size every unit of time and that has 30 individuals at time 0.

26. Find the recursion for a population that has a reproductive rate of 1/3 and that has 63 individuals at time 0.

2.1.3

In Problems 27–30, graph the functions $f(x) = a^x$, $x \in [0, \infty)$, and $N_t = R^t$, $t \in \mathbf{N}$, together in one coordinate system for the indicated values of a and R.

27. $a = R = 2$

28. $a = R = 3$

29. $a = R = 1/2$

30. $a = R = 1/3$

In Problems 31–42, find the population sizes for $t = 0, 1, 2, \ldots, 5$ for each recursion. Then write the equation for N_t as a function of t.

31. $N_{t+1} = 2N_t$ with $N_0 = 3$

32. $N_{t+1} = 2N_t$ with $N_0 = 5$

33. $N_{t+1} = 3N_t$ with $N_0 = 2$

34. $N_{t+1} = 3N_t$ with $N_0 = 7$

35. $N_{t+1} = 5N_t$ with $N_0 = 1$

36. $N_{t+1} = 7N_t$ with $N_0 = 4$

37. $N_{t+1} = \frac{1}{2}N_t$ with $N_0 = 640$

38. $N_{t+1} = \frac{3}{2}N_t$ with $N_0 = 32$

39. $N_{t+1} = \frac{1}{3}N_t$ with $N_0 = 1215$

40. $N_{t+1} = \frac{1}{3}N_t$ with $N_0 = 2430$

41. $N_{t+1} = \frac{1}{5}N_t$ with $N_0 = 31250$

42. $N_{t+1} = \frac{1}{4}N_t$ with $N_0 = 8192$

In Problems 43–50, graph the line $N_{t+1} = RN_t$ in the N_t–N_{t+1} plane for the indicated value of R and locate the points (N_t, N_{t+1}), $t = 0, 1,$ and 2, for the given value of N_0.

43. $R = 2, N_0 = 2$

44. $R = 2, N_0 = 3$

45. $R = 3, N_0 = 1$

46. $R = 4, N_0 = 2$

47. $R = \frac{1}{2}, N_0 = 16$

48. $R = \frac{1}{2}, N_0 = 64$

49. $R = \frac{1}{3}, N_0 = 81$

50. $R = \frac{1}{4}, N_0 = 16$

In Problems 51–58, graph the reproductive rate $\left(\frac{N_{t+1}}{N_t} - 1\right)$ against N_t for the indicated value of R and locate the points $\left(N_0, \frac{N_1}{N_0} - 1\right)$ on the plot for the given value of N_0. Figure 2.6 is an example of this kind of plot.

51. $R = 2, N_0 = 2$

52. $R = 2, N_0 = 4$

53. $R = 3, N_0 = 2$

54. $R = 4, N_0 = 1$

55. $R = \frac{1}{2}, N_0 = 16$

56. $R = \frac{1}{2}, N_0 = 128$

57. $R = \frac{1}{3}, N_0 = 27$

58. $R = \frac{1}{4}, N_0 = 64$

Figure 2.9 shows the dependence of reproductive rate $N_{t+1}/N_t - 1$ upon N_t in three different populations. In Problems 59–61 you will identify which relationship best represents each of the different population growth scenarios.

59. Which of the three possible relations shown in the Figure 2.9 is the most likely to represent reproductive rate for the three following populations? Explain your answers.

(a) A population of microbes grown in a flask with a limited supply of sugar to feed on.

(b) A population of rodents that is introduced into a new, large, habitat.

(c) A population of bacteria growing on a catheter. Small numbers of bacteria are destroyed at high rates by the body's immune cells, but in larger numbers bacteria can build biofilms and are not destroyed.

60. Which of the three possible relations shown in the Figure 2.9 is the most likely to represent reproductive rate for the three following populations? Explain your answers.

(a) A population of microbes growing in a flask with ample space and nutrients.

(b) A population of lions competing for prey in a national park.

(c) An ant colony. Above a certain population size, worker ants specialize to perform different roles (like defense, offspring-rearing, and farming) and the ant colony starts to outcompete nearby colonies.

61. Which of the three possible relations shown in Figure 2.9 is the most likely to represent reproductive rate for the three following populations? Explain your answers.

(a) A population of cancerous cells growing in a patient's body. Above a certain size the cancer cells start to spread through the body by metastasis.

(b) The number of sick individuals in the first few days of a flu outbreak.

(c) Plants growing in a region of California Forest that was cleared by a wildfire. The plants start to compete for sunlight.

62. You are an ecologist measuring the size of the population of plants living on an island. You measure the size of the population each year at the beginning of the breeding season. Denote the population size after t years by N_t.

The first year that you measure the population size there are 16 birds, that is $N_1 = 16$. Each year you find 50% more birds than the year before.

(a) Write down the recursion for N_t and solve it.

(b) According to your formula, when will the size of the population first exceed 100 birds?

(c) When will the population size first exceed 1000 birds?

(d) Use your formula for N_t to predict when the population size will first exceed one million birds. Should you believe this prediction?

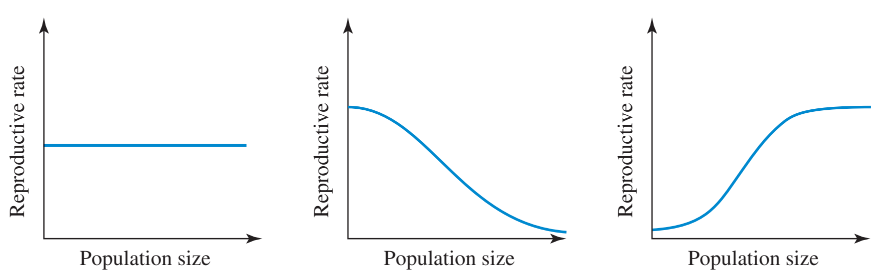


Figure 2.9 Three possible relationships between reproductive rate and population size.

2.2 Sequences

2.2.1 What Are Sequences?

In Section 2.1 we introduced the idea of a discrete function. This was a function whose domain is $\mathbf{N}_0$ (the set of natural numbers plus 0, i.e., $\{0, 1, 2, 3, \dots\}$).

$$\begin{aligned} f : \mathbf{N}_0 &\to \mathbf{R} \\ n &\to f(n) \end{aligned}$$

Here n represents the number of generations, and we frequently use t, rather than n, to represent the number of generations, with the idea being that n represents the amount of time that has elapsed measured in units of generations. We can represent these functions by tables or graphs.

EXAMPLE 1 Let

$$\begin{aligned} f : \mathbf{N}_0 &\to \mathbf{R} \\ n \to f(n) &= \frac{1}{n+1} \end{aligned}$$

Produce a table for $n = 0, 1, 2 \ldots, 5$ and graph the function.

Solution The table is

n	0	1	2	3	4	5
$\frac{1}{n+1}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$

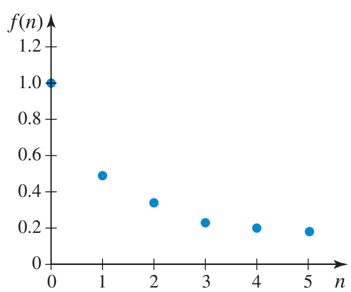


Figure 2.10 The graph of the function $f(n) = \frac{1}{n+1}$ in Example 1.

The graph of this function consists of discrete points (Figure 2.10). On the horizontal axis, we display the variable n; on the vertical axis, the function $f(n)$. Note that we did not connect the points with lines or curves because $f(n)$ is not defined except at the discrete points shown.

Functions of the form

$$\begin{aligned} f : \mathbf{N}_0 &\to \mathbf{R} \\ n &\to f(n) \end{aligned}$$

are defined by the list of their values $f(0)$, $f(1)$, $f(2)$, $f(3)$, and so on. We refer to this list as a **sequence**. We will write $\{f(n) : n \in \mathbf{N}_0\}$ to refer to the sequence. More generally a sequence is any ordered list of real numbers:

$$a_0, a_1, a_2, a_3, \ldots$$

Rather than write out every term of the sequence when we want to refer to it we use the expression $\{a_n : n \in \mathbf{N}_0\}$, or even more concisely $\{a_n\}$. The order of the terms in the sequence matters:

$$\{a_n : n \in \mathbf{N}_0\} = 3, 1, 4, 1, 5, 9, \ldots$$

and

$$\{a_n : n \in \mathbf{N}_0\} = 1, 3, 4, 1, 5, 9, \ldots$$

are different sequences even though they have the same terms, because the order of the terms must also be identical for two sequences to be identical.

How do we define the terms in a sequence? The most common methods are either to give an explicit formula for a_n as a function of n, or through a **recurrence equation** that allows a_1 to be calculated from a_0, a_2 to be calculated from a_1, a_3, to be calculated from a_2, and so on.

Here n is called the **index** of the sequence. Different values of n correspond to different terms of the sequence. The choice of letter is not important. For example:

$\{a_m : m \in \mathbf{N}_0\}$ is the same sequence as $\{a_n : n \in \mathbf{N}_0\}$ Both sequences consist of terms $a_0, a_1, a_2, a_3, \ldots$

Also it is not necessary that our indices run over $0, 1, 2, 3, \ldots$. Instead we could have indices that run over $1, 2, 3, 4, \ldots$, or $2, 3, 4, 5, \ldots$ and so on.

EXAMPLE 2 The sequence

$$a_n = (-1)^n, \quad n = 0, 1, 2, \ldots$$

takes on values

$$1, -1, 1, -1, 1, \ldots$$

When we see a sequence and recognize a pattern, we can often write an expression for a_n.

EXAMPLE 3 Find a_n for the sequence $0, 1, 4, 9, 16, 25, \ldots$

Solution We recognize the terms in a_n as perfect squares. We guess that:

$$a_n = n^2, \quad n = 0, 1, 2, \ldots$$

and we can check that this formula matches with all of the terms that are given in the sequence.

EXAMPLE 4 Find an explicit formula for the nth term of the sequence:

$$-1, \frac{1}{2}, -\frac{1}{3}, \frac{1}{4}, -\frac{1}{5}, \ldots$$

Solution This sequence has alternating signs. The first term is negative, the second term is positive, the third is negative and so on. From Example 2 we know that alternating signs can be produced if our formula includes a term like $(-1)^n$. Now $(-1)^n$ produces a sequence that goes positive-negative-positive-…, which is the opposite of what we want. So we take $-(-1)^n = (-1)\cdot(-1)^n = (-1)^{n+1}$. The sizes of the terms also change. Each term in the sequence can be written as 1 (the numerator) divided by a denominator. The denominators in the sequence are:

$$1, 2, 3, 4, 5, \ldots$$

The first denominator is 1 because $1 = \frac{1}{1}$.

We recognize this sequence as:

$$a_n = (n + 1), \quad n = 0, 1, 2, \ldots$$

Putting the two ingredients of sign and denominator together, we obtain our explicit formula for a_n:

$$a_n = \frac{(-1)^{n+1}}{n + 1}, \quad n = 0, 1, 2, \ldots$$

Our sequence terms do not need to start at $n = 0$, however. We could instead define the sequence by $a_n = \frac{(-1)^n}{n}$, $n = 1, 2, 3, \ldots$

The exponential growth model we considered in the previous section is an example of a sequence. We gave two descriptions, one explicit and the other recursive. These two descriptions can be used for sequences in general. An explicit description is of the form

$$a_n = f(n), \quad n = 0, 1, 2, \ldots$$

where $f(n)$ is a function of n.

A recursive description is of the form

$$a_{n+1} = g(a_n), \quad n = 0, 1, 2, \ldots$$

where $g(a_n)$ is a function of a_n. If, as is shown here, the value of a_{n+1} depends only on the value one time step back, namely a_n, then the recursion is called a **first-order recurrence equation**. In the review problems, we will see an example of a **second-order recurrence equation**, in which the value of a_{n+1} depends on the values a_n and a_{n-1}—that is, on the values one and two time steps back. To determine the values of successive members of a sequence given in recursive form, we need to specify an initial value a_0 if we start the sequence at $n = 0$ (or a_1 if we start the sequence at $n = 1$).

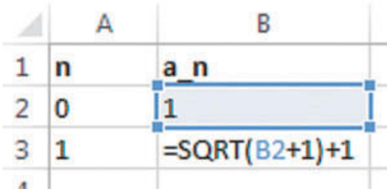
In the notation of this section, the exponential growth of the previous section is given explicitly by

$$a_n = N_0 R^n, \quad n = 0, 1, 2, \ldots$$

and recursively by

$$a_{n+1} = R a_n, \quad n = 0, 1, 2, \ldots \qquad \text{with } a_0 = N_0.$$

Note that, in the recursive definition, the initial value a_0 needs to be specified, as well as the recursion formula.



	A	B
1	n	a_n
2	0	1
3	1	=SQRT(B2+1)+1

Figure 2.11 Using a spreadsheet to calculate the terms in the sequence defined by Equation (2.7).

2.2.2 Using Spreadsheets to Calculate a Recursive Sequence

Using a spreadsheet it is possible to quickly calculate many terms in any sequence that is defined by a recurrence equation. We will explain how to do this calculation, using the specific example of the sequence:

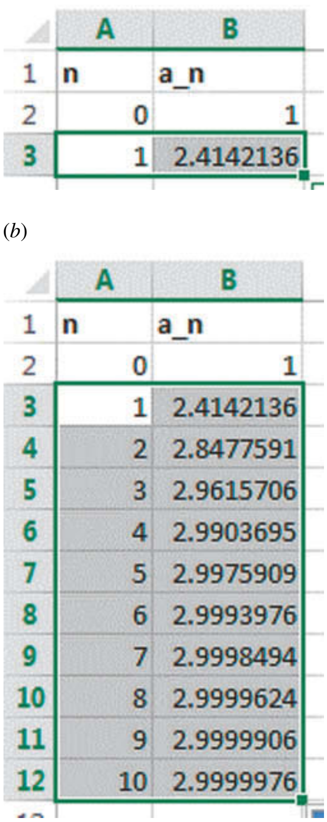
$$a_0 = 1, \quad a_{n+1} = \sqrt{a_n + 1} + 1. \tag{2.7}$$

We will use the column A of the spreadsheet to store the values of the index n for each term in the sequence and column B to store the values of the sequence a_n. Use the cells A1 and B1 to label the columns **n** and **a_n** respectively, and cells A2 and B2 to enter the index (0) and value (1) for the first term in the sequence. To generate the next row we need to use the recursion equation (2.7). In cell A3 enter 1 (the index) and in cell $B3$ enter **= *SQRT*(B2+1)+1**, as shown in Figure 2.11. The value of $B3$ (a_1) will then be computed from the value of B2 (a_0) as the recurrence equation requires.

We can then use the spreadsheet's Autofill command to generate the further terms in the sequence. Select the last row of your table (i.e., the cells $A3$ and $B3$). When you select them, these two cells will be highlighted and surrounded by a colored outline. In the bottom right corner of the outline is a small colored square. Click and hold on the square, and then drag down several rows, as shown in Figure 2.12.

The spreadsheet will automatically fill the new rows using the recursion formula. Specifically it fills A4 with the index 2, A5 with 3, and so on. More importantly, it will put the formula **= *SQRT*(B3 + 1) + 1** in B4. Since B3 holds the value a_1, $B4$ will hold the value $\sqrt{a_1 + 1} + 1$, which is our formula for a_2; $B5$ gets filled with the formula **= *SQRT*(B4+1)+1**, which gives $\sqrt{a_2 + 1} + 1$, the formula for a_3. The number of terms that are calculated in the sequence is the number of rows that we pull down the fill-box.

(*a*)



	A	B
1	n	a_n
2	0	1
3	1	2.4142136

(*b*)

	A	B
1	n	a_n
2	0	1
3	1	2.4142136
4	2	2.8477591
5	3	2.9615706
6	4	2.9903695
7	5	2.9975909
8	6	2.9993976
9	7	2.9998494
10	8	2.9999624
11	9	2.9999906
12	10	2.9999976

Figure 2.12 After entering the formula for a_1 in B3 we use autofill to compute more terms in the sequence (2.7). (a) Select the row containing the recurrence equation from (2.7). (b) Click on the colored box on the bottom right hand corner of the selected cells and drag it down several rows.

2.2.3 Limits

When studying populations over time, we are often interested in their **long-term behavior**. Specifically, if N_t is the population size at time t, we want to know how N_t behaves as t increases. More generally, for a sequence a_n, we want to know the behavior of a_n as n tends to infinity. We say that "we take the limit of the sequence a_n as n goes to infinity" and use the shorthand notation

$$\lim_{n\to\infty} a_n \qquad \text{or} \qquad \lim\nolimits_{n\to\infty} a_n$$

If the limit exists, the sequence is called **convergent** and we say that a_n **converges** to a as n tends to infinity. If the sequence has no limit, it is called **divergent**. Let's first discuss limits informally to get an idea of what can happen.

EXAMPLE 5 Let $a_n = \dfrac{1}{n+1}$, $n = 0, 1, 2, \ldots$ Find $\lim_{n\to\infty} a_n$.

Solution Plugging successive values of n into a_n, we find that a_n is the sequence

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots$$

and we guess that the terms will approach 0 as n tends to infinity.

Since the limiting value is a unique number, we say that the limit exists. •

Note that plugging in successive values of n into a_n only allows us to guess how a_n behaves as $n \to \infty$. We will give rules for calculating limits later in this section.

EXAMPLE 6 Let $a_n = (-1)^n$, $n = 0, 1, 2, \ldots$ Find $\lim_{n\to\infty} a_n$.

Solution The sequence is of the form

$$1, -1, 1, -1, 1, \ldots$$

and we see that its terms alternate between 1 and -1. There is thus no single number we could assign as the limit of a_n as $n \to \infty$. The limit does not exist. •

EXAMPLE 7 Let $a_n = 2^n$, $n = 0, 1, 2, \ldots$ Find $\lim_{n\to\infty} a_n$.

Solution Successive terms of a_n, namely,

$$1, 2, 4, 8, 16, 32, \ldots$$

indicate that the terms continue to grow. Hence, a_n goes to infinity as $n \to \infty$, and we can write $\lim_{n\to\infty} a_n = \infty$. Since infinity ($\infty$) is not a real number, we say that the limit does not exist. •

Let's look at one more example of a limit that exists before we give a formal definition.

EXAMPLE 8 Find $\displaystyle\lim_{n\to\infty} \frac{n+1}{n}$

Solution Starting with $n = 1$ and computing successive terms, we find the first few terms of the sequence are:

$$2, \frac{3}{2}, \frac{4}{3}, \frac{5}{4}, \frac{6}{5}, \ldots$$

Or in decimal values:

$$2, 1.5, 1.33, 1.25, 1.2, \ldots$$

We see that the terms get closer and closer to 1. Indeed, it can be shown that $\lim_{n\to\infty} \frac{n+1}{n} = 1$. •

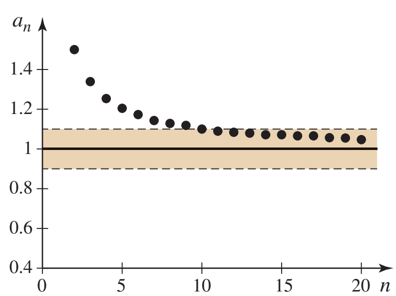


Figure 2.13 Convergence of the sequence $a_n = \frac{n+1}{n}$ to 1. For $n > 10$ all points lie in the interval (0.9, 1.1) shaded in the figure.

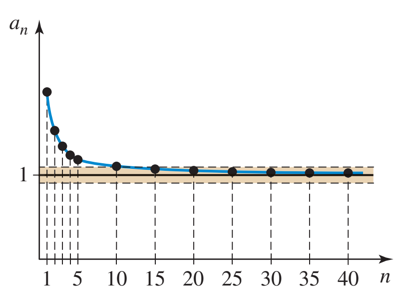


Figure 2.14 For $n > 20$ all points $a_n = \frac{n+1}{n}$ lie in the interval (0.95, 1.05) shaded in the figure. For clarity we show only selected terms from the sequence.

The way we solved the first four examples is unsatisfying: We guessed the limiting values. How do we know that our guesses are correct? For example, the sequence in Example 8 could be converging to 1.05 or 1 or 0.95. It is impossible to be sure what the limit is from just a few computed terms. There is a formal definition of limits that can be used to compute them. However, except in the simplest cases, the formal definition is quite cumbersome to use. Fortunately, there are mathematical laws that build on simple limits. We will first discuss the formal definition as an optional topic and then introduce the limit laws.

Formal Definition of Limits. Example 8 will motivate the formal definition of limits. We guessed in Example 8 that successive terms approached 1. This means that no matter how small an interval about 1 we choose, all points must lie in this interval for all sufficiently large values of n. Let's guess a specific interval size, say 0.1. Then if a_n tends to 1 as $n \to \infty$, all terms in the sequence must lie in the interval $(1-0.1, 1+0.1) = (0.9, 1.1)$ for sufficiently large n. The points of the graph of a_n must lie in the shaded region shown in Figure 2.13.

There is nothing special about the interval (0.9, 1.1); given any small interval around 1, the points in our sequence must all eventually lie in the interval. Suppose, for example, that we took an interval of half the size, namely (0.95, 1.05). For a_n to tend to 1, all terms in the sequence must lie in that interval for sufficiently large n. We find from Figure 2.14 that a_n lies in the smaller interval provided that $n > 20$.

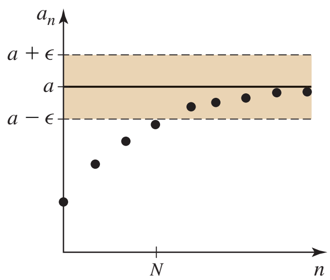


Figure 2.15 The sequence $\{a_n\}$ converges to a as $n \to \infty$. For all $n > N$, a_n lies in the strip of width 2ϵ and centered at a.

Given any interval around 1, a_n must lie in this interval for sufficiently large n. This idea inspires the formal definition of the limit.

> **Definition** **Formal Definition of Limits** The sequence $\{a_n\}$ has **limit** a, written as $\lim_{n\to\infty} a_n = a$, if, given any $\epsilon > 0$, there exists a number N such that
>
> $$|a_n - a| < \epsilon \qquad \text{whenever } n > N$$

The value of N will typically depend on ϵ: The smaller ϵ is, the larger N is. We illustrate the concept of a converging sequence in Figure 2.15. The interval $a - \epsilon < a_n < a + \epsilon$ is shaded. It forms a strip of width 2ϵ centered at a. Points a_n within this shaded strip satisfy the inequality $|a_n - a| < \epsilon$. For a sequence to be convergent, we require that *all* points a_n lie in this strip for *all* n sufficiently large (namely, larger than some N).

EXAMPLE 9 Show that $\lim\limits_{n\to\infty} \dfrac{1}{n} = 0$.

Solution Before we show this for any arbitrary ϵ, let's try to find N for a particular choice of ϵ, say, $\epsilon = 0.03$. We need to find an integer N such that

$$\left|\frac{1}{n} - 0\right| < 0.03 \quad \text{whenever } n > N$$

Solving the inequality $|\frac{1}{n} - 0| < 0.03$ for n positive, we find that

$$\left|\frac{1}{n}\right| < 0.03, \qquad \text{or} \qquad n > \frac{1}{0.03} \approx 33.33$$

Absolute value can be ignored since $\frac{1}{n}$ is positive.

So for $\epsilon = 0.03$, $|\frac{1}{n} - 0| < \epsilon$ whenever $n > 33.33$.

But to show that $a_n = \frac{1}{n}$ converges to 0, we need to do the same calculation for any arbitrary ϵ. That is, we need to show that, for every $\epsilon > 0$, we can find an N such that

$$\left|\frac{1}{n} - 0\right| < \epsilon \quad \text{whenever } n > N$$

To find a candidate for N, we solve the inequality $|\frac{1}{n}| < \epsilon$.

$$\frac{1}{n} < \epsilon, \qquad \text{or} \qquad n > \frac{1}{\epsilon}$$

Since $\frac{1}{n} > 0$ we can drop the absolute value signs.

Hence if $N = \frac{1}{\epsilon}$ then:

$$\left|\frac{1}{n} - 0\right| < \epsilon \quad \text{whenever } n > N,$$

so the formal definition of a limit is satisfied. •

We will often encounter sequences that do not have finite limits. It is often useful to deal with cases where $a_n \to \infty$ as $n \to \infty$, such as in Example 7.

As a formal definition, we say that

$$\lim_{n\to\infty} a_n = \infty,$$

if, given any $A > 0$, there exists a number N such that $a_n > A$ whenever $n > N$. Informally this means that the terms in a_n get larger and larger without bound as $n \to \infty$.

We also need to discriminate between sequences whose limits are ∞ and $-\infty$. The sequence a_n converges to $-\infty$ if, given any $A < 0$, there exists a number $N > 0$ such that $a_n < A$ for all $n > N$. Informally the terms in $\{a_n\}$ are negative but get larger and larger in absolute value as n increases.

Limit Laws. The formal definition of limits is cumbersome when we want to compute limits in specific examples. Fortunately, there are mathematical laws that we can use:

Limit Laws If $\lim_{n\to\infty} a_n$ and $\lim_{n\to\infty} b_n$ exist and c is a constant, then

1. $\lim\limits_{n\to\infty}(a_n + b_n) = \lim\limits_{n\to\infty} a_n + \lim\limits_{n\to\infty} b_n$
2. $\lim\limits_{n\to\infty}(ca_n) = c \lim\limits_{n\to\infty} a_n$
3. $\lim\limits_{n\to\infty}(a_n b_n) = (\lim\limits_{n\to\infty} a_n)(\lim\limits_{n\to\infty} b_n)$
4. $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{\lim\limits_{n\to\infty} a_n}{\lim\limits_{n\to\infty} b_n}$, provided $\lim\limits_{n\to\infty} b_n \neq 0$

We can also modify our limit laws if $\lim_{n\to\infty} a_n = \infty$ or $\lim_{n\to\infty} b_n = \infty$

Limit Laws Involving Infinite Limits

5. If $\lim\limits_{n\to\infty} a_n = \infty$ and $\lim\limits_{n\to\infty} b_n$ exists and is finite, then $\lim\limits_{n\to\infty}(a_n + b_n) = \infty$.
6. If $\lim\limits_{n\to\infty} a_n = \infty$ and $c > 0$ is a constant, then $\lim\limits_{n\to\infty} c \cdot a_n = \infty$.
   If $c < 0$ is a constant, then $\lim\limits_{n\to\infty} c \cdot a_n = -\infty$.
7. If $\lim\limits_{n\to\infty} a_n = \infty$ and $\lim\limits_{n\to\infty} b_n = b$ and $b > 0$, then $\lim\limits_{n\to\infty}(a_n \cdot b_n) = \infty$.
   If $b < 0$ then $\lim\limits_{n\to\infty}(a_n \cdot b_n) = -\infty$.
8. If $\lim\limits_{n\to\infty} a_n = \infty$ and $\lim\limits_{n\to\infty} b_n = b$ with $b > 0$, then $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = \infty$.
   If $b < 0$, then $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = -\infty$.
9. If $\lim\limits_{n\to\infty} b_n = \infty$ and $\lim\limits_{n\to\infty} a_n$ exists and is finite, then $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = 0$.

In Example 9 we proved using the formal definition of limit that $\lim_{n\to\infty} \frac{1}{n} = 0$. We can now see that this is a consequence of limit law rule (9); if we define a sequence $a_n = 1$ (i.e., a_n is the sequence $1, 1, 1, \ldots$) and $b_n = n$, then $\lim_{n\to\infty} b_n = \infty$, so $\lim_{n\to\infty} \frac{a_n}{b_n} \equiv \lim_{n\to\infty} \frac{1}{n} = 0$. We will now use this result to calculate the limits of other sequences.

EXAMPLE 10 Find $\lim\limits_{n\to\infty} \dfrac{n+1}{n}$

Solution We cannot use rule (4) or rule (9) for this sequence because both numerator and denominator converge to ∞ as $n \to \infty$. Instead we break the sequence into the sum of two sequences, noting that $\frac{n+1}{n} = 1 + \frac{1}{n}$; so

$$\begin{aligned} \lim_{n\to\infty} \frac{n+1}{n} &= \lim_{n\to\infty} 1 + \lim_{n\to\infty} \frac{1}{n} && \text{Rule (1) with } a_n = 1,\ b_n = \frac{1}{n}. \\ &= 1 + 0 = 1 && \text{Using Example 9 to calculate the limit of } \frac{1}{n}. \end{aligned}$$

•

EXAMPLE 11 Find $\lim\limits_{n\to\infty} \dfrac{4n^2 - 1}{n^2}$.

Solution We rewrite a_n:

$$a_n = \frac{4n^2 - 1}{n^2} = 4 - \frac{1}{n^2} = 4 - \frac{1}{n} \cdot \frac{1}{n}$$

Since $\lim_{n\to\infty} 4$ and $\lim_{n\to\infty} \frac{1}{n}$ exist, we have

$$\begin{aligned}
\lim_{n\to\infty} \frac{4n^2-1}{n^2} &= \lim_{n\to\infty} 4 - \lim_{n\to\infty} \frac{1}{n^2} && \text{Rule (1) with } a_n = 4,\ b_n = \frac{1}{n^2}. \\
&= 4 - \lim_{n\to\infty} \frac{1}{n} \cdot \lim_{n\to\infty} \frac{1}{n} && \text{Rule (3) with } a_n = \frac{1}{n},\ b_n = \frac{1}{n}. \\
&= 4 - 0 \cdot 0 && \text{Use Example 9 to calculate the limit of } \frac{1}{n}. \\
&= 4.
\end{aligned}$$

•

EXAMPLE 12 Without proof, we will state the long-term behavior of exponential growth. For $R > 0$, exponential growth is given by

$$a_n = a_0 R^n,\ n = 0, 1, 2, \ldots$$

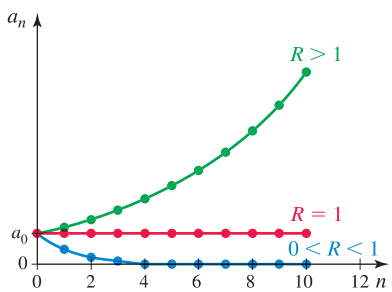


Figure 2.16 Exponential growth for three different values of R. As a guide to the eye we have added lines to connect the discrete points $n = 0, 1, 2, 3, \ldots$, where the sequence is defined.

Figure 2.16 indicates that

$$\lim_{n\to\infty} a_n = \begin{cases} 0 & \text{if } 0 < R < 1 \\ a_0 & \text{if } R = 1 \\ \infty & \text{if } R > 1 \end{cases}$$

This conclusion can also be shown rigorously by using the formal definition of limits.

•

2.2.4 Recurrence Equations

In the previous subsection, we learned how to find $\lim_{n\to\infty} a_n$ when a_n is given explicitly as a function of n. We will now discuss how to find such a limit when a_n is defined recursively.

When we define a first-order sequence $\{a_n\}$ recursively, we express a_{n+1} in terms of a_n and specify a value for a_0. By calculating a few terms $a_0, a_1, a_2, \ldots$ we may be able to guess the limit if it exists. In some cases (as in the next example), we can find a solution of the recurrence equation and then determine the limit (if it exists), using the rules given in Subsection 2.2.3.

EXAMPLE 13 Compute a_n for $n = 1, 2, \ldots, 5$ when

$$a_{n+1} = \frac{1}{4}a_n + \frac{3}{4} \quad \text{with } a_0 = 2 \tag{2.8}$$

Find a solution of the recurrence equation, and then take a guess at the limiting behavior of the sequence.

Solution By repeatedly applying the recurrence equation, we find that

$$\begin{aligned}
a_1 &= \frac{1}{4}a_0 + \frac{3}{4} = \frac{1}{4}\cdot 2 + \frac{3}{4} = \frac{5}{4} = 1.25 \\
a_2 &= \frac{1}{4}a_1 + \frac{3}{4} = \frac{1}{4}\cdot\frac{5}{4} + \frac{3}{4} = \frac{17}{16} = 1.0625 \\
a_3 &= \frac{1}{4}a_2 + \frac{3}{4} = \frac{1}{4}\cdot\frac{17}{16} + \frac{3}{4} = \frac{65}{64} \approx 1.0156 \\
a_4 &= \frac{1}{4}a_3 + \frac{3}{4} = \frac{1}{4}\cdot\frac{65}{64} + \frac{3}{4} = \frac{257}{256} \approx 1.0039 \\
a_5 &= \frac{1}{4}a_4 + \frac{3}{4} = \frac{1}{4}\cdot\frac{257}{256} + \frac{3}{4} = \frac{1025}{1024} \approx 1.0010.
\end{aligned}$$

There seems to be a pattern, namely, that the denominators are powers of 4 and the numerators are just 1 larger than the denominators. We guess that the nth term is given by the explicit formula:

$$a_n = \frac{4^n + 1}{4^n} \tag{2.9}$$

and check whether this is indeed a solution of the recursion. First, we need to check the initial condition:

$$a_0 = \frac{4^0 + 1}{4^0} = \frac{2}{1} = 2.$$

So $a_0 = 2$, which was the initial condition given. Next, we need to check whether a_n satisfies the recursion. If we substitute $a_n = \frac{4^n+1}{4^n}$ into the recursion relation, we obtain:

$$\begin{aligned} a_{n+1} &= \frac{1}{4}a_n + \frac{3}{4} = \frac{1}{4} \cdot \left(\frac{4^n + 1}{4^n}\right) + \frac{3}{4} \\ &= \frac{4^n + 1}{4^{n+1}} + \frac{3}{4} && \frac{1}{4} \cdot \frac{1}{4^n} = \frac{1}{4^{n+1}}. \\ &= \frac{4^n + 1 + 3 \cdot 4^n}{4^{n+1}} && \text{Use a common denominator, } \frac{3}{4} = \frac{3 \cdot 4^n}{4^{n+1}} \\ &= \frac{4 \cdot 4^n + 1}{4^{n+1}} = \frac{4^{n+1} + 1}{4^{n+1}} \end{aligned}$$

And this agrees with our formula for a_{n+1}. Showing that the formula (2.9) satisfies both the recursion relation and the initial condition proves that it is the solution of the recursion relation (2.8).

We can now use (2.9) to find the limit of the sequence. We have

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{4^n + 1}{4^n} = \lim_{n\to\infty} \left(1 + \frac{1}{4^n}\right) = 1 \qquad \text{Use rule (1) then rule (9).}$$

Finding an explicit expression for a_n as in Example 13 is often not a feasible strategy, because solving recursions can be very difficult or even impossible. How, then, can we say anything about the limiting behavior of a recursively defined sequence?

The following procedure will allow us to identify *candidates* for limits:

Definition Given a recursion relation: $a_{n+1} = g(a_n)$, the point a is a **fixed point** of the recursion relation if:

$$a = g(a)$$

If a is a fixed point and $a_0 = a$ then $a_1 = g(a_0) = g(a) = a$, and $a_2 = g(a_1) = g(a) = a$, and so on. So if $a_0 = a$, every subsequent term of the sequence will also be equal to a.

In many cases, if the limit of the sequence $\lim_{n\to\infty} a_n = a$ exists, then a must be a fixed point of the recursion relation. (The function $g(a_n)$ must be continuous at $a_n = a$. We will discuss continuous functions in Chapter 3.) So we can find candidate values for the limit by finding fixed points by solving $a = g(a)$. Let's apply this technique to Example 13. The recursion formula is $a_{n+1} = g(a_n)$, where $g(a_n) = \frac{1}{4} \cdot a_n + \frac{3}{4}$. Fixed points for the recursion satisfy

$$a = \frac{1}{4}a + \frac{3}{4} \qquad g(a) = \frac{1}{4}a + \frac{3}{4}$$

Solving this equation for a, we find that $a = 1$. If the limit exists, it must be a fixed point. Since there is only one fixed point ($a = 1$), we deduce that the limit, if it exists, is $\lim_{n\to\infty} a_n = 1$. A fixed point is only a *candidate* for a limit; a sequence does not have to converge to a given fixed point as the next two examples show.

EXAMPLE 14 Assume that $\lim_{n\to\infty} a_n$ exists for $a_{n+1} = \sqrt{3a_n}$ with $a_0 = 2$. Find $\lim_{n\to\infty} a_n$.

Solution Since the problem tells us that the limit exists, we don't have to worry about existence. The problem that remains is to identify the limit. To do this, we compute the fixed points. We solve

$$a = \sqrt{3a}$$

which has two solutions, namely, $a = 0$ and $a = 3$. Since the recursion relation has two fixed points there are two candidates for the limit, $a = \lim_{n\to\infty} a_n$, namely, $a = 0$ or $a = 3$. We must eliminate one of these points.

Notice that the first three terms of the sequence are $a_0 = 2$, $a_1 = \sqrt{6} = 2.45$, $a_2 = \sqrt{3 \cdot 2.45} = 2.71$. The terms in $\{a_n\}$ appear to be increasing; in particular $a_n \geq 2$ for all n. If we can show that to be true, then we can exclude $a = 0$ as a limit.

Now if $a_n \geq 2$, then $a_{n+1} = \sqrt{3a_n} \geq \sqrt{3 \times 2} > 2$. So a_n cannot converge to 0 as $n \to \infty$. This rules out one of the possible limits, and we conclude that $\lim_{n\to\infty} a_n = 3$.

Using a calculator, we calculate the first few terms in the sequence (accurate to 2 decimal places):

n	0	1	2	3	4	5	6	7
a_n	2	2.45	2.71	2.85	2.92	2.96	2.98	2.99

The tabulated values suggest that the limit is indeed 3. •

EXAMPLE 15 Let $a_{n+1} = \dfrac{3}{a_n}$. Find the fixed points of this recursion. Does the sequence converge to either fixed point if $a_0 = 2$?

Solution To find the fixed points, we need to solve

$$a = \frac{3}{a} \qquad g(a) = 3/a$$

This equation is equivalent to $a^2 = 3$; hence, $a = \sqrt{3}$ or $a = -\sqrt{3}$. These are the two fixed points. If $a_0 = \sqrt{3}$, then $a_1 = \sqrt{3}$, $a_2 = \sqrt{3}$, and so on, and likewise, if $a_0 = -\sqrt{3}$, then $a_1 = -\sqrt{3}$, $a_2 = -\sqrt{3}$, and so on.

If $a_0 = 2$, then using the recursion, we find that:

$$\begin{aligned} a_1 &= \frac{3}{a_0} = \frac{3}{2} \\ a_2 &= \frac{3}{a_1} = \frac{3}{\frac{3}{2}} = 3 \cdot \frac{2}{3} = 2 \qquad \frac{1}{a/b} = \frac{b}{a} \\ a_3 &= \frac{3}{a_2} = \frac{3}{2} \\ a_4 &= \frac{3}{a_3} = \frac{3}{\frac{3}{2}} = 2 \end{aligned}$$

and so on. That is, successive terms alternate between 2 and 3/2 so the limit of the sequence doesn't exist. This behavior is not special to our choice of initial conditions. Let's try another initial value, say $a_0 = -3$, then:

$$\begin{aligned} a_1 &= \frac{3}{a_0} = \frac{3}{-3} = -1 \\ a_2 &= \frac{3}{a_1} = \frac{3}{-1} = -3 \\ a_3 &= \frac{3}{a_2} = \frac{3}{-3} = -1 \\ a_4 &= \frac{3}{a_3} = \frac{3}{-1} = -3 \end{aligned}$$

and so on. Successive terms now alternate between -3 and -1. Alternating between two values, one of which is the initial value, happens with any initial value that is not one of the fixed points. Specifically, we have

$$a_1 = \frac{3}{a_0} \qquad \text{and} \qquad a_2 = \frac{3}{a_1} = \frac{3}{\frac{3}{a_0}} = a_0$$

Thus, a_3 is the same as a_1, a_4 is the same as a_2 and hence a_0, and so on. •

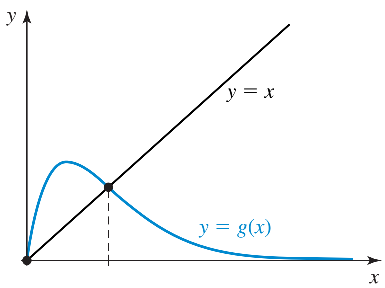


Figure 2.17 A graphical way to find fixed points. Fixed points are points at which $y = g(x)$ and $y = x$ intersect.

The last two examples illustrate that fixed points are only *candidates* for the limit and that, depending on the initial condition, the sequence $\{a_n\}$ may or may not converge to a given fixed point. If we know, however, that a sequence $\{a_n\}$ does converge, then the limit of the sequence must be one of the fixed points, provided $g(a_n)$ is a continuous function.

There is a graphical method for finding fixed points, which we will mention briefly here: If the recursion is of the form $a_{n+1} = g(a_n)$, then a fixed point satisfies $a = g(a)$. This suggests that if we graph $y = g(x)$ and $y = x$ on the same axes, then fixed points are located where the two graphs intersect, as shown in Figure 2.17.

We will return to the relationship between fixed points and limits in Section 5.7, where we will learn methods to determine whether a sequence converges to a particular fixed point.

2.2.5 Using Σ Notation to Represent Sums of Sequences

It will be convenient to have a shorthand notation for sums that involve a large number of terms:

> **Definition Sigma Notation for Finite Sums** Let $a_1, a_2, \cdots, a_n$ be a sequence of real numbers and n be a positive integer. Then
>
> $$\sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n$$

The letter Σ is the capital Greek letter sigma, and the symbol $\sum_{k=1}^{n}$ means that we sum from $k = 1$ to $k = n$, where k is called the *index of summation*, the number 1 is the lower limit of summation, and the number n is the upper limit of summation. In text, instead of writing $\displaystyle\sum_{k=1}^{n}$ we will write $\sum_{k=1}^{n}$.

EXAMPLE 16

(a) Write each sum in expanded form:

(i) $\displaystyle\sum_{k=1}^{4} k$ **(ii)** $\displaystyle\sum_{k=3}^{6} k^2$ **(iii)** $\displaystyle\sum_{k=1}^{n} \frac{1}{k}$ **(iv)** $\displaystyle\sum_{k=1}^{5} 1$

(b) Write each sum in sigma notation:

(i) $2 + 3 + 4 + 5$

(ii) $1^3 + 2^3 + 3^3 + 4^3$

(iii) $1 + 3 + 5 + 7 + \cdots + (2n + 1)$

(iv) $x + 2x^2 + 3x^3 + \cdots + nx^n$

Solution

(a) **(i)** $1 + 2 + 3 + 4$ **(ii)** $3^2 + 4^2 + 5^2 + 6^2$

(iii) $1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$ **(iv)** $1 + 1 + 1 + 1 + 1$

(b) **(i)** $\displaystyle\sum_{k=2}^{5} k$ **(ii)** $\displaystyle\sum_{k=1}^{4} k^3$ **(iii)** $\displaystyle\sum_{k=0}^{n} (2k + 1)$ **(iv)** $\displaystyle\sum_{k=1}^{n} kx^k$ •

In general Σ-notation gives a shorthand for writing sums, but it does not tell us what the sum evaluates to. We have to write out the terms long hand to evaluate the sum.

The following rules are useful in evaluating finite sums:

Rules for Evaluating Sequences Using Σ-Notation

1. Constant-value rule: $\sum_{k=1}^{n} 1 = n$
2. Constant-multiple rule: $\sum_{k=1}^{n} c \cdot a_k = c \sum_{k=1}^{n} a_k$, where c is a constant that does not depend on k
3. Sum rule: $\sum_{k=1}^{n} (a_k + b_k) = \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k$

EXAMPLE 17 Use the algebraic rules to simplify the following sums:

(a) $\sum_{k=2}^{4} (3+k) = \sum_{k=2}^{4} 3 + \sum_{k=2}^{4} k = (3+3+3) + (2+3+4) = 18$ Use the sum rule

(b) $\sum_{k=2}^{4} (1+k^2) = \sum_{k=2}^{4} 1 + \sum_{k=2}^{4} k^2 = (1+1+1) + (2^2+3^2+4^2) = 32$ Use the sum rule

Section 2.2 Problems

2.2.1

In Problems 1–16, determine the values of the sequence $\{a_n\}$ for $n = 0, 1, 2, \ldots, 5$.

1. $a_n = n + 1$

2. $a_n = 3n^2$

3. $a_n = \frac{n+2}{n}$

4. $a_n = \frac{n}{n+2}$

5. $f(n) = \frac{1}{(1+n)^2}$

6. $a_n = \frac{1}{\sqrt{n+1}}$

7. $f(n) = (n+1)^2$

8. $f(n) = (n+4)^{1/3}$

9. $a_n = (-1)^n + (-1)^{n+1}$

10. $a_n = (-1)^n + 1$

11. $a_n = \frac{n^2}{n+1}$

12. $a_n = n^3\sqrt{n+1}$

13. $f(n) = e^{n/2}$

14. $f(n) = \log(n+1)$

15. $f(n) = \left(\frac{1}{3}\right)^n$

16. $f(n) = 2^{0.2n}$

In Problems 17–24, find the next four values of the sequence $\{a_n\}$ on the basis of the values of a_0, a_1, a_2, a_3, a_4.

17. $1, 4, 9, 16, 25$

18. $0, 1, \sqrt{2}, \sqrt{3}, \sqrt{4}$

19. $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}$

20. $-1, \frac{1}{4}, -\frac{1}{9}, \frac{1}{16}, -\frac{1}{25}$

21. $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}$

22. $5, 10, 17, 26, 37$

23. $\sqrt{1+e}, \sqrt{2+e^2}, \sqrt{3+e^3}, \sqrt{4+e^4}, \sqrt{5+e^5}$

24. $1, 3, 9, 27, 81$

In Problems 25–36, find an expression for a_n on the basis of the values of $a_0, a_1, a_2, \ldots$.

25. $0, 1, 2, 3, 4, \ldots$

26. $0, 2, 4, 6, 8, \ldots$

27. $1, 2, 4, 8, 16, \ldots$

28. $1, 3, 5, 7, 9, \ldots$

29. $1, \frac{1}{3}, \frac{1}{9}, \frac{1}{27}, \frac{1}{81}, \ldots$

30. $\frac{1}{3}, \frac{2}{5}, \frac{3}{7}, \frac{4}{9}, \frac{5}{11}, \ldots$

31. $-1, 2, -3, 4, -5, \ldots$

32. $9, 16, 25, 36, 49$

33. $5, 7, 9, 11, 13$

34. $8, 18, 32, 50, 72$

35. $2, 0, 2, 0, 2$

36. $0, 1, 2, 0, 1, 2$

2.2.2

T ***In Problems 37–44, use a spreadsheet to calculate the specified term of each recursively defined sequence.***

37. If $a_{n+1} = \sqrt{a_n + 1}$ and $a_0 = 1$, find a_{11}.

38. If $a_{n+1} = \frac{1}{a_n+1}$ and $a_0 = 2$, find a_{13}.

39. If $a_{n+1} = a_n - \frac{1}{a_n}$ and $a_0 = 3$, find a_{11}.

40. If $a_{n+1} = a_n + \frac{1}{a_n}$ and $a_0 = 1$, find a_{13}.

41. If $a_{n+1} = \sqrt{\sqrt{a_n} + 1}$ and $a_0 = 6$ find a_{12}.

42. If $a_{n+1} = a_n + \frac{1}{a_n^2}$ and $a_0 = 1$ find a_{12}.

43. If $a_{n+1} = \frac{1}{4}a_n + 1$ and $a_0 = 0$, find a_{14}.

44. If $a_{n+1} = \sqrt{a_n^2 + 1}$ and $a_0 = 1$, find a_{16}.

2.2.3

In Problems 45–52, write the first five terms of the sequence $\{a_n\}$, $n = 0, 1, 2, 3, \ldots$, and find $\lim_{n\to\infty} a_n$.

45. $a_n = \dfrac{1}{n+1}$

46. $a_n = \dfrac{3}{n+3}$

47. $a_n = \dfrac{n}{n+1}$

48. $a_n = \dfrac{2n}{(n+2)^2}$

49. $a_n = \dfrac{n^2+5}{n^2+1}$

50. $a_n = \dfrac{1}{\sqrt{n+1}}$

51. $a_n = \dfrac{(-1)^{n+1}}{n+1}$

52. $a_n = \dfrac{(-1)^n}{n^3+3}$

In Problems 53–60, write the first five terms of the sequence $\{a_n\}$, $n = 0, 1, 2, 3, \ldots$, and determine whether $\lim_{n\to\infty} a_n$ exists. If the limit exists, find it.

53. $a_n = \dfrac{n^2}{n+1}$

54. $a_n = \dfrac{n+1}{n+2}$

55. $a_n = \sqrt{n}$

56. $a_n = n^2 + 3$

57. $a_n = 2^n$

58. $a_n = 2^{n+3}$

59. $a_n = 3^n$

60. $a_n = 3^{-2n}$

Formal Definition of Limits: ***In Problems 61–72, $\lim_{n\to\infty} a_n = a$. Find the limit a, and determine N so that $|a_n - a| < \epsilon$ for all $n > N$ for the given value of ϵ.***

61. $a_n = \dfrac{1}{n}$, $\epsilon = 0.01$

62. $a_n = \dfrac{1}{n}$, $\epsilon = 0.02$

63. $a_n = \dfrac{1}{n^2}$, $\epsilon = 0.01$

64. $a_n = \dfrac{1}{n^2}$, $\epsilon = 0.001$

65. $a_n = \dfrac{1}{\sqrt{n}}$, $\epsilon = 0.1$

66. $a_n = \dfrac{1}{\sqrt{n}}$, $\epsilon = 0.05$

67. $a_n = \dfrac{(-1)^n}{n}$, $\epsilon = 0.01$

68. $a_n = e^{-n}$, $\epsilon = 0.01$

69. $a_n = e^{-3n}$, $\epsilon = 0.001$

70. $a_n = \ln\left(1 + \dfrac{1}{n}\right)$, $\epsilon = 0.1$

71. $a_n = 2^{-n}$, $\epsilon = 0.01$

72. $a_n = \log\left(1 + \dfrac{2}{n^2}\right)$, $\epsilon = 0.05$

Formal Definition of Limits: ***In Problems 73–78, use the formal definition of limits to show that $\lim_{n\to\infty} a_n = a$; that is, find N such that for every $\epsilon > 0$, there exists an N such that $|a_n - a| < \epsilon$ whenever $n > N$.***

73. $\lim\limits_{n\to\infty} \dfrac{3}{n} = 0$

74. $\lim\limits_{n\to\infty} \dfrac{1}{n+1} = 0$

75. $\lim\limits_{n\to\infty} \dfrac{1}{n^2} = 0$

76. $\lim\limits_{n\to\infty} e^{-2n} = 0$

77. $\lim\limits_{n\to\infty} 2^{-3n} = 0$

78. $\lim\limits_{n\to\infty} \dfrac{n}{n+1} = 1$

In Problems 79–90, use the limit laws to determine $\lim\limits_{n\to\infty} a_n = a$.

79. $\lim\limits_{n\to\infty} \left(\dfrac{1}{n} + \dfrac{2}{n^2}\right)$

80. $\lim\limits_{n\to\infty} \left(\dfrac{2}{n} - \dfrac{3}{n^2+1}\right)$

81. $\lim\limits_{n\to\infty} \left(\dfrac{n+1}{n}\right)$

82. $\lim\limits_{n\to\infty} \left(\dfrac{2n-3}{\sqrt{n}}\right)$

83. $\lim\limits_{n\to\infty} \left(\dfrac{n^2+1}{n^2}\right)$

84. $\lim\limits_{n\to\infty} \left(\dfrac{3n^2-5}{n}\right)$

85. $\lim\limits_{n\to\infty} \left(\dfrac{n+1}{n^2-1}\right)$

86. $\lim\limits_{n\to\infty} \left(\dfrac{n^2}{n^2+4}\right)$

87. $\lim\limits_{n\to\infty} \left[\left(\dfrac{1}{3}\right)^n + 2^n\right]$

88. $\lim\limits_{n\to\infty} (3^{-n} - 4^{-n})$

89. $\lim\limits_{n\to\infty} \dfrac{n + 2^{-n}}{n}$

90. $\lim\limits_{n\to\infty} \dfrac{1 + e^{-n}}{n}$

2.2.4

In Problems 91–100, the sequence $\{a_n\}$ is recursively defined. Compute a_n for $n = 1, 2, \ldots, 5$.

91. $a_{n+1} = 2a_n$, $a_0 = 1$

92. $a_{n+1} = 2a_n$, $a_0 = 3$

93. $a_{n+1} = -2a_n$, $a_0 = 1$

94. $a_{n+1} = -2a_n$, $a_0 = 2$

95. $a_{n+1} = 1 + 2a_n$, $a_0 = 0$

96. $a_{n+1} = 4 - 2a_n$, $a_0 = \dfrac{4}{3}$

97. $a_{n+1} = \dfrac{a_n}{1 + a_n}$, $a_0 = 1$

98. $a_{n+1} = \sqrt{a_n}$, $a_0 = 16$

99. $a_{n+1} = a_n + \dfrac{1}{a_n}$, $a_0 = 1$

100. $a_{n+1} = 2a_n{}^2$, $a_0 = 1$

In Problems 101–110, the sequence $\{a_n\}$ is recursively defined. Find all fixed points of $\{a_n\}$.

101. $a_{n+1} = \dfrac{1}{2}a_n + 2$

102. $a_{n+1} = \dfrac{1}{3}a_n + \dfrac{4}{3}$

103. $a_{n+1} = \dfrac{5}{2} - \dfrac{1}{2}a_n$

104. $a_{n+1} = a_n{}^2 - a_n$

105. $a_{n+1} = \dfrac{4}{a_n}$

106. $a_{n+1} = \dfrac{4}{a_n - 3}$

107. $a_{n+1} = \dfrac{2}{a_n + 2}$

108. $a_{n+1} = \dfrac{8}{\sqrt{a_n}}$

109. $a_{n+1} = \sqrt{5a_n}$

110. $a_{n+1} = \sqrt{a_n + 2}$

In Problems 111–118, assume that $\lim_{n\to\infty} a_n$ exists. Find all fixed points of $\{a_n\}$, and use a table or other reasoning to guess which fixed point is the limiting value for the given initial condition.

111. $a_{n+1} = \dfrac{1}{3}a_n + \dfrac{4}{3}$, $a_0 = 0$

112. $a_{n+1} = \dfrac{1}{3}\left(a_n + \dfrac{1}{9}\right)$, $a_0 = 1$

113. $a_{n+1} = \sqrt{2a_n}$, $a_0 = 1$

114. $a_{n+1} = \dfrac{3}{a_n + 2}$, $a_0 = 0$

115. $a_{n+1} = 2a_n(1 - a_n)$, $a_0 = 0.1$

116. $a_{n+1} = 2a_n(1 - a_n)$, $a_0 = 0$

117. $a_{n+1} = \dfrac{1}{3}\left(a_n + \dfrac{2}{a_n}\right)$, $a_0 = 3$

118. $a_{n+1} = \dfrac{1}{2}\left(a_n + \dfrac{9}{a_n}\right)$, $a_0 = -1$

2.2.5

In Problems 119–124, write each sum in expanded form.

119. $\displaystyle\sum_{k=1}^{4} \sqrt{k}$

120. $\displaystyle\sum_{k=3}^{5} (k-1)^2$

121. $\displaystyle\sum_{k=2}^{6} 3^k$

122. $\displaystyle\sum_{k=1}^{3} \dfrac{k^2}{k^2+1}$

123. $\sum_{n=0}^{3} a_n$ where $a_0 = 1$ and $a_{n+1} = 2a_n$

124. $\sum_{n=0}^{4} a_n$ where $a_0 = 2$ and $a_{n+1} = a_n + 2$

In Problems 125–132, write each sum in sigma notation.

125. $2 + 4 + 6 + 8 + \cdots + 2n$

126. $\dfrac{1}{\sqrt{1}} + \dfrac{1}{\sqrt{2}} + \dfrac{1}{\sqrt{3}} + \dfrac{1}{\sqrt{4}}$

127. $\ln 2 + \ln 3 + \ln 4 + \ln 5$

128. $\dfrac{3}{5} + \dfrac{4}{6} + \dfrac{5}{7} + \dfrac{6}{8} + \dfrac{7}{9}$

129. $-\dfrac{1}{4} + \dfrac{1}{6} + \dfrac{2}{7} + \dfrac{3}{8}$

130. $\dfrac{1}{1} + \dfrac{1}{2} + \dfrac{1}{4} + \dfrac{1}{8} + \dfrac{1}{16} + \cdots + \dfrac{1}{2^n}$

131. $1 + q + q^2 + q^3 + q^4 + \cdots + q^{n-1}$

132. $1 - a + a^2 - a^3 + a^4 - a^5 + \cdots + (-1)^n a^n$

2.3 Modeling with Recurrence Equations

In Section 2.1 we introduced recurrence equations as a method for modeling populations that are growing or decaying and in which the population size is measured at a discrete set of times (for example, a population of dividing bacteria that is counted once per hour). Recursion relations can be used to model many different biological phenomena. In this section we will introduce models for the density-dependent growth of populations and for the passage of drugs through a body.

2.3.1 Density-Dependent Population Growth

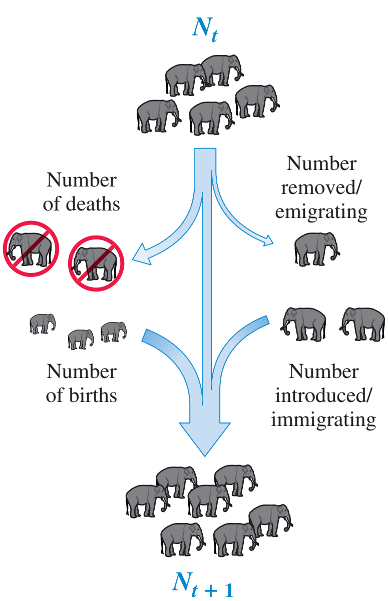


Figure 2.18 Using births, deaths, introductions and removals to calculate the change in a population of organisms.

Often it is only possible to measure population size at discrete times; for example, we might count the number of bacteria growing in a flask once per hour, or the number of red-tailed hawks nesting on the UCLA campus once per month, or the number of elephants in a Sri Lankan national park once per year. It is often useful to have a mathematical model that enables the population size in the next measurement to be predicted from the current measurement of the population size. Specifically, if the sequence of measured population sizes is $\{N_t, t = 0, 1, 2, \ldots\}$, this model will take the form of a function $f(N)$ defined so that:

$$N_{t+1} = f(N_t)$$

What information do we need to calculate this function? Consider a population of elephants living in a national park in Sri Lanka. If there are no births, deaths, and no elephants enter or leave the population, then the number of elephants present at time $t + 1$ will be identical to the number present at time t; that is, $N_{t+1} = N_t$.

How might the population size change? Any births that occur and any elephants that are introduced into or immigrate into the park must be added to N_t (see Figure 2.18). Any deaths that occur and any elephants that are removed or that emigrate out of the park must be subtracted from N_t. In words we can write:

$$\begin{aligned} N_{t+1} = N_t &+ \text{number of organisms born} + \text{number of organisms introduced or immigrated into population.} \\ &- \text{number of organisms that die} - \text{number of organisms removed or emigrated out of population.} \end{aligned}$$

Often the number of births, deaths, immigrations, or removals can be related to the population size N_t. We will consider in this section different models for these effects. Because the models can be used to predict the change in size, or $N_{t+1} - N_t$, they are often called **difference equations**.

EXAMPLE 1

Elephant Population A national park contains a population of elephants that is counted once per year. In a population of N_t elephants, half of the elephants will be female and half of the females will be of reproductive age. Among the female elephants of reproductive age, one-fifth will give birth to calves in a given year. Assume that one elephant in 30 will die in each year, and that there are no migrations into or out of the park. Starting with a population of 120 elephants in year 0, compute the population size after 1, 2, 3, 4, and 5 years.

Solution We first use the information to compute the recursion relation between N_{t+1} and N_t:

$$N_{t+1} = N_t + \text{number of births} - \text{number of deaths}$$

There are no emigrations or introductions, so these effects do not need to be included.

We are told that $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ of the elephants are female and of reproductive age, and $\frac{1}{5}$ of these females will give birth to a calf each year, so the total number of calves born in a year is $\frac{1}{5} \cdot \frac{1}{4} N_t = \frac{N_t}{20}$.

We are told that one in thirty elephants will die during the year. So the number of deaths is $\frac{N_t}{30}$.

Thus:

$$\begin{aligned} N_{t+1} &= N_t + \frac{N_t}{20} - \frac{N_t}{30} \\ &= N_t\left(1 + \frac{1}{20} - \frac{1}{30}\right) = \frac{61}{60} N_t, \qquad N_0 = 120 \end{aligned}$$

Don't forget to include the initial condition.

From the recurrence relation we calculate

$$\begin{aligned} N_1 &= \frac{61}{60} \cdot 120 = 122 \\ N_2 &= \frac{61}{60} \cdot 122 = 124.03 \end{aligned}$$

And we make a table of the population sizes for years $t = 0, 1, \ldots 5$, accurate to 2 decimal places, below:

t	0	1	2	3	4	5
N_t	120	122	124.03	126.10	128.20	130.34

This recursion relation is an example of exponential growth, with reproductive rate $R - 1 = 61/60 - 1 = 1/60$. We learned in Section 2.1 that the explicit formula for the number of elephants in year t is:

$$N_t = \left(\frac{61}{60}\right)^t N_0$$

Notice that the model predicts after year 1 that there will be a noninteger number of elephants in the park. How should we interpret $N_2 = 124.03$? What does 0.03 of an elephant look like? The numbers in the problem statement should be interpreted as averages. For example, on average, half of the elephants will be female, and half of those, on average, will be of reproductive age. So our formula for N_t predicts only the average behavior of the population. If we were to consider many similar national parks, each of which contained a starting population of 120 elephants, then after 2 years some would have 124 elephants, some 125, and some more or less. 124.03 is the average number over all of the parks. We will discuss averages in more depth in Chapter 12, but for now it is sufficient to imagine that different parks would have slightly different numbers of elephants, and our model should be interpreted as predicting what the average behavior of the population would be.

EXAMPLE 2 **Elephant Population** In an effort to maintain elephant populations and generic diversity across Sri Lanka, conservationists remove a certain number, r, of elephants from the park each year, and move them to other parks. No new elephants are introduced into the park.

(a) Assuming $N_0 = 120$, what is the largest number of elephants that can be removed without causing the population to decrease?

(b) What happens if the number of elephants removed exceeds this value?

Solution **(a)** We need to modify the recursion relation from Example 1 to account for the number of elephants removed. Now:

$$N_{t+1} = N_t + \text{number of births} - \text{number of deaths} - \text{number of elephants removed}$$

No elephants are added to the population except for births.

So:

$$\begin{aligned} N_{t+1} &= N_t + \frac{1}{20}N_t - \frac{1}{30}N_t - r \\ &= \frac{61}{60}N_t - r, \quad \text{and} \quad N_0 = 120 \end{aligned} \tag{2.10}$$

Don't forget to specify the initial condition.

The function $f(N_t)$ is not proportional to N_t in this case. We also see that there is a problem if we apply formula (2.10) to same values of N_t; suppose, for example, that $r = 6$ and $N_t = 3$; then:

$$N_{t+1} = \frac{61}{60} \cdot 3 - 6 = 3.05 - 6 = -2.95$$

So our model predicts that $N_{t+1} < 0$. Although we explained that a noninteger number of elephants is admissible, there cannot be a negative number of elephants! The problem in this example is with the removal term. We cannot remove 6 elephants if there are only 3 elephants in the park. So we will add the rule that if removing elephants, it is not allowed to take the population size down below 0. When would this situation occur? When $\frac{61}{60}N_t - r < 0$ or equivalently when $N_t < \frac{60r}{61}$.

Thus our recurrence equation involves two different formulas:

$$N_{t+1} = \begin{cases} \dfrac{61}{60}N_t - r & \text{if } N_t \geq \dfrac{60r}{61} \\ 0 & \text{if } N_t < \dfrac{60r}{61} \end{cases} \tag{2.11}$$

We can still write (2.11) in the form of an equation, $N_{t+1} = f(N_t)$, where $f(N_t)$ is the function on the right-hand side of (2.11), so the population size is still given by a recurrence equation.

For the population to not decrease we need:

$$N_{t+1} \geq N_t$$

or equivalently

$$\frac{61}{60}N_t - r \geq N_t \tag{2.12}$$

Replace N_{t+1} using the recursion formula (2.11).

which is satisfied if and only if:

$$r \leq \frac{1}{60}N_t.$$

Add r and subtract N_t from both sides of (2.12).

Since N_t changes with time, this inequality does not give a fixed bound for r. But if $r \leq \frac{1}{60}N_0 = 2$, then the inequality holds for $r = 0$. Then since $N_1 \geq N_0$, it also holds for $r = 1$. Similarly for $r = 2$, $r = 3$, and so on; the population size will never decrease.

(b) What happens if $r > 2$? Let's take a specific value to start off. Let $r = 5$. Then the first five terms in the sequence N_t are given in the table below:

t	0	1	2	3	4
N_t	120	117	113.95	110.85	107.70

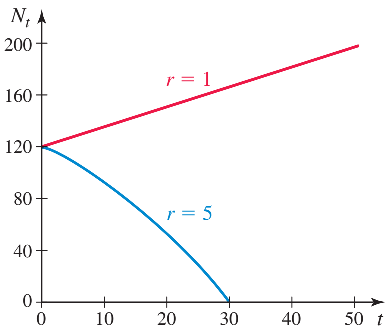


Figure 2.19 Number of elephants in a national park. If $r < 2$ the population increases each year. If $r > 2$, then the population decreases each year until it reaches $N_t = 0$.

The population shrinks from year to year. In fact, if we make a plot of N_t against t for enough terms we see that eventually N_t reaches 0 (see Figure 2.19).

The population size decreases year after year because if $r > \frac{1}{60}N_t$ then our analysis from part (a) implies that $N_{t+1} < N_t$. Now $\frac{1}{60}N_0 = 2$. So if $r > 2$, then $N_1 < N_0$, (i.e., the population size decreases at $r = 0$). But then since $N_1 < N_0$, $r > \frac{1}{60}N_1$, so the population size decreases at $r = 1$. Similarly for $r = 2$, $r = 3$, and so on.

If $r > 2$ then $N_{t+1} \leq N_t$ in all future years, so the population size decreases from year to year. However, the population size cannot drop below 0, so N_t cannot decrease without bound (we cannot have $\lim_{t\to\infty} N_t = -\infty$). Does it necessarily decrease to 0? (Another possibility is that it decreases to some nonzero limit.) Suppose that $\lim_{t\to\infty} N_t = N$; then N must be a fixed point of the recursion relation. To identify all fixed points we must consider the two cases: (i) $N \geq 60r/61$ and (ii) $N < 60r/61$

(i) If $N \geq \frac{60r}{61}$ then N is a fixed point if and only if:

$$N = \frac{61}{60}N - r \Rightarrow N = 60r > 120 \qquad \text{Since } r > 2.$$

(ii) If $N < \frac{60}{61}r$, then $N = 0$ is the only fixed point.

$N_t \leq 120$ for all t, so the fixed point from case (i) cannot be the limit of the sequence—the elephant population must drop to 0. •

2.3.2 Density-Dependent Population Growth: The Beverton-Holt Model

In Section 2.3.1 we considered models in which the rate at which organisms are born and the rate at which they die are simply proportional to N_t, the population size. In this case, if we assume no migration:

$$N_{t+1} = N_t + \text{number of births} - \text{number of deaths}$$

If there are b births for every organism present at time t and m deaths (In Example 1 we used $b = \frac{1}{20}$ and $m = \frac{1}{30}$.) then:

$$N_{t+1} = N_t + b \cdot N_t - m \cdot N_t = (1 + b - m) \cdot N_t$$

If we define a new parameter $R = 1 + b - m$, then $N_{t+1} = R \cdot N_t$. The parameter $R - 1 = b - m$ is the reproductive rate; it represents the number of new organisms added per organism present at time t. We showed in Section 2.1 that this recursion relation gives rise to exponential growth or decay, depending on whether $R > 1$ or $R < 1$, that is, depending on whether the reproduction rate is positive or negative. Alternatively, we have growth if $b > m$ (births outnumber deaths) or decay if $b < m$ (there are fewer births than deaths).

If the reproductive rate is positive, then this model predicts that the population will grow without bound.

In practice many populations do not grow without bound. In Section 2.1 we examined the growth of a real population of yeast cells, and we found that the reproductive rate changed with the population size, that is, $R = R(N_t)$. We call this phenomenon density-dependent growth. In this section we will introduce a mathematical model for this type of growth. In the yeast population, R decreased as N_t increased. More generally, real habitats typically have carrying capacities; that is, each can only support populations up to a maximum size. If the population exceeds the carrying capacity, then competition for space, food, or other resources will mean that the organisms either die off or migrate out of the habitat. A function $R(N_t)$ that describes this behavior was proposed by ***Beverton and Holt (1957)***.

Beverton and Holt proposed that instead of being a constant, $R(N_t)$ is given by a function:

Definition Beverton-Holt Model

$$R(N_t) = \frac{R_0}{1 + aN_t} \tag{2.13}$$

This function contains two parameters, R_0 and a, both of which must be positive.

The two parameters R_0 and a will be different for different species and different populations. But for a single population they remain constant as the population grows. You will see how this works in Example 3.

These parameters control different features of the population growth, and we will start by explaining how they should be interpreted. If the population size N_t is very small, then we can approximate the denominator by 1, because aN_t will be much smaller than 1, so $1 + aN_t \approx 1$. Then $N_{t+1} \approx R_0 N_t$. We recognize this as the recursion relation for a population with exponential growth. So R_0 tells us how quickly the organisms reproduce if the population size is small.

To see the role of the parameter a, let's calculate the possible limits for N_t as $t \to \infty$. If $\lim_{t\to\infty} N_t = N$, then N must be a fixed point of the equation, so:

$$N = R(N) \cdot N = \frac{R_0 N}{1 + aN}$$ In $N_{t+1} = R(N_t) \cdot N_t$ replace N_t and N_{t+1} by N.

so

$$N \cdot (1 + aN) = R_0 \cdot N$$ Multiply out by the denominator $(1 + aN)$.

$$N \cdot (1 + aN - R_0) = 0$$ Bring all terms to the right-hand side.

so either $N = 0$ or $N = \dfrac{R_0 - 1}{a}$.

$N = 0$ is a possible limit for the population size because if $N_0 = 0$, then $N_1 = 0$, $N_2 = 0$, and so on. It can be shown (though the proof will have to wait until Chapter 5) that if $N_0 > 0$, the population size will converge to the second fixed point, that is, $N_t \to \frac{R_0 - 1}{a}$. We will therefore identify $(R_0 - 1)/a$ as the carrying capacity of the habitat. If R_0 is known, then we still need to know the value of a to determine the carrying capacity. The next example shows possible dynamics in a population obeying the Beverton-Holt model.

EXAMPLE 3 **Fish Population** The population of trout in a lake is measured at yearly intervals. Let N_t be the population size at the end of the tth year. The population obeys the Beverton-Holt model, with parameters $R_0 = 5$ and $a = 10^{-3}$. Show that if

(a) $N_0 = 300$ or **(b)** $N_0 = 6000$

the population size will ultimately converge to the same limit.

Solution We show plots of N_t against t for both of the different initial conditions in Figure 2.20.

Starting with either initial condition, the population size seems to converge to the same value $N = \frac{R_0 - 1}{a} = \frac{4}{10^{-3}} = 4000$, which is the carrying capacity of the lake. In Chapter 5 we will develop the mathematics needed to prove that the trout population will always converge to this value. •

2.3.3 The Discrete Logistic Equation

We introduced the Beverton-Holt model by assuming that the rate of reproduction in a population decreases as population size increases. The decrease in reproduction rate occurs because larger populations compete for limited resources, like territory or food.

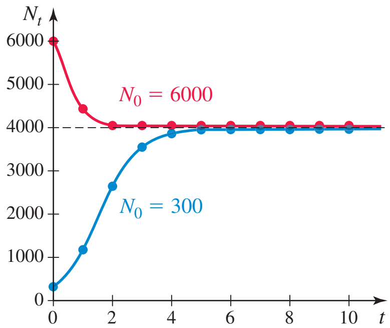


Figure 2.20 Size of a trout population modeled by the Beverton-Holt formula, starting with two different initial conditions. In both cases the population size converges to 4000.

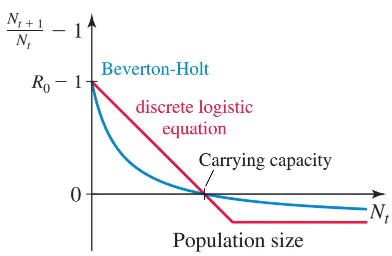


Figure 2.21 Reproductive rate, $\frac{N_{t+1}}{N_t} - 1$, varies with population size N_t in both the Beverton-Holt model and the logistic equation model for population growth. In both cases the carrying capacity of the habitat corresponds to the value of N_t at which the reproductive rate drops to zero.

The **discrete logistic equation** is another model for density-dependent population growth. The population size N_t is modeled using the recursive formula:

Definition Discrete Logistic Equation

$$N_{t+1} = \begin{cases} R_0 \cdot N_t - b \cdot N_t^2 & \text{if } N_t < \dfrac{R_0}{b} \\ 0 & \text{otherwise} \end{cases} \tag{2.14}$$

where R_0 and b are positive constants.

Before we discuss how to interpret the constants R_0 and b, let's make a plot of the reproductive rate as a function of N_t (Figure 2.21).

Just as with the Beverton-Holt model, the discrete logistic model predicts that the reproductive rate decreases as the population size increases, due to overcrowding and competition between organisms. It also has a carrying capacity at which the reproductive rate drops to 0. If the population size is less than the carrying capacity, then the population will increase (i.e., the reproductive rate is positive, so $N_{t+1} > N_t$), and if the population size is greater than the carrying capacity, then the population will decrease (i.e., the reproductive rate is negative, so $N_{t+1} < N_t$). For the discrete logistic model, we can calculate this carrying capacity by looking for a fixed point of the discrete map.

$$N = R_0 N - bN^2 \Rightarrow N(R_0 - 1 - bN) = 0 \tag{2.15}$$

Substitute $N_t = N$ and $N_{t+1} = N$ in (2.14) and factorize the equation.

So either $N = 0$ (which was also a fixed point for the Beverton-Holt equation) or $N = \frac{R_0 - 1}{b}$. If $N_t = \frac{R_0 - 1}{b}$, then $N_{t+1} = N_t$, so the reproductive ratio is equal to 0. $\frac{R_0 - 1}{b}$ is therefore the carrying capacity of the population.

But the discrete logistic model has a second critical size, above the carrying capacity; namely, if $N_t \geq R_0/b$, then $N_{t+1} = 0$, meaning that if the critical size R_0/b is ever exceeded, the population immediately goes extinct. We call R_0/b the **extinction threshold** of the population.

To interpret the parameters in the model we note that for the discrete logistic equation the graph of the reproductive rate against population size is a straight line with y-intercept $R_0 - 1$. So $R_0 - 1$ is the reproductive rate when N_t is close to zero, meaning that the population is far enough below the carrying capacity that organisms do not compete for food or other resources. The slope of the line is $-b$. So if the population grows by 1, the reproductive rate decreases by b. Reproductive rate decreases with population size because larger populations compete more strongly for resources and space. The parameter b represents the strength of this competition; larger values of b mean more competition (larger decrease in reproductive rate each time the population increases by 1) than small values of b. In particular, for larger values of b both the carrying capacity $\frac{R_0 - 1}{b}$ and the extinction threshold $\frac{R_0}{b}$ are smaller.

Although the differences between the logistic model and the Beverton-Holt model seem quite subtle, the logistic model shows many kinds of dynamics that are not seen in the Beverton-Holt model. Recall from Section 2.3.2 that if $N \neq 0$, a population described by the Beverton-Holt model will either grow or decay until it reaches the carrying capacity of the habitat. Although the logistic equation also has a carrying capacity, not all populations described by the model will converge to this carrying capacity. Before considering some examples, we will show that although there are two unknown constants in Equation (2.14), the equation can be written in a form containing only a single unknown constant.

We showed previously that the population becomes extinct when $N_t = R_0/b$. Suppose that, instead of measuring N_t directly, we measure how far the population is from its extinction threshold, that is we define a variable:

$$x_t = \frac{N_t}{\frac{R_0}{b}} = \frac{bN_t}{R_0}$$

So if $x_t = 1$, then the population is at the extinction threshold; if $x_t = 0.5$, it is at half this threshold, and so on. Then $N_t = \frac{R_0 \cdot x_t}{b}$ and $N_{t+1} = \frac{R_0 \cdot x_{t+1}}{b}$, so our recurrence equation can be rewritten as:

$$\frac{R_0}{b} \cdot x_{t+1} = \begin{cases} R_0 \cdot \left(\frac{R_0 \cdot x_t}{b}\right) - b \cdot \left(\frac{R_0 \cdot x_t}{b}\right)^2 & \text{if } x_t \leq 1 \\ 0 & \text{if } x_t > 1 \end{cases}$$

Replace N_t by $R_0 \cdot x_t / b$.
$N_t \leq R_0/b$ is equivalent to $x_t \leq 1$.

We can multiply both sides by b/R_0 and simplify:

$$x_{t+1} = \begin{cases} R_0 x_t (1 - x_t) & \text{if } x_t \leq 1 \\ 0 & \text{if } x_t > 1 \end{cases}$$

As it is now written, the logistic equation includes only a single constant, R_0. We call this form of the equation the **dimensionless** form of the discrete logistic equation. The original variable, N_t, has units (or dimensions) of number of organisms. x_t, by contrast, has no units; it gives the population size as a fraction of the extinction threshold, and is a dimensionless number between 0 and 1.

The dimensionless form of the equation contains only R_0 as an unknown parameter. Depending on the value of R_0 it can exhibit different behaviors. We will describe some of these behaviors below, although a more detailed explanation of how these behaviors arise is beyond the scope of this book.

Before beginning to look at different values of R_0, we need to determine what is the range of allowed values for R_0: Since $R_0 - 1$ represents the reproductive rate when N_t is far below the carrying capacity, we need $R_0 - 1 > 0$, or equivalently $R_0 > 1$. Otherwise, the population size would be decreasing even for very small populations in which there is little competition between organisms. Additionally we can determine an upper bound on R_0 as follows. If $x_{t+1} > 1$ for any time t, then $x_{t+2} = 0$; the population is predicted to immediately go extinct. To avoid this occurring, we choose R_0 so that for any $x_t \in [0, 1]$ we also have $x_{t+1} \in [0, 1]$; that is, the population never leaves the interval $N_t \in [0, R_0/b]$. We can see the constraints that this condition imposes on R_0 by making a sketch of the function $f(x_t) = R_0 \cdot x_t \cdot (1 - x_t)$ (Figure 2.22).

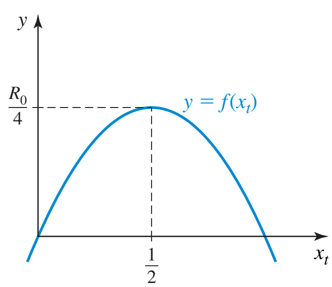


Figure 2.22 The dimensionless discrete logistic equation can be written in the form $x_{t+1} = f(x_t)$. The function $f(x_t) = R_0 \cdot x_t \cdot (1 - x_t)$ takes its maximum value at $x_t = \frac{1}{2}$.

The function $f(x_t)$ is a parabola. If $x_t \in [0, 1]$, then certainly, $f(x_t) \geq 0$, so the only constraint that may be violated is that we also need $f(x_t) \leq 1$. $f(x_t)$ takes its maximum value when $x_t = \frac{1}{2}$. $f(\frac{1}{2}) = R_0 \cdot (\frac{1}{2}) \cdot (1 - \frac{1}{2}) = \frac{R_0}{4}$. So to ensure that $f(x_t) \leq 1$, it is both necessary and sufficient that $\frac{R_0}{4} \leq 1$, that is, $R_0 \leq 4$.

We need $1 \leq R_0 \leq 4$, but different values of R_0 in this interval lead to different behaviors, as the following summary shows:

Summary of Possible Logistic Model Population Dynamics

Case 1:	$1 < R_0 < 3$	$x_t \to 1 - \dfrac{1}{R_0}$ as $t \to \infty$
Case 2:	$3 < R_0 < 3.57$	x_t oscillates in time
Case 3:	$R_0 > 3.57$	x_t is chaotic.

We consider these cases separately.

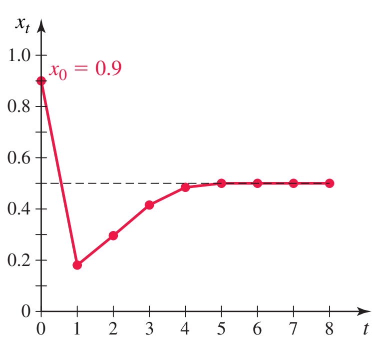


Figure 2.23 If $1 < R_0 < 3$, then x_t converges to the fixed point $1 - \frac{1}{R_0}$. Here, $R_0 = 2$, $x_0 = 0.9$, and $x_t \to 1 - \frac{1}{2} = 0.5$ as $t \to \infty$.

***Case 1: If* $1 < R_0 < 3$** We showed above that the logistic model has two fixed points: $N_t = 0$ and $N_t = \frac{R_0 - 1}{b}$. The corresponding fixed points for the dimensionless logistic model are $x_t = 0$ and $x_t = \frac{b}{R_0} \cdot (\frac{R_0 - 1}{b})$, or equivalently $x_t = 0$ and $x_t = 1 - \frac{1}{R_0}$. If $1 < R < 3$, then the solution converges to the second fixed point, as shown in Figure 2.23.

Case 2: For larger values of R_0 the logistic equation starts to develop oscillating behavior.

If $3 < R_0 < 3.449$, then x_t doesn't converge to a single value, but instead oscillates between two values, as shown in Figure 2.24. We call these dynamics a period 2 cycle, because x_t cycles between different values, rather than converging to a single limit. Period 2 means that after an interval of 2 it repeats itself, that is, $x_{t+2} = x_t$. (Compare with our previous definition of periodic functions in Chapter 1: there we said that $g(t)$ is periodic with period T if $g(t + T) = g(t)$.)

If we increase R_0 to a value in the interval $3.449 < R_0 < 3.544$, then x_t starts to oscillate between four different values. Since $x_{t+4} = x_t$ we call this a cycle with period 4 (see Figure 2.25), and since x_t changes from cycling between 2 values (period 2) and cycling between 4 value (period 4), we say that a period doubling has occurred.

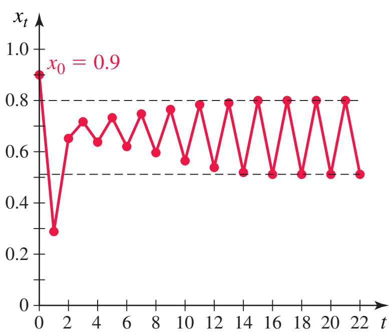


Figure 2.24 If $3 < R_0 < 3.449$, then x_t settles into a period 2 cycle. Here with $R_0 = 3.2$, $x_0 = 0.9$, x_t cycles between 0.513 and 0.800.

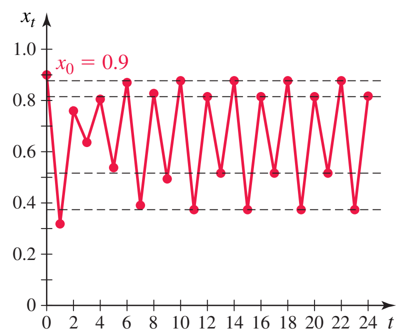


Figure 2.25 When $3.449 < R_0 < 3.554$, x_t settles into a cycle with period 4. Here with $R_0 = 3.52$ and $x_0 = 0.9$, x_t eventually cycles between the values 0.3731, 0.8233, 0.5121, and 0.8795.

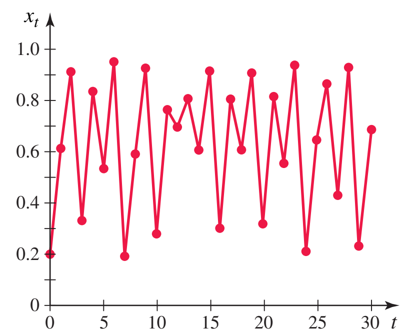


Figure 2.26 When $R > 3.57$, x_t starts to bounce around, apparently randomly, without repeating itself. Here $R_0 = 3.8$ and $x_0 = 0.2$.

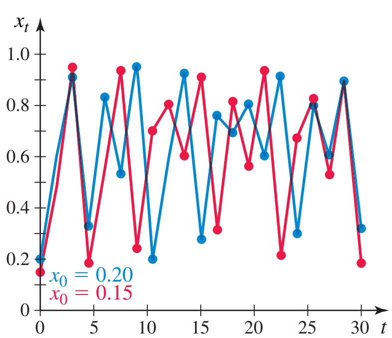


Figure 2.27 When the dynamics of the system are chaotic, small changes to the initial condition can create very different sequences $\{x\}$. Here $R_0 = 3.8$ and we show x_t starting with initial conditions $x_0 = 0.2$ and $x_0 = 0.15$.

As we continue to increase R_0, more period doublings occur. When R_0 passes through 3.544 a period 8 cycle emerges, when R_0 passes through 3.567 a period 16 cycle emerges, and so on. This process does not continue indefinitely, however.

Case 3: When R_0 exceeds 3.57, x_t stops oscillating periodically and starts to bounce around apparently randomly without ever repeating itself (Figure 2.26). Its dynamics for $3.57 < R_0 < 4$ are aperiodic or **chaotic**. In addition to there being no pattern to the sequence x_t, the dynamics are sensitive to even small changes in initial conditions, so a small change in x_0 leads to very different values for x_1, x_2, x_3, and so on (see Figure 2.27). This combination of factors—the lack of patterns in the values of x_t and the sensitivity to initial conditions—makes the dynamics of chaotic models like the logistic model very difficult to predict. In this book we will not encounter chaos again, but you should be aware that even simple-looking recurrence equations can give rise to very complicated dynamics.

2.3.4 Modeling Drug Absorption

Another important use for recursion relations is to build models for how a drug or medication is absorbed and eliminated from the body. These models can then be used to design treatment plans or to create safe dosing guidelines for how frequently the medication should be taken and in what amounts. To build these models we start by writing a word description for the processes that lead to changes in the amount of the drug present. This process is similar to how we started from word equations in Section 2.3.1 to create models for how populations grow.

Suppose that we measure the amount of the drug present in the body or in a particular part of the body, at a set of regularly spaced time intervals. We measure a sequence

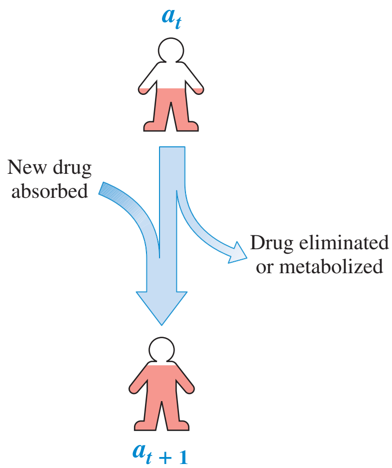


Figure 2.28 Using absorption and elimination to calculate the change in the amount of drug in a person's body.

of amounts a_t, $t = 0, 1, 2, \ldots$. a_t might be the concentration of the drug (measured, for example, in units of grams or milligrams of drug per gram of tissue) or it might be the total amount of the drug (measured, for example, in grams or international units (IU)). The time interval between measurements could be anything—1 hour, a few hours, or 1 day—and will depend on how frequently measurements of drug amount can be made, and on what questions the model is being set up to answer. We will explore these questions in the examples below.

To derive a model for the sequence a_t we start by noting that between the t-th measurement and the $(t+1)$-th measurement the amount of medication will increase if more is absorbed into the body or organ that we are considering and will decrease as medication is eliminated (i.e., removed from the organ or body, for example by the kidneys) or metabolized (i.e., converted into some other form by biochemical processes). See Figure 2.28. We can therefore write a word equation:

$$a_{t+1} = a_t + \begin{array}{c}\text{amount of new medication}\\ \text{entering the body or organ}\end{array} - \begin{array}{c}\text{amount of drug eliminated}\\ \text{or metabolized}\end{array} \tag{2.16}$$

The amount of drug entering the body will depend on whether the patient takes more medication during the time interval, and the form in which the medication is taken. We will give some examples of models for different processes by which a drug may enter the body below.

The amount of drug that is eliminated or metabolized depends on the time interval as well as the way that the drug is used by the body. Broadly there are two classes of drugs:

> **Kinetics of Drug Elimination** If a drug has **zeroth order elimination kinetics**, then the body eliminates a fixed amount of the drug per hour, independent of the amount present. It continues to eliminate the same amount of drug per hour until the drug is completely gone.
>
> If a drug has **first order elimination kinetics**, then the body eliminates a fixed fraction of the drug per hour. So if the amount of drug is doubled, then the amount eliminated per hour will also be doubled.

Some medications have zeroth order kinetics, and others have first order kinetics. This information has to be given before you can build the model.

EXAMPLE 4

Alcohol in the Body In high doses ethanol (alcohol) has zeroth order elimination kinetics; that is, a fixed amount of the drug is removed from the body each hour. Legal measurements of levels of intoxication are based on the blood alcohol concentration (abbreviated BAC, and measured in grams of ethanol per liter of blood). The rate of elimination of alcohol depends on many variables, including the sex of the person and on how active they are during the period when the alcohol is being eliminated, but a typical elimination rate is 0.186 g/(liter h).[1] The data in this question are taken from ***al-Lanqawi et al. (1992)***.

(a) A patient's blood alcohol concentration at time $t = 0$ hrs is 1.2 g/liter. Find a recursion relation for the blood alcohol concentration measured 1 hour later, 2 hours later, and so on and solve it to deduce an explicit formula for BAC as a function of time. When does the blood alcohol concentration reach 0 g/liter?

(b) A police officer pulls over a driver who is driving erratically. The driver refuses to take a breathalyzer test, but at the police station two hours later a blood test is performed and the driver's BAC is measured to be 0.61 g/liter. Using the formula from part (a), estimate what the driver's BAC was when the driver was pulled over.

(1) The calculations in this example are for illustrative purposes only; legal BAC limits vary between states, and it is never safe to drink and drive.

Solution **(a)** We are told that alcohol is eliminated from the blood at a rate of 0.186 g/(liter h). So if the patient's BAC t hours into the observation period is c_t, then because no more alcohol enters the blood after $t = 0$, our word equation (2.16) gives:

$$\begin{aligned} c_{t+1} &= c_t - \text{amount of drug elimination between hour } t \text{ and hour } t+1 \\ &= c_t - 0.186 \end{aligned}$$

with $c_0 = 1.2$. (Remember to include the initial condition on c_t.)

We need to find an explicit formula for c_t, and we get this formula by guessing based from the first few terms in the sequence:

$$\begin{aligned} c_0 &= 1.2 \\ c_1 &= c_0 - 0.186 = 1.014 \\ c_2 &= c_1 - 0.186 = (c_0 - 0.186) - 0.186 \\ &= c_0 - 2 \cdot 0.186 \\ c_3 &= c_2 - 0.186 = (c_2 - 2 \cdot 0.186) - 0.186 \\ &= c_0 - 3 \cdot 0.186 \end{aligned}$$

From these terms we can see that $c_t = c_0 - 0.186 \cdot t$. You can check that this formula satisfies both the recurrence equation and the initial condition.

c_t reaches 0 when

$$c_0 - 0.186t = 0 \qquad \text{or} \qquad t = \frac{c_0}{0.186} = \frac{1.2}{0.186} = 6.45 \text{ hours}$$

(b) Here we know the driver's BAC at time $t = 2$ hours, and we want to infer their BAC at time $t = 0$ (i.e., two hours previously). This problem is equivalent to knowing c_2, and wanting to find c_0. (It is also equivalent to having c_0 and wanting to calculate c_{-2}.)

We use the explicit formula for c_t from part (a):

$$c_t = c_0 - 0.186t \Longrightarrow c_2 = c_0 - 0.372$$

The BAC at $t = 2$ was measured to be $c_2 = 0.6$ g/liter, so the driver's BAC when they were pulled over was:

$$c_0 = c_2 + 0.372 = 0.972 \text{ g/liter}$$

•

EXAMPLE 5 **Ibuprofen in the Body** Ibuprofen (sold under the brand name Advil) is a non-steroid anti-inflammatory drug (NSAID) that is used to treat muscle pain. We want to build a mathematical model for the concentration of ibuprofen in a patient's blood. The drug concentration is measured in mg (milligrams) per liter. Assume that when a patient takes an ibuprofen pill, the drug takes about an hour to enter the blood. Each pill increases the concentration of ibuprofen in the patient's blood by 40 mg/liter.

Ibuprofen has first order elimination kinetics, meaning that a fixed proportion of the ibuprofen present in a patient's blood is eliminated in one hour. Assume that 24.25% of the ibuprofen present in the blood is eliminated in one hour. The date in this example are taken from ***Albert and Gernaat (1984)***.

A patient takes an ibuprofen pill at time $t = 0$. Initially there is no ibuprofen present in the patient's blood.

(a) The directions for ibuprofen recommend that at most one pill be taken every six hours. Suppose the patient waits exactly six hours from taking the first pill before taking a second pill. What will the drug concentration be just before the second pill is taken? One hour afterwards?

(b) The patient carefully follows the directions on the pill bottle by taking one pill every six hours, and continues to do this for many days.[2] What happens to the concentration of ibuprofen in the patient's blood?

(2) If you experience continuous pain that requires taking ibuprofen for more than 10 days, you should consult your doctor.

(c) Suppose that the patient does not follow the directions on the bottle and takes one pill every T hours. What will happen to the concentration of ibuprofen in the patient's blood?

Solution **(a)** Let c_t be the concentration of ibuprofen in the patient's blood t hours after the first pill was taken. To build a mathematical model for c_t, we start from our word equation (2.16). We are told that $c_0 = 0$, and we need to calculate c_6, the concentration 6 hours after the first pill was taken.

Over the course of the first hour, 40 mg/liter of ibuprofen enters the patient's blood from the pill taken at $t = 0$. Since there was no drug present initially there is no elimination from the blood, so:

$$\begin{aligned} c_1 &= c_0 + \text{amount absorbed into blood} - \text{amount of drug eliminated in one hour.} \\ &= c_0 + 40 - 0 = 40 \qquad c_0 = 0 \end{aligned}$$

In the next hour (between $t = 1$ and $t = 2$) no more drug enters the blood, since no pills are taken. We are told that 24.25% of the drug that was present at $t = 1$ will be eliminated by $t = 2$, so:

$$\begin{aligned} c_2 &= c_1 + \text{amount absorbed into blood} - \text{amount of drug eliminated in one hour.} \\ &= c_1 + 0 - 0.2425c_1 = 0.7575c_1 \end{aligned}$$

Similarly in the next hour (between $t = 2$ and $t = 3$) no drug enters the blood, but 24.25% of the drug that was present at $t = 2$ is eliminated, so $c_3 = 0.7575c_2$. In fact, since no pills are taken until $t = 6$, we have $c_{t+1} = 0.7575c_t$ for $t = 1, 2, \ldots 5$. Thus, between $t = 1$ and $t = 5$, c_t obeys a recurrence equation that we know gives exponential decay. We can solve for c_6 by applying this recursion relation iteratively:

$$\begin{aligned} c_3 &= 0.7575c_2 = 0.7575(0.7575c_1) = (0.7575)^2 c_1 \\ c_4 &= 0.7575c_3 = 0.7575((0.7575)^2 c_1) = (0.7575)^3 c_1 \\ &\ldots \\ c_6 &= (0.7575)^5 c_1 = (0.7575)^5 \times 40 = 9.98 \text{ mg/liter} \end{aligned}$$

is the concentration at $t = 6$, when the patient takes their second pill.

We also need to find the blood concentration one hour after the second pill was taken, that is: c_7. In the hour between $t = 6$ and $t = 7$, two processes occur: the contents of the second pill are absorbed into the blood, and some of the ibuprofen that was present in the blood at $t = 6$ will be eliminated. Using our word equation:

$$\begin{aligned} c_7 &= c_6 + \text{amount absorbed into blood} - \text{amount of drug eliminated in one hour.} \\ &= c_6 + 40 - 0.2425c_6 = 0.7575c_6 + 40 = 47.56 \text{ mg/liter} \end{aligned}$$

(b) To get a sense of what happens to the drug concentration in time, let's start by calculating the drug concentration in the period after taking the second pill, but before taking the third pill. We know that since no other drug enters the blood during this period, $c_{t+1} = 0.7575c_t$ for $t = 7, 8, \ldots 11$. So:

$$\begin{aligned} c_8 &= 0.7575c_7 \qquad c_7 \text{ was calculated in part (a).} \\ c_9 &= 0.7575c_8 = (0.7575)^2 c_7 \end{aligned}$$

and so on. By continuing to iteratively apply the recurrence equation we see that $c_t = (0.7575)^{t-7} c_7$ for $t = 7, \ldots, 12$, which is to say that the blood concentration decays exponentially in time, starting from a maximum value of 47.56 mg/liter. Just before taking the third pill, the concentration is:

$$c_{12} = (0.7575)^5 \cdot c_7 = (0.7575)^5 \cdot 47.56 = 11.86 \text{ mg/liter}$$

Then the patient takes the third pill, and in the next hour (between $t = 12$ and $t = 13$), both absorption and elimination occur, so:

$$c_{13} = c_{12} + 40 - 0.2425c_{12} = 0.7575 \times 11.86 + 40 = 48.98 \text{ mg/liter}$$

Since there is no absorption over the next 5 hours, from $t = 13$ to $t = 18$, the ibuprofen concentration will decay exponentially again. In fact we can apply our recursion relation to calculate that $c_t = (0.7575)^{t-13} c_{13}$ for $t = 13, 14, \ldots 18$. At $t = 18$ the patient takes another pill, and the blood concentration of the drug increases again. Figure 2.29 shows the behavior of the blood concentration over a 24-hour period during which the patient takes 5 pills.

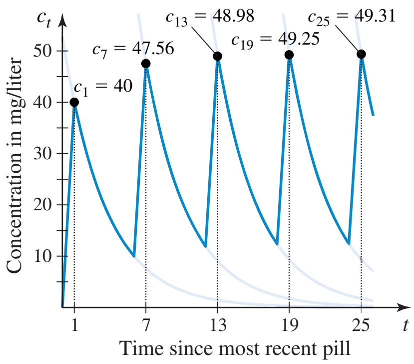


Figure 2.29 Concentration of ibuprofen in the blood of a patient taking pills at 6-hour intervals.

From Figure 2.29 we see that the blood concentration spikes one hour after each pill is taken (i.e., at times $t = 1, 7, 13, \ldots$). In the intervals between pills the drug concentration decays exponentially. We can also see from Figure 2.29 that maximum drug concentration (one hour after each pill is taken) increases over time, with: $c_1 = 40$, $c_7 = 47.56$, $c_{13} = 48.98$, $c_{19} = 49.25, \ldots$.

Does the sequence of peak concentrations (one hour after each pill is taken) increase without bound? It is important to know the answer to this question because, if the drug concentration becomes too large, then it could potentially hurt the patient. Let's consider the sequence of drug concentrations one hour after each pill is taken, $c_1, c_7, c_{13}, \ldots$. Although we can certainly compute each term in the sequence by computing the entire sequence $c_0, c_1, c_2, c_3, \ldots$, it is more efficient to use the flexibility that we have to define sequences with different time intervals. To that end, define a new sequence $C_1, C_2, C_3, \ldots$, where C_n is the blood concentration one hour after pill n is taken. So, the terms $\{C_n\}$ give the peak concentrations in Figure 2.29: $C_1 = c_1$, $C_2 = c_7$, $C_3 = c_{13}$, and so on.

We can derive a recursion relation to calculate the terms C_n directly. To find this relation, notice that between pill n and $n+1$, 6 hours elapse, so due to elimination the blood concentration falls to $(0.7575)^6$ of its starting value. Then the patient takes the $(n+1)$th pill, increasing blood concentration by 40. So, Equation (2.16) gives:

$$\begin{aligned} C_{n+1} &= C_n + \text{amount added to bloodstream} - \text{amount eliminated} \\ &= 40 + (0.7575)^6 C_n \end{aligned} \tag{2.17}$$

To compute the sequence definition we need an initial condition. For this sequence we know that $C_1 = c_1 = 40$ mg/liter. (Sequences can start at either $n = 0$ or $n = 1$. $n = 1$ makes sense here because C_n is the concentration after taking pill n, and the sequence starts with the first pill.)

From Equation (2.17) we can generate the peak concentration values and check that they agree with what we computed previously.

$$\begin{aligned} C_1 &= 40 = c_1 \\ C_2 &= 47.56 = c_7 \\ C_3 &= 48.98 = c_{13} \\ C_4 &= 49.25 = c_{19} \end{aligned}$$

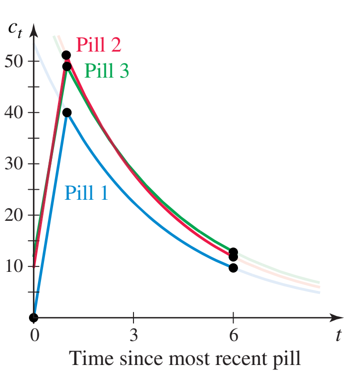


Figure 2.30 Concentration of ibuprofen in the blood of a patient taking pills at 6-hour intervals, now plotted against time since most recent pill was taken. The hourly concentration values converge to a period 6 cycle.

We see that the terms in C_n get larger and larger, meaning that the concentration spikes get larger and larger. But the increase from C_3 to C_4 is much smaller than the increase from C_2 to C_3. This, in turn, is smaller than the increase from C_1 to C_2, so the increase in maximum concentration seems to get smaller after each successive pill. In fact, it looks like the sequence is converging to a limit. To calculate this limit, use the fixed point method. The fixed point method tells us that, if the limit is C, then:

$$\begin{aligned} C &= (0.7575)^6 \cdot C + 40 \\ \Rightarrow C &= \frac{40}{1 - (0.7575)^6} = 49.317 \text{ mg/liter} \end{aligned}$$

Substitute $C_{n+1} = C$ and $C_n = C$ in Equation (2.17) to find the fixed points.

So, the patient's blood concentration one hour after taking each pill converges to 49.317 mg/liter. Between pills it decays exponentially, through the values: C, $C \cdot 0.7575 = 37.358$, $C \cdot (0.7575)^2 = 28.299$, $C \cdot (0.7575)^3 = 21.436$, and so on. The entire sequence of concentrations converges to a cycle with period 6, since every six hours the formulas are reset. We can see this convergence in the data from Figure 2.29 if, instead of plotting c_t against absolute time, we plot c_t against the time taken since the last pill was taken (Figure 2.30).

(c) Now suppose that the patient takes a pill every T hours. We can follow the same steps as in parts **(a)** and **(b)** to predict the concentration of ibuprofen in the blood. One hour after the first pill the patient's blood concentration is $c_1 = 40$ mg/liter, because the pill has been absorbed, but no elimination has yet taken place. Until they take the second pill, the blood concentration decays exponentially, so $c_{t+1} = 0.7575c_t$ for $t = 1, 2, \ldots T - 1$.

Just as in part **(b)**, we can define a sequence containing only the maximum drug concentrations (i.e., one hour after each pill is taken, starting at $t = 1$ and measured every T hours). We call this sequence $C_1, C_2, C_3, \ldots$, where C_n is the concentration one hour taking the nth pill. So: $C_1 = c_1, C_2 = c_{T+1}, C_3 = c_{2T+1}, C_4 = c_{3T+1}$, and so on. Just as in part (b), we can derive a recursion relation for the terms of the sequence C_n. Specifically, in the T hours between C_n and C_{n+1}, elimination of the drug will cause the concentration to drop to $(0.7575)^T$ of its starting value. But 40 mg/liter is also added to the blood by each pill, so:

$$C_{n+1} = (0.7575)^T C_n + 40 \qquad (2.18)$$

If $T = 6$, we recover Equation (2.17).

with initial condition $C_1 = c_1 = 40$ mg/liter. Just as in **(b)**, the values of $\{C_n\}$, which give the maximum drug concentrations one hour after each pill is taken, will form an increasing sequence. Just as in part **(b)** the sequence does not grow without bound, but converges to a limit, C, where:

$$C = (0.7575)^T \cdot C + 40$$

Substitute $C_n = C$ and $C_{n+1} = C$ into Equation (2.18) to find any fixed points of the sequence.

and so: $C = \dfrac{40}{1 - (0.7575)^T}$

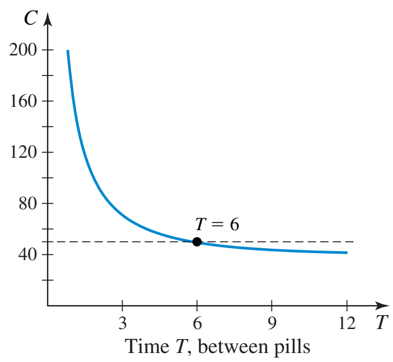


Figure 2.31 The maximum concentration of ibuprofen in a patient's blood, one hour after taking a pill, increases as the time between pills, T, decreases. For $C \le 50$ the time between pills should not exceed 6 hrs.

We plot the limiting concentration as a function of the time T between pills in Figure 2.31.

Safe doses for drugs are sometimes expressed in the form of maximum safe blood concentrations. If the maximum safe blood concentration for this patient were 50 mg/liter, then the patient should avoid taking more than one pill every T hours where:

$$\frac{40}{1 - 0.7575^T} = 50$$

$$\frac{40}{50} = 1 - 0.7575^T \qquad \text{Multiply through by denominator.}$$

$$0.7575^T = 1 - \frac{4}{5} = \frac{1}{5} \qquad \text{Isolate term in } T$$

So: $$T = \frac{\log(\frac{1}{5})}{\log(0.7575)} = \frac{-0.699}{-0.121} = 5.79 \qquad \text{Take logs of both sides.}$$

So the patient should follow the directions on the pill box and space pills apart by around 6 hours.

Section 2.3 Problems

2.3.1

1. You are building a mathematical model for the population of cod fish in a North Atlantic fishery. Write a word equation relating the population N_t in one year to the population N_{t+1} in the next year. Your word equation should include the following terms:

- Number of cod fish born during the year
- Cod fish dying of old age during the year
- Cod fish killed by predators during the year
- Cod fish removed by fishing boats during the year

2. You are building a mathematical model for the human population of a small Southern California town. Write a word equation relating the population N_t in one year to the population N_{t+1} in the next year. Your word equation should include the following terms:

- Number of children born during the year
- People dying from any cause during the year
- People moving into the town from other towns during the year
- People leaving the town to live in other towns during the year

3. You are building a math model for the size of the wild population of kakapo (rare ground dwelling flightless parrots) in New Zealand. Write a word equation relating the population N_t in one year to the population N_{t+1} in the next year. Your word equation should include the following terms:

- Number of kakapo births in the wild during the year
- Kakapo removed for captive breeding during the year
- Kakapo reintroduced into the wild from captive breeding during the year
- Kakapo killed by predators during the year
- Kakapo deaths from disease during the year

4. You are building a mathematical model for the spread of Sudden Oak Death Syndrome—a disease that has wiped out over one million oak and tanoak trees in Coastal California. Write a word equation relating number of oak trees in one year, N_t, to the number, N_{t+1}, in the next year. Your word equation will include the following terms:

- Number of trees seeded in the wild during the year
- Trees planted by people during the year
- Trees killed by the disease during the year
- Trees cut down by loggers during the year

5. You are building a mathematical model for the size of coral reefs in a patch of the Pacific Ocean. Rather than directly measuring the number of corals, you measure the area of living coral in each coral reef. Write a word equation relating the area, A_t, of living coral in one year, to the area, A_{t+1}, in the next year. Your word equation will include the following terms:

- Area killed by ocean acidification during the year
- Area killed by fishing during the year
- Area of reef restored or rebuilt during the year

6. You are a conservation ecologist trying to build a mathematical model for the size of the rhino population in a national park. Write a word equation relating the population N_t in one year to the population N_{t+1} in the previous year. Your equation will include **some, but not necessarily all** of the following terms:

- Number of rhinos born in the park during the year
- Rhinos born in other parks during the year
- Rhinos introduced into the park from captive breeding programs during the year
- Rhinos moved out of this park to other parks during the year
- Rhinos relocated to this park from other parks during the year
- Rhinos that become ill or injured during the year
- Rhinos that die during the year
- Number of female rhinos
- Number of rhinos that become pregnant during the year

7. You are trying to build a mathematical model for the population of amoeba cells growing on a petri plate. This amoeba feeds on bacteria. Write down a word equation relating the population N_t in one hour to the population N_{t+1} in the next hour. Your equation will include **some, but not necessarily all** of the following terms:

- Number of amoeba cells that divide into two cells in one hour
- Number of bacteria on the plate
- Number of bacteria that are eaten in one hour
- Number of bacteria that divide into two cells in one hour
- Number of amoeba that die in one hour

8. You are trying to develop a mathematical model for the number of students on the UCLA campus that are sick with flu. Write a word equation relating the number of students that are ill with the flu and on campus on one day, N_t, with the number N_{t+1} that are ill and on campus on the next day. Your equation will include **some, but not necessarily all** of the following terms:

- Total number of students on the campus
- Students that are not sick
- Students who catch the flu
- Students who recover from the flu
- Students who return home to recuperate from the flu
- Number of doctors or nurses treating the students in one day

9. Saving the Kakapo You are modeling the size of the population of kakapo (a rare flightless parrot) in an island reserve in New Zealand. You want to use the mathematical model to predict the size of the population. The data in this question are taken from ***Elliot et al. (2001)***

(a) You start by writing a word equation relating the population size N_t in year t, that is, t years after the study began, to the population size N_{t+1} in the next year.

$$N_{t+1} = N_t + \begin{array}{c}\text{number of birds}\\ \text{born in one year}\end{array} - \begin{array}{c}\text{number of birds that}\\ \text{die in one year}\end{array}$$

We will derive together formulas for each of these terms.

(i) To estimate the number of birds born, assume that half of the birds are female. A female bird lays one egg every four years. However, because of the large numbers of predators (mostly rats) only 29% of hatchlings survive their first year. Explain how based on this data our prediction for the number of births is: $N_t \cdot 0.5 \cdot 0.25 \cdot 0.29 = 0.03625 \cdot N_t$.

(ii) Kakapo life expectancy is not well understood, but we will assume that they live around 50 years. That is, in a given year, one in fifty kakapo will die. What is the corresponding number of deaths?

(iii) Assume that the starting population size on this island is 50 birds (i.e., $N_0 = 50$). Calculate the predicted population size over the next five years (i.e., calculate $N_1, N_2, \ldots, N_5$).

(iv) When (if ever) will the population size reach 100 birds? What about 200 birds? (You will find it helpful to derive an explicit formula for the size of the population N_t).

(b) Using your model from part (a) you want to evaluate the effectiveness of two different conservation strategies:

(Strategy 1) If the kakapo are given supplementary food, then they will breed more frequently. If given supplementary food, then rather than laying an egg every four years, a female will lay an egg every two years.

(Strategy 2) By hand-rearing kakapo chicks, it is possible to increase their one year survival rate from 29% to 75%.

(i) Write down a recurrence equation for the population size N_t if strategy 1 is implemented. Assuming $N_0 = 50$, calculate N_1, $N_2, \ldots, N_5$.

(ii) Write down a recurrence equation for the population size N_t if strategy 2 is implemented. Assuming $N_0 = 50$, calculate N_1, $N_2, \ldots, N_5$.

(iii) Which conservation strategy gives the biggest increase in population size?

10. Mountain Gorilla Conservation You are trying to build a mathematical model for the size of the population of mountain gorillas in a national park in Uganda. The data in this question are taken from ***Robbins et al. (2009)***.

(a) You start by writing a word equation relating the population of gorillas t years after the study begins, N_t, to the population N_{t+1} in the next year:

$$N_{t+1} = N_t + \begin{array}{c}\text{number of gorillas}\\ \text{born in one year}\end{array} - \begin{array}{c}\text{number of gorillas}\\ \text{that die in one year}\end{array}$$

We will derive together formulas for the number of births and the number of deaths.

(i) Around half of gorillas are female, 75 % of females are of reproductive age, and in a given year 22 % of the females of reproductive age will give birth. Explain why the number of births is equal to:

$$0.5 \cdot 0.75 \cdot 0.22 \cdot N_t = 0.0825 \cdot N_t$$

(ii) In a given year 4.5 % of gorillas will die. Write down a formula for the number of deaths.

(iii) Write down a recurrence equation for the number of gorillas in the national park. Assuming that there are 300 gorillas initially (that is $N_0 = 300$), derive an explicit formula for the number of gorillas after t years.

(iv) Calculate the population size after 1, 2, 5, and 10 years.

(v) According to the model, how long will it take for population size to double to 600 gorillas?

(b) In reality the population size is almost totally stagnant (i.e., N_t changes very little from year to year). Robbins et al. (2009) consider three different explanations for this effect:

(i) Increased mortality: Gorillas are dying sooner than was thought. What percentage of gorillas would have to die each year for the population size to not change from year to year?

(ii) Decreased female fecundity: Gorillas are having fewer offspring than was thought. Calculate the female birth rate (percentage of reproductive age females that give birth) that would lead to the population size not changing from year to year. Assume that all other values used in part (a) are correct.

(iii) Emigration: Gorillas are leaving the national park. What number of gorillas would have to leave the national park each year for the population to not change from year to year?

2.3.2

In Problems 11–16, assume that the population growth is described by the Beverton–Holt model. Find all fixed points.

11. $N_{t+1} = \dfrac{4N_t}{1+N_t/30}$ **12.** $N_{t+1} = \dfrac{2N_t}{1+N_t/60}$

13. $N_{t+1} = \dfrac{2N_t}{1+N_t/90}$ **14.** $N_{t+1} = \dfrac{3N_t}{1+N_t/100}$

15. $N_{t+1} = \dfrac{3N_t}{1+N_t/30}$ **16.** $N_{t+1} = \dfrac{5N_t}{1+N_t/240}$

In Problems 17–22, assume that the population growth is described by the Beverton–Holt recruitment curve with parameters R_0 and a. Find the population sizes for $t = 1, 2, \ldots, 5$ and find $\lim_{t\to\infty} N_t$ for the given initial value N_0.

17. $R_0 = 2, a = 0.01, N_0 = 2$

18. $R_0 = 2, a = 0.1, N_0 = 2$

19. $R_0 = 3, a = 1/20, N_0 = 7$

20. $R_0 = 3, a = 1/10, N_0 = 3$

21. $R_0 = 4, a = 1/40, N_0 = 2$

22. $R_0 = 4, a = 1/60, N_0 = 2$

23. A population obeys the Beverton-Holt model. You know that $R_0 = 3$ for this population. As $t \to \infty$ you observe that $N_t \to 100$. What value of a is needed in the model to fit it to these data?

24. A population obeys the Beverton-Holt model. You know that $R_0 = 5$ for this population. As $t \to \infty$ you observe that $N_t \to 200$. What value of a is needed in the model to fit it to these data?

25. A population obeys the Beverton-Holt model. You know that $R_0 = 2$ for this population. One year you measure $N_t = 20$. The next year you measure that $N_{t+1} = 30$. What value of a is needed in the model to fit these data?

26. A population obeys the Beverton-Holt model. You know that $R_0 = 4$ for this population. One year you measure $N_t = 50$. The next year you measure that $N_{t+1} = 40$. What value of a is needed in the model to fit these data?

2.3.3

In Problems 27–32, assume that the discrete logistic equation is used with parameters R_0 and b. Write the equation in the dimensionless form $x_{t+1} = R_0 x_t(1-x_t)$, and determine x_t in terms of N_t.

27. $R_0 = 1, b = \dfrac{1}{10}$ **28.** $R_0 = 1, b = \dfrac{1}{20}$

29. $R_0 = 2, b = \dfrac{1}{15}$ **30.** $R_0 = 2, b = \dfrac{1}{20}$

31. $R_0 = 2.5, b = \dfrac{1}{30}$ **32.** $R_0 = 2.5, b = \dfrac{1}{50}$

In Problems 33–38, we will investigate the advantage of dimensionless variables.

33. A population obeys the discrete logistic equation:

$$N_{t+1} = R_0 \cdot N_t - bN_t^2$$

Find the possible fixed points of the population size (one fixed point will depend on the unknown parameters R_0 and b).

34. You are studying a population that obeys the discrete logistic equation. You know that $R_0 = 2$. Also you observe that as $t \to \infty$, $N_t \to 50$. What value of b is needed in the model to fit these data?

35. You are studying a population that obeys the discrete logistic equation. You know that $R_0 = 2.5$. Also you observe that as $t \to \infty$, $N_t \to 40$. What value of b is needed in the model to fit these data?

36. You are studying a population that obeys the discrete logistic equation. You know that $R_0 = 2$. One year you measure $N_t = 10$. The next year you measure that $N_{t+1} = 15$. What value of b is needed in the model to fit these data?

37. You are studying a population that obeys the discrete logistic equation. You know that $R_0 = 3$. One year you measure $N_t = 15$. The next year you measure that $N_{t+1} = 30$. What value of b is needed in the model to fit these data?

38. You are studying a population that obeys the discrete logistic equation. You know that $b = \frac{1}{10}$. One year you measure $N_t = 15$.

The next year you measure that $N_{t+1} = 20$. What value of R_0 is needed in the model to fit these data?

In Problems 39–50, we investigate the behavior of the discrete logistic equation

$$x_{t+1} = R_0 x_t(1 - x_t).$$

Compute x_t for $t = 0, 1, 2, \ldots, 20$ for the given values of r and x_0, and graph x_t as a function of t.

39. $R_0 = 2, x_0 = 0.2$

40. $R_0 = 2, x_0 = 0.1$

41. $R_0 = 2, x_0 = 0.9$

42. $R_0 = 2, x_0 = 0$

43. $R_0 = 3.1, x_0 = 0.5$

44. $R_0 = 3.1, x_0 = 0.1$

45. $R_0 = 3.1, x_0 = 0.9$

46. $R_0 = 3.1, x_0 = 0$

47. $R_0 = 3.8, x_0 = 0.5$

48. $R_0 = 3.8, x_0 = 0.1$

49. $R_0 = 3.8, x_0 = 0.9$

50. $R_0 = 3.8, x_0 = 0$

2.3.4

51. Adderall in the Body Adderall is a proprietary combination of amphetamine salts that is used to treat ADHD (Attention-Deficit/Hyperactivity Disorder). The patient takes one pill before 8 am, and the drug has completely entered the blood by 8 am. At 8 am the blood concentration of the drug is 33.8 ng/ml. Adderall has first order elimination kinetics with 7.7 % of the drug being removed from the blood in each hour. The data in this equation are taken from ***Greenhill et al. (2003)***.

(a) The concentration of drug in the patient's blood t hours after 8 am is measured to be C_t. Write a recursive relation for C_{t+1} in terms of C_t. Assume that the patient takes no other Adderall pills.

(b) Solve your recurrence equation to derive an explicit formula for C_t as a function of t.

(c) When does the concentration of drug first drop below 0.1 ng/ml?

In Problems 52 and 53 we model painkillers that are absorbed into the blood from a slow release pill. Our mathematical model for the amount, a_t, of drug in the blood t hours after the pill is taken must include the amount absorbed from the pill each hour. Our model starts with the word equation.

$$a_{t+1} = a_t + \begin{array}{c}\textit{amount absorbed}\\ \textit{from the pill}\end{array} - \begin{array}{c}\textit{amount eliminated}\\ \textit{from the blood}\end{array}$$

52. Assume the amount absorbed from the pill between time t and time $t+1$ is $10 \cdot (0.4)^t$.

(a) The drug has first order elimination kinetics. 10 % of the drug is eliminated from the blood each hour. Write down the recursion relation for a_{t+1} in terms of a_t.

(b) Assuming that $a_0 = 0$, meaning that no drug is present in the blood initially, calculate the amount of drug present at times $t = 1, 2, \ldots, 6$.

(c) What is the maximum amount of drug present at any time in this interval? At what time is this maximum amount reached?

(d) Use a spreadsheet to calculate the amount of drug present in hourly intervals from $t = 0$ up to $t = 24$.

(e) Show when t is large, the amount of drug present in the blood decreases approximately exponentially with t. *Hint*: Plot the values that you computed for a_t against t on semilogarithmic axes.

53. Assume the amount absorbed from the pill between time t and time $t+1$ is $20 \cdot (0.2)^t$.

(a) The drug has first order elimination kinetics. 40 % of the drug is eliminated from the blood each hour. Write down the recursion relation for a_{t+1} in terms of a_t.

(b) Assuming that $a_0 = 0$, meaning that no drug is present in the blood initially, calculate the amount of drug present at times $t = 1, 2, \ldots, 6$.

(c) What is the maximum amount of drug present at any time in this interval? At what time is this maximum amount reached?

(d) Use a spreadsheet to calculate the amount of drug present in hourly intervals from $t = 0$ up to $t = 24$.

(e) Show that, when t is large, the amount of drug present in the blood decreases approximately exponentially with t. *Hint*: Plot the values that you computed for a_t against t on semilogarithmic axes.

54. A drug has zeroth order elimination kinetics. At time $t = 0$ an amount $a_0 = 20$ mg is present in the blood. One hour later, at $t = 1$, an amount $a_1 = 14$ mg is present.

(a) Assuming that no drug is added to the blood between $t = 0$ and $t = 1$, calculate the amount of drug that is removed from the blood each hour.

(b) Write a recursion relation for the amount of drug a_t that is present at time t. Assume no extra drug is added to the blood.

(c) Find an explicit formula for a_t as a function of t.

(d) When does the amount of drug present in the blood first drop to 0?

55. A drug has first order elimination kinetics. At time $t = 0$ an amount $a_0 = 20$ mg is present in the blood. One hour later, at $t = 1$, an amount $a_1 = 14$ mg is present.

(a) Assuming that no drug is added to the blood between $t = 0$ and $t = 1$, calculate the percentage of drug that is removed each hour.

(b) Write a recursion relation for the amount of drug a_t present at time t. Assume no extra drug.

(c) Find an explicit formula for a_t as a function of t.

(d) Will the amount of drug present ever drop to 0 according to your model?

In Problems 56 and 57 you do not know whether a drug has zeroth order or first order elimination kinetics. You will use data to determine which type of kinetics it has.

56. You measure the concentration of the drug (in mg/ml) at time $t = 0$ and at time $t = 1$. No drug is added to the blood between $t = 0$ and $t = 1$. You measure the following data:

t	c_t
0	40
1	32

(a) Assume that the drug has zeroth order kinetics. What amount is eliminated from the blood each hour?

(b) Assume that the drug has zeroth order kinetics and no more drug is added to the blood. Write a recursion relation for c_t and predict c_2.

(c) Now assume the drug has first order elimination kinetics. What percentage of drug is eliminated from the blood each hour?

(d) Assume that the drug has first order kinetics and no more drug is added. Write a recursion relation for c_t and predict c_2.

(e) You measure the concentration at time $t = 2$ and find $c_2 = 25.6$. By comparing with your predictions from (b) and (d), decide: Does the drug have zeroth or first order kinetics?

57. You measure the concentration of the drug (in mg/ml) at time $t = 0$ and at time $t = 1$. No drug is added to the blood in this interval. You measure the following data:

t	c_t
0	50
1	35

(a) Assume that the drug has zeroth order kinetics. What amount is eliminated from the blood each hour?

(b) Assume that the drug has zeroth order kinetics and no more drug is added. Write a recursion relation for c_t and predict c_2.

(c) Now assume the drug has first order elimination kinetics. What percentage of drug is eliminated from the blood each hour?

(d) Assume that the drug has first order kinetics and no more drug is added to the blood. Write a recursion relation for c_t and predict c_2.

(e) You measure the concentration at time $t = 2$ and find $c_2 = 20$. By comparing with your predictions from (b) and (d), decide: Does the drug have zeroth or first order kinetics?

58. Tylenol in the Body A patient is taking Tylenol (a painkiller that contains acetaminophen) to treat a fever. The data in this question is taken from ***Rawlins, Henderson, and Hijab (1977)***.

At $t = 0$ the patient takes their first pill. One hour later the drug has been completely absorbed and the blood concentration, measured in μg/ml, is 15. Acetaminophen has first order elimination kinetics; in one hour, 23% of the acetaminophen present in the blood is eliminated.

(a) Write a recursion relation for the concentration c_t of drug in the patient's blood. For $t \geq 1$ you may assume for now that no other pills are taken after the first one.

(b) Find an explicit formula for c_t as a function of t.

(c) Suppose that the patient follows the directions on the pill box and takes another Tylenol pill 4 hours after the first (at time $t = 4$). What is the concentration at the time at which the second pill is taken? In others words, what is c_4?

(d) Over the next hour 15 μg/ml of drug enter the patient's bloodstream. So, c_5 can be calculated from c_4 using the word equation:

$$c_5 = c_4 + \begin{matrix}\text{amount added}\\ \text{to bloodstream}\end{matrix} - \begin{matrix}\text{amount eliminated}\\ \text{from bloodstream}\end{matrix}$$

Given that the amount added is 15 μg/ml, and the amount eliminated is $0.23 \cdot c_4$, calculate c_5.

(e) For $t = 5, 6, 7, 8$ the drug continues to be eliminated at a rate of 23% per hour. No pills are taken and no extra drug enters the patient's blood. Compute c_8.

(f) At time $t = 8$, the patient takes another pill. Calculate c_9. Do not forget to include elimination of drug between $t = 8$ and $t = 9$.

(g) We want to calculate the maximum concentration of drug in the patient's blood. We know that concentrations are highest in the hour after a pill is taken, namely at time $t = 1$, $t = 5$, $t = 9, \ldots$. Define a sequence C_n representing the concentration of the drug one hour after the nth pill is taken.

(h) What terms of the original sequence $\{c_t : t = 1, 2, \ldots\}$ are C_1, C_2, and C_3?

(i) Explain why

$$C_{n+1} = (0.77)^4 \cdot C_n + 15$$

and $c_1 = 15$

(j) From the recursion relation, assuming that the patient continues to take Tylenol pills at 4-hour intervals, calculate C_1, C_2, C_3, C_4, C_5, and C_6.

(k) Does C_n increase indefinitely, or do you think that it converges?

(l) By looking for fixing point of the recursion relation in (h), find the limit of C_n as $n \to \infty$.

Because of complex interactions with other drugs, some drugs have zeroth order elimination kinetics in some circumstances, and first order kinetics in other circumstances, depending on what other drugs are in the patient's system, as well as on age and preexisting medical conditions. In Problems 59–62 you will use the data on how concentration varies with time to determine whether the drug has zeroth or first order kinetics.

59. Given the following sequence of measurements for drug concentration, determine whether the drug has zeroth or first order kinetics.

t (Hours)	0	1	2	3
c_t (mg/liter)	16	12	9	6.75

60. Given the following sequence of measurements of drug concentration, determine whether the drug has zeroth or first order kinetics.

t (Hours)	1	2	3	4
c_t (μg/ml)	20	18	16	14

61. Given the following sequence of measurements of drug concentration, determine whether the drug has zeroth or first order kinetics.

t (Hours)	1	2	3	4
c_t (μg/ml)	40	36	32	28

62. Given the following sequence of measurements for drug concentration, determine whether the drug has zeroth or first order kinetics.

t (Hours)	0	2	4	6
c_t (mg/ml)	40	36	32.4	29.16

63. Hormone Implant You are studying an implanted contraceptive that releases hormone continuously into a patient's blood. The data in this question are from ***Nilsson et al. (1986)***.

The device adds 20 μg of hormone to the blood each day. In the blood the hormone has first order elimination kinetics; 4% of the hormone is eliminated each day.

(a) Let the amount of hormone in the blood on day t be a_t. Write a word equation for the change in a_t over one day.

(b) Put in mathematical formulas for each of the terms in your word equation from (a).

(c) Assume that on day 0 no hormone is present in the patient's blood, in other words, $a_0 = 0$. Use your equation from (b) to compute the amount of hormone in the blood on days 1, 2, 3, 4, 5, 6.

(d) Over time the level of hormone in the blood converges to a limit. Find the value of this limit by looking for a fixed point of your recurrence relation in (b).

64. For some drugs dosing must start small, but increases over time. Suppose a patient is on such a drug. Her treatment begins on day 0. On day 0 she receives a dose of 10 mg. On day 1 the dose increases to 20 mg. On day 3, the dose increases to 30 mg, and so on, with the dose increasing by 10 mg each day. The drug has first order kinetics in the blood; each day half of the drug is eliminated from the blood.

(a) Let a_t be the amount of drug in the patient's blood on day t. Write a word equation for the change in a_t over one day.

(b) Put in formulas for each term in your word equation in part (a) to derive from it a recurrence relation for a_{t+1} in terms of a_t.

(c) Assuming there is initially no drug in the patient's blood, in other words, $a_0 = 0$, calculate the amount of drug present on day $t = 1, 2, \ldots, 5$.

(d) Suppose that the drug dose stops increasing once a_t reaches 100 mg. On what day is that level reached?

Chapter 2 Review

Key Terms

Discuss the following definitions and concepts:

1. First order kinetics
2. Zeroth order kinetics
3. Population growth model
4. Reproductive rate
5. Exponential growth
6. Growth constant
7. Fixed point
8. Equilibrium
9. Recursion
10. Solution
11. Density independence
12. Sequence
13. First-order recurrence equation
14. Limit
15. Long-term behavior
16. Convergence, divergence
17. Limit laws
18. Difference equation
19. Beverton–Holt model
20. Density dependence
21. Carrying capacity
22. Growth parameter
23. Discrete logistic equation
24. Periodic behavior
25. Chaos
26. Drug absorption model
27. Drug absorption
28. Drug elimination

Review Problems

In Problems 1–10, find the limits.

1. $\lim_{n\to\infty} 2^{-n}$

2. $\lim_{n\to\infty} 3^{n}$

3. $\lim_{n\to\infty} 40(1-4^{-n})$

4. $\lim_{n\to\infty} \dfrac{2}{1+2^{-n}}$

5. $\lim_{n\to\infty} a^{n}$ when $a > 1$

6. $\lim_{n\to\infty} a^{n}$ when $0 < a < 1$

7. $\lim_{n\to\infty} \dfrac{n(n+1)}{n^2-1}$

8. $\lim_{n\to\infty} \dfrac{n^2+n-6}{n-2}$

9. $\lim_{n\to\infty} \dfrac{\sqrt{n}}{n+1}$

10. $\lim_{n\to\infty} \dfrac{n+1}{\sqrt{n}}$

In Problems 11–14, write a_n explicitly as a function of n on the basis of the first five terms of the sequence a_n, $n = 0, 1, 2, \ldots$.

11. $\frac{1}{2}, \frac{3}{4}, \frac{5}{6}, \frac{7}{8}, \frac{9}{10}$

12. $\frac{2}{2}, \frac{6}{4}, \frac{12}{8}, \frac{20}{16}, \frac{30}{32}$

13. $\frac{1}{2}, \frac{2}{5}, \frac{3}{10}, \frac{4}{17}, \frac{5}{26}$

14. $0, \frac{1}{3}, \frac{2}{4}, \frac{3}{5}, \frac{4}{6}$

15. Saving the Red Wolf You are trying to build a mathematical model for the size of the red wolf population in North Carolina. Red wolves are a critically endangered species, once found throughout the southeastern United States but now found mainly in a single North Carolina wildlife reserve. The data in this question come from ***U.S. Fish and Wildlife Service (2007)***.

(a) If the population size t years after the start of a study on red wolf numbers is N_t, start by writing a word equation so that N_{t+1} can be predicted from N_t:

$$N_{t+1} = N_t + \begin{matrix}\text{number of pups}\\\text{born in one year}\end{matrix} - \begin{matrix}\text{number of wolves}\\\text{that die}\end{matrix}$$

We will derive the formula of each of the terms in the word equation.

(i) It is possible to count the number of pups born in given year. A subset of these data is given in the following table:

Year	N_t	Number of Pups Born
1990	18	3
1993	44	16
1996	70	16
1999	126	37
2002	123	33
2005	115	41

Plot the number of pups against the population size N_t. Show that the data are consistent with the following formula for the number of pups born in one year:

$$\text{Number of pups born in one year} = 0.28N_t$$

[*Hint*: Draw the line given by this equation on your data plot.]

(ii) Independently conservationists are measuring the number of red wolf deaths in each year. Red wolves are killed by hunters and ranchers, or by being struck by vehicles. The conservationists estimate that in one year 22% of red wolves are killed. Write recurrence relation for the population size N_t.

(iii) Assuming that the current population size is $N_0 = 130$ wolves, use your formula from (ii) to predict the population size N_t for the next 10 years (i.e., calculate $N_1, N_2, \ldots, N_{10}$).

(iv) The current conservation goal for the wild red wolf population is to reach 220 individuals. When, according to your model, will that population size be reached?

(b) You want to use your model to evaluate the effectiveness of different proposed conservation strategies.

In addition to the wild red wolf population there is a captive breeding population. Captive bred wolves can be introduced into the wild. Suppose that r captive bred wolves are added to the wild population each year. However, captive bred wolves are less likely to survive in the wild than wolves born in the wild. Among introduced wolves 43% die each year. Assume that a mixed population of wild-born and introduced red wolves has a death rate of 30 % each year.

(i) Explain why your model for population size should be modified to:

$$N_{t+1} = N_t + 0.28 \cdot N_t - 0.30 \cdot N_t + r$$

(ii) Assume that $N_0 = 130$. Calculate the population size for 10 years (i.e., calculate $N_1, N_2, \ldots, N_{10}$) assuming $r = 5$.

(c) Red wolf parents will foster any newborn pups that are introduced into their dens. These fostered pups then have the same survival rate as wild-born wolves. Write down a recurrence equation for the wolf population size, assuming r pups are introduced each year. Since pups can be introduced more quickly than adult wolves, it is possible to introduce $r = 10$ pups per year by this route. If this strategy is used, will the population reach the target size of 220 individuals earlier or later than it would if strategy (b) were used?

16. The Fibonacci Sequence The Fibonacci equation is a classical (and very unrealistic) model for the growth of a population. Suppose that after t months there are N_t organisms present in the population. The Fibonacci equation provides a recurrence relation for N_{t+1}:

$$N_{t+1} = N_t + N_{t-1} \tag{2.19}$$

So to predict N_{t+1}, the population size for the next month, you need to know both the current population size, N_t, and their size in the previous month, N_{t-1}. This kind of recurrence relation is called a second order recurrence relation.

(a) Assume $N_0 = 1$ and $N_1 = 1$. Calculate N_2 using (2.19).

(b) Calculate N_3, N_4, N_5, and N_6.

(c) Use a spreadsheet to calculate N_{20}.

(d) When we introduced first order recurrence relations, we were able to show by iteratively applying the recurrence relation that its solution is exponential in time. By using your spreadsheet to make a semilog plot of N_t against t, show that the population size in the Fibonacci sequence also grows exponentially.

17. Fish Population A fish population is modeled by the discrete logistic equation. Specifically, if during month t there are N_t fish, then:

$$N_{t+1} = 2N_t - \frac{N_t^2}{200}$$

Recall that the term $2N_t$ means that the reproduction rate for a fish population far below the carrying capacity is 1.

(a) Assuming that initially there are 10 fish in the lake (in other words, $N_0 = 0$), calculate the position size after $t = 1, 2, 3, 4$ months.

(b) What size does the population converge to as $t \to \infty$?

(c) In fact, when you examine the fish population in the real lake, you find that the limiting fish population is actually equal to 160 fish. You suspect that fishing is responsible for the decrease in population size. Assume that a fraction p of the fish is caught and removed from the lake each month. You must then modify the model to:

$$N_{t+1} = 2 \cdot N_t - \frac{N_t^2}{200} - p \cdot N_t$$

Calculate the limiting population size for a population that obeys this equation. (Your answer will need to contain p as an unknown parameter.)

(d) From your model and the measured population size, calculate the proportion of fish being caught each month.

18. Extended Release Drugs Adderall XR is an extended release version of Adderall. Adderall XR pills are designed to release the drug slowly. To model the effect of taking a slowly released drug, we will not assume that the blood concentration jumps during the hour after the pill is taken, but that drug is slowly added to the blood by being absorbed from the gut. We need to build a two-compartment model in which the amount of drug in the gut and in the blood are separately modeled. The model in this equation is modified from ***J. J. McGough et al. (2003)***.

(a) First we will build a model for the total amount of drug in the gut. At time $t = 0$, 10 mg of the drug is present in the gut. From the gut it passes slowly into the blood. The passage of drug to the blood has first order kinetics; that is, every hour 42% of the drug remaining in the gut is passed from the gut into the blood. No more pills are taken, so no extra drug is added to the gut. Define a sequence a_t representing the amount of Adderall XR in the patient's gut t hours after the pill was taken. Write down a recursion relation for a_t.

(b) Find an explicit formula for a_t.

(c) Calculate the time at which the amount of drug remaining in the gut drops to 1% of its starting value; that is, find t for which

$$a_t = 0.01a_0$$

(d) Now let's build a model for the amount of drug in the patient's blood. Define another sequence b_t that represents the amount of Adderall XR in the patient's blood at time t. At time $t = 0$ there is no Adderall XR in the blood.

To model b_t we start with a word equation:

$$b_{t+1} = b_t + \begin{matrix}\text{amount of drug that} \\ \text{enters blood from the gut}\end{matrix} - \begin{matrix}\text{amount of drug that is} \\ \text{eliminated from blood}\end{matrix}$$

We need mathematical expressions for the terms in this equation.

1. Any drug that leaves the gut enters the blood. We know that 42% of the Adderall XR in the gut leaves the gut each hour. So the amount of drug that enters the blood is $0.42a_t$.

2. Adderall XR has first order elimination kinetics; 6% of the Adderall XR present in the blood leaves the blood each hour. So the amount that leaves the blood is $0.06 \cdot b_t$.

Write the recursion relation for b_t. *Hint*: Use your answer from part (b) to substitute for a_t.

(e) Calculate how the amount of drug present in the blood changes over a 6-hour interval starting at $t = 0$; that is, calculate $b_0, b_1, b_2, \ldots, b_6$. What is the maximum amount of drug in the blood? At what time is the amount of drug in the blood highest?

T **(f)** Using a spreadsheet calculate and plot the amount of drug present in the blood over a 24-hour interval (i.e., calculate b_0, $b_1, b_2, \ldots, b_{24}$).

(g) Suppose that, instead of being slowly released, the drug entered the bloodstream immediately after the pill was taken at time $t = 0$.

(i) Explain why the amount of Adderall XR in the blood would then be modeled by a recursion relation: $b_{t+1} = 0.94 \cdot b_t$, $b_0 = 10$ mg.

(ii) Calculate b_{24}, the amount of drug present after 24 hours under the new model.

(iii) Compare your answer from (ii) with your answer to (f) above. Which strategy (slow release or immediate absorption) gives the highest amount of drug present in the blood 24 hours after the pill was taken?

19. Insulin Pump You are studying an insulin pump that is being tested on patients for the first time. An insulin pump is a device that treats type I diabetes, a condition in which patients are unable to produce insulin to digest sugar. The data in this question is taken from ***Lauritzen et al. (1983)***. The pump releases insulin at a constant rate, adding an amount of 0.01 IU of insulin to the blood each minute. (IU is a unit for the total amount of a drug. For insulin 1 IU = 0.035 mg.) In the blood insulin is eliminated with first order kinetics; that is, a fixed but unknown fraction, p, of insulin is eliminated from the blood each minute.

(a) Let the amount of insulin in the blood after t minutes be a_t. Assume that a_t is measured in IU. Write a recurrence relation for a_{t+1} as a function of a_t. Your recurrence relation will include p as an unknown parameter.

(b) For patient 1 you measure the amount of insulin in the blood every hour. You obtain the following data:

t	0	1	2	3	4	5
a_t	0	0.010	0.019	0.027	0.034	0.040

Show that $p = 0.11$ is a good fit to this data. [*Hint*: Plot a_{t+1} against a_t and find a line that goes through all of the data.]

T **(c)** Using a spreadsheet, calculate the value of a_t for $t = 6, 7, 8, 9, 10$.

(d) In fact, you find the following data:

t	6	7	8	9	10
a_t **(IU)**	0.036	0.032	0.028	0.025	0.022

Show that these data suggest that, after $t = 5$, the insulin pump stops adding insulin to the patient's blood.

20. Fishing Reserve You are developing a mathematical model to aid in the management of a fishing reserve. You are trying to model the number of fish removed from the reserve each week. To do this you stop restocking the reserve on week $t = 0$, and measure N_t, the number of fish in the reserve at the beginning of week t. For the first two weeks you collect the following data for N_t:

t	1	2
N_t	1000	750

(a) You want to consider two possible models for the numbers of fish in the reserve.

(Scenario 1) The same number of fish are removed each week.

(Scenario 2) The same percentage of fish are removed each week.

Scenario 1 describes what happens if each fisher takes the same number of fish each week, regardless of population size.

Describe a scenario that is modeled by **scenario 2**.

(b) Assume **scenario 1** is true for now. Calculate from the data the number of fish that are removed each week.

(c) Assuming that fish numbers are not increased in any way, either by restocking or reproduction, write a recurrence relation for N_{t+1} in terms of N_t.

(d) From your answer to (c) predict when the number of the fish in the reserve will drop to 0.

(e) Actually, your measurements of the number of fish in the lake at any time are based on sampling and are not terribly accurate. In fact, each measurement has 10 % accuracy, so if you measure 1000 fish, then the true number is somewhere between $1000 \pm 10\%$, that is between 900 and 1100. Estimate the maximum and minimum values for the number of fish that are removed from the lake each week.

(f) When, based on your answer to (e), are the earliest and the latest weeks at which the population size will drop to zero?

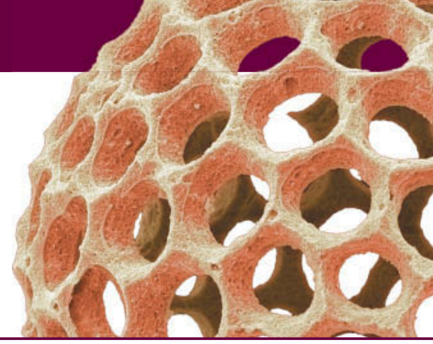
(g) You measure the following data for subsequent weeks:

t	3	4	5	6	7	8	9	10
N_t	596	480	447	355	284	213	140	137

By plotting the data for N_t against t on both standard and semilogarithmic axes, determine which scenario **(1** or **2)** most closely matches the observed data, and explain your answer.

CHAPTER

3



Limits and Continuity

The two concepts of limits and continuity are fundamental to differential calculus. Specifically, in this chapter we will learn how to

- determine the value of a function $f(x)$ at $x = c$ as x approaches c, both from graphs and mathematical expressions defining the function;
- determine whether a function is continuous or discontinuous at a point;
- identify where a function is continuous and where it is discontinuous;
- use the rigorous definitions of a limit and of continuity.

3.1 Limits

In Chapter 2, we discussed limits of the form $\lim_{n\to\infty} a_n$, where n took on integer values. In this chapter, we will consider limits of the form

$$\lim_{x\to c} f(x) \tag{3.1}$$

where x is now a continuously varying real variable that tends to a fixed value c (which may be finite or infinite). Let's look at an example that will motivate the need for limits of the form (3.1).

Ecologists are interested in quantifying the diversity of habitats. In Section 1.4 we discussed how some of the relationships that can then be explored—for example, how diversity can be related to the number of predators, or changes in climate or other conditions. There are several ways to measure diversity; one of the most commonly used is the **Shannon diversity index**. Assume there are only two types of organism present in a habitat; let's call them species 1 and species 2. Suppose that a fraction p of the organisms are of species 1, and a fraction $1 - p$ are of species 2. Then the Shannon diversity index is defined to be

$$H(p) = -p\ln p - (1-p)\ln(1-p), 0 < p < 1 \tag{3.2}$$

This function is shown in Figure 3.1.

We see from the plot that diversity is largest when $p = 1/2$, that is, when there are equal numbers of organisms of each species. The Shannon diversity index also has the symmetry property that $H(p) = H(1-p)$ because if we switch the numbering of the species, that is, if we rename species 1 as species 2, and rename species 2 as species 1, then with the new labels the proportion of species 1 is $q = 1 - p$; and the Shannon diversity index is $H(q)$. However, the habitat is still the same, and its diversity should also be the same, that is, independent of which species we regard as species 1 and which as species 2. So $H(p) = H(q) = H(1-p)$.

Equation (3.2) can only be used when $0 < p < 1$, since $\ln(0)$ is undefined. But we may want to evaluate H when $p = 0$ or when $p = 1$ (that is, when species 2 is extinct or when species 1 is extinct). To define $H(0)$ use the fact that although we cannot use

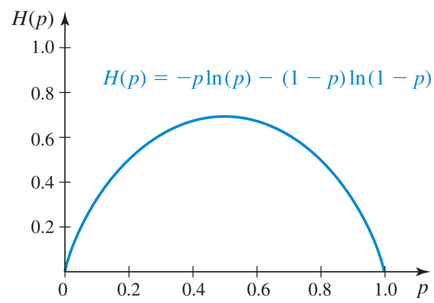


Figure 3.1 The Shannon diversity index for a habitat containing two species making up fractions p and $1 - p$ of the total number of organisms.

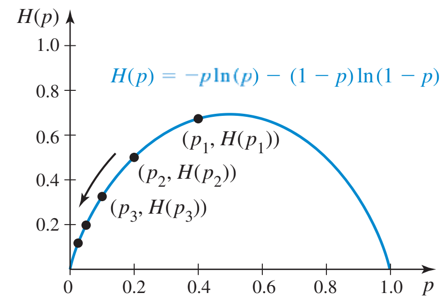


Figure 3.2 If $p_i \to 0$, then $H(p_i) \to 0$ also.

(3.2) to calculate $H(0)$, we can use it to evaluate $H(p)$ for any small values of p. We evaluate $H(p)$ for a sequence of values that gets smaller and smaller, shown in the following table.

p	0.1	0.01	0.001	10^{-4}	10^{-5}
$H(p)$	0.325	0.056	0.008	0.001	10^{-4}

As p gets smaller and smaller, so does $H(p)$. In fact, as Figure 3.2 shows, if $\{p_i\}$ is any sequence for which $p_i \to 0$ as $i \to \infty$, the corresponding sequence of values $\{H(p_i)\}$ will also converge to 0 as $i \to \infty$.

By the above argument we deduce that $H(0) = 0$, and this value makes sense; if $p = 0$, then only species 2 is present in the habitat. In that case we would want our measure of the habitat diversity to be very low. Indeed, $H(0) = 0$ is the smallest value the diversity can take.

3.1.1 A Non-Rigorous Discussion of Limits

Definition **The limit of $f(x)$, as x approaches c, is equal to L** means that $f(x)$ becomes arbitrarily close to L if x is close enough (but not equal) to c. We denote this statement by

$$\lim_{x\to c} f(x) = L$$

or $f(x) \to L$ as $x \to c$.

An alternative definition that is helpful when actually calculating the value of L is as follows:

Definition 2 $\lim_{x\to c} f(x) = L$ if for **any** sequence of numbers $\{x_i\}$, $i = 0, 1, 2, \ldots,$ for which $\lim_{i\to\infty} x_i = c$ and $x_i \neq c$, the sequence of values $f(x_i)$, $i = 0, 1, 2, \ldots$ also converges and $\lim_{i\to\infty} f(x_i) = L$.

If $\lim_{x\to c} f(x) = L$ and L is a finite number, we say that the limit **exists** and that $f(x)$ **converges** to L. If the limit does not exist, we say that $f(x)$ **diverges** as x tends to c.

To calculate the limit we choose x close, but not equal, to c. That is, when finding the limit of $f(x)$ as x approaches c, we do not simply plug c into $f(x)$. In fact, as the example of the Shannon diversity index showed, the limit may exist even if $f(x)$ is not defined at $x = c$. The value of $f(c)$ is irrelevant when we compute the value of $\lim_{x\to c} f(x)$.

Furthermore, when we say that "x is close enough to c," we mean that x approaches c from either direction. When x approaches c from only one side, we use the notation

$$\lim_{x\to c^+} f(x) \quad \text{when } x \text{ approaches } c \text{ from the right}$$

$$\lim_{x\to c^-} f(x) \quad \text{when } x \text{ approaches } c \text{ from the left}$$

and talk about right-handed and left-handed limits, respectively. The notation "$x \to c^+$" indicates that x approaches c but that x is always greater than c. How does $x \to c^+$ translate into the terms used in our definition? If $\lim_{x\to c^+} f(x) = L$, then $f(x)$ becomes arbitrarily close to L if x is both close enough to c and $x > c$. Similarly, when x approaches c from the left ($x \to c^-$), we need consider only values of x less than c.

Let's look at some examples.

EXAMPLE 1 Define $f(x) = x^2$, $x \in \mathbf{R}$. Find $\lim_{x\to 2} f(x)$.

Solution The graph of $f(x) = x^2$ (see Figure 3.3) immediately shows that the limit of x^2 is 4 as x approaches 2 (from either side). We also suspect this from the following table, where we compute values of x^2 for x close, but not equal, to 2:

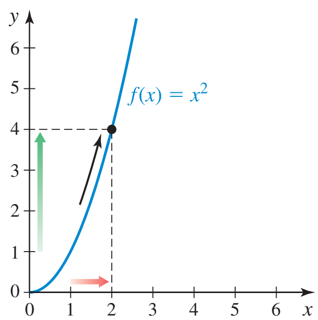


Figure 3.3 As x approaches 2, $f(x) = x^2$ approaches 4.

x	x^2	x	x^2
1.9	3.61	2.1	4.41
1.99	3.9601	2.01	4.0401
1.999	3.996001	2.001	4.004001
1.9999	3.99960001	2.0001	4.00040001

In the left half of the table we approach $x = 2$ from the left ($x \to 2^-$), whereas in the right half of the table we approach x from the right ($x \to 2^+$).

We find that

$$\lim_{x\to 2} x^2 = 4$$

Since this limit is a finite number, we say that the limit exists and that x^2 converges to 4 as x tends to 2. The fact that the limit is equal to $f(2)$ is a useful property that will be named later. Not all functions are like that. •

EXAMPLE 2 Find $\lim_{x\to 3} \dfrac{x^2-9}{x-3}$.

Solution Here both numerator and denominator tend to 0 as $x \to 3$. We define $f(x) = \frac{x^2-9}{x-3}$, $x \neq 3$. Since the denominator of $f(x)$ is equal to 0 when $x = 3$, we exclude $x = 3$ from the domain. Although we could proceed as in Example 1 by evaluating $f(x)$ for a sequence of values of x that approach $x = 3$, we can avoid this step if we notice that when $x \neq 3$, we can simplify the expression:

$$f(x) = \frac{x^2-9}{x-3} = \frac{(x-3)(x+3)}{x-3} = x+3 \qquad \text{for } x \neq 3$$

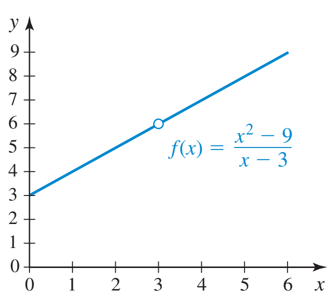


Figure 3.4 The graph of $f(x) = \frac{x^2-9}{x-3}$ is a straight line with the point (3, 6) removed.

We were able to cancel the term $x - 3$ because $x - 3 \neq 0$ for $x \neq 3$ and we assumed that $x \neq 3$. (If we allowed $x = 3$, then canceling $x - 3$ would mean dividing by 0.) The graph of $f(x)$ is a straight line with one point deleted at $x = 3$. (See Figure 3.4.) Taking the limit, we find that

$$\lim_{x\to 3} \frac{x^2-9}{x-3} = \lim_{x\to 3}(x+3)$$

Now, using either the graph of $y = x + 3$ for $x \neq 3$ or a table, we see that $\lim_{x\to 3}(x + 3) = 6$. We conclude that $\lim_{x\to 3} f(x)$ exists and that $f(x)$ converges to 6 as x tends to 3. Note that unlike Example 1, we cannot calculate the limit by substituting $x = 3$ into $f(x)$, because $f(3) = \frac{3^2-9}{3-3} = \frac{0}{0}$ is not defined. •

EXAMPLE 3 **Rate of Growth of a Population** A population of cells doubles in size every hour. At time $t = 0$ there are 1000 cells in the population. In Chapter 2 we showed that the population growth could be modeled using the formula

$$N(t) = 1000 \cdot 2^t, \quad t \geq 0 \tag{3.3}$$

In Chapter 2 we treated t as a discrete variable; that is, we only required the population to be known at integer times $t = 0, 1, 2, 3, \ldots$. But we can also treat t as a continuous variable and use equation (3.3) to evaluate the population size at any time t.

What is the growth rate of the population at time $t = 0$?

Solution We have not yet defined the growth rate; defining growth rate turns out to introduce one of the most important ideas in calculus. In Chapter 2 we calculated growth rate at time t from the difference in population size at t and at $t+1$; that is, from $N(t+1)-N(t)$. We may use the same expression when $N(t)$ is defined for all t, with

$$N(1) - N(0) = 1000 \cdot 2^1 - 1000 \cdot 2^0 = 2000 - 1000 = 1000.$$

But since t is a continuous variable there is nothing special about the time interval 1 anymore. We can measure the population size at any pair of times we want. In particular we can take $t = 0$ and $t = 0.5$. Then:

$$N(0.5) - N(0) = 1000 \cdot 2^{0.5} - 1000 = 414$$

Fewer births occur between $t = 0$ and $t = 0.5$ than between $t = 0$ and $t = 1$. It makes sense that the longer we wait between our two measurements, the more births will occur. To account for this we modify our definition of growth rate as follows:

$$\text{growth rate} = \frac{\text{change in population between } t = 0 \text{ and } t = h}{h} = \frac{N(h) - N(0)}{h}$$

that is, we divide the number of births by the length of the time interval over which births are counted. However, even using the new definition of growth rate, our data above produce different values for the growth rate.

Using the formula between $t = 0$ and $t = 1$, the growth rate is:

$$\frac{N(1) - N(0)}{1} = \frac{2000 - 1000}{1} = 1000 \text{ births/hr}$$

whereas between $t = 0$ and $t = 0.5$, the growth rate is:

$$\frac{N(0.5) - N(0)}{0.5} = \frac{1414 - 1000}{0.5} = 828 \text{ births/hr.}$$

Different values of h therefore produce different values for the growth rate. We can rationalize the discrepancy as follows; the number of births in any time interval increases as the number of cells in the population increases. Although the time intervals $0 \leq t \leq 0.5$ and $0.5 \leq t \leq 1$ have the same length, in the first interval the population starts with 1000 individuals, whereas in the second interval the population starts with 1414 individuals. Since there are more births in larger populations, there are more births in the second interval. Indeed, in any finite interval $[0, h]$ there will be more births in the second half of the interval, $h/2 \leq t \leq h$, than in the first half of the interval, $0 \leq t \leq h/2$. To make this difference as small as possible we make h as small as possible. Specifically, we can plot the growth rate estimate $\frac{N(h)-N(0)}{h}$ for different values of h; that is, we treat h as a variable (see Figure 3.5). As the figure shows, different interval lengths, h, produce different estimates for the birth rate. But as h approaches 0 these estimates converge to a limit. We consider a short sequence of values of h in the following table.

h	0.1	0.01	0.001	10^{-4}
$\dfrac{N(h) - N(0)}{h}$	717.73	695.56	693.39	693.17

Our table suggests that $\lim_{h\to 0} \frac{N(h)-N(0)}{h}$ exists and is equal to approximately 693.17.

In fact we will show in Chapter 4 how to calculate this limit exactly, and we will show that it is equal to $1000 \ln 2 = 693.15$. We identify this limit as the growth rate of

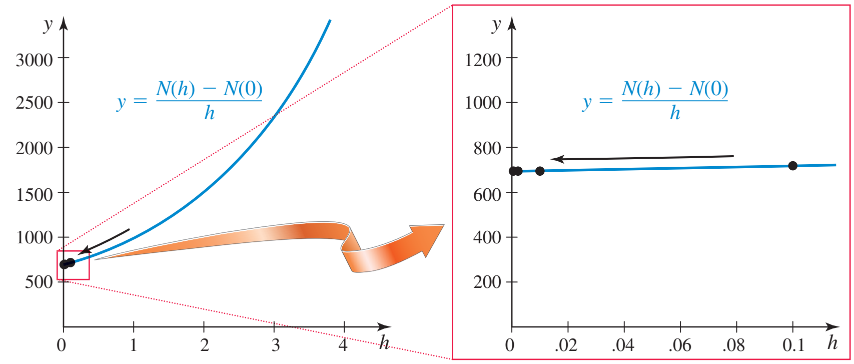


Figure 3.5 The growth rate estimate $\frac{N(h)-N(0)}{h}$ varies with h, but has a well-defined limit as $h \to 0$.

the population. This is another example where evaluating $f(x)$ at $x = c$ tells us nothing about $\lim_{x\to c} f(x)$. If we set $h = 0$ in Equation (3.3) we obtain $\frac{N(0)-N(0)}{0} = \frac{0}{0}$, which is not defined.

EXAMPLE 4 Find $\lim\limits_{x\to 0} e^{-|x|}$.

Solution We set

$$f(x) = e^{-|x|} = \begin{cases} e^{-x} & \text{for } x \geq 0 \\ e^{x} & \text{for } x < 0 \end{cases}.$$

Figure 3.6 indicates that $\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} e^{-x} = 1$ and $\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-} e^{x} = 1$. The two one-sided limits are equal so we conclude that

$$\lim_{x\to 0} e^{-|x|} = 1$$

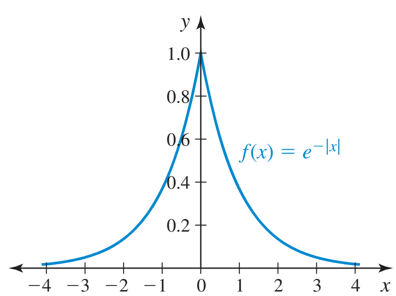


Figure 3.6 The graph of $f(x) = e^{-|x|}$ in Example 4.

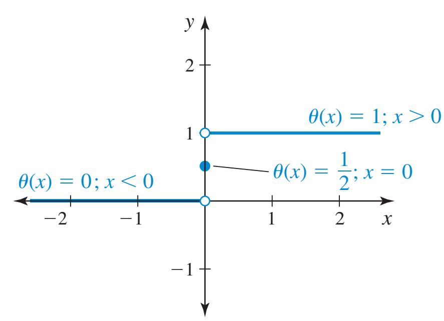


Figure 3.7 The Heaviside function $y = \theta(x)$.

EXAMPLE 5 **Biological Thresholds** Many biological processes change suddenly when a particular threshold is crossed. For example, bacteria may switch from one behavior (e.g., growth) to another (e.g., making spores) when the amount of a particular chemical crosses a threshold. When modeling these processes we often make use of the **Heaviside function**, which is defined as follows:

$$\theta(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{2} & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$$

θ is the Greek letter "theta."

The domain of $\theta(x)$ is the entire real number line. Show that $\lim_{x\to 0-} \theta(x)$ and $\lim_{x\to 0^+} \theta(x)$ exist but that $\lim_{x\to 0} \theta(x)$ does not exist.

Solution We plot $\theta(x)$ against x in Figure 3.7. We see that $\theta(x)$ converges to 1 as x tends to 0 from the right and that $\theta(x)$ converges to 0 as x tends to 0 from the left. We can write these limits using the one-sided limit notation

$$\lim_{x\to 0^+} \theta(x) = 1 \qquad \text{and} \qquad \lim_{x\to 0^-} \theta(x) = 0$$

and observe that the one-sided limits exist.

But since the right-hand limit differs from the left-hand limit, we conclude that $\lim_{x\to 0} \theta(x)$ does not exist because the phrase "x approaches 0" (or, in symbols, $\lim_{x\to 0}$) means that x approaches 0 from either direction. Specifically we could imagine the following sequence: $x_0 = 1$, $x_1 = -0.1$, $x_2 = +0.01$, $x_3 = -0.001$, $x_4 = 10^{-4}$, (in general in this sequence $x_i = (-1)^i 10^{-i}$). The terms in this sequence get smaller and smaller in absolute value, so $x_i \to 0$ as $i \to \infty$, but when we attempt to calculate the corresponding limit for the Heaviside function we get a sequence $\theta(x_0) = 1, \theta(x_1) = 0$, $\theta(x_2) = 1, \ldots$ that does not converge. •

As Example 5 shows, not all limits exist. In Example 5 both left and right limits existed but had different values. However, there are many reasons why limits may not exist. In the following Examples we explore some functions for which limits do not exist.

3.1.2 Pitfalls of Finding Limits

EXAMPLE 6 Find $\lim_{x\to 0} \dfrac{1}{x^2}$.

Solution A graph of $f(x) = 1/x^2$, $x \neq 0$, reveals that $f(x)$ increases without bound as $x \to 0$. (See Figure 3.8.) We also suspect such an increase when we plug in values close to 0. By choosing values sufficiently close to 0, we can get arbitrarily large values of $f(x)$:

First let's consider a sequence that approaches $x = 0$ from the right.

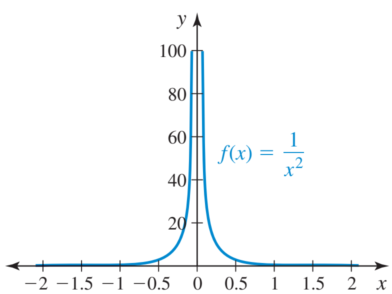


Figure 3.8 The graph of $f(x) = \frac{1}{x^2}$ in Example 6: The function grows without bound as x tends to 0.

$\boldsymbol{x}$	0.1	0.01	0.001	10^{-4}
$\boldsymbol{f(x)}$	100	10^4	10^6	10^8

So $f(x)$ gets arbitrarily large as the terms get smaller and smaller. The same occurs if our sequence approaches $x = 0$ from the left:

$\boldsymbol{x}$	-0.1	-0.01	-0.001	-10^{-4}
$\boldsymbol{f(x)}$	100	10^4	10^6	10^8

So $f(x)$ does not converge as $x \to 0$. •

When $\lim_{x\to c} f(x)$ does not exist, we say that $f(x)$ **diverges** as x tends to c. The divergence in Example 6 was such that the function grew without bound. This is an important case, and we define it in the following box:

Definition Divergent Limits

$$\lim_{x\to c} f(x) = +\infty \quad \text{if } f(x) \text{ increases without bound as } x \to c$$

$$\lim_{x\to c} f(x) = -\infty \quad \text{if } f(x) \text{ decreases without bound as } x \to c$$

Similar definitions can be given for one-sided limits, which we will need in the next example. Note that when we write $\lim_{x\to c} f(x) = +\infty$ (or $-\infty$), it is implied that $f(x)$ diverges as $x \to c$. In particular, this means that $\lim_{x\to c} f(x)$ does *not* exist. (A limit exists only if $\lim_{x\to c} f(x) = L$ and L is a real number. But $+\infty$ and $-\infty$ are not real numbers.) Nevertheless, we write $\lim_{x\to c} f(x) = +\infty$ (or $-\infty$) if $f(x)$ increases (or decreases) without bound as $x \to c$, since it is useful to know when a function does that.

EXAMPLE 7 Find $\lim_{x \to 3} \frac{1}{x-3}$.

Solution The graph of $f(x) = 1/(x-3)$, $x \neq 3$, in Figure 3.9 reveals that

$$\lim_{x \to 3^+} \frac{1}{x-3} = +\infty \quad \text{and} \quad \lim_{x \to 3^-} \frac{1}{x-3} = -\infty$$

We arrive at the same conclusion when we compute values of $f(x)$ for x close to 3. We see that if x is slightly larger than 3, then $f(x)$ is positive and increases without bound as x approaches 3 from the right. Likewise, if x is slightly smaller than 3, $f(x)$ is negative and decreases without bound as x approaches 3 from the left. We conclude that $f(x)$ diverges as x approaches 3. •

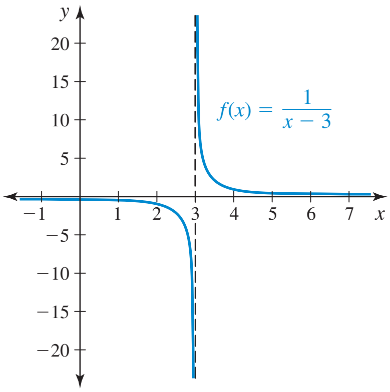


Figure 3.9 The graph of $f(x) = \frac{1}{x-3}$.

In the previous examples we have identified limits either by looking at the graph of the function or by finding a specific $\{x_i\}$ sequence that converges to $x = c$ and evaluating $\{f(x_i)\}$ for each term in that sequence. If the limit exists or is infinite this method will find it, but if the limit does not exist it can give the wrong answers, as the following example shows.

EXAMPLE 8 Find $\lim_{x \to 0} \sin \frac{\pi}{x}$.

Solution Simply using a calculator and plugging in values to find limits can yield wrong answers if we do not exercise proper caution. If we produced a table of values of $f(x) = \sin \frac{\pi}{x}$ for $x = 0.1, 0.01, 0.001, \ldots$, we would find that $\sin \frac{\pi}{0.1} = 0$, $\sin \frac{\pi}{0.01} = 0$, $\sin \frac{\pi}{0.001} = 0$, and so on. But we get a different answer if we try a different sequence that also approaches 0, say: $x = \frac{2}{11}, \frac{2}{101}, \frac{2}{1001}, \frac{2}{10001}$, (the ith term in this sequence is: $\frac{2}{10^i+1}$) and so on. Then, if we tabulate $f(x)$ for the terms in this sequence we obtain:

x	$\frac{2}{11}$	$\frac{2}{101}$	$\frac{2}{1001}$	$\frac{2}{10001}$
$f(x)$	-1	1	1	1

This sequence does not converge to 0 as x approaches 0. What is going on? In Figure 3.10 we plot the graph of $f(x)$ against x; the graph shows that the values of $f(x)$ oscillate infinitely often between -1 and $+1$ as $x \to 0$. We can see why as follows: As $x \to 0^+$, the argument in the sine function goes to infinity. (Likewise, as $x \to 0^-$, the argument goes to negative infinity.) That is,

$$\lim_{x \to 0^+} \frac{\pi}{x} = +\infty \qquad \text{and} \qquad \lim_{x \to 0^-} \frac{\pi}{x} = -\infty$$

As the argument of the sine function goes to $+\infty$ or $-\infty$, the function values oscillate between -1 and $+1$. Therefore, $\sin \frac{\pi}{x}$ continues to oscillate between -1 and $+1$ as $x \to 0$, and so the limit does not exist. •

The presence of these oscillations explains why different sequences produce different values for the limit. Our first sequence consisted only of zeros of the function; the function $f(x)$ takes the value 0 at each of these points so $f(x)$ appears to converge to 0. If we took a sequence of points corresponding to the peaks of the function then each value in the sequence will be equal to 1 (see Figure 3.10).

The behavior exhibited in Example 8 is called **divergence by oscillation**. This type of divergence shows why the definition of the limit requires that for all sequences that converge to c, $f(x)$ must converge to L. It is not enough for $f(x)$ to converge for just one sequence. However, we are aware of no functions arising in biological models that have divergence by oscillation or for which different sequences have different limits. Thus in practice checking the limit for just one sequence that approaches c from the left and one that approaches c from the right will be enough to establish a limit for any

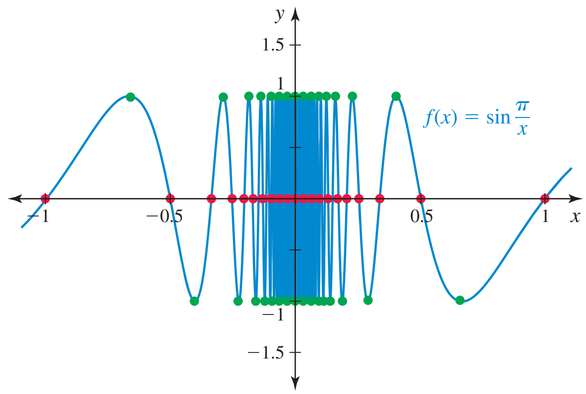


Figure 3.10 The graph of $f(x) = \sin \frac{\pi}{x}$ in Example 8. Two sequences approach $x = 0$. For each point in the (red) sequence, $f(x) = 0$. For each point in the (green) sequence, $f(x) = 1$.

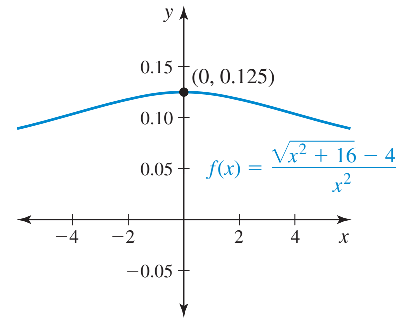


Figure 3.11 The graph of $f(x)$ in Example 9: As x tends to 0, the function approaches 0.125.

function that you are likely to encounter in a real application. However, the approach of calculating $f(x)$ for a sequence of values of x that get closer and closer to c needs to be used with some caution as the next Example shows.

EXAMPLE 9 Find $\lim_{x\to 0} \dfrac{\sqrt{x^2+16}-4}{x^2}$.

Solution The graph of $f(x) = \frac{\sqrt{x^2+16}-4}{x^2}$, $x \neq 0$, in Figure 3.11 indicates that the limit exists. So, on the basis of the graph, we conjecture that the limit is equal to 0.125. If, instead, we use a calculator to produce a table of values of $f(x)$ close to 0, something strange seems to happen:

x	0.01	0.001	0.0001	0.00001	0.000001	0.0000001
$f(x)$	0.1249998	0.125	0.125	0.125	0.1	0

As we get closer to 0, we first find that $f(x)$ gets closer to 0.125, but when we get very close to 0, $f(x)$ seems to drop to 0. What is going on? First, before you worry too much, note that $\lim_{x\to 0} f(x) = 0.125$ is the correct limit. In the next section, we will learn how to compute this limit without resorting to the (somewhat dubious) help of the calculator. The strange behavior of the calculated values happens because, when x is very small, the difference in the numerator is so close to 0 that the calculator can no longer accurately determine its value. The calculator can compute only a certain number of digits accurately, which is good enough for most cases. Here, however, we need greater accuracy. The same strange thing happens when you try to graph this function on a graphing calculator. When the x range of the viewing window is too small, the graph is no longer accurate. (Try, for instance, $-0.00001 \leq x \leq 0.00001$ and $-0.03 \leq y \leq 0.15$ as the range for the viewing window.) •

At the end of this chapter, we will discuss how limits are rigorously defined. The rigorous definition is conceptually similar to the one we used to define limits of the form $\lim_{n\to\infty} a_n$, but we will not use it to compute limits. As in Chapter 2, there are mathematical laws that will allow us to compute limits much more easily.

3.1.3 Limit Laws

We encountered limit laws of sequences in Chapter 2. Analogous laws hold for limits of the type $\lim_{x\to c} f(x)$.

Limit Laws for Functions Suppose that a is a constant and that

$$\lim_{x\to c} f(x) \quad \text{and} \quad \lim_{x\to c} g(x)$$

exist. Then the following rules hold:

1. $\lim_{x\to c} a\,f(x) = a \lim_{x\to c} f(x)$
2. $\lim_{x\to c}[f(x) + g(x)] = \lim_{x\to c} f(x) + \lim_{x\to c} g(x)$
3. $\lim_{x\to c}[f(x) \cdot g(x)] = \lim_{x\to c} f(x) \cdot \lim_{x\to c} g(x)$
4. $\lim_{x\to c} \dfrac{f(x)}{g(x)} = \dfrac{\lim_{x\to c} f(x)}{\lim_{x\to c} g(x)}$ provided that $\lim_{x\to c} g(x) \neq 0$

Let's start with the fact that:

$$\lim_{x\to c} x = c \tag{3.4}$$

In the optional discussion in Section 3.6, we will use the formal definition of limits to show that this equation is true, but you should find it evident from the graph of the function. Starting from equation (3.4), we can use the limit laws to compute limits of polynomials and rational functions.

EXAMPLE 10 Find $\lim_{x\to 2}[x^3 + 4x - 1]$.

Solution

$$\lim_{x\to 2}[x^3 + 4x - 1] = \lim_{x\to 2} x^3 + 4\lim_{x\to 2} x - \lim_{x\to 2} 1 \qquad \text{Using Rules 1 and 2.}$$

provided that the individual limits exist. For the first term, we use Rule 3,

$$\lim_{x\to 2} x^3 = \left(\lim_{x\to 2} x\right)\left(\lim_{x\to 2} x\right)\left(\lim_{x\to 2} x\right)$$

provided that $\lim_{x\to 2} x$ exists. From (3.4), $\lim_{x\to 2} x = 2$ so

$$\left(\lim_{x\to 2} x\right)\left(\lim_{x\to 2} x\right)\left(\lim_{x\to 2} x\right) = (2)(2)(2) = 8$$

To compute the second term, we use (3.4) again to obtain $\lim_{x\to 2} x = 2$. For the last term, we find that $\lim_{x\to 2} 1 = 1$. Since the individual limits exist,

$$\lim_{x\to 2}[x^3 + 4x - 1] = \lim_{x\to 2} x^3 + 4\lim_{x\to 2} x - \lim_{x\to 2} 1 = 8 + (4)(2) - 1 = 15$$

•

EXAMPLE 11 Find $\lim_{x\to 4} \dfrac{x^2+1}{x-3}$.

Solution

$$\lim_{x\to 4} \frac{x^2+1}{x-3} = \frac{\lim_{x\to 4}(x^2+1)}{\lim_{x\to 4}(x-3)} \qquad \text{Using Rule 4}$$

provided that the limits in the numerator and denominator exist and the limit in the denominator is not equal to 0. We evaluate the denominator and numerator limits separately:

$$\lim_{x\to 4}(x^2+1) = \left(\lim_{x\to 4} x^2\right) + \left(\lim_{x\to 4} 1\right) = (4)(4) + 1 = 17 \qquad \text{Using Rules 2 and 3}$$

Again, breaking up the limit into a sum of two limits is valid only because the individual limits exist. Similarly:

$$\lim_{x\to 4}(x-3) = \lim_{x\to 4} x - \lim_{x\to 4} 3 = 4 - 3 = 1 \qquad \text{Using Rule 2}$$

Again, using the limit laws is justified only after we have demonstrated that the individual limits exist. Since the limits in both the denominator and the numerator exist and the limit in the denominator is not equal to 0, we obtain

$$\lim_{x\to 4}\frac{x^2+1}{x-3} = \frac{17}{1} = 17$$

The computations in Examples 10 and 11 look somewhat awkward, and it appears that what we have done is plug 2 into the expression x^3+4x-1 in Example 10 and 4 into the expression $\frac{x^2+1}{x-3}$ in Example 11, even though we emphasized in the definition of limits that we are not allowed to simply plug c into $f(x)$ when computing $\lim_{x\to c} f(x)$. But, in essence, we did the calculation

$$\lim_{x\to 2}[x^3+4x-1] = 2^3 + (4)(2) - 1 = 15$$

in Example 10 and the calculation

$$\lim_{x\to 4}\frac{x^2+1}{x-3} = \frac{4^2+1}{4-3} = 17$$

in Example 11.

Even though we made a point that we cannot simply substitute the value c into $f(x)$ when we take the limit $x \to c$ of $f(x)$, the limit laws and (3.4) (which we will prove in Section 3.6) show that we can do just that when we take a limit of a polynomial or a rational function. Let's summarize this property and then look at two more examples that show how to compute limits of polynomials or rational functions by using these results.

> If $f(x)$ is a polynomial, then $\lim_{x\to c} f(x) = f(c)$.
>
> If $f(x)$ is a rational function $f(x) = \frac{p(x)}{q(x)}$ where $p(x)$ and $q(x)$ are polynomials, and if $q(c) \neq 0$, then
>
> $$\lim_{x\to c} f(x) = \lim_{x\to c}\frac{p(x)}{q(x)} = \frac{p(c)}{q(c)} = f(c).$$

EXAMPLE 12 Find $\lim_{x\to 3}[x^2-2x+1]$.

Solution Since $f(x) = x^2-2x+1$ is a polynomial, it follows that

$$\lim_{x\to 3}[x^2-2x+1] = 9 - 6 + 1 = 4$$

EXAMPLE 13 Find $\lim_{x\to -1}\dfrac{2x^3-x+5}{x^2+3x+1}$.

Solution Note that $f(x) = \frac{2x^3-x+5}{x^2+3x+1}$ is a rational function that is defined for $x = -1$. (The denominator is not equal to 0 when we substitute $x=-1$.) We find that

$$\lim_{x\to -1}\frac{2x^3-x+5}{x^2+3x+1} = \frac{2(-1)^3-(-1)+5}{(-1)^2+3(-1)+1} = \frac{4}{-1} = -4$$

When you use the limit laws for finding limits of the form

$$\lim_{x\to c}[f(x)+g(x)] \qquad \text{or} \qquad \lim_{x\to c}[f(x)\cdot g(x)] \qquad \text{or} \qquad \lim_{x\to c}\frac{f(x)}{g(x)}$$

you need to check first that both $\lim_{x\to c} f(x)$ and $\lim_{x\to c} g(x)$ exist and, in the case of $\lim_{x\to c} \frac{f(x)}{g(x)}$, that $\lim_{x\to c} g(x) \neq 0$. The next two examples illustrate the importance of checking the assumptions in the limit laws before applying them.

EXAMPLE 14 Find $\lim_{x\to 0} \dfrac{\frac{1}{x}}{\frac{1}{x}+1}$.

Solution We observe that neither $\lim_{x\to 0} \frac{1}{x}$ nor $\lim_{x\to 0} \left(\frac{1}{x}+1\right)$ exist. So we cannot use Rule 4 right away. Multiplying both numerator and denominator by x, however, will help:

$$\lim_{x\to 0} \frac{\frac{1}{x}}{\frac{1}{x}+1} = \lim_{x\to 0} \frac{1}{1+x}$$

Now we have a rational function on the right-hand side, and we can plug in 0 because the denominator, $1+x$, does not equal 0. We get

$$\lim_{x\to 0} \frac{1}{1+x} = \frac{1}{1+0} = 1$$

•

EXAMPLE 15 Find $\lim_{x\to 4} \dfrac{x^2-16}{x-4}$.

Solution The function $f(x) = \frac{x^2-16}{x-4}$ is a rational function, but since $\lim_{x\to 4}(x-4) = 0$, we cannot use Rule 4. Instead, we need to simplify $f(x)$ first:

$$\lim_{x\to 4} \frac{x^2-16}{x-4} = \lim_{x\to 4} \frac{(x-4)(x+4)}{x-4}$$

Because $x \neq 4$, we can cancel $x-4$ in the numerator and denominator, which yields

$$\lim_{x\to 4}(x+4) = 8$$

where we used the fact that $x+4$ is a polynomial to compute the final limit. •

Section 3.1 Problems

3.1.1 and 3.1.2

In Problems 1–32, use a table or a graph to investigate each limit.

1. $\lim_{x\to 2}(x^2-4x+4)$

2. $\lim_{x\to 2} \dfrac{x^2+3}{x+2}$

3. $\lim_{x\to -1} \dfrac{2x}{1+x^2}$

4. $\lim_{s\to 0} s(s^2-4)$

5. $\lim_{x\to \pi} 3\cos\dfrac{x}{4}$

6. $\lim_{t\to \pi/4} \sin(2t)$

7. $\lim_{x\to \pi/2} 2\sec\dfrac{x}{3}$

8. $\lim_{x\to \pi/2} \tan\dfrac{x-\pi/2}{2}$

9. $\lim_{x\to 0} e^{-x^2/2}$

10. $\lim_{x\to 0} \dfrac{e^x+1}{2x+3}$

11. $\lim_{x\to 0} \ln(x+1)$

12. $\lim_{t\to e} \ln\left(\dfrac{1}{t}\right)$

13. $\lim_{x\to 3} \dfrac{x^2-16}{x-4}$

14. $\lim_{x\to 3} \dfrac{x^2-9}{x+3}$

15. $\lim_{x\to \pi/2} \sin(2x)$

16. $\lim_{x\to \pi/2} \cos(x-\pi)$

17. $\lim_{x\to 0} \dfrac{1}{1+x^2}$

18. $\lim_{x\to 0} \dfrac{1}{x^2-1}$

19. $\lim_{x\to 0^+} \dfrac{1}{(1-e^{-x})}$

20. $\lim_{x\to 0} \dfrac{x}{1-e^{-x}}$

21. $\lim_{x\to 4^-} \dfrac{2}{x-4}$

22. $\lim_{x\to 3^+} \dfrac{1}{x-3}$

23. $\lim_{x\to 1^-} \dfrac{2}{1-x}$

24. $\lim_{x\to 2^+} \dfrac{4}{2-x}$

25. $\lim_{x\to 1^-} \dfrac{1}{1-x^2}$

26. $\lim_{x\to 2^+} \dfrac{2}{x^2-4}$

27. $\lim_{x\to 3} \dfrac{1}{(x-3)^2}$

28. $\lim_{x\to 0} \dfrac{1-x^2}{x^2}$

29. $\lim_{x\to 0^+} x\ln x$

30. $\lim_{x\to 0^+} x^2\ln x$

31. $\lim_{x\to 0} \dfrac{1-\sqrt{1-x^2}}{x^2}$

32. $\lim_{x\to 0} \dfrac{\sqrt{2-x}-\sqrt{2}}{2x}$

33. Use a table and a graph to find out what happens to

$$f(x) = \frac{2}{x} - \frac{1}{x^2}$$

as $x \to 0^-$. What happens as $x \to 0^+$?

34. Use a table and a graph to find out what happens to

$$f(x) = \exp\left(\frac{1}{x}\right)$$

as $x \to 0$.

35. Use a graphing calculator to investigate

$$\lim_{x\to 0} \frac{\sin x}{x}$$

36. Use a graphing calculator to investigate

$$\lim_{x\to 0^+} \left(\frac{1}{x^{1/2}} - \frac{1}{x}\right)$$

37. Tumor Size The Gompertz function is used to model the growth of tumors with time. According to the Gompertz function, the number of cells in a tumor increases with time according to:

$$N(t) = A\exp(-be^{-ct})$$

where A, b, and c are all positive constants that take different values for different tumor types and depending on whether the tumor is being treated or not.

(a) Assume that $A = b = c = 1$. Use a table or a graph to calculate $\lim_{t\to 0} N(t)$.

(b) Can you explain (without evaluating $N(t)$) why doubling A will double $\lim_{t\to 0} N(t)$?

(c) Show (it is okay to calculate $N(t)$) that changing the value of b changes $\lim_{t\to 0} N(t)$ but changing c does not affect $\lim_{t\to 0} N(t)$.

38. Rate of Growth of a Tumor The rate of proliferation (that is, reproduction) for the cells in a tumor varies depending on the size of the tumor. The Gompertz growth model is sometimes used to model this growth. According to the Gompertz model the total number of divisions occurring in 1 hour, R, depends on the number of cells, N, through a formula:

$$R(N) = dN \ln\left(\frac{A}{N}\right)$$

where d and A are both positive constants that depend on the type of tumor, whether it is being treated or not, and so on.

(a) Assume that $d = A = 1$. Use a table or a graph to show that $\lim_{N\to 0} R(N) = 0$.

(b) The per cell rate of reproduction tells us how many times any cell in the tumor will divide in one hour. It is given by $r(N) = \frac{R(N)}{N}$. Show that $\lim_{N\to 0} r(N)$ does not exist (again assume that $d = A = 1$).

3.1.3

In Problems 39–56, use the limit laws to evaluate each limit.

39. $\lim_{x\to -1} (x^3 + 7x - 1)$

40. $\lim_{x\to 2} (3x^4 - 2x + 1)$

41. $\lim_{x\to -5} (4 + 2x^2)$

42. $\lim_{x\to 2} (8x^3 - 2x + 4)$

43. $\lim_{x\to 3} \left(2x^2 - \frac{1}{x}\right)$

44. $\lim_{x\to +2} \left(\frac{x^2}{2} - \frac{2}{x^2}\right)$

45. $\lim_{x\to -3} \frac{x^3 - 20}{x + 1}$

46. $\lim_{x\to 1} \frac{x^3 + 1}{x + 2}$

47. $\lim_{x\to 3} \frac{3x^2 + 1}{2x - 3}$

48. $\lim_{x\to -2} \frac{1 + x}{1 - x}$

49. $\lim_{x\to 1} \frac{1 - x^2}{1 - x}$

50. $\lim_{u\to 2} \frac{4 - u^2}{2 - u}$

51. $\lim_{x\to 3} \frac{x^2 - 2x - 3}{x - 3}$

52. $\lim_{x\to 1} \frac{(x-1)^2}{x^2 - 1}$

53. $\lim_{x\to 2} \frac{2 - x}{x^2 - 4}$

54. $\lim_{x\to -4} \frac{x + 4}{16 - x^2}$

55. $\lim_{x\to -2} \frac{2x^2 + 3x - 2}{x + 2}$

56. $\lim_{x\to 1} \frac{1 - 2x + x^2}{1 - x}$

3.2 Continuity

3.2.1 What Is Continuity?

Consider the two functions

$$f(x) = \begin{cases} x + 3 & \text{if } x \neq 3 \\ 6 & \text{if } x = 3 \end{cases} \quad \text{and} \quad g(x) = \begin{cases} x + 3 & \text{if } x \neq 3 \\ 7 & \text{if } x = 3 \end{cases}$$

We are interested in how these functions behave for x close to 3. Both functions are defined for all $x \in \mathbf{R}$ and produce the same values for $x \neq 3$. Furthermore, using the limit laws

$$\lim_{x\to 3} f(x) = \lim_{x\to 3} g(x) = \lim_{x\to 3} x + 3 = 6 \tag{3.5}$$

But the two functions differ at $x = 3$: $f(3) = 6$ and $g(3) = 7$. So:

$$\lim_{x\to 3} f(x) = f(3) \qquad \text{but} \qquad \lim_{x\to 3} g(x) \neq g(3)$$

This difference can also be seen graphically [Figures 3.12(a) and 3.12(b)]: Although the graph of $f(x)$ can be drawn without lifting the pencil, in graphing $g(x)$ we need to lift the pencil at $x = 3$, since $\lim_{x\to 3} g(x) \neq g(3)$. We say that the function $f(x)$ is **continuous** at $x = 3$, whereas $g(x)$ is **discontinuous** at $x = 3$.

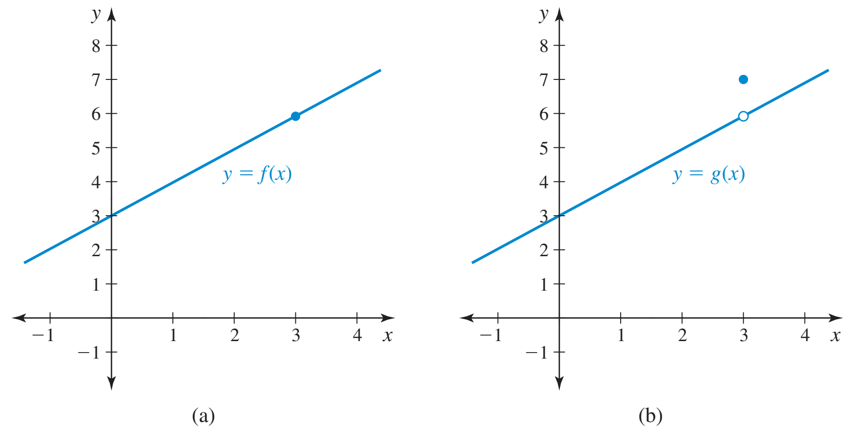


Figure 3.12 (a) The graph of $y = f(x)$ is continuous at $x = 3$. (b) The graph of $y = g(x)$ is discontinuous at $x = 3$.

Definition A function f is said to be **continuous at the point** $x = c$ if

$$\lim_{x \to c} f(x) = f(c)$$

To check whether a function is continuous at $x = c$, we need to check the following three conditions:

Conditions for $f(x)$ to be continuous at $x = c$.

1. $f(x)$ is defined at $x = c$.
2. $\lim_{x \to c} f(x)$ exists.
3. $\lim_{x \to c} f(x)$ is equal to $f(c)$.

If any of these three conditions fails, the function is **discontinuous** at $x = c$.

EXAMPLE 1 Show that $f(x) = 2x - 3$, $x \in \mathbf{R}$, is continuous at $x = 1$.

Solution We must check all three conditions:

1. $f(x)$ is defined at $x = 1$, since $f(1) = 2 \cdot 1 - 3 = -1$.
2. We use the fact that $\lim_{x \to c} x = c$ to conclude that $\lim_{x \to 1} f(x)$ exists.
3. Using the limit laws, we find that $\lim_{x \to 1} f(x) = -1$. This is the same as $f(1)$.

Since all three conditions are satisfied, $f(x) = 2x - 3$ is continuous at $x = 1$. •

EXAMPLE 2 Let $f(x) = \begin{cases} \dfrac{x^2 - x - 6}{x - 3} & \text{if } x \neq 3 \\ a & \text{if } x = 3 \end{cases}$ and find a so that $f(x)$ is continuous at $x = 3$.

Solution To compute $\lim_{x \to 3} \frac{x^2 - x - 6}{x - 3}$ we factor the numerator: $x^2 - x - 6 = (x - 3)(x + 2)$. Hence, for $x \neq 3$, $f(x) = x + 2$

$$\lim_{x \to 3} f(x) = \lim_{x \to 3} (x + 2) = 5$$

To ensure that $f(x)$ is continuous at $x = 3$, we require that $\lim_{x \to 3} f(x) = f(3)$. Since $f(3) = a$, we need to choose $a = 5$. This is the only choice for a that will make $f(x)$ continuous. Any other value of a would result in $f(x)$ being discontinuous. •

If we did not assign a value to $f(3)$, then the function $y = \frac{x^2-x-6}{x-3}$, $x \neq 3$, is not defined at $x = 3$ and is therefore automatically discontinuous there. (Condition 1 does not hold.) But Example 2 shows that we can *remove the discontinuity* by appropriately defining the function at $x = 3$.

EXAMPLE 3 **Species Diversity** In Section 3.1 we introduced the Shannon diversity index as a measure for the diversity of a particular habitat. If the habitat contains two species that are present in proportions p and $1 - p$, respectively, then the Shannon diversity index is defined to be:

$$H(p) = \begin{cases} -p \ln p - (1-p)\ln(1-p) & \text{if } 0 < p < 1 \\ 0 & \text{if } p = 0 \quad \text{or} \quad p = 1 \end{cases}$$

Show that $H(p)$ is continuous at $p = 0$ and $p = 1$.

Solution We check that $H(p)$ satisfies all three continuity criteria:

1. $H(0)$ and $H(1)$ are defined.
2. We saw in Section 3.1 that $\lim_{p\to 0^+} H(p)$ exists
3. $\lim_{p\to 0^+} H(p) = 0 = H(0)$

Note that since $H(p)$ is only defined for $0 \leq p \leq 1$, we can only approach $p = 0$ from the right, which is why $\lim_{p\to 0^+} H(p)$ was calculated in **3.** Since all three conditions are satisfied, $H(p)$ is continuous at $p = 0$. We can similarly show it is continuous at $p = 1$, or we may note that $H(p)$ is symmetric about $p = \frac{1}{2}$, that is, $H(p) = H(1-p)$. So since $H(p)$ is continuous at $p = 0$, it is also continuous at $p = 1 - 0 = 1$. •

It is not always possible to remove discontinuities, as the next three examples will show. In the first two, the discontinuity is a jump; that is, both the left-hand and the right-hand limits exist at the point where the jump occurs, but the limits differ. In the third example, the function grows without bound where it is discontinuous.

EXAMPLE 4 The **floor function**:

$$f(x) = \lfloor x \rfloor = \text{the largest integer less than or equal to } x$$

is graphed in Figure 3.13. The closed circles in the figure correspond to endpoints that are contained in the graph of the function, whereas the open circles correspond to endpoints that are not contained in the graph of the function. To explain this function, we compute a few values: $f(2.1) = 2$, $f(2) = 2$, and $f(1.9999) = 1$. The function jumps whenever x is an integer. Let k be an integer; then $f(k) = k$ and

$$\lim_{x\to k^+} f(x) = k, \quad \lim_{x\to k^-} f(x) = k - 1$$

That is, only when x approaches an integer from the right is the limit equal to the value of the function. The function is therefore discontinuous at integer values, and the discontinuity cannot be removed. If c is *not* an integer, then $f(x)$ is continuous at $x = c$. •

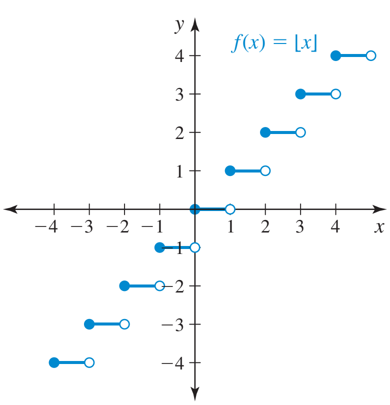


Figure 3.13 The floor function $f(x) = \lfloor x \rfloor$.

Example 4 motivates the definition of one-sided continuity:

Definition A function f is said to be **continuous from the right** at $x = c$ if

$$\lim_{x\to c^+} f(x) = f(c)$$

and **continuous from the left** at $x = c$ if

$$\lim_{x\to c^-} f(x) = f(c).$$

The function $f(x) = \lfloor x \rfloor$, $x \in \mathbf{R}$, of Example 4, is continuous from the right at $x = 1$ and at $x = 2$, and so on, but it is not continuous from the left. In the next example, the discontinuity is again a jump; however, this time we do not even have one-sided continuity.

EXAMPLE 5 Show that

$$\theta(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0 \\ \frac{1}{2} & \text{if } x = 0 \end{cases}$$

is discontinuous at $x = 0$ and that the discontinuity cannot be removed. The Heaviside function $\theta(x)$ was introduced in Section 3.1, Example 5, and is used to model biological processes that switch on when a threshold is crossed.

Solution The graph of $\theta(x)$ is shown in Figure 3.14. We see from the figure that

$$\lim_{x \to 0^+} \theta(x) = 1 \qquad \text{and} \qquad \lim_{x \to 0^-} \theta(x) = 0$$

The one-sided limits exist, but they are not equal [which implies that $\lim_{x \to 0} \theta(x)$ does not exist]. A jump occurs at $x = 0$. (See Figure 3.14.) This function does not exhibit even one-sided continuity because $\theta(x)$ is neither 1 nor 0 at $x = 0$. There is no way that we could assign a value to $\theta(0)$ such that the function would be continuous at $x = 0$. •

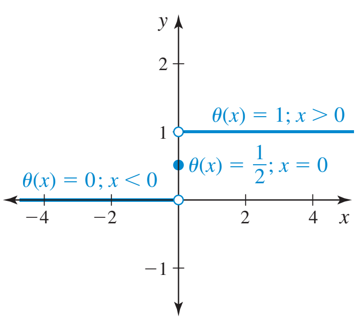


Figure 3.14 The Heaviside function $y = \theta(x)$ is discontinuous at $x = 0$.

EXAMPLE 6 At which point is the function $f(x) = \frac{1}{(x-4)^2}$ discontinuous? Can the discontinuity be removed?

Solution The graph of $f(x)$ is shown in Figure 3.15. The function $f(x)$ cannot be defined for $x = 4$, since $f(x)$ is of the form $\frac{1}{0}$ when $x = 4$. The function is defined for all other values of x. Therefore, we look at $x = 4$. We find that

$$\lim_{x \to 4} \frac{1}{(x-4)^2} = \infty \quad \text{(limit does not exist)}$$

Because ∞ is not a real number, we cannot assign a value to $f(4)$ such that $f(x)$ would be continuous at $x = 4$. We therefore conclude that $f(x)$ is discontinuous at $x = 4$ and the discontinuity cannot be removed. We will discuss this kind of discontinuity at more length in Section 5.6. •

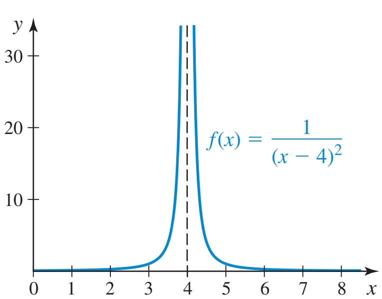


Figure 3.15 The function $f(x) = \frac{1}{(x-4)^2}$ is discontinuous at $x = 4$. Note that we have shifted the origin of the y-axis to coincide with the discontinuity at $x = 4$.

3.2.2 Combinations of Continuous Functions

Using the limit laws, we find that the following statements hold for combinations of continuous functions:

> **Combinations of Continuous Functions** Suppose that a is a constant and the functions f and g are continuous at $x = c$. Then the following functions are continuous at $x = c$:
>
> **1.** $a \cdot f$
>
> **2.** $f + g$
>
> **3.** $f \cdot g$
>
> **4.** $\dfrac{f}{g}$ provided that $g(c) \neq 0$

Proof We will prove only the second statement. We must show that conditions 1–3 from Section 3.2.1 all hold:

1. Note that $[f+g](x) = f(x)+g(x)$. Therefore, since $f(x)$ and $g(x)$ are defined at $x = c$, $f + g$ is defined at $x = c$ and $[f+g](c) = f(c)+g(c)$.
2. We assumed that f and g are continuous at $x = c$. This means, in particular, that

$$\lim_{x\to c} f(x) \qquad \text{and} \qquad \lim_{x\to c} g(x)$$

both exist. That is, the hypothesis in the limit laws holds, and so:

$$\lim_{x\to c}[f+g](x) = \lim_{x\to c}[f(x)+g(x)] = \lim_{x\to c} f(x) + \lim_{x\to c} g(x) \qquad \text{Using Rule 2 of the limit laws} \tag{3.6}$$

In other words, $\lim_{x\to c}[f+g](x)$ exists and condition 2 holds.
3. Since f and g are continuous at $x = c$, it follows that

$$\lim_{x\to c} f(x) = f(c) \qquad \text{and} \qquad \lim_{x\to c} g(x) = g(c) \tag{3.7}$$

Therefore, combining (3.6) and (3.7), we obtain

$$\lim_{x\to c}[f+g](x) = \lim_{x\to c} f(x) + \lim_{x\to c} g(x) = f(c)+g(c)$$

which is equal to $[f+g](c)$ and hence condition 3 holds.

Since all three conditions hold, it follows that $f + g$ is continuous at $x = c$. The other statements are shown in a similar way, using the limit laws. •

We say that a function f is **continuous on an interval I** if f is continuous for all $x \in I$. Note that if I is a closed interval, then continuity at the left (and, respectively, right) endpoint of the interval means continuous from the right (and, respectively, left). For example in Example 3 $H(p)$, is continuous on the interval $0 \le p \le 1$, because it is continuous from the right at $p = 0$ and from the left at $p = 1$. Many of the elementary functions are indeed continuous wherever they are defined. For polynomials and rational functions continuity follows immediately from the fact that certain combinations of continuous functions are continuous. We give a list of the most important cases:

> The following functions are continuous wherever they are defined:
>
> 1. polynomial functions
> 2. rational functions
> 3. power functions
> 4. trigonometric functions
> 5. exponential functions of the form a^x, $a > 0$ and $a \neq 1$
> 6. logarithmic functions of the form $\log_a x$, $a > 0$ and $a \neq 1$

The phrase "wherever they are defined" is crucial. It helps us to identify points where a function might be discontinuous. For instance, the power function $1/x^2$ is defined only for $x \neq 0$, and the logarithmic function $\log_a x$ is defined only for $x > 0$. We will illustrate the six cases cited in the preceding box in the next example, paying particular attention to the phrase "wherever they are defined."

EXAMPLE 7 For which values of $x \in \mathbf{R}$ are the following functions continuous?

(a) $f(x) = 2x^3 - 3x + 1$ **(b)** $f(x) = \dfrac{x^2+x+1}{x-2}$ **(c)** $f(x) = x^{1/4}$

(d) $f(x) = 3\sin x$ **(e)** $f(x) = \tan x$ **(f)** $f(x) = 3^x$

(g) $f(x) = 2\ln(x+1)$

Solution **(a)** $f(x)$ is a polynomial and is defined for all $x \in \mathbf{R}$; it is therefore continuous for all $x \in \mathbf{R}$.

(b) $f(x)$ is a rational function defined for all $x \neq 2$; it is therefore continuous for all $x \neq 2$.

(c) $f(x) = x^{1/4} = \sqrt[4]{x}$ is a power function defined for $x \geq 0$; it is therefore continuous for $x \geq 0$.

(d) $f(x)$ is a trigonometric function. $\sin x$ is defined for all $x \in \mathbf{R}$, so $3 \sin x$ is continuous for all $x \in \mathbf{R}$, by rule 1 for combinations of continuous functions.

(e) $f(x)$ is a trigonometric function. The tangent function is defined for all $x \neq \frac{\pi}{2} + k\pi$, where k is an integer; it is therefore continuous for all $x \neq \frac{\pi}{2} + k\pi$, where k is an integer.

(f) $f(x)$ is an exponential function. $f(x) = 3^x$ is defined for all $x \in \mathbf{R}$ and is therefore continuous for all $x \in \mathbf{R}$.

(g) $f(x)$ is a logarithmic function. $f(x) = 2\ln(x+1)$ is defined so long as $x + 1 > 0$ or $x > -1$; $f(x)$ is therefore continuous for all $x > -1$. •

The following result is useful in determining whether a composition of functions is continuous:

Theorem Continuity of Composed Functions If $g(x)$ is continuous at $x = c$ with $g(c) = L$ and $f(x)$ is continuous at $x = L$, then $(f \circ g)(x)$ is continuous at $x = c$. In particular,

$$\lim_{x \to c} (f \circ g)(x) = \lim_{x \to c} f[g(x)] = f[\lim_{x \to c} g(x)] = f[g(c)] = f(L)$$

To explain this theorem, recall what it means to compute $(f \circ g)(c) = f[g(c)]$. When we compute $f[g(c)]$, we input c into $g(x)$ to calculate $g(c)$, and then take the result $g(c)$ and plug $x = g(c)$ into the function $f(x)$ to obtain $f[g(c)]$. If, at each step, the functions are continuous, the resulting function will be continuous.

EXAMPLE 8 Determine where the following functions are continuous:

(a) $h(x) = e^{-x^2}$ **(b)** $h(x) = \sin \dfrac{\pi}{x}$ **(c)** $h(x) = \dfrac{1}{1 + 2x^{1/3}}$

Solution **(a)** Set $g(x) = -x^2$ and $f(x) = e^x$, then $h(x) = (f \circ g)(x)$. Since $g(x)$ is a polynomial, it is continuous for all $x \in \mathbf{R}$, and the range of $g(x)$ is $(-\infty, 0]$. $f(x)$ is continuous for all values in the range of $g(x)$. [In fact, $f(x)$ is continuous for all $x \in \mathbf{R}$.] It therefore follows that $h(x)$ is continuous for all $x \in \mathbf{R}$.

(b) Set $g(x) = \frac{\pi}{x}$ and $f(x) = \sin x$, then $h(x) = (f \circ g)(x)$. $g(x)$ is continuous for all $x \neq 0$. The range of $g(x)$ is the set of all real numbers, excluding 0. $f(x)$ is continuous for all x in the range of $g(x)$. Hence, $h(x)$ is continuous for all $x \neq 0$. Recall that we showed in Example 8 of Section 3.1 that

$$\lim_{x \to 0} \sin \frac{\pi}{x}$$

does not exist. That is, $h(x)$ is discontinuous at $x = 0$.

(c) Set $g(x) = x^{1/3}$ and $f(x) = \frac{1}{1+2x}$, then $h(x) = (f \circ g)(x)$. $g(x)$ is continuous for all $x \in \mathbf{R}$, since $g(x) = x^{1/3} = \sqrt[3]{x}$ and 3 is an odd integer. The range of $g(x)$ is $(-\infty, \infty)$. $f(x)$ is continuous for all real x different from $-1/2$. Since $g(-\frac{1}{8}) = -\frac{1}{2}$, $h(x)$ is continuous for all real x different from $-1/8$. Another way to see that we need to exclude $-\frac{1}{8}$ from the domain of $h(x)$ is by looking directly at the denominator of $h(x)$: we have $1 + 2x^{1/3} = 0$ when $x = -\frac{1}{8}$. •

If $f(x)$ is continuous at $x = c$, then $\lim_{x \to c} f(x) = f(c)$. The next three examples illustrate use of this property to calculate $\lim_{x \to c} f(x)$.

EXAMPLE 9 Find $\lim_{x\to 3} \sin\left(\frac{\pi(x^2-1)}{4}\right)$.

Solution The function $f(x) = \sin\left(\frac{\pi(x^2-1)}{4}\right)$ is continuous at $x = 3$. Hence,

$$\lim_{x\to 3} \sin\left(\frac{\pi(x^2-1)}{4}\right) = \sin\left(\frac{\pi(9-1)}{4}\right) = \sin(2\pi) = 0$$

•

EXAMPLE 10 Find $\lim_{x\to 1} \sqrt{2x^3 - 1}$.

Solution The function $f(x) = \sqrt{2x^3 - 1}$ is continuous at $x = 1$. Thus,

$$\lim_{x\to 1} \sqrt{2x^3-1} = \sqrt{(2)(1)^3 - 1} = \sqrt{1} = 1$$

•

EXAMPLE 11 Find $\lim_{x\to 0} e^{x-1}$.

Solution The function $f(x) = e^{x-1}$ is continuous at $x = 0$. Therefore,

$$\lim_{x\to 0} e^{x-1} = e^{0-1} = e^{-1}$$

•

We conclude this section by calculating the limit of the expression in Example 9 of Section 3.1.

EXAMPLE 12 Find $\lim_{x\to 0} \frac{\sqrt{x^2+16}-4}{x^2}$.

Solution We cannot apply Rule 4 of Section 3.1, since $f(x) = (\sqrt{x^2+16}-4)/x^2$ is not defined for $x = 0$. (If we plug in 0, we get the expression 0/0.) We use a trick that will allow us to find the limit: We rationalize the numerator. For $x \neq 0$, we find that

$$\begin{aligned}
\frac{\sqrt{x^2+16}-4}{x^2} &= \frac{(\sqrt{x^2+16}-4)}{x^2}\,\frac{(\sqrt{x^2+16}+4)}{(\sqrt{x^2+16}+4)} \\
&= \frac{x^2+16-16}{x^2(\sqrt{x^2+16}+4)} = \frac{x^2}{x^2(\sqrt{x^2+16}+4)} \\
&= \frac{1}{\sqrt{x^2+16}+4}
\end{aligned}$$

Note that we are allowed to divide by x^2 in the last step, since we are assuming that $x \neq 0$. Then

$$\lim_{x\to 0} \frac{\sqrt{x^2+16}-4}{x^2} = \lim_{x\to 0} \frac{1}{\sqrt{x^2+16}+4} = \frac{1}{8} = 0.125 \qquad \text{Using Rule 4 of the limit laws}$$

as we saw in Example 9 of Section 3.1. In Section 5.5, we will learn a different method for finding the limit of expressions of the form 0/0. •

Section 3.2 Problems

3.2.1

In Problems 1–4, show that each function is continuous at the given value.

1. $f(x) = 2x$, $c = 1$

2. $f(x) = -x$, $c = 0$

3. $f(x) = x^3 + 2x + 1$, $c = 2$

4. $f(x) = x^2 + 1$, $c = -1$

5. Show that

$$f(x) = \begin{cases} \frac{x^2-x-2}{x-2} & \text{if } x \neq 2 \\ 3 & \text{if } x = 2 \end{cases}$$

is continuous at $x = 2$.

6. Show that

$$f(x) = \begin{cases} \dfrac{2x^2+x-6}{x+2} & \text{if } x \neq -2 \\ -7 & \text{if } x = -2 \end{cases}$$

is continuous at $x = -2$.

7. Let

$$f(x) = \begin{cases} \dfrac{x^2-9}{x-3} & \text{if } x \neq 3 \\ a & \text{if } x = 3 \end{cases}$$

Which value must you assign to a so that $f(x)$ is continuous at $x = 3$?

8. Let

$$f(x) = \begin{cases} \dfrac{3+2x-x^2}{x-3} & \text{if } x \neq 3 \\ a & \text{if } x = 3 \end{cases}$$

Which value must you assign to a so that $f(x)$ is continuous at $x = 3$?

In Problems 9–12, determine at which points $f(x)$ is discontinuous.

9. $f(x) = \dfrac{1}{x-3}$

10. $f(x) = \dfrac{1}{x^2-1}$

11. $f(x) = \begin{cases} x^2-1 & \text{if } |x| \geq 1 \\ x-1 & \text{if } |x| < 1 \end{cases}$

12. $f(x) = \begin{cases} x^2-1 & \text{if } x \leq 0 \\ x-1 & \text{if } x > 0 \end{cases}$

13. Show that the floor function $f(x) = \lfloor x \rfloor$ is continuous at $x = 5/2$ but discontinuous at $x = 3$.

14. Show that the floor function $f(x) = \lfloor x \rfloor$ is continuous from the right at $x = 2$.

3.2.2

In Problems 15–24, find the values of $x \in \mathbf{R}$ for which the given functions are both defined and continuous.

15. $f(x) = 3x^4 - x^2 + 4$

16. $f(x) = \sqrt{x^2-1}$

17. $f(x) = \dfrac{x^2+1}{x-1}$

18. $f(x) = \cos(2x)$

19. $f(x) = e^{-|x|}$

20. $f(x) = \ln(x-2)$

21. $f(x) = \dfrac{x}{x+1}$

22. $f(x) = \exp[\sqrt{x-1}]$

23. $f(x) = \tan(2\pi x)$

24. $f(x) = \cos\left(\dfrac{2x}{3+x}\right)$

25. Let

$$f(x) = \begin{cases} x^2+2 & \text{for } x \leq 0 \\ x+c & \text{for } x > 0 \end{cases}$$

(a) Graph $f(x)$ when $c = 1$, and determine whether $f(x)$ is continuous for this choice of c.

(b) How must you choose c so that $f(x)$ is continuous for all $x \in (-\infty, \infty)$?

26. Let

$$f(x) = \begin{cases} \dfrac{1}{x} & \text{for } x \geq 1 \\ 2x+c & \text{for } x < 1 \end{cases}$$

(a) Graph $f(x)$ when $c = 0$, and determine whether $f(x)$ is continuous for this choice of c.

(b) How must you choose c so that $f(x)$ is continuous for all $x \in (-\infty, \infty)$?

27. **(a)** Show that

$$f(x) = \sqrt{x-1}, \quad x \geq 1$$

is continuous from the right at $x = 1$.

(b) Graph $f(x)$.

(c) Does it make sense to look at continuity from the left at $x = 1$?

28. **(a)** Show that

$$f(x) = \sqrt{x^2-4}, \quad |x| \geq 2$$

is continuous from the right at $x = 2$ and continuous from the left at $x = -2$.

(b) Graph $f(x)$.

(c) Does it make sense to look at continuity from the left at $x = 2$ and at continuity from the right at $x = -2$?

In Problems 29–48, find the limits.

29. $\lim\limits_{x\to\pi/3} \sin\left(\dfrac{x}{2}\right)$

30. $\lim\limits_{x\to-\pi/2} \cos(2x)$

31. $\lim\limits_{x\to\pi/2} \dfrac{\cos^2 x}{1-\sin^2 x}$

32. $\lim\limits_{x\to-\pi/2} \dfrac{1+\tan^2 x}{\sec^2 x}$

33. $\lim\limits_{x\to-1} \sqrt{4+5x^4}$

34. $\lim\limits_{x\to-2} \sqrt{6+x}$

35. $\lim\limits_{x\to-1} \sqrt{x^2+2x+2}$

36. $\lim\limits_{x\to1} \sqrt{x^3+4x-1}$

37. $\lim\limits_{x\to0} e^{-x^2/3}$

38. $\lim\limits_{x\to0} e^{3x+2}$

39. $\lim\limits_{x\to3} e^{x^2-9}$

40. $\lim\limits_{x\to-1} e^{x^2/2-1}$

41. $\lim\limits_{x\to0} \dfrac{e^{2x}-1}{e^x-1}$

42. $\lim\limits_{x\to0} \dfrac{e^{-x}-e^x}{e^{-x}+1}$

43. $\lim\limits_{x\to-2} \dfrac{1}{\sqrt{5x^2-4}}$

44. $\lim\limits_{x\to1} \dfrac{1}{\sqrt{3-2x^2}}$

45. $\lim\limits_{x\to0} \dfrac{\sqrt{x^2+9}-3}{x^2}$

46. $\lim\limits_{x\to0} \dfrac{5-\sqrt{25+x^2}}{2x^2}$

47. $\lim\limits_{x\to0} \ln(1-x)$

48. $\lim\limits_{x\to1} \ln[e^x \cos(x-1)]$

49. **Tumor Growth Rate** The rate of growth of a tumor (number of cells added in one day) depends on the current size of the tumor. If the number of cells added to the tumor in one day is R, and the number of cells in the tumor is N, then a commonly used model is the Gompertz growth model, which predicts that growth rate R will vary with size, N, according to

$$R(N) = dN \ln\left(\frac{A}{N}\right) \qquad N > 0.$$

where d and A are both positive constants that depend on the type of tumor and whether the tumor is being treated or not.

(a) Assume that $d = 1$ and $A = 10^6$. Show that if $R(0) = 0$, then $R(N)$ is continuous at $N = 0$ (you should use a table or graph to calculate $\lim\limits_{N\to0+} R(N)$).

(b) You notice that large tumors neither grow nor shrink, that is, $R(N) = 0$ for any tumor whose size exceeds a critical threshold, $N > N_c$. Accordingly, you propose to modify the Gompertz growth model to the following:

$$R(N) = \begin{cases} 0 & \text{if } N = 0 \\ dN \ln\left(\frac{A}{N}\right) & \text{if } 0 < N \le N_c \\ 0 & \text{if } N \ge N_c. \end{cases}$$

What value should be assigned to N_c to make $R(N)$ a continuous function of N for all $N \ge 0$?

50. Fungal Growth As a fungus grows, its rate of growth changes. Young fungi grow exponentially, while in larger fungi growth slows, and the total dimensions of the fungus increase as a linear function of time. You want to build a mathematical model that describes the two phases of growth. Specifically if $R(t)$ is the rate of growth given as a function of time, t, then you model

$$R(t) = \begin{cases} 2e^t & \text{if } 0 \le t \le t_c \\ a & \text{if } t > t_c. \end{cases}$$

where t_c is the time at which the fungus switches from exponential to linear growth and a is a constant.

(a) For what value of a is the function $R(t)$ continuous at $t = t_c$? (Your answer will include the unknown constant t_c).

(b) Assume that $t_c = 2$. Draw the graph of $R(t)$ as a function of t.

51. Panting in Animals Animals use different strategies to control their internal temperature depending on how hot they are. When the core temperature of a dog, duck, or cat exceeds a critical value, it will start to pant (make quick, gasping breaths that increase evaporation of water from the tongue and mouth). ***Vieth (1989)*** studied heat loss as a function of the ducks' core temperature, T. She found that different functions described heat loss below the temperature at which the ducks started to pant and above this temperature. If $H(T)$ is the rate of heat loss:

$$H(T) = \begin{cases} 0.6 & \text{if } T \le T_c \\ 4.3T - 183 & \text{if } T > T_c \end{cases}$$

(here T is measured in °C and $H(T)$ in watts per kg of body mass)

(a) Calculate the value of T_c that makes $H(T)$ continuous for all T.

(b) Draw the graph of the function $H(T)$ over the normal body temperature range for ducks: $41°\text{C} \le T \le 44°\text{C}$.

3.3 Limits at Infinity

The limit laws discussed in Section 3.1 also hold as x tends to ∞ (or $-\infty$). To calculate these limits we will often have to make use of Rule 4 from Section 3.1. To make it easier to use this result we will reproduce the statement of Rule 4 here:

> **Rule 4 of the Limit Laws** If $\lim_{x\to c} f(x)$ and $\lim_{x\to c} g(x)$ both exist, then $\lim_{x\to c} \frac{f(x)}{g(x)}$ exists and $\lim_{x\to c} \frac{f(x)}{g(x)} = \frac{\lim_{x\to c} f(x)}{\lim_{x\to c} g(x)}$ provided that $\lim_{x\to c} g(x) \ne 0$.

EXAMPLE 1 Find $\displaystyle\lim_{x\to\infty} \frac{x}{x+1}$.

Solution First let's try to determine the behavior of the function in the limit by evaluating it for some large values of x, shown in the following table.

x	1	10	100	1000
$\frac{x}{1+x}$	0.5	0.909	0.990	0.999

From the table we see that as $x \to \infty$, $\frac{x}{1+x}$ appears to converge to 1. Could we find the limit without using a table?

If we set $f(x) = x$ and $g(x) = x + 1$ we would like to use Rule 4 to calculate the limit. But neither $\lim_{x\to\infty} f(x)$ nor $\lim_{x\to\infty} g(x)$ exists. However, we can divide both numerator and denominator by x. When we do, we find that

$$\lim_{x\to\infty} \frac{x}{x+1} = \lim_{x\to\infty} \frac{1}{1+\frac{1}{x}}$$

Since $\lim_{x\to\infty} 1 = 1$ and $\lim_{x\to\infty}(1 + \frac{1}{x}) = 1$, both limits exist. Furthermore, $\lim_{x\to\infty}(1 + \frac{1}{x}) \ne 0$. We can now apply Rule 4 of Section 3.1 after having done the algebraic manipulation:

$$\lim_{x\to\infty} \frac{x}{x+1} = \lim_{x\to\infty} \frac{1}{1+\frac{1}{x}} = \frac{\lim_{x\to\infty} 1}{\lim_{x\to\infty}(1+\frac{1}{x})} = \frac{1}{1} = 1$$

We show a plot of the function $\frac{x}{1+x}$ in Figure 3.16.

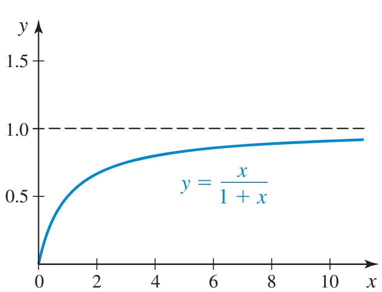


Figure 3.16 The function $y = \frac{x}{1+x}$ converges to 1 as $x \to \infty$.

In Example 1, we computed the limit of a rational function as x tended to infinity. Rational functions are ratios of polynomials. To find out how the limit of a rational function behaves as x tends to infinity, we will first compare the relative growth of functions of the form $y = x^n$: If $n > m$, then x^n dominates x^m for large x, in the sense that

$$\lim_{x\to\infty} \frac{x^n}{x^m} = \infty \qquad \text{and} \qquad \lim_{x\to\infty} \frac{x^m}{x^n} = 0$$

The preceding statement follows immediately if we simplify the fractions $\frac{x^n}{x^m} = x^{n-m}$ with $n - m > 0$ and $\frac{x^m}{x^n} = \frac{1}{x^{n-m}}$ with $n - m > 0$.

This limiting behavior is important when we compute limits of rational functions as $x \to \infty$.

EXAMPLE 2 Calculate the following three limits:

(a) $\lim_{x\to\infty} \dfrac{x^2 + 2x - 1}{x^3 - 3x + 1}$ **(b)** $\lim_{x\to\infty} \dfrac{2x^3 - 4x + 7}{3x^3 + 7x^2 - 1}$ **(c)** $\lim_{x\to\infty} \dfrac{x^4 + 2x - 5}{x^2 - x + 2}$

Solution To determine whether the numerator or the denominator dominates, we look at each of their leading terms. (The leading term is the term with the largest exponent.) The leading term of a polynomial tells us how quickly the polynomial increases as x increases.

(a) The leading term in the numerator is x^2, and the leading term in the denominator is x^3. As $x \to \infty$, the denominator grows much faster than the numerator. We therefore expect the limit to be equal to 0. We can show this by dividing both numerator and denominator by the higher of the two powers, namely, x^3. We get

$$\lim_{x\to\infty} \frac{x^2 + 2x - 1}{x^3 - 3x + 1} = \lim_{x\to\infty} \frac{\frac{1}{x} + \frac{2}{x^2} - \frac{1}{x^3}}{1 - \frac{3}{x^2} + \frac{1}{x^3}}$$

Since $\lim_{x\to\infty}(\frac{1}{x} + \frac{2}{x^2} - \frac{1}{x^3})$ exists (it is equal to 0), and $\lim_{x\to\infty}(1 - \frac{3}{x^2} + \frac{1}{x^3})$ exists and is not equal to 0 (it is equal to 1), we can apply Rule 4 to find that

$$\lim_{x\to\infty} \frac{\frac{1}{x} + \frac{2}{x^2} - \frac{1}{x^3}}{1 - \frac{3}{x^2} + \frac{1}{x^3}} = \frac{\lim_{x\to\infty}(\frac{1}{x} + \frac{2}{x^2} - \frac{1}{x^3})}{\lim_{x\to\infty}\left(1 - \frac{3}{x^2} + \frac{1}{x^3}\right)} = \frac{0}{1} = 0$$

(b) The leading term in both the numerator and the denominator is x^3, so we divide numerator and denominator by x^3 and obtain

$$\lim_{x\to\infty} \frac{2x^3 - 4x + 7}{3x^3 + 7x^2 - 1} = \lim_{x\to\infty} \frac{2 - \frac{4}{x^2} + \frac{7}{x^3}}{3 + \frac{7}{x} - \frac{1}{x^3}} = \frac{2}{3}$$

In the last step, we used the facts that the limits in both the numerator and the denominator exist and that the limit in the denominator is not equal to 0. Applying Rule 4 yields the limiting value. Note that the limiting value is equal to the ratio of the coefficients of the leading terms in the numerator and the denominator.

(c) The leading term in the numerator is x^4 and the leading term in the denominator is x^2. Since the leading term in the numerator grows much more quickly than the leading term in the denominator, we expect the limit to be undefined. This is indeed the case and can be seen if we divide the numerator by the denominator. We find that

$$\lim_{x\to\infty} \frac{x^4 + 2x - 5}{x^2 - x + 2} = \lim_{x\to\infty} \left(x^2 + x - 1 - \frac{x + 3}{x^2 - x + 2}\right) \qquad \text{which does not exist.}$$

It is often useful to determine whether the limit tends to $+\infty$ or $-\infty$. Since $x^2 + x - 1$ tends to $+\infty$ as $x \to \infty$ and the ratio $\frac{x+3}{x^2-x+2}$ tends to 0 as $x \to \infty$, $\frac{x^4+2x-5}{x^2-x+2}$ tends to $+\infty$ as $x \to \infty$.

Let's summarize our findings:

Limits of Rational Functions If $f(x)$ is a rational function of the form $f(x) = p(x)/q(x)$, where $p(x)$ and $q(x)$ are both polynomials. If the leading term of $p(x)$ is ax^m and the leading term of $q(x)$ is bx^n then:

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \frac{ax^m}{bx^n} = \frac{a}{b} \cdot \lim_{x\to\infty} x^{m-n} = \begin{cases} 0 & \text{if } m < n. \\ \dfrac{a}{b} & \text{if } m = n. \\ \text{does not exist} & \text{if } m > n. \end{cases}$$

The same behavior holds as $x \to -\infty$.

When the limit is divergent (that is, $m > n$) we may want to know whether $f(x) \to +\infty$ or $f(x) \to -\infty$. To make that determination, note that

Divergent Limits of Polynomials If $m > n$, $\lim_{x\to\infty} x^{m-n} = \infty$ and

$$\lim_{x\to-\infty} x^{m-n} = \begin{cases} \infty & \text{if } m-n \text{ is even.} \\ -\infty & \text{if } m-n \text{ is odd.} \end{cases}$$

EXAMPLE 3 Compute each limit if it exists

(a) $\displaystyle\lim_{x\to-\infty} \frac{1-x+2x^2}{3x-5x^2}$ **(b)** $\displaystyle\lim_{x\to\infty} \frac{1-x^3}{1+x^5}$

(c) $\displaystyle\lim_{x\to\infty} \frac{2-x^2}{1+2x}$ **(d)** $\displaystyle\lim_{x\to-\infty} \frac{4+3x^2}{1-7x}$.

Solution

(a) $$\lim_{x\to-\infty} \frac{\overbrace{1-x+2x^2}^{\text{Leading term } 2x^2}}{\underbrace{3x-5x^2}_{\text{Leading term: } -5x^2}} = \lim_{x\to-\infty} \frac{2x^2}{-5x^2} = \lim_{x\to-\infty} \frac{2}{-5} = \frac{-2}{5}.$$

(b) $$\lim_{x\to\infty} \frac{\overbrace{1-x^3}^{\text{Leading term: } -x^3}}{\underbrace{1+x^5}_{\text{Leading term } x^5}} = \lim_{x\to\infty} \frac{-x^3}{x^5} = \lim_{x\to\infty} \frac{-1}{x^2} = 0.$$

(c) $$\lim_{x\to\infty} \frac{\overbrace{2-x^2}^{\text{Leading term: } -x^2}}{\underbrace{1+2x}_{\text{Leading term: } 2x}} = \lim_{x\to\infty} \frac{-x^2}{2x} = \lim_{x\to\infty} \frac{-x}{2} = -\infty \quad \text{(limit does not exist).}$$

In the last step we used the fact that $\lim_{x\to\infty} x = \infty$ so $\lim_{x\to\infty} \frac{-x}{2} = -\frac{1}{2} \cdot \lim_{x\to\infty} x = -\infty$

(d) $$\lim_{x\to-\infty} \frac{\overbrace{4+3x^2}^{\text{Leading term: } 3x^2}}{\underbrace{1-7x}_{\text{Leading term: } -7x}} = \lim_{x\to-\infty} \frac{3x^2}{-7x} = \lim_{x\to-\infty} \frac{-3x}{7} = \infty \quad \text{(limit does not exist).}$$

$\lim_{x\to-\infty} x = -\infty$, so $\lim_{x\to-\infty} \frac{-3x}{7} = -\frac{3}{7} \cdot \lim_{x\to-\infty} x = \infty$.

Rational functions are not the only functions that involve limits as $x \to \infty$ (or $x \to -\infty$). Many important applications in biology involve exponential functions. We will use the following result repeatedly—it is one of the most important limits:

Limit of e^{-x} as $x \to \infty$

$$\lim_{x\to\infty} e^{-x} = 0$$

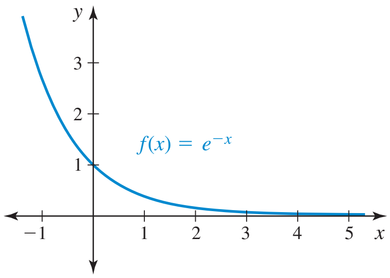


Figure 3.17 The graph of $f(x) = e^{-x}$.

The graph of $f(x) = e^{-x}$ is given in Figure 3.17. You should familiarize yourself with the basic shape of the function $f(x) = e^{-x}$ and its behavior as $x \to \infty$.

Unlike the rational functions that we studied in Example 3, $e^{-x} \to 0$ only if $x \to \infty$ and not if $x \to -\infty$. If $x \to -\infty$ then, as the graph shows, $e^{-x} \to +\infty$.

EXAMPLE 4 **Logistic Growth of a Population** The logistic curve describes the growth of a population over time, where the rate of growth depends on the population size. We will discuss this function in more detail in coming chapters. The logistic equation is one model for density-dependent growth, and models the growth of a population whose reproductive rate decreases as the population size increases because of competition among organisms for territory or resources. If $N(t)$ denotes the size of the population at time t, then the logistic curve is given by

$$N(t) = \frac{K}{1 + ae^{-rt}} \quad \text{for } t \geq 0$$

The parameters K and r are positive numbers that describe the population dynamics and $a > -1$ is a coefficient that determines the initial population size. The graph of $N(t)$ is shown in Figure 3.18. We will interpret K now; the interpretation of r, and a must wait until future chapters.

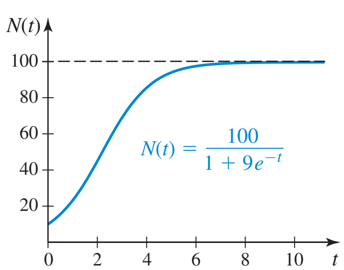


Figure 3.18 The graph of the logistic curve with $K = 100$, $a = 9$, and $r = 1$.

If we are interested in the long-term behavior of the population as it evolves in accordance with the logistic growth curve, we need to investigate what happens to $N(t)$ as $t \to \infty$. We find that

$$\lim_{t\to\infty} \frac{K}{1 + ae^{-rt}} = K \qquad \text{Using Rule 4 since } \lim_{t\to\infty}(1 + ae^{-rt}) \to 1 \text{ for } r > 0$$

That is, as $t \to \infty$, the population size approaches K, which is called the **carrying capacity** of the population; i.e., the maximum number of organisms that can be sustained by the habitat.

Section 3.3 Problems

Evaluate the limits in Problems 1–24.

1. $\lim_{x\to\infty} \frac{2x^2 - 3x + 5}{x^4 - 2x + 1}$

2. $\lim_{x\to\infty} \frac{2x^2 + 3}{2x + 1 - 5x^2}$

3. $\lim_{x\to-\infty} \frac{x^3 - 3}{x - 2}$

4. $\lim_{x\to-\infty} \frac{2x + 1}{3 - 4x}$

5. $\lim_{x\to\infty} \frac{1 - x^3 + 2x^4}{2x^2 - x^4}$

6. $\lim_{x\to\infty} \frac{3 - 5x^3}{1 + 3x^4}$

7. $\lim_{x\to\infty} \frac{x^2 + 2}{2x + 3}$

8. $\lim_{x\to-\infty} \frac{3 - x^2}{2 - 2x^2}$

9. $\lim_{x\to-\infty} \frac{x^2 - 3x + 1}{4 - x}$

10. $\lim_{x\to-\infty} \frac{1 + x^4}{2 + x^2}$

11. $\lim_{x\to-\infty} \frac{2 + x^2}{1 - x^2}$

12. $\lim_{x\to-\infty} \frac{x + x^2}{3x + 2}$

13. $\lim_{x\to\infty} \frac{4e^{-x}}{1 + e^{-2x}}$

14. $\lim_{x\to\infty} \frac{e^{-x}}{1 + e^{-x}}$

15. $\lim_{x\to\infty} \frac{e^x}{e^x + 2}$

16. $\lim_{x\to\infty} \frac{1 - e^x}{2 - e^x}$

17. $\lim_{x\to-\infty} \exp[x]$

18. $\lim_{x\to\infty} \exp[-x^2]$

19. $\lim_{x\to\infty} \frac{3e^{2x} + 1}{2e^{2x} - e^x}$

20. $\lim_{x\to\infty} \frac{3e^{2x}}{2e^{2x} - e^x}$

21. $\lim_{x\to\infty} \frac{3}{2 + e^{-x}}$

22. $\lim_{x\to-\infty} \frac{4}{1 + e^{-x}}$

23. $\lim_{x\to\infty} \frac{2}{e^x(1 + x)}$

24. $\lim_{x\to-\infty} \frac{e^x}{1 + x}$

25. **Bacterial Growth** The Monod model is used to describe how the rate of reproduction of organisms depends on the amount of nutrients that are available. ***Monod (1949)*** studied how the rate of division of *E. coli* cells depended upon the amount of sugar added to their growth flask. If r is the rate of reproduction (number of divisions in one hour) and C is the amount of glucose sugar added to the growth medium measured in moles then:

$$r(C) = \frac{1.35\,C}{C + 0.22 \times 10^{-4}}$$

(a) Show that the reproduction rate goes to zero when the sugar level is low; that is:

$$\lim_{C \to 0} r(C) = 0$$

(b) Show that if more and more sugar is added, the reproductive rate plateaus; that is, $\lim_{C \to \infty} r(C)$ exists and calculate this limit.

26. Hemoglobin-Oxygen Binding The Hill equation is used to model how hemoglobin in blood binds to oxygen. If the proportion of hemoglobin molecules that are bound to oxygen is h and the concentration of oxygen (measured as a partial pressure, that varies from 0 to ∞) is P, then a common model is:

$$h(P) = \frac{aP^k}{30^k + P^k}$$

where $k \geq 1$ and $a > 0$ are constants that depend on the species of animal and its environment (e.g., whether it lives at sea-level or at altitude).

(a) Show that no matter what the values of a and k are, the amount of bound oxygen goes to zero as the oxygen concentration goes to 0; that is:

$$\lim_{P \to 0} h(P) = 0.$$

(b) It is known that as P increases, the amount of bound oxygen plateaus. Since $h = 1$ when all hemoglobin molecules are bound to oxygen, we want our model to reflect that:

$$\lim_{P \to \infty} h(P) = 1.$$

This is called the saturation value for oxygen binding. Explain what value of a must be chosen for this condition to be satisfied.

(c) The half-saturation constant, $P_{1/2}$, is defined to be the concentration of oxygen at which the proportion of bound hemoglobin molecules reaches half its saturation value, that is:

$$h(P_{1/2}) = \frac{1}{2} \lim_{P \to \infty} h(P).$$

Show that $P_{1/2} = 30$.

(d) In a patient with carbon monoxide poisoning carbon monoxide binds preferentially to the hemoglobin instead of oxygen, stopping the blood from effectively transporting oxygen around the body. For a patient with acute carbon monoxide poisoning, the relationship between proportion of bound hemoglobin molecules and oxygen concentration can be modeled by:

$$h(P) = \frac{0.9P^3}{P^3 + 26^3} \qquad \text{(we have assumed that } k = 3\text{)}$$

Show that both the saturation level for oxygen binding and the half-saturation constant are both changed from your answers in (b) and (c).

27. Population Growth Suppose the size of a population at time t is given by

$$N(t) = \frac{500t}{3 + t}, \quad t \geq 0.$$

(a) Use a graphing calculator to sketch the graph of $N(t)$.

(b) Determine the size of the population as $t \to \infty$. We call this the **limiting population size**.

(c) Show that, at time $t = 3$, the size of the population is half its limiting size.

28. Logistic Growth Suppose that the size of a population at time t is given by

$$N(t) = \frac{50}{1 + 6e^{-2t}}, \quad t \geq 0.$$

(a) Use a graphing calculator to sketch the graph of $N(t)$.

(b) Determine the size of the population as $t \to \infty$, using the basic rules for limits. Compare your answer with the graph that you sketched in (a).

29. Logistic Growth Suppose that the size of a population at time t is given by

$$N(t) = \frac{100}{1 + 3e^{-t}}, \quad t \geq 0.$$

(a) Use a graphing calculator to sketch the graph of $N(t)$.

(b) Determine the size of the population as $t \to \infty$, using the basic rules for limits. Compare your answer with the graph that you sketched in (a).

3.4 Trigonometric Limits and the Sandwich Theorem

The trigonometric functions, sine, cosine, tangent, show up throughout biology, in particular when one attempts to model a periodic phenomenon, such as rainfall over the course of one year, or the rate of growth of a population of cells that have a circadian rhythm (that is, whose growth changes predictably over one day). In Chapter 4 we will have to calculate the rate of change of these functions. But to do that it will be necessary to know the following trigonometric limits:

$$\boxed{\lim_{x \to 0} \frac{\sin x}{x} = 1 \qquad \text{and} \qquad \lim_{x \to 0} \frac{1 - \cos x}{x} = 0.}$$

We will start by giving a nonrigorous geometric argument for these limits. Proving the limits requires that we use the sandwich theorem. In 3.4.2 we state the sandwich theorem and then use it to prove the limits rigorously.

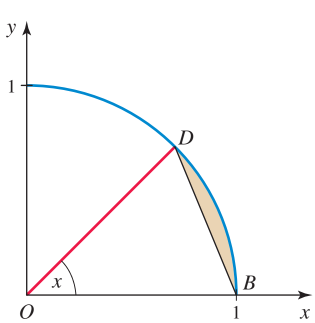


Figure 3.19 The unit circle with triangle ODB.

3.4.1 Geometric Argument for Trigonometric Limits

On Figure 3.19 we draw a circle of radius 1 and a triangle ODB whose sides OD and OB are radii of the circle and in which the angle between OD and OB is x.

Since the sides OD and OB of the triangle both have length 1, the area of the triangle ODB is:

$$\text{area of triangle} = \frac{1}{2}\sin x.$$

Area of an isosceles triangle with side r and angle x is $\frac{1}{2}r^2 \sin x$

We may compare this to the area of the circular sector ODB, which has the same vertices, but in which DB are connected not by a straight line but by the arc of a circle. Since this sector also has angle x, its area is:

$$\text{area of sector} = \frac{1}{2}x.$$

Area of a sector with radius r, and angle x is $\frac{1}{2}r^2 x$

The difference between the two areas is the space between the triangle and the wedge that is shaded in Figure 3.20. If we let x get smaller and smaller, this space becomes a smaller and smaller fraction of the wedge (see Figure 3.20).

So, since:

$$\text{area of triangle} = \text{area of sector } ODB - \text{shaded area}$$

$$\frac{\text{triangle area}}{\text{sector area}} = 1 - \frac{\text{shaded area}}{\text{sector area}}$$

$$\frac{\cancel{\frac{1}{2}}\sin x}{\cancel{\frac{1}{2}}x} = 1 - \frac{\text{shaded area}}{\text{sector area}}$$

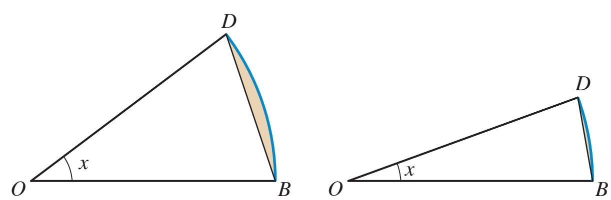


Figure 3.20 As x becomes smaller, the areas of the wedge ODB and of the triangle ODB become closer.

Now since the shaded area occupies a smaller and smaller fraction of the sector as x is made smaller and smaller, we may deduce by taking the limit as $x \to 0$:

$$\lim_{x\to 0} \frac{\sin x}{x} = 1$$

Since we have not proven that the area of the shaded region is negligible in the limit as $x \to 0$, this argument does not constitute a proof. In Section 3.4.2 we will introduce the sandwich theorem, which allows the limit to be proven rigorously.

Proof that $\boldsymbol{\lim_{x\to 0} \frac{1-\cos x}{x} = 0}$ The other important trigonometric limit can be deduced from $\lim_{x\to 0} \frac{\sin x}{x} = 1$. Multiply both numerator and denominator of $f(x) = (1-\cos x)/x$ by $1+\cos x$ to obtain:

$$\begin{aligned}
\lim_{x\to 0} \frac{1-\cos x}{x} &= \lim_{x\to 0} \frac{1-\cos x}{x} \cdot \frac{1+\cos x}{1+\cos x} \\
&= \lim_{x\to 0} \frac{1-\cos^2 x}{x(1+\cos x)} && \text{Multiply out numerator and denominator.} \\
&= \lim_{x\to 0} \frac{\sin^2 x}{x(1+\cos x)} && \sin^2 x + \cos^2 x = 1. \\
&= \lim_{x\to 0} \left(\frac{\sin x}{x} \cdot \frac{\sin x}{1+\cos x}\right) && \text{Factorize numerator and denominator.} \\
&= \lim_{x\to 0} \frac{\sin x}{x} \cdot \lim_{x\to 0} \frac{\sin x}{1+\cos x} && \text{By Limit Law 3 since both limits exist} \\
&= 1 \cdot 0 = 0 && \lim_{x\to 0} \frac{\sin x}{x} = 1 \text{ and } \lim_{x\to 0} \frac{\sin x}{1+\cos x} = \frac{0}{1+1} = 0 \text{ since } \cos x, \sin x \text{ are continuous.}
\end{aligned}$$

■

EXAMPLE 1 Find the following limits.

(a) $\lim_{x\to 0} \frac{\sin 3x}{5x}$ **(b)** $\lim_{x\to 0} \frac{\sin^2 x}{x^2}$ **(c)** $\lim_{x\to 0} \frac{\sec x - 1}{x \sec x}$

Solution **(a)** We cannot apply the first trigonometric limit directly. The trick is to substitute $z = 3x$ and observe that $z \to 0$ as $x \to 0$. Then

$$\lim_{x\to 0} \frac{\sin 3x}{5x} = \lim_{z\to 0} \frac{\sin z}{5z/3} = \frac{3}{5} \lim_{z\to 0} \frac{\sin z}{z} = \frac{3}{5}$$

(b) We note that

$$\lim_{x\to 0} \frac{\sin^2 x}{x^2} = \lim_{x\to 0} \left(\frac{\sin x}{x}\right)^2 = \left(\lim_{x\to 0} \frac{\sin x}{x}\right)^2 = 1$$

Here, we used the fact that the limit of a product is the product of the limits, provided that the individual limits exist.

(c) We first write $\sec x = 1/\cos x$

$$\lim_{x\to 0} \frac{\sec x - 1}{x \sec x} = \lim_{x\to 0} \frac{\frac{1}{\cos x} - 1}{\frac{x}{\cos x}} = \lim_{x\to 0} \frac{\left(\frac{1}{\cos x} - 1\right)\cos x}{\frac{x}{\cos x}\cos x}$$

$$= \lim_{x\to 0} \frac{1 - \cos x}{x} = 0 \qquad \text{Multiply numerator and denominator by } \cos x.$$

3.4.2 The Sandwich Theorem

What happens during bungee jumping? The jumper is tied to an elastic rope, jumps off a bridge, and experiences damped oscillations until she comes to rest and will be hauled in to safety. The trajectory over time might resemble the function (Figure 3.21)

$$g(x) = e^{-x} \cos(10x), \qquad x \geq 0.$$

We suspect from the graph that

$$\lim_{x\to\infty} e^{-x} \cos(10x) = 0.$$

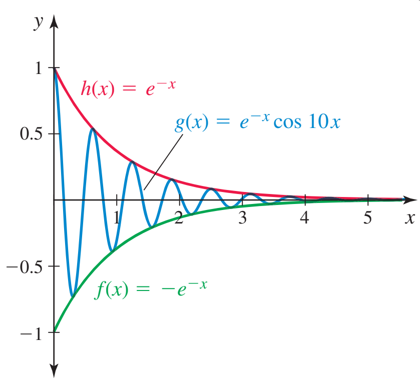


Figure 3.21 $f(x) = e^{-x} \cos x$, is sandwiched between the two functions e^{-x} and $-e^{-x}$.

If we wanted to calculate this limit, we would quickly see that none of the rules we have learned so far apply. Although $\lim_{x\to\infty} e^{-x} = 0$, we find that $\lim_{x\to\infty} \cos(10x)$ does not exist: The function $\cos(10x)$ oscillates between -1 and 1. But because $\cos(10x)$ cannot be larger than 1 or smaller than -1 we can sandwich function $g(x) = e^{-x} \cos(10x)$ between $f(x) = -e^{-x}$ and $h(x) = e^{-x}$ (see Figure 3.21). To do so, we note that because

$$-1 \leq \cos(10x) \leq 1$$

it follows that:

$$-e^{-x} \leq e^{-x} \cos(10x) \leq e^{-x} \qquad \text{Multiply by } e^{-x}$$

Then, since

$$\lim_{x\to\infty} (-e^{-x}) = \lim_{x\to\infty} e^{-x} = 0$$

our function $g(x) = e^{-x} \cos(10x)$ gets squeezed in between the two functions $f(x) = -e^{-x}$ and $h(x) = e^{-x}$, which both go to 0 as x tends to infinity. Therefore,

$$\lim_{x\to\infty} e^{-x} \cos(10x) = 0$$

This useful method is known as the sandwich theorem. We will not prove it.

Sandwich Theorem If $f(x) \leq g(x) \leq h(x)$ for all x in an open interval that contains c (except possibly at c) and

$$\lim_{x\to c} f(x) = \lim_{x\to c} h(x) = L$$

then

$$\lim_{x\to c} g(x) = L$$

The theorem is called the sandwich theorem because we "sandwich" the function $g(x)$ between the two functions $f(x)$ and $h(x)$. Since $f(x)$ and $h(x)$ converge to the same value as $x \to c$, $g(x)$ also must converge to that value as $x \to c$, because it is squeezed in between $f(x)$ and $h(x)$. The sandwich theorem also applies to one-sided limits. We demonstrate how to use the sandwich theorem in the next example.

EXAMPLE 2 Show that $\lim_{x\to 0} x^2 \sin \frac{1}{x} = 0$.

Solution First, note that we cannot use Rule 3—which says that the limit of a product is equal to the product of the limits—because it requires that the limits of both factors exist. The limit of $\sin(1/x)$ as $x \to 0$ does not exist; instead, it diverges by oscillating. (See Example 8 of Section 3.1 for a similar limit.) However, we know that

$$-1 \le \sin \frac{1}{x} \le 1$$

for all $x \neq 0$, so

$$-x^2 \le x^2 \sin \frac{1}{x} \le x^2 \qquad \text{Multiply all three parts by } x^2$$

Since $\lim_{x\to 0}(-x^2) = \lim_{x\to 0} x^2 = 0$, we can apply the sandwich theorem to obtain $\lim_{x\to 0} x^2 \sin \frac{1}{x} = 0$. This limit is illustrated in Figure 3.22. •

EXAMPLE 3 Show that $\lim_{x\to 0} x \sin \frac{1}{x} = 0$.

Solution At first sight this limit looks almost identical to Example 2, but we will see that care must be taken to find the bounding functions that sandwich $x \sin \frac{1}{x}$ for both $x > 0$ and $x < 0$.

As in Example 2, note that we cannot use Rule 3 of the Limit Laws because it requires that the limits of both factors exist. The limit of $\sin(1/x)$ as $x \to 0$ does not exist. However, we know that

$$-1 \le \sin \frac{1}{x} \le 1$$

for all $x \neq 0$. To go from this set of inequalities to one that involves $x \sin \frac{1}{x}$, we need to multiply all three parts by x.

If $x > 0$, then

$$-x \le x \sin \frac{1}{x} \le x \qquad \text{Multiply by } x.$$

Since $x > 0$ there is no need to reverse the inequality signs. Because $\lim_{x\to 0^+}(-x) = \lim_{x\to 0^+} x = 0$, we can apply the sandwich theorem to obtain $\lim_{x\to 0^+} x \sin \frac{1}{x} = 0$.

We can repeat the same steps for $x < 0$:

$$-x \ge x \sin \frac{1}{x} \ge x \qquad \text{Multiply by } x \text{, since } x < 0 \text{ reverse inequality signs.}$$

Because $\lim_{x\to 0^-}(-x) = \lim_{x\to 0^-} x = 0$, we can again apply the sandwich theorem and get $\lim_{x\to 0^-} x \sin \frac{1}{x} = 0$.

To apply the sandwich theorem using the same functions, independent of whether $x > 0$ or $x < 0$, we could have combined our inequalities to find:

$$-|x| \le x \sin \frac{1}{x} \le |x|$$

Since $\lim_{x\to 0} |x| = 0$ and $\lim_{x\to 0}(-|x|) = 0$, by the sandwich theorem $\lim_{x\to 0} x \sin \frac{1}{x} = 0$.

The left-hand and right-hand limits are the same. Therefore, combining the two results, we find that

$$\lim_{x\to 0} x \sin \frac{1}{x} = 0$$

This limit is illustrated in Figure 3.23. •

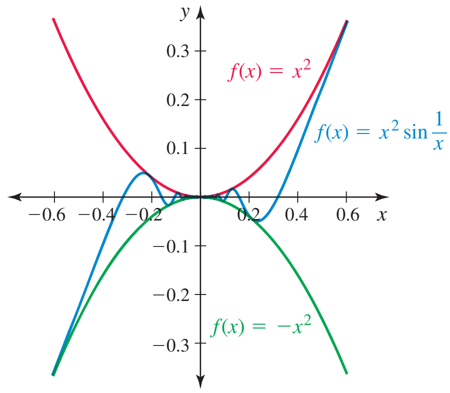


Figure 3.22 The graph of $f(x) = x^2 \sin \frac{1}{x}$.

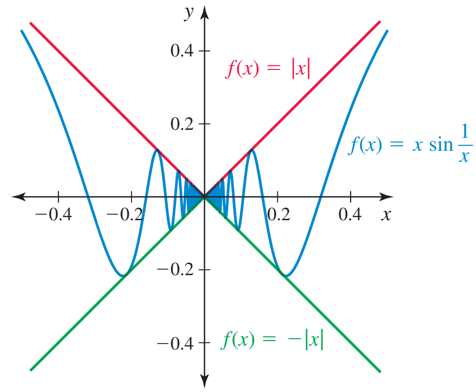


Figure 3.23 The sandwich theorem illustrated on $\lim_{x\to 0} x \sin(1/x)$.

We will now revisit the result from Section 3.4.1, that $\lim_{x\to 0} \frac{\sin x}{x} = 1$. Using the sandwich theorem we may actually prove this limit.

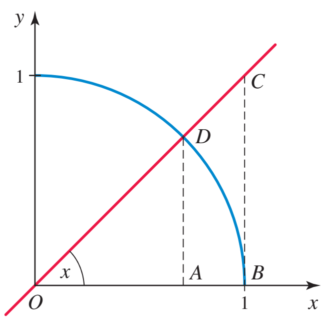


Figure 3.24 The unit circle with the triangles OAD and OBC.

Proof that $\lim_{x\to 0} \frac{\sin x}{x} = 1$ In Figure 3.24, we draw the unit circle together with the triangles ODA and OCB, both of which are right-angled triangles with angle x at the point O. The angle x is measured in radians. From the figure we find:

$$\begin{aligned} \overline{BD} &= x \qquad \text{$\overline{BD}$ means length of the segment BD} \\ \overline{OB} &= 1 \\ \overline{OA} &= \cos x \\ \overline{AD} &= \sin x \\ \overline{BC} &= \tan x \end{aligned} \tag{3.8}$$

Furthermore, using the symbol $\triangle$ to denote a triangle, we obtain

$$\text{area of } \triangle ODA \le \text{area of sector } ODB \le \text{area of } \triangle OCB$$

The area of a sector of central angle x (measured in radians) and radius r is $\frac{1}{2}r^2x$. Therefore,

$$\frac{1}{2}\overline{OA}\cdot\overline{AD} \le \frac{1}{2}\overline{OB}^2\cdot x \le \frac{1}{2}\overline{OB}\cdot\overline{BC} \qquad \text{Area of triangle} = \tfrac{1}{2}\text{ base} \times \text{height}$$

or using our list of arc lengths (3.8).

$$\frac{1}{2}\cos x \sin x \le \frac{1}{2}\cdot 1^2 \cdot x \le \frac{1}{2}\cdot 1 \cdot \tan x$$

Dividing this pair of inequalities by $\frac{1}{2}\sin x$

$$\cos x \le \frac{x}{\sin x} \le \frac{1}{\cos x} \qquad \text{Assuming } \sin x > 0\text{, i.e., } 0 < x < \pi.$$

On the rightmost part, we used the fact that $\tan x = \frac{\sin x}{\cos x}$. Taking reciprocals gives:

$$\frac{1}{\cos x} \ge \frac{\sin x}{x} \ge \cos x \qquad \text{When a reciprocal is taken, inequality signs are reversed.}$$

We can now take the limit as $x \to 0^+$. We find that

$$\lim_{x\to 0^+} \cos x = 1 \qquad \text{and} \qquad \lim_{x\to 0^+} \frac{1}{\cos x} = \frac{1}{\lim_{x\to 0^+}\cos x} = 1 \qquad \text{Using Limit Law 4}$$

We now apply the sandwich theorem, which yields

$$\lim_{x\to 0^+} \frac{\sin x}{x} = 1 \qquad \text{Sandwich theorem with } f(x) = \cos x,\ h(x) = \tfrac{1}{\cos x}$$

We have only considered the limit $x \to 0^+$ because in our geometric argument we assumed that $x > 0$. However, $\sin x$ is an odd function of x, so $\frac{\sin x}{x}$ is an even function of x, which means that $\lim_{x\to 0^-} \frac{\sin x}{x} = \lim_{x\to 0^+} \frac{\sin x}{x}$, and so $\lim_{x\to 0} \frac{\sin x}{x} = 1$; that is, the limit exists however x approaches 0.

Section 3.4 Problems

3.4.1

In Problems 1–16, evaluate the trigonometric limits.

1. $\lim_{x\to 0} \frac{\sin(3x)}{3x}$

2. $\lim_{x\to 0} \frac{\sin(2x)}{4x}$

3. $\lim_{x\to 0} \frac{\sin(5x)}{x}$

4. $\lim_{x\to 0} \frac{\sin x}{-x}$

5. $\lim_{x\to 0} \frac{\sin(\pi x)}{x}$

6. $\lim_{x\to 0} \frac{\sin(\pi x/2)}{2x}$

7. $\lim_{x\to 0} \frac{\sin(\pi x)}{\sqrt{x}}$

8. $\lim_{x\to 0} \frac{\sin^2(2x)}{x}$

9. $\lim_{x\to 0} \frac{\sin x \cos x}{x(1-x)}$

10. $\lim_{x\to 0} \frac{1-\cos^2 x}{x^2}$

11. $\lim_{x\to 0} \frac{1-\cos 2x}{x}$

12. $\lim_{x\to 0} \frac{1-\cos(x)^2}{3x}$

13. $\lim_{x\to 0} \frac{1-\cos(5x)}{2x}$

14. $\lim_{x\to 0} \frac{1-\cos(x/2)}{x}$

15. $\lim_{x\to 0} \frac{\sin x(1-\cos x)}{x^2}$

16. $\lim_{x\to 0} \frac{1-\cos^2 x}{x}$

3.4.2

T **17.** Let $f(x) = x^2 \cos \frac{1}{x}, \quad x \neq 0$

(a) Use a graphing calculator to sketch the graph of $y = f(x)$.

(b) Show that $-x^2 \le x^2 \cos \frac{1}{x} \le x^2$ holds for $x \neq 0$.

(c) Use your result in (b) and the sandwich theorem to show that $\lim_{x\to 0} x^2 \cos \frac{1}{x} = 0$

T **18.** Let $f(x) = x^3 \cos \frac{1}{x}, \quad x \neq 0$

(a) Use a graphing calculator to sketch the graph of $y = f(x)$.

(b) Use the sandwich theorem to show that $\lim_{x\to 0} x^3 \cos \frac{1}{x} = 0$

T **19.** Let $f(x) = \frac{\ln x}{x}, \quad x > 0$

(a) Use a graphing calculator to graph $y = f(x)$.

(b) Use a graphing calculator to investigate the values of x for which

$$\frac{1}{x} \le \frac{\ln x}{x} \le \frac{1}{\sqrt{x}}$$

holds.

(c) Use your result in (b) to explain why: $\lim_{x\to\infty} \frac{\ln x}{x} = 0$

T **20.** Let $f(x) = \frac{\sin^2 x}{x}, \quad x > 0$

(a) Use a graphing calculator to graph $y = f(x)$.

(b) Explain why you cannot use the basic rules for finding limits to compute $\lim_{x\to\infty} \frac{\sin^2 x}{x}$

(c) Show that $0 \le \frac{\sin^2 x}{x} \le \frac{1}{x}$ holds for $x > 0$, and use the sandwich theorem to compute $\lim_{x\to\infty} \frac{\sin^2 x}{x}$

T **21.** **(a)** Use a graphing calculator to sketch the graph of

$$f(x) = e^{ax} \sin x, \quad x \ge 0$$

for $a = -0.1, -0.01, 0, 0.01$, and 0.1.

(b) Which part of the function $f(x)$ produces the oscillations that you see in the graphs sketched in (a)?

(c) Describe in words the effect that the value of a has on the shape of the graph of $f(x)$.

(d) Graph $f(x) = e^{ax} \sin x$, $g(x) = -e^{ax}$, and $h(x) = e^{ax}$ together in one coordinate system for **(i)** $a = 0.1$ and **(ii)** $a = -0.1$. [Make separate graphs for (i) and (ii).] Explain what you see in each case. Show that

$$-e^{ax} \le e^{ax} \sin x \le e^{ax}$$

Use this pair of inequalities to determine the values of a for which $\lim_{x\to\infty} f(x)$ exists, and find the limiting value.

3.5 Properties of Continuous Functions

3.5.1 The Intermediate-Value Theorem and The Bisection Method

As you hike up a mountain, the temperature decreases with increasing elevation. Suppose the temperature at the bottom of the mountain is 70°F and the temperature at the top of the mountain is 40°F. How do you know that at some time during your hike you must have crossed a point where the temperature was exactly 50°F? Your answer will probably be something like the following: "To go from 70°F to 40°F, I must have passed through 50°F, since 50°F is between 40°F and 70°F and the temperature changed continuously as I hiked up the mountain." This statement represents the content of the intermediate-value theorem.

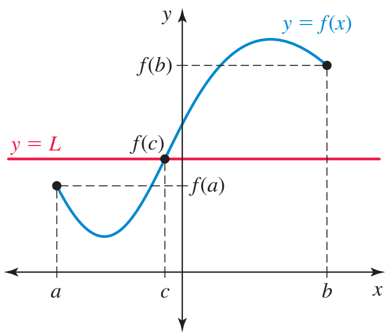


Figure 3.25 Illustration of the intermediate-value theorem.

The Intermediate-Value Theorem Suppose that f is continuous on the closed interval $[a, b]$. If L is any real number with $f(a) < L < f(b)$ or $f(b) < L < f(a)$, then there exists at least one number c on the open interval (a, b) such that $f(c) = L$.

We will not prove this theorem, but Figure 3.25 should convince you that it is true. In the figure, f is continuous and defined on the closed interval $[a, b]$. Let L be any real number between $f(a)$ and $f(b)$ (that is, $f(a) < L < f(b)$). Then the graph of $f(x)$ must intersect the line $y = L$ at least once on the open interval (a, b).

EXAMPLE 1 Let $f(x) = 3 + \sin x$ for $0 \le x \le \frac{3\pi}{2}$. Show that there exists at least one point c in $(0, 3\pi/2)$ such that $f(c) = 5/2$.

Solution The graph of $f(x)$ is shown in Figure 3.26. First, note that $f(x)$ is defined on a closed interval and is continuous on $[0, 3\pi/2]$. Furthermore, we find that

$$f(0) = 3 + \sin 0 = 3 + 0 = 3$$

$$f\left(\frac{3\pi}{2}\right) = 3 + \sin\frac{3\pi}{2} = 3 + (-1) = 2$$

Given that

$$2 < \frac{5}{2} < 3$$

we conclude from the intermediate-value theorem that there exists a number c such that $f(c) = 5/2$. Note that the theorem does not tell us where c is or whether there is just one possible value of c or multiple possible values for c. •

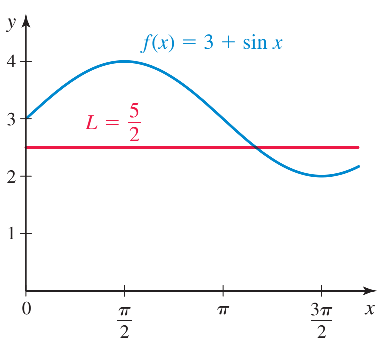


Figure 3.26 Illustration of the intermediate-value theorem for $f(x) = 3 + \sin x$ and $L = 5/2$.

In applying the intermediate-value theorem, it is important to check that f is continuous. Discontinuous functions can easily miss values; for example, the floor function in Example 4 of Section 3.2 misses all numbers that are not integers. As mentioned in Example 1, the intermediate-value theorem gives us only the existence of a number c; it does not tell us how many such points there are or where they are located.

You might wonder how such a result can be of any use. One important application is that the theorem can be used to find approximate roots (or solutions) of equations of the form $f(x) = 0$. We show how in the next example.

EXAMPLE 2 Find a root of the equation $x^5 - 7x^2 + 3 = 0$.

Solution Let $f(x) = x^5 - 7x^2 + 3 = 0$. Because $f(x)$ is a polynomial, it is continuous for all $x \in \mathbf{R}$. Furthermore,

$$\lim_{x\to-\infty} f(x) = -\infty \quad \text{and} \quad \lim_{x\to\infty} f(x) = \infty \qquad \lim_{x\to\pm\infty} f(x) = \lim_{x\to\pm\infty} x^5 \text{(the leading term)}$$

That is, if we choose a large enough interval $[a, b]$, then $f(a) < 0$ and $f(b) > 0$ and, therefore, there must be a number $c \in (a, b)$ such that $f(c) = 0$. This number c is a root of the equation $f(x) = 0$. The existence of c is guaranteed by the intermediate-value theorem.

To find a number c for which $f(c) = 0$, we use the **bisection method**. We start by finding a and b such that $f(a) < 0$ and $f(b) > 0$. For example,

$$f(-1) = -5 \qquad \text{and} \qquad f(2) = 7$$

The intermediate-value theorem then tells us that there must be a number in $(-1, 2)$ for which $f(c) = 0$. To locate this root with more precision, we take the midpoint of $(-1, 2)$, which is 0.5, and evaluate the function at $x = 0.5$. [The midpoint of the interval (a, b) is $(a + b)/2$.] Now, $f(0.5) = 1.28$ (to two decimals). We thus have

$$f(-1) = -5 \qquad f(0.5) = 1.28 \qquad f(2) = 7$$

Using the intermediate-value theorem again, we can now guarantee a root in $(-1, 0.5)$, since $f(-1) < 0$ and $f(0.5) > 0$. Bisecting the new interval and computing the respective values of $f(x)$, we find that

$$f(-1) = -5 \qquad f(-0.25) = 2.562 \qquad f(0.5) = 1.28$$

Using the intermediate-value theorem yet again, we can guarantee a root in $(-1, -0.25)$, since $f(-1) < 0$ and $f(-0.25) > 0$. Repeating this procedure of bisecting and selecting a new (smaller) interval will eventually produce an interval that is small enough that we can locate the root to any desired accuracy. The first several steps are summarized in Table 3-1.

TABLE 3-1 Bisection Method

a	$\frac{a+b}{2}$	b	$f(a)$	$f(\frac{a+b}{2})$	$f(b)$
-1	0.5	2	-5	1.28	7
-1	-0.25	0.5	-5	2.562	1.28
-1	-0.625	-0.25	-5	0.170	2.562
-1	-0.8125	-0.625	-5	-1.975	0.170
-0.8125	-0.71875	-0.625	-1.975	-0.808	0.170
-0.71875	-0.671875	-0.625	-0.808	-0.297	0.170
-0.671875	-0.6484375	-0.625	-0.297	-0.0579	0.170
-0.6484375	-0.63671875	-0.625	-0.0579	0.0575	0.170
-0.6484375	-0.642578125	-0.63671875	-0.0579	9.9×10^{-5}	0.0575
-0.6484375		-0.642578125			

After nine steps, we find that there exists a root in

$$(-0.6484375, -0.642578125)$$

The length of this interval is 0.005859375. If we are satisfied with that level of precision, we can stop here and choose, for instance, the midpoint of the last interval as an approximate value for a root of the equation $x^5 - 7x^2 + 3 = 0$. The midpoint is

$$\begin{aligned}\frac{-0.642578125 + (-0.6484375)}{2} &= -0.6455078125 \\ &= -0.646\end{aligned}$$

(rounded to three decimals).

Note that the length of the interval decreases by a factor of 1/2 at each step. That is, after nine steps, the length of the interval is $(1/2)^9$ of the length of the original interval. In this example, the length of the original interval was 3; hence, the length of the interval after nine steps is

$$3 \cdot \left(\frac{1}{2}\right)^9 = \frac{3}{512} = 0.005859375$$

as we saw. The bisection method is fairly slow when we need high accuracy. For instance, to reduce the length of the interval to 10^{-6}, we would need at least 22 steps, since after 22 steps, the width of the interval in which we know the root lies is:

$$3 \cdot \left(\frac{1}{2}\right)^{22} = 0.7 \times 10^{-6}$$

In Section 5.8 we will learn a faster method.

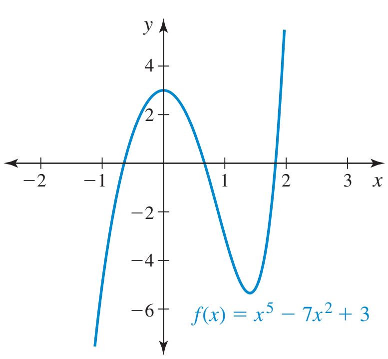


Figure 3.27 The graph of $f(x) = x^5 - 7x^2 + 3$.

Figure 3.27 shows the graph of $f(x) = x^5 - 7x^2 + 3$. We see that the graph intersects the x-axis three times. We found an approximation of the leftmost root of the equation

$x^5 - 7x^2 + 3 = 0$. If we had used another starting interval—say, (1, 2)—we would have located an approximation of the rightmost root of the equation. •

3.5.2 Using a Spreadsheet to Implement the Bisection Method

The bisection method identifies an interval in which the root of $f(x)$ is known to be located. Each time we evaluate the function f at the midpoint of this interval, we can reduce the length of the interval by half. We can use a spreadsheet to automatically update the interval and perform this sequence of calculations much quicker than we could perform them using a calculator.

Suppose we want to solve $f(x) = 0$, and we know that the root lies somewhere in the interval $x_{\min} < x < x_{\max}$, because $f(x_{\min})$ and $f(x_{\max})$ have opposite signs. Let's further assume that $f(x_{\min}) < 0$ and $f(x_{\max}) > 0$. [If on the other hand $f(x_{\min}) > 0$ and $f(x_{\max}) < 0$ then we let $F(x) = -f(x)$ and find the root of $F(x)$ instead.]

For example, suppose that we wanted to find the second root of $f(x) = x^5 - 7x^2 + 3$. From Figure 3.27 we see that this root lies somewhere between $x = 0$ and $x = 1.5$, so we will let these be the starting endpoints for our bisection search; $x_{\min} = 0$ and $x_{\max} = 1.5$. Then $f(x_{\min}) = 3$ and $f(x_{\max}) = -5.16$ (to two decimals). For this interval $f(x_{\min}) > 0$ and $f(x_{\max}) < 0$, so we will instead look for a root of $F(x) = -f(x) = -x^5 + 7x^2 - 3$. Then $F(x) = 0$ has the same root as $f(x) = 0$, but $F(x_{\min}) = -3 < 0$ and $F(x_{\max}) = 5.16 > 0$.

We use the first row of our spreadsheet for column labels. The first three columns will store our values for $x_{\min}$, the midpoint of the interval, and $x_{\max}$ respectively. In cells A1, B1, C1 enter the following labels:

A1 : ***x*_min**
B1 : ***x*_mid**
C1 : ***x*_max**

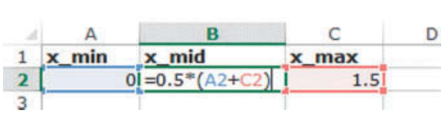
In cell A2 enter **0** (our starting value for $x_{\min}$) and in cell C2 enter **1.5** (our starting value for $x_{\max}$). B2 will hold the starting value of x_{mid} (the mid-point of the interval on which we know that the root is located). We can have the spreadsheet calculate that value for us. In B2 enter the formula **= 0.5 ∗ (A2 + C2)**. At this stage your spreadsheet should look like Figure 3.28.

We will use the columns D, E, and F to store the value of the function evaluated at $x_{\min}$, x_{mid}, and $x_{\max}$ respectively. In D1, E1, and F1 enter the following labels:

D1 : **F(*x*_min)**
E1 : **F(*x*_mid)**
F1 : **F(*x*_max)**

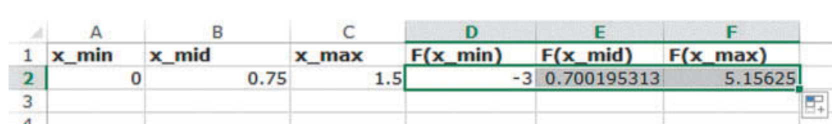
D2 will store the value of $F(x)$ evaluated at $x_{\min} = 0$. Use the spreadsheet to evaluate the function. In D2 enter the formula: **= −A2ˆ5 + 7 ∗ A2ˆ2 − 3.** We want E2 to contain this same formula with A2 replaced by B2, and F2 should contain the same formula with A2 replaced by C2. But it is not necessary to enter these formulas by hand—use AutoFill to copy the formulas over from D2. To use AutoFill select cell D2. Cell D2 will be highlighted (surrounded by a colored box). In the bottom right corner of this box there will be a small square. Click on this square, hold, and drag two cells to the right (i.e., over E2 and F2). At this stage your spreadsheet should look like Figure 3.29.

In our spreadsheet we see that $F(x_{\text{mid}}) = 0.70 > 0$, so according to the Intermediate Value Theorem the root $F(x) = 0$ lies between $x = x_{\min}$ and $x = x_{\text{mid}}$. That is,



	A	B	C	D
1	x_min	x_mid	x_max	
2	0	=0.5*(A2+C2)	1.5	
3				

Figure 3.28 Defining the initial interval for bisection search.



	A	B	C	D	E	F
1	x_min	x_mid	x_max	F(x_min)	F(x_mid)	F(x_max)
2	0	0.75	1.5	-3	0.700195313	5.15625
3						

Figure 3.29 Using AutoFill to calculate $F(x_{\text{mid}})$ and $F(x_{\max})$.

our new interval will be $[x_{\text{min}}, x_{\text{mid}}] = [0, 0.75]$. We want to put those values in A3 and C3. But **do not** enter these values by hand.

We want the spreadsheet to decide whether the new interval is $[x_{\text{min}}, x_{\text{mid}}]$ or $[x_{\text{mid}}, x_{\text{max}}]$. Note that we would choose the first interval if $F(x_{\text{mid}}) > 0$ and the second interval if $F(x_{\text{mid}}) < 0$.

So the value in A3 should be the same as A2 if $F(x_{\text{mid}}) > 0$ and the same as B2 if $F(x_{\text{mid}}) < 0$. We use the IF command to implement this choice in the spreadsheet. Specifically in A3 enter the formula:

$$\mathbf{= IF(E2 > 0, A2, B2)}$$

What does this command do? The first part is a statement that the spreadsheet needs to evaluate: Is E2 > 0? (Remember, E2 stores the value of $F(x_{\text{mid}})$ from the first interval so the spreadsheet is evaluating whether or not $F(x_{\text{mid}}) > 0$.) If this statement is true, then the spreadsheet puts the second value in the IF command in this cell (that is, it puts A2 in the cell). If the statement is not true, that is, if E2 ≤ 0, then the spreadsheet puts the third value (B2) in the cell. For this function E2 $= 0.70 > 0$, so the statement evaluates as true and the spreadsheet puts A2$(= 0)$ in the cell. We need to take similar steps to calculate the new right endpoint, x_{max}. In cell C3 enter the formula:

$$\mathbf{= IF(E2 > 0, B2, C2)}$$

Since E2 > 0, this puts the value of B2$(= 0.75)$ in the cell. Your spreadsheet will now look like Figure 3.30.

	A	B	C	D	E	F
1	x_min	x_mid	x_max	F(x_min)	F(x_mid)	F(x_max)
2	0	0.75	1.5	-3	0.700195313	5.15625
3	0		=IF(E2>0,B2,C2)			
4						

Figure 3.30 Using the IF function to calculate the new endpoints of the bisection search interval.

	A	B	C	D	E	F
1	x_min	x_mid	x_max	F(x_min)	F(x_mid)	F(x_max)
2	0	0.75	1.5	-3	0.700195313	5.15625
3	0	0.375	0.75	-3	-2.02304077	0.700195313
4						
5						

Figure 3.31 Using AutoFill to calculate $F(x_{\text{min}})$, $F(x_{\text{mid}})$, $F(x_{\text{max}})$ in the second round of the bisection search.

We then want to calculate the new value of x_{mid} ($x_{\text{mid}} = \frac{1}{2}(x_{\text{min}} + x_{\text{max}})$) as well as $F(x_{\text{min}}), F(x_{\text{mid}}), F(x_{\text{max}})$. The formulas for all of these terms were previously entered on row 2, so we can AutoFill down from cell B2 to B3, and from cells D2, E2, and F2 to D3, E3 and F3. At this stage your spreadsheet should now look like Figure 3.31.

At this stage you may rightly feel that we have done a lot more work to do two rounds of the bisection search using the spreadsheet than if we had just proceeded the way we did in Section 3.5.1. But we have set everything up to do all further rounds automatically, by AutoFilling down the rows of our spreadsheet. Specifically we can select the cells A3, B3, ..., F3 and then AutoFill multiple rows down to calculate the third, fourth, fifth rounds etc., of the bisection search (see Figure 3.32). For example, AutoFill puts the formula

$$\mathbf{= IF(E3 > 0, A3, B3)}$$

in cell A4, which sets A4 equal to A3 (the old x_{min}) or B3 (the old x_{mid}) depending on whether $F(x_{\text{mid}}) > 0$ or $F(x_{\text{mid}}) \leq 0$. So we can quickly run more rounds of the bisection search by AutoFilling more rows. In Figure 3.32 we calculate 22 rounds of the bisection search, and find that the root is approximately $x = 0.66913$.

	A	B	C	D	E	F
1	x_min	x_mid	x_max	F(x_min)	F(x_mid)	F(x_max)
2	0	0.75	1.5	-3	0.700195313	5.15625
3	0	0.375	0.75	-3	-2.02304077	0.700195313
4	0.375	0.5625	0.75	-2.02304077	-0.84146976	0.700195313
5	0.5625	0.65625	0.75	-0.84146976	-0.10706726	0.700195313
6	0.65625	0.703125	0.75	-0.10706726	0.288838151	0.700195313
7	0.65625	0.6796875	0.703125	-0.10706726	0.088766104	0.288838151
8	0.65625	0.66796875	0.6796875	-0.10706726	-0.00970253	0.088766104
9	0.667969	0.673828125	0.6796875	-0.00970253	0.039396505	0.088766104
10	0.667969	0.670898438	0.67382813	-0.00970253	0.014812826	0.039396505
11	0.667969	0.669433594	0.67089844	-0.00970253	0.002546566	0.014812826
12	0.667969	0.668701172	0.66943359	-0.00970253	-0.00358013	0.002546566
13	0.668701	0.669067383	0.66943359	-0.00358013	-0.00051732	0.002546566
14	0.669067	0.669250488	0.66943359	-0.00051732	0.001014489	0.002546566
15	0.669067	0.669158936	0.66925049	-0.00051732	0.000248551	0.001014489
16	0.669067	0.669113159	0.66915894	-0.00051732	-0.00013439	0.000248551
17	0.669113	0.669136047	0.66915894	-0.00013439	5.70771E-05	0.000248551
18	0.669113	0.669124603	0.66913605	-0.00013439	-3.8658E-05	5.70771E-05
19	0.669125	0.669130325	0.66913605	-3.8658E-05	9.20928E-06	5.70771E-05
20	0.669125	0.669127464	0.66913033	-3.8658E-05	-1.4725E-05	9.20928E-06
21	0.669127	0.669128895	0.66913033	-1.4725E-05	-2.7576E-06	9.20928E-06
22	0.669129	0.66912961	0.66913033	-2.7576E-06	3.22582E-06	9.20928E-06
23	0.669129	0.669129252	0.66912961	-2.7576E-06	2.34089E-07	3.22582E-06
24						

Figure 3.32 Using AutoFill we can run an additional 20 rounds of the bisection search.

3.5.3 A Final Remark on Continuous Functions

Many functions in biology are in fact discontinuous. For example, the size of a population of cells growing in a flask must take integer values (there can be 0, 1, 2, 3, 4, ..., etc., cells in the flask, but never 3.2 cells). However, as we discussed in Section 2.3, typically the models that we create to describe biological processes, such as the growth of a population, are valid only as descriptions of the average behavior of the process. For example, we may have a function $N(t)$ representing the average number of cells in the flask; that is, if we had many flasks starting with the same number of cells initially, then $N(t)$ tells us the number of cells at time t if we averaged all the flasks (we will discuss averages in more detail in Chapter 12). So if $N(t) = 1700.5$ at some time, t, some of our flasks may contain 1700 cells, some 1701 cells, and others more or fewer cells. Even though each flask contains an integer number of cells, the average over all flasks might not be an integer, and $N(t)$ will generally be a continuous function of t.

Section 3.5 Problems

3.5.1

1. Let $f(x) = x^2 - 2, \quad 0 \le x \le 2$.

(a) Graph $y = f(x)$ for $0 \le x \le 2$.

(b) Show that $f(0) < 0 < f(2)$ and use the Intermediate Value Theorem to conclude that there exists a number $c \in (0, 2)$ such that $f(c) = 0$.

2. Let $f(x) = x^3 + 3, \quad -3 \le x \le -1$.

(a) Graph $y = f(x)$ for $-3 \le x \le -1$.

(b) Use the Intermediate Value Theorem to conclude that $x^3 + 3 = 0$ has a solution in $(-3, -1)$.

3. Let $f(x) = \sqrt{x} + x, \quad 1 \le x \le 2$.

(a) Graph $y = f(x)$ for $1 \le x \le 2$.

(b) Use the Intermediate Value Theorem to conclude that $\sqrt{x} + x = 3$ has a solution in $(1, 2)$.

4. Let $f(x) = \sin x - x, \quad -1 \le x \le 1$.

(a) Graph $y = f(x)$ for $-1 \le x \le 1$.

(b) Use the Intermediate Value Theorem to conclude that $\sin x = x$ has a solution in $(-1, 1)$.

5. Use the Intermediate Value Theorem to show that $e^{-x} = x^2$ has a solution in $(0, 1)$.

6. Use the Intermediate Value Theorem to show that $\cos x = x$ has a solution in $(0, 1)$.

7. Show that $e^{-x} = x^2$ has a solution in $(0.5, 1)$. Use the bisection method to find a solution that is accurate to two decimal places.

8. Show that $\cos x = x$ has a solution in $(0.5, 1)$. Use the bisection method to find a solution that is accurate to two decimal places.

9. (a) Use the bisection method to find a solution of $3x^3 - 4x^2 - x + 2 = 0$ that is accurate to two decimal places.

T **(b)** Graph the function $f(x) = 3x^3 - 4x^2 - x + 2$.

(c) Which solution did you locate in (a)? Is it possible in this case to find the other solution by using the bisection method together with the Intermediate Value Theorem?

10. In Example 2, how many steps are required to guarantee that the approximate root is within 0.0001 of the true value of the root?

11. Explain why a polynomial of degree 3 has at least one root.

12. Explain why a polynomial of degree n, where n is an odd number, has at least one root.

13. Explain why $y = x^2 - 5$ has at least two roots.

14. On the basis of the Intermediate Value Theorem, what can you say about the number of roots of a polynomial of even degree?

3.5.2

T **15. (a)** Use the Intermediate Value Theorem to show that $e^x = 2 - x$ has a solution in $(0, 2)$.

(b) Find this solution to an accuracy of 10^{-4} using the bisection search method, implemented as a spreadsheet.

T **16.** Use the Intermediate Value Theorem to show that $e^x - e^{-1.5x} - 1 = 0$ has a solution in the interval $(0, 1)$. Use a spreadsheet to calculate the value of this root to an accuracy of 10^{-5}.

For each of the following equations show that the equation has a root in the given interval. Then use the bisection search method, implemented as a spreadsheet, to find this root to an accuracy of 10^{-5}.

	Equation	Interval
17.	$x^3 - 2x + 1 = 0$	$(-2, -1)$
18.	$x^5 + x^2 - x + 1 = 0$	$(-2, -1)$
19.	$\sin^2 x = \frac{x}{10}$	$(2, 3)$
20.	$\cos^2 x = x$	$(0, 1)$

3.6 A Formal Definition of Limits

In Section 3.1 we defined the limit of a function in an informal, nonrigorous way. The definition that we used in Section 3.1 is perfectly adequate for calculating any of the limits that you will encounter in the mathematical modeling of biological processes. But it is not the definition generally used by mathematicians. In this section we will give a rigorous definition of limits. This material is optional.

Before we write the formal definition, let's return to the informal one. In that definition, we stated that $\lim_{x\to c} f(x) = L$ means that the value of $f(x)$ can be made arbitrarily close to L whenever x is close enough to c. But just how close is close enough? Take Example 1 from Section 3.1: Suppose we wish to show that

$$\lim_{x\to 2} x^2 = 4$$

without using the continuity of $y = x^2$, which itself was based on $\lim_{x\to c} x = c$ [Equation (3.4)]. What would we have to do? We would need to show that x^2 can be made arbitrarily close to 4 for all values of x close enough, but not equal, to 2. (In what follows, we will always exclude $x = 2$ from the discussion, since the value of x^2 at $x = 2$ is irrelevant in finding the limit.) Suppose we wish to make x^2 within 0.01 of 4; that is, we want $|x^2 - 4| < 0.01$. Does this inequality hold for all x close enough, but not equal, to 2? We begin with

$$|x^2 - 4| < 0.01$$

which is equivalent to

$$-0.01 < x^2 - 4 < 0.01$$

$$3.99 < x^2 < 4.01$$

$$\sqrt{3.99} < |x| < \sqrt{4.01}$$

Now, $\sqrt{3.99} = 1.997498\ldots$ and $\sqrt{4.01} = 2.002498\ldots$. We therefore find that values of $x \neq 2$ in the interval $(1.998, 2.002)$ satisfy $|x^2 - 4| < 0.01$. (We chose a somewhat smaller interval than indicated, to get an interval that is symmetric about 2.) That is, for all values of x within 0.002 of 2 but not equal to 2 (i.e., $0 < |x - 2| < 0.002$), x^2 is within the prescribed precision—that is, within 0.01 of 4.

You might think about this example in the following way: Suppose that you wish to stake out a square of area 4 m^2. Each side of your square is 2 m long. You bring along a stick, which you cut to a length of 2 m. We can then ask: How accurately do we need to cut the stick so that the area will be within a prescribed precision? Our prescribed precision was 0.01, and we found that if we cut the stick within 0.002 of 2 m, we would be able to obtain the prescribed precision.

There is nothing special about 0.01; we could have chosen any other degree of precision and would have found a corresponding interval of x-values. We translate this procedure into a formal definition of limits. (See Figure 3.33.)

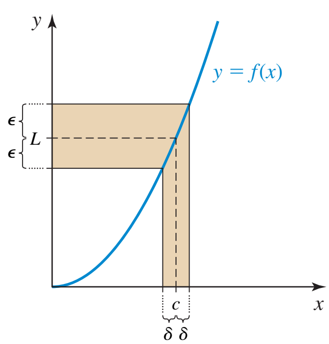


Figure 3.33 The ϵ–δ definition of limits.

> **Definition The $\epsilon - \delta$ definition of a limit** The statement
>
> $$\lim_{x\to c} f(x) = L$$
>
> means that, for every $\epsilon > 0$, there exists a number $\delta > 0$ such that
>
> $$|f(x) - L| < \epsilon \quad \text{whenever} \quad 0 < |x - c| < \delta$$

Note that, as in the informal definition of limits, we exclude the value $x = c$ from the statement. (This is done in the inequality $0 < |x - c|$.) To apply the formal definition, we first need to guess the limiting value L. We then choose an $\epsilon > 0$, the prescribed precision, and try to find a $\delta > 0$ such that $f(x)$ is within ϵ of L whenever x is within δ of c but not equal to c. [In our example, $f(x) = x^2$, $c = 2$, $L = 4$, $\epsilon = 0.01$, and $\delta = 0.002$.]

EXAMPLE 1 Show that $\lim\limits_{x\to 1}(2x - 3) = -1$.

Solution We let $f(x) = 2x - 3$. Our guess for the limiting value is $L = f(1) - 1$. Then:

$$\begin{aligned} |f(x) - L| &= |2x - 3 - (-1)| \\ &= |2x - 2| \\ &= 2|x - 1|. \end{aligned}$$

Now if we are given any $\epsilon > 0$, our goal is to find a $\delta > 0$ such that $2|x-1| < \epsilon$ whenever x is within δ of 1 but not equal to 1; that is, $0 < |x-1| < \delta$. (ϵ is arbitrary, and we do not specify it because our statement needs to hold for all $\epsilon > 0$.) The value of δ will typically depend on our choice of ϵ. Since $|x-1| < \delta$ implies that $2|x-1| < 2\delta$, we should try $2\delta = \epsilon$. If we choose $\delta = \epsilon/2$, then, indeed,

$$|f(x) - L| = 2|x-1| < 2\delta = 2\frac{\epsilon}{2} = \epsilon$$

This means that, for every $\epsilon > 0$, we can find a number $\delta > 0$ (namely, $\delta = \epsilon/2$) such that

$$|f(x) - (-1)| < \epsilon \qquad \text{whenever} \qquad 0 < |x-1| < \delta$$

But this is exactly the definition of $\lim_{x\to 1}(2x-3) = -1$. •

EXAMPLE 2 We promised in Section 3.2 that we would show that $\lim_{x\to c} x = c$, use the rigorous definition of a limit to prove this statement.

Solution Let $f(x) = x$. We need to show that, for every $\epsilon > 0$, there corresponds a number $\delta > 0$ such that

$$|f(x) - L| = |x - c| < \epsilon \qquad \text{whenever} \qquad 0 < |x - c| < \delta \qquad f(x) = x \text{ and } L = c \qquad (3.9)$$

We should choose $\delta = \epsilon$, and, indeed, if $\delta = \epsilon$, then (3.9) holds. •

Let's look at an example in which $f(x)$ is not linear.

EXAMPLE 3 Use the formal definition of limits to show that $\lim_{x\to 0} x^3 = 0$.

Solution We need to show that, for every $\epsilon > 0$, there corresponds a number $\delta > 0$ such that

$$|f(x) - L| = |x^3| < \epsilon \quad \text{whenever} \quad 0 < |x| < \delta \qquad f(x) = x^3 \text{ and } L = 0 \qquad (3.10)$$

Now, $|x^3| < \epsilon$ is equivalent to

$$-\epsilon < x^3 < \epsilon$$

$$-\epsilon^{1/3} < x < \epsilon^{1/3}$$

This pair of inequalities suggests that we set $\delta = \epsilon^{1/3}$. Accordingly, if $0 < |x| < \epsilon^{1/3}$, then

$$-\epsilon^{1/3} < x < \epsilon^{1/3}$$

or

$$-\epsilon < x^3 < \epsilon$$

which is the same as $|x^3| < \epsilon$. •

We can also use the formal definition to show that a limit does not exist.

EXAMPLE 4 Show that $\lim_{x\to 0}\theta(x)$ does not exist. $\theta(x)$ is the Heaviside function defined by

$$\theta(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{2} & \text{if } x = 0 \\ 1 & \text{if } x > 0. \end{cases}$$

Solution Showing that this limit does not exist is tricky. (See Figure 3.34.) The approach is as follows: we assume that the limit exists and try to find a contradiction.[1] Suppose, then, that there exists an L such that

$$\lim_{x\to 0}\theta(x) = L$$

(1) This approach is called "indirect proof" or "*reductio ad absurdum*." We assume the opposite of what we wish to prove, and then we show that assuming the opposite leads to a contradiction. Therefore, what we originally sought to prove must be true.

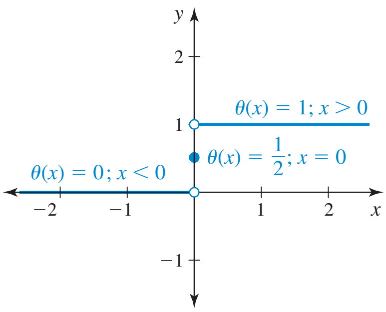


Figure 3.34 The graph of the Heaviside function, $\theta(x)$, in Example 4: The limit of $\theta(x)$ as x tends to 0 does not exist.

If we look at Figure 3.34, we see that if, for instance, we choose $L = 1$, then we cannot get close to L when x is less than 0. Similarly, we see that, for any value of L, either the distance to $+1$ exceeds 1/2, or the distance to 0 exceeds 1/2. So whatever the value of L, if $\epsilon = \frac{1}{4}$ we will not be able to find a value of δ such that if $0 < |x| < \delta$, then $|\theta(x) - L| < \epsilon$, since $\theta(x)$ takes on both the values $+1$ and 0 for $0 < |x| < \delta$. Therefore, $\lim_{x \to 0} \theta(x)$ does not exist. •

In the previous section, we considered an example in which $\lim_{x \to c} f(x) = \infty$. This statement can be made precise as well.

> **Definition Divergent Limits** The statement $\lim_{x \to c} f(x) = \infty$ means that, for every $M > 0$, there exists a $\delta > 0$ such that
>
> $$f(x) > M \quad \text{whenever} \quad 0 < |x - c| < \delta.$$

Similar definitions hold for the case when $\lim_{x \to c} f(x) = -\infty$ and for one-sided limits. We will not give definitions for all possible cases; rather, we illustrate how we would use such a definition.

EXAMPLE 5 Show that $\lim_{x \to 0} \dfrac{1}{x^2} = \infty$.

Solution The graph of $f(x) = 1/x^2$, $x \neq 0$, is shown in Figure 3.35. Given any $M > 0$, we need to find a $\delta > 0$ such that $f(x) > M$ whenever $0 < |x| < \delta$. (Again, M is arbitrary, because our solution must hold for all $M > 0$.) We start with the inequality $f(x) > M$ and try to determine how to choose δ.

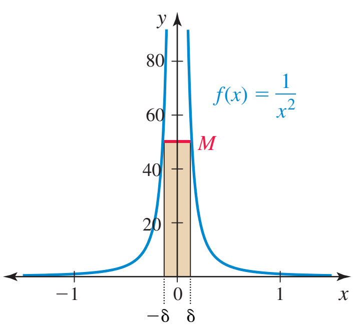


Figure 3.35 The function $f(x) = \frac{1}{x^2}$ in Example 5: The limit of $\frac{1}{x^2}$ as x tends to 0 does not exist.

$$\frac{1}{x^2} > M \qquad \text{is the same as} \qquad x^2 < \frac{1}{M}$$

Taking square roots on both sides, we find that

$$|x| < \frac{1}{\sqrt{M}}$$

This suggests that we should choose $\delta = 1/\sqrt{M}$. Let's try that value: Given $M > 0$, we choose $\delta = 1/\sqrt{M}$. If $0 < |x| < \delta$, then

$$x^2 < \delta^2, \qquad \text{or} \qquad \frac{1}{x^2} > \frac{1}{\delta^2} = M$$

That is, $1/x^2 > M$ whenever $0 < |x| < 1/\sqrt{M}$. •

There is also a formal definition when $x \to \infty$ (and a similar one for $x \to -\infty$). This definition is analogous to that in Chapter 2.

> **Definition Limits of Infinity** The statement $\lim_{x \to \infty} f(x) = L$ means that, for every $\epsilon > 0$, there exists an $x_0 > 0$ such that
>
> $$|f(x) - L| < \epsilon \quad \text{whenever } x > x_0.$$

EXAMPLE 6 Show that $\lim_{x \to \infty} \dfrac{x}{0.5 + x} = 1$.

Solution This limit is illustrated in Figure 3.36. You can see that $f(x) = x/(0.5 + x)$, $x \geq 0$, is in the strip of width 2ϵ and centered at the limiting value $L = 1$ for all values of x greater

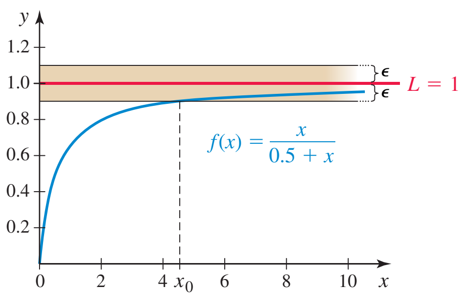


Figure 3.36 The function $f(x) = \frac{x}{0.5+x}$ in Example 6: The limit of $f(x)$ as x tends to infinity is 1.

than x_0. We now determine x_0. To do this, we try to solve

$$\left|\frac{x}{0.5+x} - 1\right| < \epsilon$$

for any $\epsilon > 0$. Equivalently

$$-\epsilon < \frac{x}{0.5+x} - 1 < \epsilon \qquad \text{Remove the absolute values.}$$

or

$$1 - \epsilon < \frac{x}{0.5+x} < 1 + \epsilon \qquad \text{Add 1 to all three parts.}$$

Since $\frac{x}{0.5+x} < 1$ for $x > 0$, the right-hand inequality always holds. We therefore need only consider

$$1 - \epsilon < \frac{x}{0.5+x}$$

Because we are interested in the behavior of $f(x)$ as $x \to \infty$, we need only look at large values of x. Multiplying by $0.5 + x$ (and noticing that we can assume that $0.5 + x > 0$, because we let $x \to \infty$), we obtain

$$(1 - \epsilon)(0.5 + x) < x$$

or

$$(1 - \epsilon)(0.5) < x - x(1 - \epsilon) \qquad \text{Isolate terms in } x.$$

$$(1 - \epsilon)(0.5) < \epsilon x \qquad \text{Simplify right-hand side}$$

$$\frac{1 - \epsilon}{2\epsilon} < x \qquad \text{Solve for } x.$$

For instance, if $\epsilon = 0.1$ (as in Figure 3.36), then for all

$$x > \frac{0.9}{0.2} = 4.5$$

$|f(x) - 1| < 0.1$. That is, we would set $x_0 = 4.5$ and conclude that, for $x > 4.5$, $|f(x) - 1| < 0.1$.

More generally, we find that, for any $\epsilon > 1$, we may define $x_0 = \frac{1-\epsilon}{2\epsilon}$ and then

$|f(x) - 1| < \epsilon$ whenever $x > x_0$ •

Section 3.6 Problems

1. Find the values of x such that

$$|2x + 1| < 0.01$$

2. Find the values of x such that

$$|3x - 9| < 0.01$$

3. Find the values of x such that

$$|x^2 - 9| < 0.1$$

4. Find the values of x such that

$$|2\sqrt{x} - 5| < 0.01$$

5. Let

$$f(x) = 2x - 1, \ x \in \mathbf{R}$$

(a) Graph $y = f(x)$ for $-3 \le x \le 5$.

(b) For which values of x is $y = f(x)$ within 0.1 of 3? [*Hint*: Find values of x such that $|(2x - 1) - 3| < 0.1$.]

(c) Illustrate your result in (b) on the graph that you obtained in (a).

6. Let

$$f(x) = \sqrt{x}, \quad x \ge 0.$$

(a) Graph $y = f(x)$ for $0 \le x \le 6$.

(b) For which values of x is $y = f(x)$ within 0.2 of 2? (*Hint*: Find values of x such that $|\sqrt{x} - 2| < 0.2$.)

(c) Illustrate your result in (b) on the graph that you obtained in (a).

7. Let

$$f(x) = \frac{1}{x}, \quad x > 0.$$

(a) Graph $y = f(x)$ for $0 < x \le 4$.

(b) For which values of x is $y = f(x)$ greater than 5?

(c) Illustrate your result in (b) on the graph that you obtained in (a).

8. Let

$$f(x) = \frac{e^{-x}}{2}, \quad x \ge 0.$$

(a) Graph $y = f(x)$ for $0 \le x \le 6$.

(b) For which values of x is $y = f(x)$ less than 0.1?

(c) Illustrate your result in (b) on the graph that you obtained in (a).

In Problems 9–22, use the formal definition of limits to prove each statement.

9. $\lim_{x\to 2}(2x-1) = 3$.

10. $\lim_{x\to 0} 2x^3 = 0$.

11. $\lim_{x\to 0} x^5 = 0$.

12. $\lim_{x\to 1} \frac{1}{x} = 1$.

13. $\lim_{x\to 0} \frac{3}{x^2} = \infty$.

14. $\lim_{x\to 0} \frac{-2}{x^2} = -\infty$.

15. $\lim_{x\to 0} \frac{1}{x^4} = \infty$.

16. $\lim_{x\to 3} \frac{1}{(x-3)^2} = \infty$.

17. $\lim_{x\to 0^+} \frac{1}{x} = \infty$.

18. $\lim_{x\to 0^-} \frac{1}{x} = -\infty$.

19. $\lim_{x\to 3^+} \frac{1}{3-x} = -\infty$.

20. $\lim_{x\to 0^-} \frac{-4}{x} = \infty$.

21. $\lim_{x\to c}(mx) = mc$, where m is a constant

22. $\lim_{x\to c}(mx+b) = mc+b$, where m and b are constants

Chapter 3 Review

Key Terms

Discuss the following definitions and concepts:

1. Limit of $f(x)$ as x approaches c

2. One-sided limits

3. Infinite limits

4. Divergence by oscillations

5. Convergence

6. Divergence

7. Limit laws

8. Continuity

9. One-sided continuity

10. Continuous function

11. Removable discontinuity

12. Sandwich theorem

13. Trigonometric limits

14. Intermediate-value theorem

15. Bisection method

16. ϵ–δ definition of limits

Review Problems

In Problems 1–4, determine where each function is defined and continuous. Investigate the behavior as $x \to \pm\infty$. Use a graphing calculator to sketch the corresponding graphs.

1. $f(x) = e^{-|x|}$.

2. $f(x) = \begin{cases} \dfrac{1-\cos x}{x} & \text{if } x \ne 0 \\ 0 & \text{if } x = 0 \end{cases}$.

3. $f(x) = \dfrac{2}{e^x + e^{-x}}$.

4. $f(x) = \dfrac{1}{\sqrt{x^2-1}}$.

5. Sketch the graph of a function that is discontinuous from the left and continuous from the right at $x = 1$.

6. Sketch the graph of a function $f(x)$ that is continuous on $[0,2]$, except at $x = 1$, where $f(1) = 3$, $\lim_{x\to 1^-} f(x) = 2$, and $\lim_{x\to 1^+} f(x) = 4$.

7. Sketch the graph of a continuous function on $[0,\infty)$ with $f(0) = 0$ and $\lim_{x\to\infty} f(x) = 1$.

8. Sketch the graph of a continuous function on $(-\infty,\infty)$ with $f(0) = 1$, $f(x) \ge 0$ for all $x \in \mathbf{R}$, and $\lim_{x\to\pm\infty} f(x) = 0$.

9. Show that the floor function

$$f(x) = \lfloor x \rfloor$$

is continuous from the right, but discontinuous from the left at $x = -2$.

10. Suppose $f(x)$ is continuous on the interval $[1,3]$. If $f(1) = 0$ and $f(3) = 3$ explain why there must be a number $c \in (1,3)$ such that $f(c) = 1$.

11. Population Size Assume that the size of a population at time t is

$$N(t) = \frac{at}{k+t}, \quad t \ge 0$$

where a and k are positive constants. Suppose that the limiting population size is

$$\lim_{t\to\infty} N(t) = 1.24 \times 10^6,$$

and that, at time $t = 5$, the population size is half the limiting population size. Use the preceding information to determine the constants a and k.

12. Population Size Suppose that

$$N(t) = 10 + 2e^{-0.3t}\sin t, \quad t \ge 0$$

describes the size of a population (in millions) at time t (measured in weeks).

T **(a)** Use a graphing calculator to sketch the graph of $N(t)$, and describe in words what you see.

(b) Give lower and upper bounds on the size of the population; that is, find N_1 and N_2 such that, for all $t \ge 0$,

$$N_1 \le N(t) \le N_2$$

(c) Use the sandwich theorem to find $\lim_{t\to\infty} N(t)$.

13. Physiology Suppose that an organism reacts to a stimulus only when the stimulus exceeds a certain threshold. Assume that the stimulus is a function of time t and that it is given by

$$s(t) = \sin(\pi t), \quad t \ge 0$$

The organism reacts to the stimulus and shows a certain reaction when $s(t) \ge 1/2$. Define a function $g(t)$ such that $g(t) = 0$ when the organism shows no reaction at time t and $g(t) = 1$ when the organism shows the reaction.

(a) Plot $s(t)$ and $g(t)$ in the same coordinate system.

(b) Is $s(t)$ continuous? Is $g(t)$ continuous?

14. Quorum Sensing For some species of growth and other behaviors change depending on how many bacteria are present in the population. This effect is known as quorum sensing because a certain quorum or critical population size must be exceeded for the bacterial behavior to change. ***Nealson et al. (1970)*** studied how the amount of light produced by glow-in-the-dark bacterium *Aliivibrio fischeri* depends on population size. They found that when the population size was fewer than 10^8 cells the bacteria glowed very little, but above this threshold bacteria started to glow intensely. If the amount of light produced by the bacteria is L, and the population size is N, then ***Nealson et al.*** measured:

$$L(N) = \begin{cases} 0.006 \times 10^{10} & \text{if } N \le 10^8 \\ \dfrac{30 \times 10^{10} N^2}{N^2 + a} & \text{if } N > 10^8 \end{cases}$$

where a is a positive constant. Light output is measured in number of photons produced per second.

(a) Calculate the value of a that makes $L(N)$ a continuous function for all $N \ge 0$.

(b) Show that as population size grows the light output from the bacteria saturates at some level; that is: $\lim_{N\to\infty} L(N)$ exists, and find the limit.

(c) Sketch the graph of $L(N)$ against N.

15. Circadian Rhythms Many microorganisms alter their growth rate in response to the time of day, for example, growing fastest in the daytime and more slowly at night. The fungus *Neurospora crassa* uses light levels to tell time—it grows faster in light (day) and slower in the dark (night). However, even when the fungus is moved to an environment with constant light (no day–night cycle) it continues to vary its growth rate with time for some days because it has its own circadian clock that enables it to tell time even without cues from its environment.

Up to time $t = 3$ a *Neurospora crassa* fungus is grown in conditions with a regular day–night cycle. At time $t = 3$ it is transferred to new conditions with constant light. It is measured that the growth rate $r(t)$ measured in mm/hr varies with time according to:

$$r(t) = \begin{cases} 3 + 1.2\sin(2\pi t) & \text{if } t \le 3 \\ 3.6 + Ae^{-t}\sin(2\pi t) & \text{if } t > 3. \end{cases}$$

where A is a constant.

(a) Is there any value of A that can be chosen to make $r(t)$ a continuous function of t at $t = 3$?

(b) Assume $A = 0.6$. Calculate $\lim_{t\to\infty} r(t)$.

(c) Make a sketch of the function $r(t)$ against t.

16. Medication in the Body In Section 2.3 we built mathematical models for how drugs and medications pass through the human body. There we introduced two models for the dependence of the rate E at which medication would be eliminated from the body upon the amount a, of medication present in the body. We defined a medication to have *zeroth order kinetics* if r is a constant, independent of a, and *first order kinetics* if $r \propto a$. In reality most medications are somewhere between the two; they have first order kinetics if the amount present in the body is not too high, and zeroth order kinetics if the amount is higher. This is because at very high doses the body reaches the limit of how quickly it can eliminate the medication.

In this question we will consider two models for such a medication.

Assume that we know that if the amount of a particular medication is small then 20% of it is removed from the body each hour; that is, $r(a) = 0.2a$. If the amount is large, then it is removed at a constant rate of 20 units/hr, so $r(a) = 2$.

(a) For a first model consider the following function:

$$r(a) = \begin{cases} 0.2a & \text{if } 0 \le a \le c \\ 20 & \text{if } a > c \end{cases}$$

where c is a constant that needs to be determined as part of your model. What value should be assigned to c if we want $r(a)$ to be a continuous function of a for all $a \ge 0$?

(b) An alternate model for the changing rate of elimination is to assume that the drug elimination has Michaelis-Menten kinetics. That is:

$$r(a) = \frac{20\,a}{a + 100}, \quad a \ge 0$$

You will show that this particular version of the Michaelis-Menten equation captures the small a and large a behavior of $r(a)$.

(i) Show that $\lim_{a\to 0} \frac{r(a)}{a} = 0.2$, and explain why this limiting behavior is consistent with $r(a)$ having approximately first order kinetics for small a.

(ii) Show that $\lim_{a\to\infty} r(a) = 20$ and explain why this limiting behavior is consistent with $r(a)$ having approximately zeroth order kinetics for large a.

17. Predator–Prey Model There are a number of mathematical models that describe predator–prey interactions. Typically, they share the feature that the number of prey eaten per predator increases with the number of prey available. In the simplest version, the number of encounters with prey per predator is proportional to the product of the total number of prey and the period over which the predators search for prey. That is, if we let N be the number of prey, P be the number of predators, T be the period available for searching, and E be the number of encounters with prey, then

$$\frac{E}{P} = aTN, \tag{3.11}$$

where a is a positive constant. The quantity E/P is the number of prey encountered per predator.

(a) Set $f(N) = aTN$, and sketch the graph of $f(N)$ when $a = 0.1$ and $T = 2$ for $N \ge 0$.

(b) Predators usually spend some time eating the prey that they find. Therefore, not all of the time T can be used for searching. Each time a predator catches prey it must spend a time T_h eating or handling it (for example, spiders and some birds store dead prey to eat them later). Thus, if a predator catches $\frac{E}{P}$ prey, then it must spend a total time $\frac{T_h E}{P}$ handling those prey, and it therefore has time only $T - \frac{T_h E}{P}$ available to search for new prey.

Show that if $T - T_h \frac{E}{P}$ is substituted for T in (3.11), then

$$\frac{E}{P} = \frac{aTN}{1 + aT_h N}. \tag{3.12}$$

Define

$$g(N) = \frac{aTN}{1 + aT_h N},$$

and graph $g(N)$ for $N \ge 0$ when $a = 0.1$, $T = 2$, and $T_h = 0.1$.

(c) Show that (3.12) reduces to (3.11) when $T_h = 0$.

(d) Find

$$\lim_{N\to\infty} \frac{E}{P}$$

in the cases when $T_h = 0$ and when $T_h > 0$. Explain, in words, the difference between the two cases.

18. **CO_2 Storage** Since increasing levels of man-made CO_2 in the atmosphere are known to affect climate there is increasing interest in trying to remove CO_2 from the atmosphere by planting trees and other plants. Plants remove CO_2 from the air during photosynthesis, as CO_2 molecules are broken down to make sugars and starches that the plant then stores. But plants can also produce CO_2 when they respire (break down sugars for energy) just like humans and other animals. Whether or not a plant ecosystem can or cannot remove CO_2 from the air depends on whether the rate at which CO_2 is stored (S) exceeds or is less than the rate of respiration (R).

Duarte and Agustí (1998) investigated the CO_2 balance of aquatic ecosystems. They related the community respiration rates (R) to the gross storage rates (S) of aquatic ecosystems. They summarize their results in the following quote:

> The relation between community respiration rate and gross production is not linear. Community respiration is scaled as the approximate two-thirds power of gross storage.

(a) Use the preceding quote to explain why

$$R = aS^b,$$

can be used to describe the relationship between the community respiration rates (R) and the gross storage (S). What value would you assign to b on the basis of their quote?

(b) Suppose that you obtained data on the gross production and respiration rates of a number of freshwater lakes. How would you display your data graphically to quickly convince an audience that the exponent b in the power equation relating R and S is indeed approximately $2/3$? (*Hint*: Use an appropriate log transformation.)

(c) The ratio R/S for an ecosystem is important in assessing the global CO_2 budget. If respiration exceeds storage (i.e., $R > S$), then the ecosystem acts as a carbon dioxide source, whereas if storage exceeds respiration (i.e., $S > R$), then the ecosystem acts as a carbon dioxide sink. Assume now that the exponent in the power equation relating R and S is $2/3$. Show that the ratio R/S, as a function of P, is continuous for $P > 0$. Furthermore, show that

$$\lim_{P\to 0^+} \frac{R}{S} = \infty$$

and

$$\lim_{P\to\infty} \frac{R}{S} = 0.$$

(d) Use your results in (c) and the intermediate-value theorem to conclude that there exists a value S^* such that the ratio R/S at S^* is equal to 1.

(e) Assume that $a = 1$, so:

$$R = S^{2/3},$$

Calculate the value of S at which $R = S$. Explain why the plant ecosystems having highest respiration rates (that is, large R) are actually the best candidates for CO_2 storage.

19. Hyperbolic functions are used in the sciences. We take a look at the following three examples: the hyperbolic sine, $\sinh x$; the hyperbolic cosine, $\cosh x$; and the hyperbolic tangent, $\tanh x$, defined respectively as

$$\sinh x = \frac{e^x - e^{-x}}{2}, \quad x \in \mathbf{R}$$

$$\cosh x = \frac{e^x + e^{-x}}{2}, \quad x \in \mathbf{R}$$

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad x \in \mathbf{R}$$

T **(a)** Show that these three hyperbolic functions are continuous for all $x \in \mathbf{R}$. Use a graphing calculator to sketch the graphs of all three functions.

(b) Find

$$\lim_{x\to\infty} \sinh x, \qquad \lim_{x\to-\infty} \sinh x,$$

$$\lim_{x\to\infty} \cosh x, \qquad \lim_{x\to-\infty} \cosh x,$$

$$\lim_{x\to\infty} \tanh x, \quad \text{and} \quad \lim_{x\to-\infty} \tanh x.$$

(c) Prove the two identities

$$\cosh^2 x - \sinh^2 x = 1,$$

and

$$\tanh x = \frac{\sinh x}{\cosh x}.$$

(d) Show that $\sinh x$ and $\tanh x$ are odd functions and that $\cosh x$ is even.

(e) The function $f(x) = \frac{1}{2} + \frac{1}{2}\tanh x$ is often used in models to approximate the Heaviside function introduced in Section 3.1. Show that

(i) $f(0) = \frac{1}{2}$

(ii) $\lim_{x\to-\infty} f(x) = 0$

(iii) $\lim_{x\to+\infty} f(x) = 1$

(iv) and that $f(x)$ is continuous for all $x \in \mathbb{R}$.

CHAPTER

4

Differentiation

This chapter presents the fundamentals of differentiation. Specifically, we will learn how to

- calculate the rate of change of a function;
- formally define a derivative;
- interpret rate of change of a function;
- differentiate specific functions;
- approximate a function by a linear function;
- calculate how a measurement error propagates.

Differential calculus allows us to solve two of the basic problems that we mentioned in Chapter 1: constructing a tangent line to a curve (Figure 4.1) and finding maxima and minima of a curve (Figure 4.2). The solutions to these two problems, by themselves, cannot explain the impact calculus has had on the sciences. Many mathematical models, including models used to predict the growth of populations, the spread of diseases, the working of neurons, and the evolution of organisms, take the form of equations for the rate of change of a quantity; e.g., the rate of growth of a population. We need the tools of calculus to write these models down and to solve them.

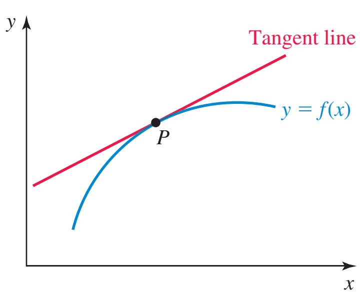


Figure 4.1 Tangent line to a curve at a point.

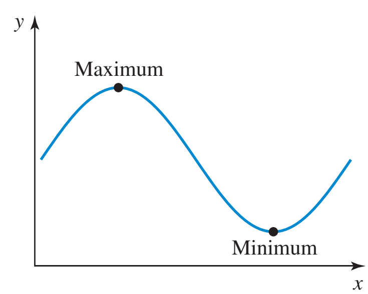


Figure 4.2 Maxima and minima of a curve.

Population growth will be of particular interest to us. Let's revisit Example 2 from Section 3.1, in which we looked at a population whose size at time t is given by $N(t)$. The average growth rate during the time interval $0 \leq t \leq h$ is equal to

$$\text{average growth rate} = \frac{\text{change in population size}}{\text{length of time interval}} = \frac{\Delta N}{\Delta t}$$

where

$$\Delta N = N(h) - N(0) \qquad \text{and} \qquad \Delta t = h - 0 = h$$

Thus,

$$\frac{\Delta N}{\Delta t} = \frac{N(h) - N(0)}{h}$$

The **instantaneous rate of growth** is defined as the limit of $\Delta N/\Delta t$ as $\Delta t \to 0$ (or $h \to 0$), or

$$\lim_{\Delta t \to 0} \frac{\Delta N}{\Delta t} = \lim_{h \to 0} \frac{N(h) - N(0)}{h}$$

provided that this limit exists.

In Example 2 from Section 3.1 we considered only the rate of growth of the population at $t = 0$, based on the change in population size between $t = 0$ and $t = h$. But we could do this same calculation to calculate the rate of growth at any time t; that is, we define the instantaneous rate of growth at time t, based on the change of population size between time t and time $t + h$:

$$\begin{aligned}\text{rate of growth} = \lim_{\Delta t \to 0} \frac{\Delta N}{\Delta t} &= \lim_{h \to 0} \frac{N(t+h) - N(t)}{(t+h) - t} \\ &= \lim_{h \to 0} \frac{N(t+h) - N(t)}{h}\end{aligned}$$

If $N(t)$ is continuous, then as $h \to 0$, $N(t+h) \to N(t)$, so $N(t+h) - N(t) \to 0$. But $h \to 0$ also, so the limit of the ratio, if it exists, may be non-zero.

We are interested in the geometric interpretation of the limit when it exists. When we draw a straight line through the points $(t, N(t))$ and $(t + h, N(t + h))$, we obtain the **secant line**. The slope of this line is given by the quantity $\Delta N/\Delta t$ (Figure 4.3). In the limit as $\Delta t \to 0$, the secant line converges to the line that touches the graph at the point $(t, N(t))$. This line is called the **tangent line** at the point $(t, N(t))$ (Figure 4.4). The limit of $\Delta N/\Delta t$ as the length of the time interval $[t, t+h]$ goes to 0 (i.e., $\Delta t \to 0$ or $h \to 0$) will therefore be equal to the slope of the tangent line at $(t, N(t))$.

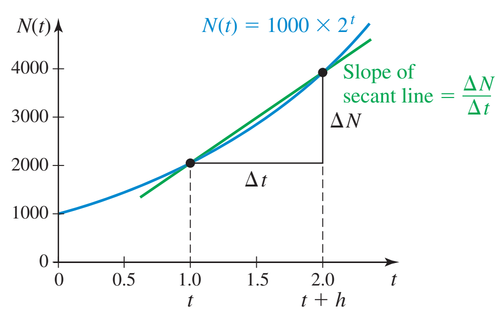


Figure 4.3 The average growth rate $\Delta N/\Delta t$ is equal to the slope of the secant line.

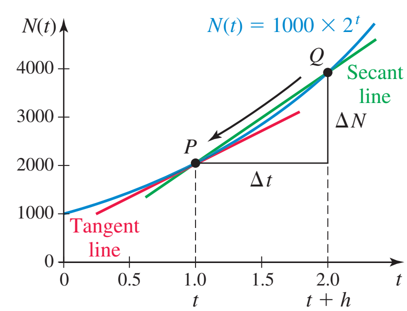


Figure 4.4 The instantaneous growth rate is the limit $\Delta N/\Delta t$ as $\Delta t \to 0$. Geometrically, the point Q moves toward the point P on the graph of $N(t)$, and the secant line through P and Q becomes the tangent line at P. The instantaneous growth rate is then equal to the slope of the tangent line.

4.1 Formal Definition of the Derivative

We have shown that $\lim_{\Delta t \to 0} \frac{\Delta N}{\Delta t}$ can be interpreted geometrically as the slope of a tangent line. It is convenient to have a special term for this quantity. We call it the **derivative** and denote it by $N'(t)$ (read as "N prime t"). More generally:

Definition The **derivative** of a function f at x, denoted by $f'(x)$, is

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

provided that the limit exists.

If the limit exists, then we say that f is **differentiable** at x. The phrase "provided that the limit exists" is crucial: If we take an arbitrary function, f, the limit may not exist. In fact, we saw many examples in the previous chapter in which limits did not exist. The geometric interpretation will help us to understand when the limit exists and under which conditions we cannot expect the limit to exist. Notice that $\lim_{h\to 0}$ is a two-sided limit (i.e., we approach 0 from both the negative and the positive side). The quotient

$$\frac{f(x+h)-f(x)}{h}$$

is called the **difference quotient**, and we denote it by $\frac{\Delta f}{\Delta x}$. (See Figure 4.5.)

We say that f is differentiable on (a,b) if f is differentiable at every $x \in (a,b)$. (Since the limit in the definition is two sided, we exclude the endpoints of the interval. At endpoints, only one-sided limits can be computed, which yield one-sided derivatives.)

If we want to compute the derivative at $x = c$, we can also write

$$f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c}$$

which emphasizes that the point $(x, f(x))$ converges to the point $(c, f(c))$ as we take the limit as $x \to c$ (Figure 4.6). This approach will be important when we discuss the geometric interpretation of the derivative in the next subsection.

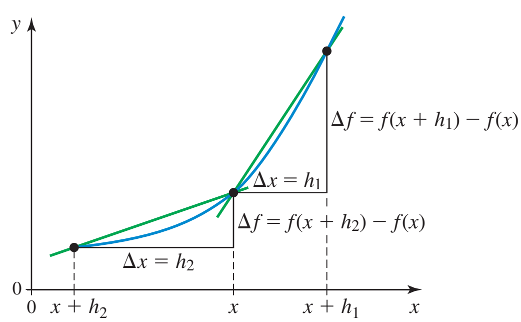


Figure 4.5 The difference quotient $\frac{f(x+h)-f(x)}{h}$ when $h = h_1 > 0$ and $h = h_2 < 0$.

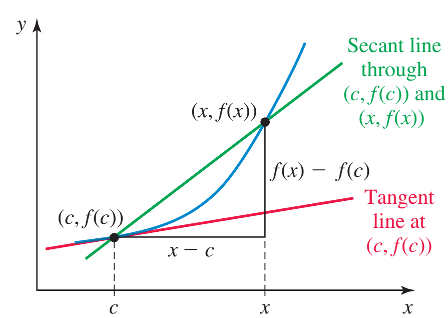


Figure 4.6 The derivative $f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c}$ is the slope of the tangent line at $(c, f(c))$.

There is more than one way to write the derivative of a function $y = f(x)$. The following expressions are equivalent:

$$y' = \frac{dy}{dx} = f'(x) = \frac{df}{dx} = \frac{d}{dx} f(x)$$

The notation $\frac{df}{dx}$ dates back 350 years to Leibniz's early work on calculus and is called **Leibniz notation**. It should remind you that we take the limit of $\Delta f/\Delta x$ as Δx approaches 0.

If we wish to emphasize that we evaluate the derivative of $f(x)$ at $x = c$, we write

$$f'(c) = \left.\frac{df}{dx}\right|_{x=c}$$

Newton used different notation to denote the derivative of a function. He wrote $\dot{y}$ (read "y dot") for the derivative of y. This notation is still common in physics when derivatives are taken with respect to a variable that represents time. We will use either Leibniz notation or the notation $f'(x)$.

For some functions we can calculate the derivative by making use of the geometric interpretation of the derivative, while for others we use the limit laws from Chapter 3.

Let's look at $f(x) = x^2$, $x \in \mathbf{R}$. (Refer to Figures 4.7 and 4.8 as we go along.) To compute the derivative of f at, say, $x = 1$ from the definition, we first compute the difference quotient at $x = 1$.

$$\begin{aligned}\frac{\Delta f}{\Delta x} &= \frac{f(1+h)-f(1)}{h} = \frac{(1+h)^2-1^2}{h} && \text{Assume } h \neq 0.\\ &= \frac{1+2h+h^2-1}{h} = \frac{2h+h^2}{h} && \text{Expand } (1+h)^2.\\ &= 2+h && \text{Cancel } h.\end{aligned}$$

The difference quotient $\Delta f/\Delta x$ is the slope of the secant line through the points $(x, y) = (1, 1)$ and $(x, y) = (1 + h, (1 + h)^2)$ (Figure 4.7).

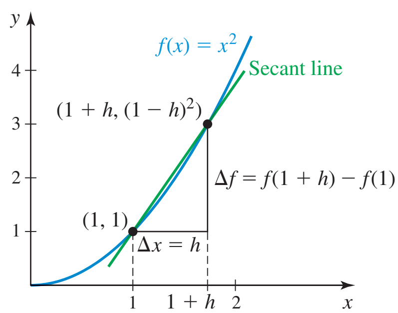


Figure 4.7 The slope of the secant line through $(1, 1)$ and $(1 + h, (1 + h)^2)$ is $\Delta f/\Delta x$.

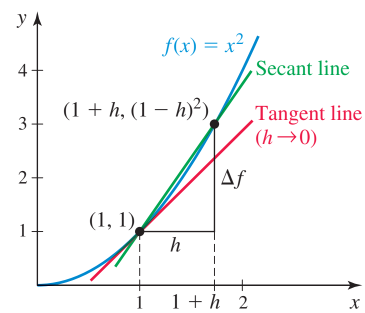


Figure 4.8 Taking the limit $h \to 0$, the secant line converges to the tangent line at $(1, 1)$.

To find the derivative $f'(1)$, we need to take the limit as $h \to 0$ (Figure 4.8):

$$f'(1) = \lim_{h\to 0} \frac{f(1+h)-f(1)}{h} = \lim_{h\to 0}(2+h) = 2$$

Taking the limit as $h \to 0$ means that the point $(1 + h, (1 + h)^2)$ approaches the point $(1, 1)$. [The limit as $h \to 0$ is a two-sided limit; in Figure 4.8, we only drew one point $(1 + h, (1 + h)^2)$ for a specific value of $h > 0$.] As $h \to 0$, the secant line through the points $(1, 1)$ and $(1+h, (1+h)^2)$ converges to the line that touches the graph at $(1, 1)$. As mentioned earlier, the limiting line is called the tangent line. Since $f'(1)$ is the limit of the slope of the secant line as the point $(1 + h, (1 + h)^2)$ approaches $(1, 1)$, we find that $f'(1) = 2$ is the slope of the tangent line at the point $(1, 1)$.

Motivated by this example, we define the tangent line formally:

> **Definition of the Tangent Line** If the derivative of a function f exists at $x = c$, then the tangent line at $x = c$ is the line going through the point $(c, f(c))$ with slope
>
> $$f'(c) = \lim_{h\to 0} \frac{f(c+h)-f(c)}{h}$$

Knowing the derivative at a point (which is the slope of the tangent line at that point) and the coordinates of that point allows us to find the equation of the tangent line by using the point–slope form of a straight line, namely,

$$y - y_0 = m(x - x_0)$$

where (x_0, y_0) is the point and m is the slope. Going back to the function $y = x^2$, we see that the point at $c = x_0 = 1$ has coordinates $(x_0, y_0) = (1, 1)$ while the derivative at $c = 1$ is $m = 2$. The equation of the tangent line is then given by

$$y - 1 = 2(x - 1), \qquad \text{or} \qquad y = 2x - 1$$

(Figure 4.9).

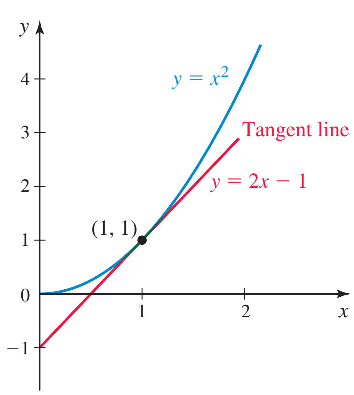


Figure 4.9 The slope of $f(x) = x^2$ at $(1, 1)$ is $m = 2$. The equation of the tangent line at $(1, 1)$ is $y = 2x - 1$.

> **Equation of the Tangent Line** If the derivative of a function f exists at $x = c$, then $f'(c)$ is the slope of the tangent line at the point $(c, f(c))$. The equation of the tangent line is given by
>
> $$y - f(c) = f'(c)(x - c)$$

The geometric interpretation will help us to compute derivatives in the next two examples.

EXAMPLE 1

The Derivative of a Constant Function The graph of $f(x) = a$ is a horizontal line that intersects the y-axis at $(0, a)$ (Figure 4.10). Since the graph is a straight line, the tangent line at x coincides with the graph of $f(x)$ and, therefore, the slope of the tangent line at x is equal to the slope of the straight line. The slope of a horizontal line is 0; we thus expect that $f'(x) = 0$. Indeed, using the formal definition with $f(x) = a$ and $f(x + h) = a$, we find that

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{a - a}{h} = \lim_{h\to 0} \frac{0}{h} = \lim_{h\to 0} 0 = 0.$$

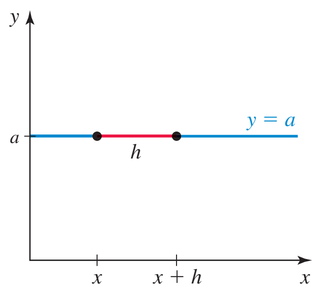


Figure 4.10 The slope of a horizontal line is $m = 0$.

It is important to remember that when we take the limit as $h \to 0$, h approaches 0 (from both sides) but is not equal to 0. Since $h \neq 0$, the expression $0/h = 0$. This property was used in going from $\lim_{h\to 0} \frac{0}{h}$ to $\lim_{h\to 0} 0$. •

EXAMPLE 2

The Derivative of a Linear Function The graph of $f(x) = mx + b$ is a straight line with slope m and y-intercept b (Figure 4.11). The derivative of $f(x)$ is the slope of the tangent line at x. Since the graph is a straight line, the tangent line at x coincides with the graph of $f(x)$ and, therefore, the slope of the tangent line at x is equal to the slope of the straight line. We thus expect that $f'(x) = m$. Using the formal definition, we can confirm this expectation:

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{\overbrace{m(x+h) + b}^{f(x+h)} - \overbrace{(mx+b)}^{f(x)}}{h}$$

$$= \lim_{h\to 0} \frac{mx + mh + b - mx - b}{h} = \lim_{h\to 0} \frac{mh}{h} = \lim_{h\to 0} m = m.$$

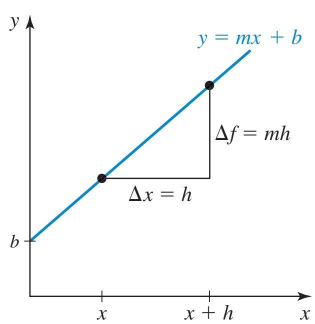


Figure 4.11 The slope of the line $y = mx + b$ is m.

This reasoning yields the following result: If $f(x) = mx + b$, then $f'(x) = m$. This includes the special case of a constant function, for which $m = 0$ (Example 1). •

EXAMPLE 3

Derivative of a power Find the derivative of

$$f(x) = \frac{1}{x} \quad \text{for } x \neq 0.$$

Solution We cannot use the geometric interpretation of the derivative for this Example. We will use the formal definition of the derivative to compute $f'(x)$ (Figure 4.12).

The main algebraic step is the computation of $f(x + h) - f(x)$;

$$f(x+h) - f(x) = \overbrace{\frac{1}{x+h}}^{f(x+h)} - \overbrace{\frac{1}{x}}^{f(x)}$$

$$= \frac{x - (x+h)}{x(x+h)} = \frac{-h}{x(x+h)} \qquad \text{Put over a common denominator.}$$

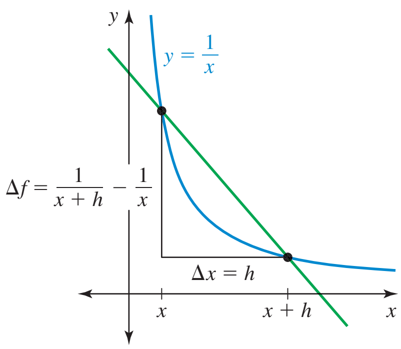


Figure 4.12 The graph of $f(x) = 1/x$ for Example 3.

To compute $f'(x)$, we need to divide both sides of this equation by h and take the limit as $h \to 0$:

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{\frac{-h}{x(x+h)}}{h}$$
$$= \lim_{h\to 0} \left(-\frac{h}{x(x+h)} \frac{1}{h}\right) = \lim_{h\to 0} \left(-\frac{1}{x(x+h)}\right) = -\frac{1}{x^2}.$$

That is, if $f(x) = \frac{1}{x}$, $x \neq 0$, then

$$f'(x) = -\frac{1}{x^2}, \ x \neq 0.$$

Looking back at the first three examples, we see that in order to compute $f'(x)$ from the formal definition of the derivative, we evaluate the limit

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

Since both $\lim_{h\to 0}[f(x+h) - f(x)]$ and $\lim_{h\to 0} h$ are equal to 0, we cannot simply evaluate the limits in the numerator and the denominator separately, because this would result in the undefined expression $0/0$. It is important to simplify the difference quotient before we take the limit. In Sections 4.3 and 4.4, we will learn methods to differentiate other functions.

Section 4.1 Problems

In Problems 1–8, find the derivative at the indicated point from the graph of $y = f(x)$.

1. $f(x) = 5; x = 1$

2. $f(x) = -3x; x = -2$

3. $f(x) = 2x - 3; x = -1$

4. $f(x) = -5x + 1; x = 0$

5. $f(x) = x^2; x = 0$

6. $f(x) = (x+2)^3; x = -2$

7. $f(x) = \cos x; x = 0$

8. $f(x) = \sin x; x = \dfrac{3\pi}{2}$

In Problems 9–16, find c so that $f'(c) = 0$.

9. $f(x) = -x^2 + 1$

10. $f(x) = -x^2 + 4$

11. $f(x) = (x+2)^2$

12. $f(x) = (x+3)^2$

13. $f(x) = x^2 + 6x + 9$

14. $f(x) = x^2 + 4x + 4$

15. $f(x) = \sin\left(\dfrac{\pi}{2}x\right)$

16. $\cos(\pi - x)$

In Problems 17–20, compute $f(c+h) - f(c)$ at the indicated point. Your answers will contain h as an unknown variable.

17. $f(x) = -2x + 1; c = 2$

18. $f(x) = 3x^2; c = 1$

19. $f(x) = \sqrt{x}; c = 4$

20. $f(x) = \dfrac{1}{x}; c = -2$

21. (a) Use the formal definition of the derivative to find the derivative of $y = 2x^2$ at $x = -1$.

(b) Show that the point $(-1, 2)$ is on the graph of $y = 2x^2$, and find the equation of the tangent line at the point $(-1, 2)$.

(c) Graph $y = 2x^2$ and the tangent line at the point $(-1, 2)$ in the same coordinate system.

22. (a) Use the formal definition to find the derivative of $y = x^2 + 1$ at $x = 1$.

(b) Show that the point $(1, 2)$ is on the graph of $y = x^2 + 1$, and find the equation of the tangent line at the point $(1, +2)$.

(c) Graph $y = x^2 + 1$ and the tangent line at the point $(1, 2)$ in the same coordinate system.

23. (a) Use the formal definition to find the derivative of $y = 1 - x^3$ at $x = 2$.

(b) Show that the point $(2, -7)$ is on the graph of $y = 1 - x^3$, and find the equation of the normal line at the point $(2, -7)$.

(c) Graph $y = 1 - x^3$ and the tangent line at the point $(2, -7)$ in the same coordinate system.

24. (a) Use the formal definition to find the derivative of $y = \frac{1}{x}$ at $x = 2$.

(b) Show that the point $(2, \frac{1}{2})$ is on the graph of $y = \frac{1}{x}$, and find the equation of the normal line at the point $(2, \frac{1}{2})$.

(c) Graph $y = \frac{1}{x}$ and the tangent line at the point $(2, \frac{1}{2})$ in the same coordinate system.

25. Use the formal definition to find the derivative of $y = \sqrt{x}$ at $x = 2$.

26. Use the formal definition to find the derivative of $f(x) = \frac{1}{x+1}$ at $x = 0$.

27. Find the equation of the tangent line to the curve $y = 3x^2 + 1$ at the point $(0, 1)$.

28. Find the equation of the tangent line to the curve $y = 2/x$ at the point $(2, 1)$.

29. Find the equation of the tangent line to the curve $y = \sqrt{x}$ at the point $(4, 2)$.

30. Find the equation of the tangent line to the curve $y = x^2 - 3x + 1$ at the point $(2, -1)$.

31. Find the equation of the normal line to the curve $y = -3x^2 + x$ at the point $(-1, -4)$.

32. Find the equation of the normal line to the curve $y = 4/x$ at the point $(-1, -4)$.

33. Find the equation of the normal line to the curve $y = 2x^2 - 1$ at the point $(1, 1)$.

34. Find the equation of the normal line to the curve $y = \sqrt{x} - 1$ at the point $(4, 1)$.

35. The following limit represents the derivative of a function $f(x)$ at the point $x = a$:

$$\lim_{h \to 0} \frac{(a+h)^2 - a^2}{h}.$$

Find $f(x)$.

36. The following limit represents the derivative of a function $f(x)$ at the point $x = a$:

$$\lim_{h \to 0} \frac{4(a+h)^3 - 4a^3}{h}.$$

Find $f(x)$.

37. The following limit represents the derivative of a function $f(x)$ at the point $x = a$:

$$\lim_{h \to 0} \frac{\frac{1}{(2+h)^2+1} - \frac{1}{5}}{h}.$$

Find f and a.

38. The following limit represents the derivative of a function f at the point $(a, f(a))$:

$$\lim_{h \to 0} \frac{\sin(\frac{\pi}{6} + h) - \sin \frac{\pi}{6}}{h}.$$

Find f and a.

4.2 Properties of the Derivative

We introduced the derivative through the example of measuring the instantaneous rate of growth of a population. In general, the derivative gives the instantaneous rate of change of a function. In this section we will describe some ways in which that information can be used.

First we will discuss how the derivative of the function can be interpreted using velocity, rate of growth of a population and the rate of a chemical reaction as examples. Then we will show how the geometric interpretation of the derivative as the slope of the tangent line can be used to identify the conditions under which the derivative does not exist.

4.2.1 Interpreting the Derivative

Velocity. Suppose that you ride your bike on a straight road. Your position (in miles) at time t (in hours) is given by (Figure 4.13)

$$s(t) = -t^3 + 6t^2 \quad \text{for } 0 \le t \le 6$$

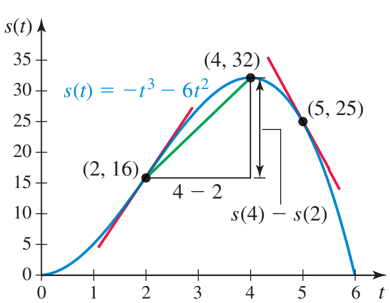


Figure 4.13 The average velocity $\frac{s(4)-s(2)}{4-2}$ is the slope of the secant line through $(2, 16)$ and $(4, 32)$. The velocity at time t is the slope of the tangent line at t: At $t = 2$, the velocity is positive; at $t = 5$, the velocity is negative.

What is the average velocity during a time interval (say $[2, 4]$)? This velocity is defined as the net change in position during the interval, divided by the length of the interval. To compute the average velocity, find the position at time 2 and at time 4, and take the difference of these two quantities. Then divide this difference by the length of the time intervals. Hence, the average velocity is

$$\frac{s(4) - s(2)}{4 - 2} = \frac{32 - 16}{4 - 2} = 8 \text{ mph.} \qquad s(2) = -8 + 24 = 16, \quad s(4) = -64 + 96 = 32$$

We recognize this ratio as the difference quotient

$$\frac{\Delta s}{\Delta t} = \frac{s(t+h) - s(t)}{h}$$

and call the difference quotient $\Delta s/\Delta t$ the **average velocity**, which is an average rate of change.

It is also useful to know how fast you are travelling at any moment in time, t. The **instantaneous velocity** at time t is defined as the limit of $\frac{\Delta s}{\Delta t}$ as $\Delta t \to 0$, or

$$\lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = \lim_{h \to 0} \frac{s(t+h) - s(t)}{h}$$

provided that the limit exists. This quantity is the derivative of $s(t)$ at time t, which we denote by

$$\frac{ds}{dt} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t}$$

Note that $\frac{ds}{dt}$ is an instantaneous rate of change. The instantaneous velocity (or, simply, the velocity) is the slope of the tangent line of the position function $s(t)$, provided that the derivative exists.

Let's look at two points on the graph of $s(t)$, namely, $(2, 16)$ and $(5, 25)$. Looking at the graph we find that the slope of the tangent line is positive at $(2, 16)$ and negative at $(5, 25)$. The velocity is therefore positive at time $t = 2$ and negative at time $t = 5$. At $t = 2$ we are travelling away from our starting point, whereas at $t = 5$ we are travelling back toward our starting point. At these two times, we move in opposite directions.

There is a difference between *velocity* and *speed*. If you had a speedometer on your bike, it would tell you the speed and not the velocity. Speed is the absolute value of velocity; it ignores direction.

Population Growth. At the beginning of this chapter, we described the growth of a population at time t by the continuous function $N(t)$. If the derivative of $N(t)$ exists at time t, we can define the instantaneous growth rate of the population by

$$\text{instantaneous growth rate} = N'(t) = \frac{dN}{dt}$$

We are frequently interested in the **per capita growth rate**. This is the growth rate per individual. This is a useful quantity to focus on because under some conditions the growth rate increases in proportion to the number of organisms in the population; so a population of 200 organisms will grow at twice the rate of a population of 100 organisms. Under these conditions the per capita growth rate would be a constant. The per capita growth rate is found by dividing the growth rate by the current population size. That is,

$$\text{per capita growth rate} = \frac{1}{N(t)}\frac{dN}{dt}$$

We can also write the expression as: $\frac{1}{N}\frac{dN}{dt}$

The Rate of a Chemical Reaction. In Example 5 of Section 1.3, we discussed the reaction rate of an irreversible chemical reaction

$$\mathrm{A} + \mathrm{B} \to \mathrm{AB}$$

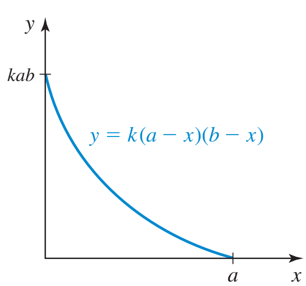


Figure 4.14 The reaction rate for $a \le b$.

Here the reaction rate gives the rate at which molecules of the compound AB are produced by the chemical reaction. It is often (but not always) the case that the reaction rate for such a reaction is proportional to the concentrations of the two reacting chemicals A and B. Let's use the notation, from chemistry, that [A] denotes the concentration of A, [B] denotes the concentration of B, and so on. Then AB is produced at a rate $k[\mathrm{A}][\mathrm{B}]$ where the constant of proportionality k is defined to be the rate constant of this reaction. So:

$$\frac{d}{dt}[\mathrm{AB}] = k[\mathrm{A}][\mathrm{B}]$$

Rate of change of [AB] is equal to rate of the reaction

Since A and B molecules are used up by the reaction we can also write:

$$\frac{d[\mathrm{A}]}{dt} = -k[\mathrm{A}][\mathrm{B}]$$

where the minus sign is needed because the reaction removes rather than produces A. Similarly we can write:

$$\frac{d[\mathrm{B}]}{dt} = -k[\mathrm{A}][\mathrm{B}]$$

Now, as time increases, A and B molecules are removed by the reaction, so the reaction rate will decrease, because [A] and [B] both decrease. Suppose that at time $t = 0$ there is no AB present, and at time t there is concentration $x(t)$ present. The concentration of AB is a function, $x(t)$, that changes with time.

Then, since producing one molecule of AB removes one molecule of A and one of B, if the initial concentration of A is a and the initial concentration of B is b, at time t, $[\mathrm{A}] = a - x$ and $[\mathrm{B}] = b - x$. Then:

$$\frac{dx}{dt} = k(a - x)(b - x) \tag{4.1}$$

Rewriting the equation $\frac{d[\mathrm{AB}]}{dt} = k[\mathrm{A}][\mathrm{B}]$

Equation (4.1) is an example of a **differential equation**—an equation that contains the derivative of a function. We will discuss such equations extensively in Chapters 5 and 8.

Another important type of reaction is when a molecule spontaneously breaks down into two smaller molecules:

$$\mathrm{AB} \to \mathrm{A} + \mathrm{B}$$

The rate of this reaction will also be proportional to the concentration [AB] of the molecule that is breaking down. So if the rate constant for the reaction is c, the two molecules A and B are produced at rates

$$\frac{d[\mathrm{A}]}{dt} = c[\mathrm{AB}], \qquad \frac{d[\mathrm{B}]}{dt} = c[\mathrm{AB}].$$

Since AB molecules are removed by the reaction, [AB] has rate of change $\frac{d[\mathrm{AB}]}{dt} = -c[\mathrm{AB}]$. The concentration of AB can be written as a function $x(t)$ that varies with time, and is given by a differential equation

$$\frac{dx}{dt} = -cx \tag{4.2}$$

The above examples show that the rate of change in a chemical reaction is described by a differential equation. If we wanted to know the amount of the chemical AB present at time t (i.e. function $x(t)$) we would need to solve the differential equation, meaning find a function $x(t)$ whose derivative satisfies (4.1) or (4.2). We discuss how to do this in Chapters 5 and 8.

From this point on, when we say "rate of change," we will always mean "instantaneous rate of change." When we are interested in the average rate of change, we will always state this explicitly.

4.2.2 Differentiability and Continuity

Recall that:

$$f'(c) = \lim_{h\to 0} \frac{f(c+h) - f(c)}{h}.$$

The limit may not always be defined. Generally we may have to use the methods introduced in Chapter 3 to decide whether or not the limit exists. In many cases, however, we can use the geometric interpretation introduced in Section 4.1, that the derivative $f'(c)$ gives the slope of the tangent to $y = f(x)$ at $x = c$, to decide whether or not the derivative exists.

We can find situations in which $f'(c)$ does not exist at one or more values of c.

EXAMPLE 1 **A Function with a "Corner"** Let

$$f(x) = |x| = \begin{cases} x & \text{for } x \geq 0 \\ -x & \text{for } x < 0 \end{cases}$$

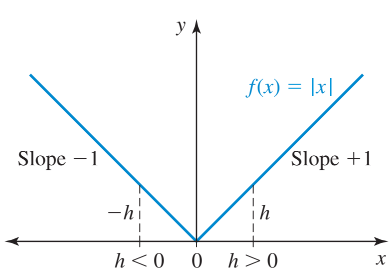


Figure 4.15 f is not differentiable at $x = 0$.

The graph of $f(x)$ is shown in Figure 4.15. Looking at the graph, we realize that there is no tangent line at $x = 0$ and therefore we do not expect that $f'(0)$ exists. We can define the slope of the secant line when we approach 0 from the right and also when we approach 0 from the left; however, the slopes converge to different values in the limit. The former is $+1$, the latter is -1. In this example, we can read off the slopes from the graph. But we can also find the slopes formally by taking appropriate limits. When $h > 0$, $f(h) = |h| = h$ and

$$\lim_{h\to 0+} \frac{f(0+h) - f(0)}{h} = \frac{h-0}{h} = 1$$

When $h < 0$, $f(h) = |h| = -h$ and

$$\lim_{h\to 0-} \frac{f(0+h) - f(0)}{h} = \frac{-h-0}{h} = -1.$$

So

$$f'(0) = \lim_{h\to 0} \frac{f(0+h) - f(0)}{h}$$

does not exist since the left and right limits are different.

At all other points, the derivative exists. We can find the derivative by looking at the graph. We see that

$$f'(x) = \begin{cases} +1 & \text{for } x > 0 \\ -1 & \text{for } x < 0 \end{cases}$$

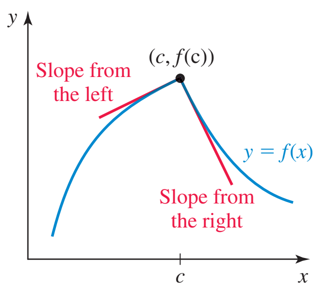


Figure 4.16 $f(x)$ is continuous at $x = c$ but not differentiable at $x = c$: The slopes from the left and the right are not equal.

Example 1 shows that continuity alone is not enough for a function to be differentiable: The function $f(x) = |x|$ is continuous at all values of x, but it is not differentiable at $x = 0$. Any function that has a corner at a point $x = c$ will be continuous but not differentiable at that point (Figure 4.16).

However, if a function is differentiable, it is also continuous. We say that continuity is a necessary, but not a sufficient, condition for differentiability. This result is important enough that we will formulate it as a theorem and prove it:

Theorem If f is differentiable at $x = c$, then f is also continuous at $x = c$.

Proof Since f is differentiable at $x = c$, we know that the limit

$$\lim_{x\to c} \frac{f(x) - f(c)}{x - c} \tag{4.3}$$

exists and is equal to $f'(c)$. To show that f is continuous at $x = c$, we must show that

$$\lim_{x\to c}[f(x) - f(c)] = 0 \qquad \text{equivalently: } \lim_{x\to c} f(x) = f(c),$$

First, note that f is defined at $x = c$, otherwise, we could not have computed the difference quotient $\frac{f(x)-f(c)}{x-c}$. Now,

$$\lim_{x\to c}[f(x) - f(c)] = \lim_{x\to c}\left[\left(\frac{f(x) - f(c)}{x - c}\right)(x - c)\right]$$

Given that $\lim_{x\to c} \frac{f(x)-f(c)}{x-c}$ exists and is equal to $f'(c)$, and that $\lim_{x\to c}(x - c)$ exists (it is equal to 0), we can apply the product rule for limits and find that

$$\lim_{x\to c}\left[\left(\frac{f(x) - f(c)}{x - c}\right)(x - c)\right] = \lim_{x\to c} \frac{f(x) - f(c)}{x - c} \cdot \lim_{x\to c}(x - c) = f'(c) \cdot 0 = 0.$$

This set of equations shows that $\lim_{x\to c}[f(x) - f(c)] = 0$ and consequently that f is continuous at $x = c$.

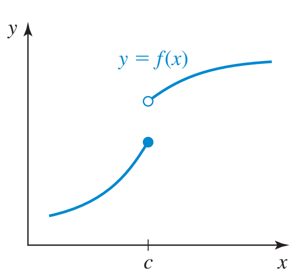


Figure 4.17 The function $y = f(x)$ is not differentiable at $x = c$.

It follows from the preceding theorem that if a function f is not continuous at $x = c$, then f is not differentiable at $x = c$. The function $y = f(x)$ in Figure 4.17 is discontinuous at $x = c$; so we cannot draw a tangent line there either.

Functions can have vertical tangent lines, but since the slope of a vertical line is not defined, the function would not be differentiable at any point where the tangent line is vertical. This situation is illustrated in the next example.

EXAMPLE 2 **Vertical Tangent Line** Show that $f(x) = x^{1/3}$ is not differentiable at $x = 0$.

Solution We see from the graph of $f(x)$ in Figure 4.18 that $f(x)$ is continuous at $x = 0$. Using the formal definition, we find that

$$f'(0) = \lim_{h\to 0} \frac{f(h) - f(0)}{h} = \lim_{h\to 0} \frac{h^{1/3} - 0}{h}$$

$$= \lim_{h\to 0} \frac{1}{h^{2/3}} = \infty \qquad \text{does not exist}$$

Since the limit does not exist, $f(x)$ is not differentiable at $x = 0$. We see from Figure 4.18 that the tangent line at $x = 0$ is vertical. •

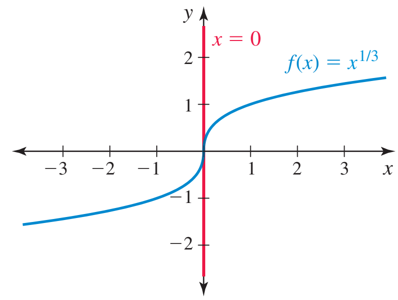


Figure 4.18 The function $f(x) = x^{1/3}$ has a vertical tangent line at $x = 0$. It is therefore not differentiable at $x = 0$.

Section 4.2 Problems

4.2.1

In the following examples quantities x and y are given. Interpret the role of change dy/dx in words.

1. y is the length of a fish and x is the age of the fish.

2. y is the number of cells in a petri dish, x is time.

3. y is the heart rate of a mammal, x is the mammal's body mass.

4. y is the amount of chemical produced by a chemical reaction, x is the amount of catalyst added.

5. y is the number of cars leaving a freeway in one minute, x is the number of cars on the freeway.

6. y is the bite strength of a mammal, x is its body mass.

7. y is the body mass of a mammal, x is its age.

8. y is the temperature of the Pacific Ocean at Santa Monica beach, x is the time of day.

9. y is the height of water in a rain collecting column, x is time.

10. y is the energy required by a tree in one day, x is its height.

11. Velocity A car moves along a straight road. Its location at time t is given by

$$s(t) = 10t^2, \ 0 \le t \le 2$$

where t is measured in hours and $s(t)$ is measured in kilometers.

(a) Graph $s(t)$ for $0 \le t \le 2$.

(b) Find the average velocity of the car between $t = 0$ and $t = 2$. Illustrate the average velocity on the graph of $s(t)$.

(c) Use calculus to find the instantaneous velocity of the car at $t = 1$. Illustrate the instantaneous velocity on the graph of $s(t)$.

12. Velocity A train moves along a straight line. Its location at time t is given by

$$s(t) = \frac{100}{t}, \qquad 1 \le t \le 5$$

where t is measured in hours and $s(t)$ is measured in kilometers.

(a) Graph $s(t)$ for $1 \le t \le 5$.

(b) Find the average velocity of the train between $t = 1$ and $t = 5$. Where on the graph of $s(t)$ can you find the average velocity?

(c) Use calculus to find the instantaneous velocity of the train at $t = 2$. Where on the graph of $s(t)$ can you find the instantaneous velocity? What is the speed of the train at $t = 2$?

13. Velocity If $s(t)$ denotes the position of an object that moves along a straight line, then $\Delta s/\Delta t$, called the average velocity, is the average rate of change of $s(t)$, and $v(t) = ds/dt$, called the (instantaneous) velocity, is the instantaneous rate of change of $s(t)$. The speed of the object is the absolute value of the velocity, $|v(t)|$.

Suppose now that a car moves along a straight road. The location at time t is given by

$$s(t) = \frac{160}{3}t^2, \qquad 0 \le t \le 1$$

where t is measured in hours and $s(t)$ is measured in kilometers.

(a) Where is the car at $t = 3/4$, and where is it at $t = 1$?

(b) Find the average velocity of the car between $t = 3/4$ and $t = 1$.

(c) Find the velocity and the speed of the car at $t = 3/4$.

14. Velocity Suppose a particle moves along a straight line. The position at time t is given by

$$s(t) = 3t - t^2, \qquad t \ge 0$$

where t is measured in seconds and $s(t)$ is measured in meters.

(a) Graph $s(t)$ for $t \ge 0$.

(b) Use the graph in (a) to answer the following questions:

(i) Where is the particle at time 0?

(ii) Is there another time at which the particle visits the location where it was at time 0?

(iii) How far to the right on the straight line does the particle travel?

(iv) How far to the left on the straight line does the particle travel?

(v) Where is the velocity positive? negative? equal to 0?

(c) Find the velocity of the particle.

(d) When is the velocity of the particle equal to 1 m/s?

15. Equation of a Projectile A projectile that is launched directly upward will, unless it is caught by the wind, fall back directly to Earth. If the projectile is shot upward at speed v initially then its height at time t is given by:

$$s(t) = vt - \frac{1}{2}gt^2$$

where $g = 9.81\ \mathrm{m\,s^{-2}}$ is the deceleration of the projectile due to gravity.

(a) At some time $t > 0$, the projectile will return to Earth; that is, $s(t) = 0$. Find this time. Your answer will depend on v and g.

(b) Show, using the definition of the derivative, that at time $t = 0$ (i.e., when the projectile is launched) its velocity is v.

(c) Let $v = 10\ \mathrm{m\,s^{-1}}$. Plot $s(t)$ against t.

(d) Use the graph in (c) to answer the following questions.

(i) What is the maximum height reached by the projectile?

(ii) Over what time interval(s) does the projectile have positive (upward) velocity?

(iii) Over what time interval(s) does the projectile have negative (downward) velocity?

(iv) What is the velocity of the particle when it reaches its maximum height, given by your answer to (i)?

16. **Population Growth** Assume that $N(t)$ denotes the size of a population at time t, and that in some conditions $N(t)$ satisfies the differential equation

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

where r is a constant.

(a) Find the per capita growth rate.

(b) Assume that $r < 0$ and that $N(0) = 20$. Is the population size at time 1 greater than 20 or less than 20? Explain your answer.

17. **Population Growth** $N(t)$ is the size of the population at time t. The population is modeled using a differential equation

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

where r is a constant.

(a) According to the differential equation what is the growth rate of the population?

(b) Assume that $r > 0$. Explain why the population size at time $t = 1$ will be larger than the population size at time $t = 0$.

(c) If $r > 0$, will the growth rate at time $t = 1$ be larger or smaller than the growth rate at time $t = 0$?

(d) Answer (c) again but comparing the per capita growth rates at times $t = 0$ and $t = 1$.

18. **Chemical Reaction** Consider the chemical reaction

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{AB}$$

Assume k is the rate constant for the reaction.

(a) Explain why, if [A] is the amount of the chemical A and [B] is the amount of the chemical B present, then

$$\frac{d[\mathrm{A}]}{dt} = -k[\mathrm{A}][\mathrm{B}]$$

(b) What is the corresponding differential equation describing the rate of change of [B]?

19. **Chemical Reaction** Consider the chemical reaction

$$\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{AB}$$

If $x(t)$ denotes the concentration of AB at time t, and k is the rate constant for of the reaction, explain why:

$$\frac{dx}{dt} = k(a - x)(b - x)$$

where k is a positive constant and a and b denote the concentrations of A and B, respectively, at time 0.

20. **Chemical Reaction** Chemicals A and X react through an autocatalytic reaction:

$$\mathrm{A} + \mathrm{X} \longrightarrow 2\mathrm{X}$$

(a) If the rate constant for the reaction is k, and the amounts of A and X present are denoted by [A] and [X], then explain why

$$\frac{d[\mathrm{A}]}{dt} = -k[\mathrm{A}][\mathrm{X}]$$

(b) Find a differential equation that describes the rate of change of the amount of [X].

21. **Chemical Reaction** Chemical A spontaneously decomposes into chemicals B and C: $\mathrm{A} \rightarrow \mathrm{B} + \mathrm{C}$. The rate constant for this reaction is k.

(a) Explain why, if the amounts of A, B, and C present are denoted by [A], [B], and [C] respectively, then [B] obeys the differential equation

$$\frac{d[\mathrm{B}]}{dt} = k[\mathrm{A}]$$

(b) What differential equations represent the rate of change of [C] and [A]?

(c) Suppose that at time $t = 0$ the initial amount of A present is a, and that there is no B or C present. Explain why

$$[\mathrm{A}] = a - [\mathrm{B}].$$

22. **Chemical Reaction** Molecules of A and B react to produce products C and D

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{C} + \mathrm{D}$$

The rate constant for this reaction is k. We denote by [A] the amount of A present, and so on.

(a) Explain why the amount of C present obeys the differential equation

$$\frac{d[\mathrm{C}]}{dt} = k[\mathrm{A}][\mathrm{B}]$$

(b) Find similar differential equations for [A], [B], and [D].

4.2.2

23. Which of the following statements is true?

(A) If $f(x)$ is continuous, then $f(x)$ is differentiable.

(B) If $f(x)$ is differentiable, then $f(x)$ is continuous.

24. Sketch the graph of a function that is continuous at all points in its domain and differentiable in the domain except at one point.

25. If $f(x)$ is differentiable for all $x \in \mathbf{R}$ except at $x = c$, is it true that $f(x)$ must be continuous at $x = c$? Justify your answer.

In Problems 26–39, graph each function and, on the basis of the graph, guess where the function is not differentiable. (Assume the largest possible domain.)

26. $y = |x - 2|$

27. $y = -|x + 5|$

28. $y = |x + 2| - 1$

29. $y = 2 - |x - 3|$

30. $y = \dfrac{1}{2 + x}$

31. $y = \dfrac{1}{x - 3}$

32. $y = \dfrac{x-1}{x+1}$

33. $y = \dfrac{3-x}{3+x}$

34. $y = |x^2 - 3|$

35. $y = |2x^2 - 1|$

36. $f(x) = \begin{cases} x & \text{for } x \le 0 \\ x+1 & \text{for } x > 0 \end{cases}$

37. $f(x) = \begin{cases} 2x & \text{for } x \le 1 \\ x+1 & \text{for } x > 1 \end{cases}$

38. $f(x) = \begin{cases} x^2 & \text{for } x \le -1 \\ 2-x^2 & \text{for } x > -1 \end{cases}$

39. $f(x) = \begin{cases} x^4+1 & \text{for } x \le 0 \\ e^{-x} & \text{for } x > 0 \end{cases}$

40. Suppose the function $f(x)$ is defined by different expressions in two different intervals; that is, $f(x) = f_1(x)$ for $x \le a$ and $f(x) = f_2(x)$ for $x > a$. Assume that $f_1(x)$ is continuous and differentiable for $x < a$ and that $f_2(x)$ is continuous and differentiable for $x > a$. Sketch graphs of $f(x)$ for the following three cases:

(a) $f(x)$ is continuous and differentiable at $x = a$.

(b) $f(x)$ is continuous, but not differentiable, at $x = a$.

(c) $f(x)$ is neither continuous nor differentiable at $x = a$.

4.3 The Power Rule, the Basic Rules of Differentiation, and the Derivatives of Polynomials

In this section, we will begin a systematic treatment of the computation of derivatives. Knowing how to differentiate is an essential skill for this book. Although computer software is now available to compute derivative functions, being able to calculate the derivatives without software is important for applying differentiation to build models or graph functions as we shall see in Chapter 5.

The power rule is the most fundamental of the differentiation rules. It allows us to compute the derivative of a function of the form $y = x^a$, where a is any constant.

> **Power Rule** Let a be a constant; then
> $$\frac{d}{dx}(x^a) = ax^{a-1}$$

In this section we will prove the power rule when a is a non-negative integer (that is, when $a = 0, 1, 2, \ldots$). We already met the special cases of the power rule, $a = 0$ and $a = 1$, in Section 4.1.

$$\frac{d}{dx}(1) = 0 \quad \text{and} \quad \frac{d}{dx}(x) = 1$$

$x^0 = 1, x^1 = x.$

We prove the power rule first for $n = 2$—that is, for $f(x) = x^2$ (Figure 4.19). In Section 4.1, we computed the derivative of $y = x^2$ at $x = 1$. Now we will compute the difference quotient $\frac{\Delta f}{\Delta x}$ at any arbitrary x:

$$\frac{\Delta f}{\Delta x} = \frac{f(x+h) - f(x)}{h} = \frac{(x+h)^2 - x^2}{h}$$

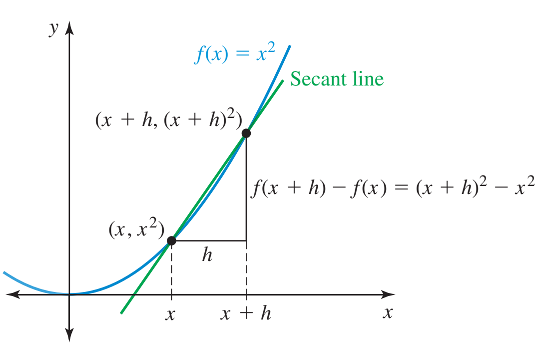


Figure 4.19 The slope of the secant line through (x, x^2) and $(x+h, (x+h)^2)$ is $\frac{(x+h)^2 - x^2}{h}$.

Using the expansion $(x+h)^2 = x^2 + 2xh + h^2$, we find that

$$\frac{\Delta f}{\Delta x} = \frac{x^2 + 2xh + h^2 - x^2}{h} = \frac{2xh + h^2}{h} = 2x + h \tag{4.4}$$

after canceling h in both the numerator and the denominator. To find the derivative, we need to let $h \to 0$:

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{h \to 0}(2x + h) = 2x.$$

This sequence of steps proves the power rule for $n = 2$. The proof of the rule for other positive integers of n is conceptually no different from the case $n = 2$, but it gets algebraically much more involved. For general n, we need the expansion of $(x+h)^n$, given by the **binomial theorem**, which we will state but not prove.

Binomial Theorem If n is a positive integer, then

$$\begin{aligned}(x+y)^n = x^n + nx^{n-1}y &+ \frac{n(n-1)}{2\cdot 1}x^{n-2}y^2 \\ &+ \frac{n(n-1)(n-2)}{3\cdot 2\cdot 1}x^{n-3}y^3 \\ &+ \cdots + \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 2\cdot 1}x^{n-k}y^k \\ &+ \cdots + nxy^{n-1} + y^n\end{aligned}$$

The expansion of $(x+y)^n$ is thus a sum of terms of the form

$$C_{n,k}x^{n-k}y^k, \quad k = 0, 1, \ldots, n$$

where $C_{n,k}$ is a coefficient that depends on n and k. The exact form of the coefficients $C_{n,k}$ will not be important in the proof of the power rule, except for the two terms $C_{n,0} = 1$ and $C_{n,1} = n$, which are the coefficients for x^n and $x^{n-1}y$, respectively.

Proof of the Power Rule We use the binomial theorem to expand and then compute the numerator of the difference quotient:

$$\begin{aligned}\Delta f &= f(x+h) - f(x) = (x+h)^n - x^n \\ &= \left[x^n + nx^{n-1}h + \frac{n(n-1)}{2}x^{n-2}h^2 + \cdots + nxh^{n-1} + h^n\right] - x^n\end{aligned}$$

The x^n terms cancel. We can then factor h out of the remaining terms and find that

$$f(x+h) - f(x) = h\left[nx^{n-1} + \frac{n(n-1)}{2}x^{n-2}h + \cdots + nxh^{n-2} + h^{n-1}\right]$$

When we divide by h and let $h \to 0$, we obtain

$$\begin{aligned}f'(x) &= \lim_{h\to 0}\frac{f(x+h) - f(x)}{h} \\ &= \lim_{h\to 0}\left[nx^{n-1} + n(n-1)^{n-2}h + \cdots + nxh^{n-2} + h^{n-1}\right]\end{aligned}$$

All terms except for the first have h as a factor and thus tend to 0 as $h \to 0$. (The first term does not depend on h.) We find that

$$f'(x) = nx^{n-1}$$

which proves the power rule. •

EXAMPLE 1 We apply the power rule to various functions and take the opportunity to practice the different notations.

(a) If $f(x) = x^6$, then $f'(x) = 6x^5$.

(b) If $f(x) = x^{300}$, then $f'(x) = 300x^{299}$.

(c) If $g(t) = t^5$, then $\frac{d}{dt}g(t) = 5t^4$.

(d) If $z = s^3$, then $\frac{dz}{ds} = 3s^2$.

(e) If $x = y^4$, then $\frac{dx}{dy} = 4y^3$. •

Example 1 illustrates the importance of knowing how the variables depend on each other (Figure 4.20). If $y = f(x)$, we call x the **independent variable** and y the **dependent variable**, because y depends on the variable x. For instance, in (a), y is a function of x; x is the independent variable, and y is the dependent variable. In (e),

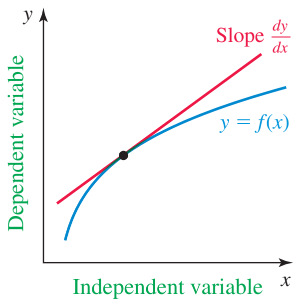


Figure 4.20 If $y = f(x)$, then x is the independent variable and y is the dependent variable.

by contrast, x is a function of y; y is now the independent variable, and x the dependent variable. The Leibniz notation $\frac{dy}{dx}$ emphasizes this dependence. When we write $\frac{dy}{dx}$, we consider y to be a function of x (i.e., y is the dependent variable, and x is the independent variable) and differentiate y with respect to x.

Since polynomials and rational functions are built up by the basic operations of addition, subtraction, multiplication, and division operating on power functions of the form $y = x^n$, $n = 0, 1, 2, \ldots$, we need differentiation rules for such operations.

Sum and multiplication rules Suppose a is a constant and $f(x)$ and $g(x)$ are differentiable at x. Then the following relationships hold:

1. $\frac{d}{dx}[af(x)] = a\frac{d}{dx}f(x)$
2. $\frac{d}{dx}[f(x) + g(x)] = \frac{d}{dx}f(x) + \frac{d}{dx}g(x)$

Rule 1 says that a constant factor can be pulled out of the derivative expression; Rule 2 says that the derivative of a sum of two functions is equal to the sum of the derivatives of the functions. Similarly, since $f(x) - g(x) = f(x) + (-1)g(x)$, the derivative of a difference of functions is the difference of the derivatives:

$$\frac{d}{dx}[f(x) - g(x)] = \frac{d}{dx}[f(x) + (-1)g(x)] = \frac{d}{dx}f(x) + \frac{d}{dx}[(-1)g(x)]$$

Using Rule 1 on the rightmost term, we find that $\frac{d}{dx}[(-1)g(x)] = (-1)\frac{d}{dx}g(x)$. Therefore,

$$\frac{d}{dx}[f(x) - g(x)] = \frac{d}{dx}f(x) - \frac{d}{dx}g(x)$$

We can use the definition of the derivative to prove these rules. We will prove the first rule here. You will be asked to prove the second rule in Problem 83.

Proof If $f(a)$ is a constant and $f(x)$ is differentiable at x, then if we define a function $p(x) = af(x)$

$$\frac{p(x+h) - p(x)}{h} = \frac{af(x+h) - af(x)}{h} = a\left(\frac{f(x+h) - f(x)}{h}\right)$$

So using the limit laws, if $f'(x) = \lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h}\right)$ exists, then $\lim_{h\to 0}\left(\frac{p(x+h)-p(x)}{h}\right)$ exists and is equal to

$$\lim_{h\to 0}\left(\frac{p(x+h) - p(x)}{h}\right) = af'(x).$$

So $p(x)$ is differentiable at x and:

$$\frac{d}{dx}(af(x)) = p'(x) = a\frac{d}{dx}f(x)$$

•

Along with the Power Rule Rules 1 and 2 allow us to differentiate polynomials, as illustrated in the next three examples.

EXAMPLE 2 Differentiate $y = 2x^4 - 3x^3 + x - 7$.

Solution

$$\frac{d}{dx}(2x^4 - 3x^3 + x - 7) = \frac{d}{dx}(2x^4) - \frac{d}{dx}(3x^3) + \frac{d}{dx}x - \frac{d}{dx}(7) \quad \text{Rule 2}$$

$$= 2\frac{d}{dx}x^4 - 3\frac{d}{dx}x^3 + \frac{d}{dx}x - \frac{d}{dx}(7) \quad \text{Rule 1}$$

$$= 2(4x^3) - 3(3x^2) + 1 - 0 = 8x^3 - 9x^2 + 1$$

The derivative of any constant is 0.

•

EXAMPLE 3

(a) $\frac{d}{dx}(-5x^7 + 2x^3 - 10) = -35x^6 + 6x^2$

(b) $\frac{d}{dt}(t^3 - 8t^2 - 3t) = 3t^2 - 16t + 3$

(c) Suppose that n is a positive integer and a is a constant. Then $\frac{d}{ds}(as^n) = ans^{n-1}$.

(d) $\frac{d}{dN}(\ln 5 + N \ln 7) = \ln 7$

(e) $\frac{d}{dr}(r^2 \sin\frac{\pi}{4} - r^3 \cos\frac{\pi}{12} + \sin\frac{\pi}{6}) = 2r \sin\frac{\pi}{4} - 3r^2 \cos\frac{\pi}{12}$ •

In the previous section, we related the derivative to the slope of the tangent line; the next example uses this interpretation.

EXAMPLE 4

Tangent and Normal Lines If $f(x) = 2x^3 - 3x + 1$, find the equations of the tangent and normal lines at $(-1, 2)$.

Solution The slope of the tangent line at $(-1, 2)$ is $f'(-1)$. We begin by calculating this derivative. First we find:

$$f'(x) = 6x^2 - 3$$

Evaluating $f'(x)$ at $x = -1$, we get $f'(-1) = 6(-1)^2 - 3 = 3$. Therefore, the equation of the tangent line at $(-1, 2)$ is

$$y - 2 = 3(x - (-1)), \qquad \text{or} \qquad y = 3x + 5$$

To find the equation of the normal line, recall that the normal line is perpendicular to the tangent line; hence, the slope m of the normal line is given by

$$m = -\frac{1}{f'(-1)} = -\frac{1}{3}$$

The normal line goes through the point $(-1, 2)$ as well. The equation of the normal line is therefore

$$y - 2 = -\frac{1}{3}(x - (-1)), \qquad \text{or} \qquad y = -\frac{1}{3}x + \frac{5}{3}$$

The graph of $f(x)$, including the tangent and normal lines at $(-1, 2)$, is shown in Figure 4.21. •

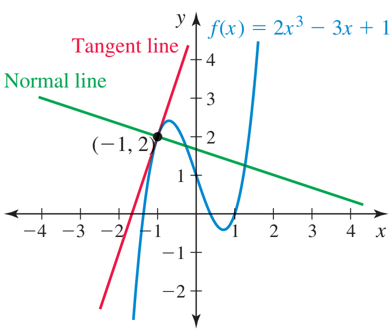


Figure 4.21 The graph of $f(x) = 2x^3 - 3x + 1$, together with the tangent and normal lines at $(-1, 2)$.

Look again at the last example: When we computed $f'(-1)$, we *first* computed $f'(x)$; the *second* step was to evaluate $f'(x)$ at $x = -1$. It makes no sense to plug -1 into $f(x)$ and then differentiate the result. The notation $f'(-1)$ means that we evaluate the function $f'(x)$ at $x = -1$.

Section 4.3 Problems

Differentiate the functions given in Problems 1–22 with respect to the independent variable.

1. $f(x) = 4x^3 - 7x + 1$

2. $f(x) = -3x^4 + 5x^2$

3. $f(x) = -2x^5 + 7x - 4$

4. $f(x) = -3x^4 + 6x^2 - 2$

5. $f(x) = 3 - 4x - 5x^2$

6. $f(x) = -1 + 3x^2 - 2x^4$

7. $g(s) = 5s^7 + 2s^3 - 5s$

8. $g(s) = 3 - 4s^2 - 4s^3$

9. $h(t) = -\frac{1}{3}t^4 + 4t$

10. $h(t) = \frac{1}{2}t^2 - 3t + 2$

11. $f(x) = x^2 \sin\frac{\pi}{3} + \tan\frac{\pi}{4}$

12. $f(x) = 2x^3 \cos\frac{\pi}{3} + \cos\frac{\pi}{6}$

13. $f(x) = -3x^4 \tan\frac{\pi}{6} - \cot\frac{\pi}{6}$

14. $f(x) = x^2 \sec\frac{\pi}{6} + 3x \sec\frac{\pi}{4}$

15. $f(t) = t^3 e^{-2} + t + e^{-1}$

16. $f(x) = \frac{1}{2}x^2 e^3 - x^4$

17. $f(s) = s^3 e^3 + 3e$

18. $f(x) = \frac{x}{e} + e^2 x + e$

19. $f(x) = 20x^3 - 4x^6 + 9x^8$

20. $f(x) = \frac{x^3}{15} - \frac{x^4}{20} + \frac{2}{15}$

21. $f(x) = \pi x^3 - \frac{1}{\pi} + \frac{x}{\pi}$

22. $f(x) = \pi^3 x - x^2 \pi$

23. Differentiate

$$f(x) = ax^3$$

with respect to x. Assume that a is a constant.

24. Differentiate

$$f(x) = x^3 + a$$

with respect to x. Assume that a is a constant.

25. Differentiate
$$f(x) = ax^2 - 2a$$
with respect to x. Assume that a is a constant.

26. Differentiate
$$f(x) = a^2x^4 - 2ax^2$$
with respect to x. Assume that a is a constant.

27. Differentiate
$$h(s) = rs^2 - r$$
with respect to s. Assume that r is a constant.

28. Differentiate
$$f(r) = rs^2 - r$$
with respect to r. Assume that s is a constant.

29. Differentiate
$$f(x) = rs^2x^3 - rx + s$$
with respect to x. Assume that r and s are constants.

30. Differentiate
$$f(x) = \frac{r+x}{rs^2} - rsx + (r+s)x - rs$$
with respect to x. Assume that r and s are nonzero constants.

31. Differentiate
$$f(N) = (b-1)N^4 - \frac{N^2}{b}$$
with respect to N. Assume that b is a nonzero constant.

32. Differentiate
$$f(N) = \frac{bN^2 + N}{K + b}$$
with respect to N. Assume that b and K are positive constants.

33. Differentiate
$$g(t) = a^3t - at^3$$
with respect to t. Assume that a is a constant.

34. Differentiate
$$h(s) = a^4s^2 - as^4 + \frac{s^2}{a^4}$$
with respect to s. Assume that a is a positive constant.

35. Differentiate
$$V(t) = V_0(1 + \gamma t)$$
with respect to t. Assume that V_0 and γ are positive constants.

36. Differentiate
$$p(T) = \frac{NkT}{V}$$
with respect to T. Assume that N, k, and V are positive constants.

37. Differentiate
$$g(N) = N\left(1 - \frac{N}{K}\right)$$
with respect to N. Assume that K is a positive constant.

38. Differentiate
$$g(N) = rN\left(1 - \frac{N}{K}\right)$$
with respect to N. Assume that K and r are positive constants.

39. Differentiate
$$g(N) = rN^2\left(1 - \frac{N}{K}\right)$$
with respect to N. Assume that K and r are positive constants.

40. Differentiate
$$g(N) = rN(a - N)\left(1 - \frac{N}{K}\right)$$
with respect to N. Assume that r, a, and K are positive constants.

41. Differentiate
$$R(T) = \frac{2\pi^5}{15}\frac{k^4}{c^2h^3}T^4$$
with respect to T. Assume that k, c, and h are positive constants.

In Problems 42–48, find the tangent line, in standard form, to $y = f(x)$ at the indicated point.

42. $y = 3x^2 - 4x + 7$, at $x = 2$

43. $y = 7x^3 + 2x - 1$, at $x = -3$

44. $y = -2x^3 - 3x + 1$, at $x = 1$

45. $y = 2x^4 - 5x$, at $x = 1$

46. $y = -x^3 - 2x^2$, at $x = 0$

47. $y = \frac{1}{\sqrt{2}}x^2 - \sqrt{2}$, at $x = 4$

48. $y = 3\pi x^5 - \frac{\pi}{2}x^3$, at $x = -1$

In Problems 49–54, find the normal line, in standard form, to $y = f(x)$ at the indicated point.

49. $y = 2 + x^2$, at $x = -1$

50. $y = 1 - 3x^2$, at $x = -2$

51. $y = \sqrt{3}x^4 - 2\sqrt{3}x^2$, at $x = -\sqrt{3}$

52. $y = -2x^2 - x$, at $x = 0$

53. $y = x^3 - 3$, at $x = 1$

54. $y = 1 - \pi x^2$, at $x = -1$

55. Find the tangent line to
$$f(x) = ax^2$$
at $x = 1$. Assume that a is a positive constant.

56. Find the tangent line to
$$f(x) = ax^3 - 2ax$$
at $x = -1$. Assume that a is a positive constant.

57. Find the tangent line to
$$f(x) = \frac{ax^2}{a^2 + 2}$$
at $x = 2$. Assume that a is a positive constant.

58. Find the tangent line to
$$f(x) = \frac{x^2 + x}{a + 1}$$
at $x = a$. Assume that a is a positive constant.

59. Find the normal line to
$$f(x) = ax^3$$
at $x = -1$. Assume that a is a positive constant.

60. Find the normal line to
$$f(x) = ax^2 - 3ax$$
at $x = 2$. Assume that a is a positive constant.

61. Find the normal line to
$$f(x) = \frac{ax^2}{a + 1}$$
at $x = 2$. Assume that a is a positive constant.

62. Find the normal line to
$$f(x) = \frac{x^3}{a+1}$$
at $x = 2a$. Assume that a is a positive constant.

In Problems 63–70, find the coordinates of all of the points of the graph of $y = f(x)$ that have horizontal tangents.

63. $f(x) = x^2$

64. $f(x) = 2 - x^2$

65. $f(x) = 3x - x^2$

66. $f(x) = 4x + 2x^2$

67. $f(x) = 3x^3 - x^2$

68. $f(x) = -4x^4 + x^3$

69. $f(x) = \frac{1}{2}x^4 - \frac{7}{3}x^3 - 2x^2$

70. $f(x) = 3x^5 - \frac{3}{2}x^4$

71. Find a point on the curve
$$y = 4 - x^2$$
whose tangent line is parallel to the line $y = 2$. Is there more than one such point? If so, find all other points with this property.

72. Find a point on the curve
$$y = (4 - x)^2$$
whose tangent line is parallel to the line $y = -3$. Is there more than one such point? If so, find all other points with this property.

73. Find a point on the curve
$$y = 2x^2 - \frac{1}{2}$$
whose tangent line is parallel to the line $y = x$. Is there more than one such point? If so, find all other points with this property.

74. Find a point on the curve
$$y = 1 - 3x^3$$
whose tangent line is parallel to the line $y = -x$. Is there more than one such point? If so, find all other points with this property.

75. Find a point on the curve
$$y = x^3 + 2x + 2$$
whose tangent line is parallel to the line $3x - y = 2$. Is there more than one such point? If so, find all other points with this property.

76. Find a point on the curve
$$y = 2x^3 - 4x + 1$$
whose tangent line is parallel to the line $y - 2x = 1$. Is there more than one such point? If so, find all other points with this property.

77. Show that the tangent line to the curve
$$y = x^2$$
at the point $(1, 1)$ passes through the point $(0, -1)$.

78. Find all tangent lines to the curve
$$y = x^2$$
that pass through the point $(0, -1)$.

79. Find all tangent lines to the curve
$$y = x^2$$
that pass through the point $(0, -a^2)$, where a is a positive number.

80. How many tangent lines to the curve
$$y = x^2 + 2x$$
pass through the point $(-\frac{1}{2}, -3)$?

81. Suppose that $P(x)$ is a polynomial of degree 4. Is $P'(x)$ a polynomial as well? If yes, what is its degree?

82. Suppose that $P(x)$ is a polynomial of degree k. Is $P'(x)$ a polynomial as well? If yes, what is its degree?

83. Prove the sum rule. That is, if $f(x)$ and $g(x)$ are both differentiable at x, show that $f(x) + g(x)$ is also differentiable at x and that
$$\frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx}f(x) + \frac{d}{dx}g(x)$$
Hint: Start by defining a function $p(x) = f(x) + g(x)$ and write the difference quotients for $p(x)$ in terms of difference quotients involving $f(x)$ and $g(x)$.

84. **Change in Heart Rate with Age** ***Tanaka et al. (2001)*** investigated how maximal heart rate depends on age. They found that if age x is given in years, then the maximum heart rate of a healthy adult can be predicted by the following formula:
$$H(x) = 208 - 0.7x$$
where $H(x)$ is the maximum number of heart beats in one minute. The data from Tanaka et al. suggests that each additional year of age decreases $H(x)$ by the same amount.

(a) Explain in words what dH/dx represents.

(b) Show that dH/dx is a constant.

85. **Estimating the Age of a Fetus** Ultrasound is often used to make images of developing fetuses. In particular, by measuring the size of a fetus ultrasound technicians can estimate its age and then predict its birthdate. To do this requires formulas for fetus size as a function of age. Verburg et al. (2008) fit data from over 6000 fetal ultrasounds. They measured the femur length, L, (in mm) as a function of the fetus age, t, (in weeks) and found the following formula:
$$L = -37.50 + 3.71t - 6.33 \times 10^{-4}t^3.$$
Calculate the rate of growth, dL/dt, at $t = 15$, 20, and 30 weeks. Does the rate of growth of the fetus increase or decrease as it ages?

86. **Physical Properties of a Cell** To measure the physical properties of cells, a piezoelectric probe is used. The force applied by the probe is compared against how much the cell deforms. If F is the force applied by the probe, and w is the distance it moves into the cell, then the stiffness of the cell can be calculated from the rate of change, dF/dw. Zhang et al. found that if F is measured in μN and w in μm then for a zebrafish embryo:
$$F = 3 \times 10^{-4}w^3 - 4.4 \times 10^{-3}w^2 + 3.93w + 0.221$$
(a) Calculate dF/dw for this sample.

(b) Stiffer cells have larger values of dF/dw when $w = 0$. Later in embryo development Zhang et al. measure:
$$F = 6 \times 10^{-4}w^3 - 5.04 \times 10^{-2}w^2 + 4.08w + 1.12$$
By calculating $\frac{dF}{dw}\int_{w=0}$, determine whether the embryo has become stiffer or softer.

4.4 The Product and Quotient Rules, and the Derivatives of Rational and Power Functions

4.4.1 The Product Rule

We saw in Section 4.3 that the derivative of a sum of differentiable functions is the sum of the derivatives of the functions. The rule for products is not so simple. Consider for example: $y = x^5 = (x^3)(x^2)$. We know that

$$\frac{d}{dx}x^5 = 5x^4, \quad \frac{d}{dx}x^3 = 3x^2 \quad \text{and} \quad \frac{d}{dx}x^2 = 2x.$$

So $\frac{d}{dx}x^5 = 5x^4$ is not equal to $\left(\frac{d}{dx}x^3\right)\left(\frac{d}{dx}x^2\right) = (3x^2)(2x) = 6x^3$

> **The Product Rule** If $h(x) = f(x)g(x)$ and both $f(x)$ and $g(x)$ are differentiable at x, then:
> $$h'(x) = f'(x)g(x) + f(x)g'(x)$$
> If we set $u = f(x)$ and $v = g(x)$, then:
> $$(uv)' = u'v + uv'$$

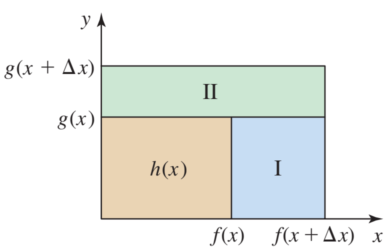


Figure 4.22 The product rule.

Proof Since $h(x) = f(x)g(x)$ is a product of two functions, we can visualize $h(x)$ as the area of a rectangle with sides $f(x)$ and $g(x)$. To compute the derivative, we need $h(x + \Delta x)$; this is given by

$$h(x + \Delta x) = f(x + \Delta x)g(x + \Delta x)$$

To compute $h'(x)$, we must compute $h(x + \Delta x) - h(x)$ (Figure 4.22). We find that

$$\begin{aligned} h(x + \Delta x) - h(x) &= \text{area of I} + \text{area of II} \\ &= [f(x + \Delta x) - f(x)]g(x) \\ &\quad + [g(x + \Delta x) - g(x)]f(x + \Delta x) \end{aligned}$$

Dividing this result by Δx and taking the limit as $\Delta x \to 0$, we obtain

$$\begin{aligned} h'(x) &= \lim_{\Delta x \to 0} \frac{h(x + \Delta x) - h(x)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{[f(x + \Delta x) - f(x)]g(x) + [g(x + \Delta x) - g(x)]f(x + \Delta x)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \left[\frac{f(x + \Delta x) - f(x)}{\Delta x} g(x) + \frac{g(x + \Delta x) - g(x)}{\Delta x} f(x + \Delta x)\right] \end{aligned}$$

Now, we need the assumption that $f'(x)$ and $g'(x)$ exist and that $f(x)$ is continuous at x [which follows from the fact that $f(x)$ is differentiable at x]. These assumptions allow us to use the rules for limits, and we write the last expression as

$$\left(\lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}\right) g(x) + \left(\lim_{\Delta x \to 0} \frac{g(x + \Delta x) - g(x)}{\Delta x}\right)\left(\lim_{\Delta x \to 0} f(x + \Delta x)\right)$$

The difference quotients become $f'(x)$ and $g'(x)$ in the limit as $\Delta x \to 0$. Using the fact that $f(x)$ is continuous at x, we find that $\lim_{\Delta x \to 0} f(x + \Delta x) = f(x)$. Therefore,

$$h'(x) = f'(x)g(x) + g'(x)f(x)$$

as claimed. •

EXAMPLE 1 Differentiate $f(x) = (3x+1)(2x^2-5)$.

Solution We write $u = 3x+1$ and $v = 2x^2-5$. The product rule says that $(uv)' = u'v + uv'$. That is, we need the derivatives of both u and v:

$$\begin{aligned} u &= 3x+1 \qquad & v &= 2x^2-5 \\ u' &= 3 & v' &= 4x \end{aligned}$$

Then

$$\begin{aligned} (uv)' &= u'v + uv' \\ &= 3(2x^2-5) + (3x+1)(4x) \\ &= 6x^2 - 15 + 12x^2 + 4x = 18x^2 + 4x - 15. \qquad \text{Collect powers of } x \end{aligned}$$

We could have gotten this result by first multiplying out $(3x+1)(2x^2-5) = 6x^3 - 15x + 2x^2 - 5$, which is a polynomial function and can be differentiated using the methods from Section 4.3. We then would have found that

$$\frac{d}{dx}(6x^3 - 15x + 2x^2 - 5) = 18x^2 - 15 + 4x$$

which is the same answer. •

EXAMPLE 2 Differentiate $f(x) = (3x^3 - 2x)^2$.

Solution Again, we could expand the square and then differentiate the resulting polynomial—but we can also use the product rule. Write

$$f(x) = \underbrace{(3x^3-2x)}_{u(x)}\underbrace{(3x^3-2x)}_{v(x)}$$

Then:

$$\begin{aligned} f'(x) &= u'v + uv' \\ &= 2uu' \qquad \text{Because } v(x) = u(x). \\ &= 2(3x^3-2x)(9x^2-2). \qquad u'(x) = 9x^2 - 2 \end{aligned}$$

•

EXAMPLE 3 Apply the product rule repeatedly to find the derivative of

$$y = (2x+1)(x+1)(3x-4).$$

Solution We need to generalize the product rule to functions having three factors: If $y = uvw$, then

$$\begin{aligned} y' = (uvw)' &= (uv)'w + (uv)w' \qquad \text{Product rule on } y = UV \text{ with } U = uv, V = w \\ &= (u'v + uv')w + uvw'. \qquad \text{Product rule on } (uv)'. \\ &= u'vw + uv'w + uvw' \end{aligned}$$

In this example:

$$\begin{aligned} u &= 2x+1, \quad & v &= x+1, \quad & w &= 3x-4 \\ u' &= 2 & v' &= 1 & w' &= 3 \end{aligned}$$

Therefore:

$$\begin{aligned} y' &= 2(x+1)(3x-4) + (2x+1)\cdot 1 \cdot (3x-4) \\ &\quad + (2x+1)(x+1)\cdot 3 \\ &= 2(3x^2 - x - 4) + (6x^2 - 5x - 4) \\ &\quad + 3(2x^2 + 3x + 1) \\ &= 18x^2 + 2x - 9. \qquad \text{Collect like powers of } x. \end{aligned}$$

•

EXAMPLE 4 **Density Dependent Population Growth** In Section 2.3 we analyzed a discrete model for the growth of a population in which the reproduction rate (that is, the number of individuals added in one time step, per individual present) depends on the size of the population. In Section 4.2 we showed that the rate of growth of a population can be represented by the derivative dN/dt.

For a population whose reproductive rate is not dependent on the population size

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

where r is a constant. The rate of growth is proportional to N because if the population size is doubled, then the number of organisms added in a fixed time (say one hour, or one day) also will double.

In a population with density-dependent growth the reproductive rate will vary with N and we may write:

$$\frac{dN}{dt} = r(N) \cdot N$$

where $r(N)$ is now a function of N. In particular, doubling N will not necessarily double the rate of population growth under a model with density-dependent reproduction.

Show that when $N = 0$, $(r(N)N)' = r(0)$. Later we will show that this result allows us to interpret $r(0)$ as the reproduction rate when the population size is very small.

Solution Using the product rule:

$$(r(N)N)' = r'(N) \cdot N + r(N) \qquad u = r(N),\ v = N,\ u' = r'(N),\ v' = 1.$$

So when $N = 0$:

$$(r(N)N)' = r'(0) \cdot 0 + r(0) = r(0)$$

4.4.2 The Quotient Rule

The quotient rule will allow us to compute the derivative of a quotient of two functions, that is a function formed by dividing one function by another. In particular, the rule will allow us to compute the derivative of a rational function, because a rational function is the quotient of two polynomial functions.

The Quotient Rule If $h(x) = \frac{f(x)}{g(x)}$, $g(x) \neq 0$, and both $f'(x)$ and $g'(x)$ exist, then

$$h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$$

Alternatively, with $u = f(x)$ and $v = g(x)$,

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$$

We could prove the quotient rule similarly to how we proved the product rule, by using the formal definition of derivatives, and you will be guided through the steps of that proof in Problem 95. However we will use the chain rule, which will be introduced in Section 4.5 to prove the quotient rule. The proof will therefore be deferred to Section 4.5. But you can start to use the quotient rule without knowing how it is proved, as the following examples show.

Note carefully the exact forms of the product and quotient rules. In the product rule we add $f'g$ and fg', whereas in the quotient rule we subtract fg' from $f'g$. As mentioned, we can use the quotient rule to find the derivative of rational functions. We illustrate this application in the next two examples.

EXAMPLE 5 Differentiate $y = \frac{x^3-3x+2}{x^2+1}$.

Solution y is defined for all $x \in \mathbf{R}$, since $x^2+1 \neq 0$. Use the quotient rule to differentiate y:

$$y = \frac{\overbrace{x^3-3x+2}^{u}}{\underbrace{x^2+1}_{v}} \qquad u' = 3x^2-3 \quad v' = 2x$$

So:

$$\begin{aligned} y' = \frac{u'v-uv'}{v^2} &= \frac{(3x^2-3)(x^2+1)-(x^3-3x+2)2x}{(x^2+1)^2} \\ &= \frac{3x^4+3x^2-3x^2-3-2x^4+6x^2-4x}{(x^2+1)^2} \qquad \text{Multiply out factors} \\ &= \frac{x^4+6x^2-4x-3}{(x^2+1)^2}. \qquad \text{Collect like powers of } x \end{aligned}$$

EXAMPLE 6 **Monod Growth Function** The Monod growth function describes how the rate of reproduction of a population of microorganisms may depend on the availability of resources. If we use R to represent the amount of a specific resource, then the Monod growth function models the reproduction rate as a function of R by

$$f(R) = \frac{aR}{k+R}, \qquad R \geq 0$$

where a and k are positive constants, which will depend on the particular species of microorganism as well as the resource that is being varied. Differentiate $f(R)$ and interpret what your answer says about the function $f(R)$.

Solution Since a and k are positive constants, $f(R)$ is defined for all $R \geq 0$. Use the quotient rule:

$$f(R) = \frac{\overbrace{aR}^{u}}{\underbrace{k+R}_{v}} \qquad u' = a,\ v' = 1$$

Hence,

$$\frac{df}{dR} = \frac{u'v-uv'}{v^2} = \frac{a(k+R)-aR\cdot 1}{(k+R)^2} = \frac{ak}{(k+R)^2}$$

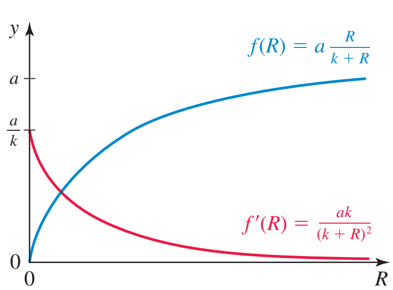


Figure 4.23 The graph of $f(R)$ and $f'(R)$ in Example 6.

In Figure 4.23, we graph both $f(R)$ and $f'(R)$. We see that the slope of the tangent line at $(R, f(R))$ is positive for all $R \geq 0$. We can also draw this conclusion from the graph of $f'(R)$, since it is positive for all $R \geq 0$. If the tangent line has a positive slope everywhere then the function must be increasing; i.e., increasing the amount of resource always increases the reproduction rate. We will return to the idea of using the derivative to deduce whether a function is increasing or decreasing in Chapter 5.

The quotient rule allows us to extend the power rule proven for non-negative integers in Section 4.3 to the case where the exponent is a negative integer:

> **Power Rule (Negative Integer Exponents)** If $f(x) = x^n$, where n is a negative integer, then
> $$f'(x) = nx^{n-1}$$

Note that the power rule for negative integer exponents works the same way as the power rule for positive integer exponents.

Proof Let $m = -n$, then $f(x) = \frac{1}{x^m}$ where m is a positive integer. Apply the quotient rule:

$$f(x) = \frac{\overbrace{1}^{u}}{\underbrace{x^m}_{v}} \qquad u' = 0,\ v' = mx^{m-1} \text{ by power rule for positive integer exponents}$$

$$f'(x) = \frac{u'v - uv'}{v^2} = \frac{0 \cdot x^m - 1 \cdot mx^{m-1}}{(x^m)^2} = -\frac{mx^{m-1}}{x^{2m}} = -mx^{-m-1} = nx^{n-1}$$

EXAMPLE 7 Use the power rule to differentiate
(a) $y = \frac{1}{x}$ **(b)** $g(x) = \frac{3}{x^4}$.

Solution
(a) $y' = \frac{d}{dx}(x^{-1}) = (-1)x^{-1-1} = -\frac{1}{x^2}$ $\quad n = -1$
(b) $g'(x) = \frac{d}{dx}(3x^{-4}) = 3\frac{d}{dx}x^{-4} = 3(-4)x^{-4-1} = -12x^{-5} = -\frac{12}{x^5}$ $\quad n = -4$

There is a general form of the power rule in which the exponent can be any real number. In Section 4.5 we will prove the power rule in the case where the exponent is a rational number:

> **Power Rule (Rational Exponents)** Let $f(x) = x^r$, where r is any rational number. Then:
> $$f'(x) = rx^{r-1}$$

EXAMPLE 8 But we can start to use the power rule without knowing how it is proved, as the following examples show. Use the power rule to differentiate
(a) $y = \sqrt{x}$
(b) $y = \sqrt[5]{x}$
(c) $g(t) = \frac{1}{\sqrt[3]{t}}$
(d) $h(s) = s^{2/3}$

Solution
(a) $y' = \frac{d}{dx}\left(x^{1/2}\right) = \frac{1}{2}x^{\frac{1}{2}-1} = \frac{1}{2}x^{-1/2} = \frac{1}{2\sqrt{x}}$ $\quad \sqrt{x} = x^{1/2}$
(b) $y' = \frac{d}{dx}\left(x^{1/5}\right) = \frac{1}{5}x^{(1/5)-1} = \frac{1}{5}x^{-4/5} = \frac{1}{5x^{4/5}}$ $\quad \sqrt[5]{x} = x^{1/5}$
(c) $g'(t) = \frac{d}{dt}\left(t^{-1/3}\right) = \left(-\frac{1}{3}\right)t^{(-1/3)-1} = \left(-\frac{1}{3}\right)t^{-4/3} = -\frac{1}{3t^{4/3}}$ $\quad \frac{1}{\sqrt[3]{t}} = t^{-1/3}$
(d) $h'(s) = \frac{2}{3}s^{-1/3} = \frac{2}{3s^{1/3}}$.

The function $f(x) = \sqrt{x}$, $x \geq 0$, appears quite frequently. It is therefore worthwhile to memorize its derivative, which is defined only for $x > 0$:

$$\boxed{\frac{d}{dx}\sqrt{x} = \frac{1}{2\sqrt{x}}}$$

EXAMPLE 9 **Combining the Rules** Differentiate $f(x) = \sqrt{x}(x^2 - 1)$.

Solution 1 Use the product rule:

$$f(x) = \underbrace{\sqrt{x}}_{u}\ \underbrace{(x^2 - 1)}_{v} \qquad u' = \frac{1}{2\sqrt{x}},\ v' = 2x$$

Hence,

$$f'(x) = u'v + uv' = \frac{1}{2\sqrt{x}}(x^2 - 1) + \sqrt{x}(2x)$$
$$= \frac{x^2 - 1 + \sqrt{x}(2x)2\sqrt{x}}{2\sqrt{x}} = \frac{x^2 - 1 + 4x^2}{2\sqrt{x}} = \frac{5x^2 - 1}{2\sqrt{x}}.$$

Solution 2 Since $f(x) = \sqrt{x}(x^2 - 1) = x^{5/2} - x^{1/2}$, we can also use the power rule. We find that

$$f'(x) = \frac{5}{2}x^{(5/2)-1} - \frac{1}{2}x^{(1/2)-1} = \frac{5}{2}x^{3/2} - \frac{1}{2\sqrt{x}} = \frac{5x^{3/2}\sqrt{x} - 1}{2\sqrt{x}} = \frac{5x^2 - 1}{2\sqrt{x}}.$$

As we saw in Example 6, functions that model biological phenomena often contain unknown coefficients; for example, a function modeling growth of microorganisms might have coefficients that reflect the fact that growth may be different between different species or between different environments. It is important to practice differentiating functions with unknown coefficients.

EXAMPLE 10 **A Function That Contains a Constant** Differentiate $h(t) = (at)^{1/3}(a + 1) - a$, where a is a positive constant.

Solution Since $h(t)$ is a function of t, we need to differentiate with respect to t, keeping in mind that a is a constant. Rewriting $h(t)$ will make this easier:

$$h(t) = a^{1/3}(a + 1)t^{1/3} - a$$

The factor $a^{1/3}(a + 1)$ in front of $t^{1/3}$ is a constant. Thus,

$$h'(t) = a^{1/3}(a + 1)\frac{1}{3}t^{-2/3} - 0 = \frac{a^{1/3}(a + 1)}{3t^{2/3}}$$

Both the product rule and the quotient rule can be used to simplify the derivatives. In both of the following examples the function to be differentiated has some unknown part $f(x)$. We will relate the derivatives to $f(x)$ and $f'(x)$.

EXAMPLE 11 **Differentiating a Function That Is Not Specified** Suppose $f(2) = 3$ and $f'(2) = 1/4$. Find $\frac{d}{dx}[xf(x)]$ at $x = 2$.

Solution Use the product rule:

$$\frac{d}{dx}[xf(x)] = f(x) + xf'(x) \qquad u = x, v = f(x)$$

Hence,

$$\left.\frac{d}{dx}[xf(x)]\right|_{x=2} = f(2) + 2f'(2) = 3 + \frac{1}{2} = \frac{7}{2}$$

EXAMPLE 12 **Differentiating a Function That Is Not Specified** Suppose that $f(x)$ is differentiable. Find an expression for the derivative of $y = \frac{f(x)}{x^2}$ in terms of $f(x)$ and $f'(x)$.

Solution Use the quotient rule

$$y = \frac{\overbrace{f(x)}^{u}}{\underbrace{x^2}_{v}} \qquad u' = f'(x), \quad v' = 2x$$

Then:

$$y' = \frac{f'(x)x^2 - f(x) \cdot 2x}{x^4} = \frac{xf'(x) - 2f(x)}{x^3}$$

Section 4.4 Problems

4.4.1

In Problems 1–16, use the product rule to find the derivative with respect to the independent variable.

1. $f(x) = (x+5)(x^2-3)$

2. $f(x) = (2x^3-1)(3+2x^2)$

3. $f(x) = (3x^4-5)(2x-5x^3)$

4. $f(x) = (3x^4-x^2+1)(2x^2-5x^3)$

5. $f(x) = \left(\frac{1}{2}x^2-1\right)(2x+3x^2)$

6. $f(x) = 2(3x^2-2x^3)(1-5x^2)$

7. $f(x) = \frac{1}{5}(x^2-1)(x^2+1)$

8. $f(x) = 3(x^2+2)(4x^2-5x^4)-3$

9. $f(x) = (3x-1)^2$

10. $f(x) = (4-2x^2)^2$

11. $f(x) = 3(1-2x)^2$

12. $f(x) = \frac{(2x^2-3x+1)^2}{4}+2$

13. $g(s) = (2s^2-5s)^2$

14. $h(t) = 4(3t^2-1)(2t+1)$

15. $g(t) = 3(2t^2-5t^4)^2$

16. $h(s) = (4-3s^2+4s^3)^2$

In Problems 17–20, apply the product rule to find the tangent line, in slope–intercept form, of $y = f(x)$ at the specified point.

17. $f(x) = (1-2x)(1+2x)$, at $x = 2$

18. $f(x) = (3x^2-2)(x-1)$, at $x = 1$

19. $f(x) = 4(2x^4+3x)(4-2x^2)$, at $x = -1$

20. $f(x) = (3x^3-3)(2-2x^2)$, at $x = 0$

In Problems 21–24, apply the product rule to find the normal line, in slope–intercept form, of $y = f(x)$ at the specified point.

21. $f(x) = (2x+1)(3x^2-1)$, at $x = 1$

22. $f(x) = (1-x)(2-x^2)$, at $x = 2$

23. $f(x) = 5(1-2x)(x+1)-3$, at $x = 0$

24. $f(x) = \frac{(2-x)(3-x)}{4}$, at $x = -1$

In Problems 25–28, apply the product rule for the product of three functions to find the derivative of $y = f(x)$.

25. $f(x) = (x-3)(2-3x)(5-x)$

26. $f(x) = (2x-1)(3x+4)(1-x)$

27. $f(x) = (x-3)(2x^2+1)(1-x^2)$

28. $f(x) = (2x+1)(4-x^2)(1+x^2)$

29. Differentiate
$$f(x) = a(x+1)(2x-1)$$
with respect to x. Assume that a is a positive constant.

30. Differentiate
$$f(x) = (a-x)(a+x)$$
with respect to x. Assume that a is a positive constant.

31. Differentiate
$$f(x) = 2a(x^2-a)+a^2$$
with respect to x. Assume that a is a positive constant.

32. Differentiate
$$f(x) = \frac{3(x-1)^2}{2+a}$$
with respect to x. Assume that a is a positive constant.

33. Differentiate
$$g(t) = (at+1)^2$$
with respect to t. Assume that a is a positive constant.

34. Differentiate
$$h(t) = \sqrt{a}(t-a)+a$$
with respect to t. Assume that a is a positive constant.

35. Suppose that $f(2) = -4$, $g(2) = 3$, $f'(2) = 1$, and $g'(2) = -2$. Find
$$(fg)'(2)$$

36. Suppose that $f(2) = -4$, $g(2) = 3$, $f'(2) = 1$, and $g'(2) = -2$. Find
$$(f^2+g^2)'(2)$$

In Problems 37–40, assume that $f(x)$ is differentiable. Find an expression for the derivative of y at $x = 1$, assuming that $f(1) = 2$ and $f'(1) = -1$.

37. $y = 2xf(x)$

38. $y = 3x^2f(x)$

39. $y = -5x^3f(x)-2x$

40. $y = \frac{xf(x)}{2}$

In Problems 41–44, assume that $f(x)$ and $g(x)$ are differentiable at x. Find an expression for the derivative of y.

41. $y = 3f(x)g(x)$

42. $y = [f(x)-3]g(x)$

43. $y = [f(x)+2g(x)]g(x)$

44. $y = [-2f(x)-3g(x)]g(x)+\frac{2g(x)}{3}$

45. Gini–Simpson Diversity Index The diversity of a population that contains two species can be measured by the Gini–Simpson diversity index. If the fraction of organisms from species 1 is p, then the diversity is given by:
$$H(p) = 2p(1-p).$$
Use the product rule to calculate $H'(p)$. For what value of p does $H'(p) = 0$?

46. Density-dependent Population Growth One widely used model of density-dependent growth (see Example 4) is the logistic equation:
$$\frac{dN}{dt} = rN(1-N/K)$$
where r and K are both positive coefficients.

Show that, if $N < K$ (remember $N \geq 0$, because populations cannot contain negative numbers of individuals), $dN/dt \geq 0$; i.e., the population is growing. What happens if $N > K$?

(b) The logistic equation may be written in the form:
$$\frac{dN}{dt} = R(N)\cdot N \quad \text{where} \quad R(N) = r(1-N/K).$$
Show explicitly that when $N = 0$, $\frac{d}{dN}(R(N)\cdot N) = R(0)$.

(c) Show that when $N = K$, $dN/dt = 0$. When the population size is equal to K, the population neither grows nor decreases in size.

(d) Show when $N = K$, $\frac{d}{dN}(R(N)\cdot N) = -r$.

We will show in Chapter 8 how to use the results from (c) and (d) to predict the size of the population as $t \to \infty$.

47. **Density-dependent Population Growth** Small populations of organisms will often find themselves outcompeted by other species. Populations do not start to grow until they exceed some critical size. This is known as the Allee effect. One model for population growth that incorporates the Allee effect is:

$$\frac{dN}{dt} = f(N) \text{ where } f(N) = rN(N-a)(1-N/K)$$

where r, a, and K are all positive constants and $K > a$.

(a) Show that if $N < a$, then $\frac{dN}{dt} \leq 0$.

(b) Show that if $a < N < K$, then $dN/dt > 0$.

(a) and (b) together imply that populations smaller than a will shrink, and populations larger than a will grow.

(c) Show that $f'(0) < 0$ and $f'(K) < 0$.

(d) Show that $f'(a) > 0$.

We will show in Chapter 8 how to use (c) and (d) to predict the size of the population as $t \to \infty$.

48. **Chemical Reaction** Consider the chemical reaction:

$$\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{AB}$$

If x denotes the concentration of AB at time t and a, b are the initial amounts of A and B present, then the reaction rate $R(x)$ is given by

$$R(x) = k(a-x)(b-x)$$

where k, a, and b are positive constants. Differentiate $R(x)$.

4.4.2

In Problems 49–70, differentiate with respect to the independent variable.

49. $f(x) = \dfrac{3x-1}{x+1}$

50. $f(x) = \dfrac{1-4x^3}{1-x}$

51. $f(x) = \dfrac{3x^2-2x+1}{2x+1}$

52. $f(x) = \dfrac{x^4+2x-1}{5x^2-2x+1}$

53. $f(x) = \dfrac{3-x^3}{1-x}$

54. $f(x) = \dfrac{1+2x^2-4x^4}{3x^3-5x^5}$

55. $h(t) = \dfrac{t^2-3t+1}{t+1}$

56. $h(t) = \dfrac{3-t^2}{(t-1)^2}$

57. $f(s) = \dfrac{4-2s^2}{1-s}$

58. $f(s) = \dfrac{2s^3-4s^2+3s-4}{(s^2-3)^2}$

59. $f(x) = \sqrt{x}(x-1)$

60. $f(x) = \sqrt{x}(x^4-x^2)$

61. $f(x) = \sqrt{3x}(x^2-1)$

62. $f(x) = \dfrac{\sqrt{5x}(1+x^2)}{\sqrt{2}}$

63. $f(x) = x^3 + \dfrac{1}{x^3}$

64. $f(x) = x^5 - \dfrac{1}{x^5}$

65. $f(x) = 2x^2 - \dfrac{x-3}{x^2}$

66. $f(x) = -x^3 + \dfrac{2x^2+3}{x^4+1}$

67. $g(s) = \dfrac{s^{1/3}-1}{s^{2/3}-1}$

68. $g(s) = \dfrac{s^{1/7}-s^{2/7}}{s^{3/7}+s^{4/7}}$

69. $f(x) = (1-2x)\left(\sqrt{2x} + \dfrac{2}{\sqrt{x}}\right)$

70. $f(x) = (x^2-1)\left(\sqrt{x} + \dfrac{1}{\sqrt{x}} - 1\right)$

In Problems 71–74, find the tangent line, in slope–intercept form, of $y = f(x)$ at the specified point.

71. $f(x) = \dfrac{x^2+3}{x^3+5}$, at $x = 0$

72. $f(x) = \dfrac{1}{x} - \dfrac{2}{\sqrt{x}} + \dfrac{4}{x^2}$, at $x = 1$

73. $f(x) = \dfrac{x+5}{x^3}$, at $x = 2$

74. $f(x) = \sqrt{x}(x^3-1)$, at $x = 1$

75. Differentiate

$$f(x) = \frac{ax}{3+x}$$

with respect to x. Assume that a is a positive constant.

76. Differentiate

$$f(x) = \frac{ax}{k+x}$$

with respect to x. Assume that a and k are positive constants.

77. Differentiate

$$f(x) = \frac{ax^2}{4+x^2}$$

with respect to x. Assume that a is a positive constant.

78. Differentiate

$$f(x) = \frac{ax^2}{k^2+x^2}$$

with respect to x. Assume that a and k are positive constants.

79. **Blood Oxygen Content** Hill's function models how the amount of oxygen bound to hemoglobin in the blood depends on oxygen concentration, P, in the surrounding tissues. In its most general form Hill's function models the fraction of hemoglobin molecules in blood that are bound to oxygen by:

$$f(P) = \frac{P^n}{k^n + P^n}$$

where k is a positive constant, and n is a positive integer.

(a) Calculate $f'(P)$.

(b) Show that $f'(P) > 0$ for all $P > 0$. This result means that increasing the oxygen concentration always increases the fraction of hemoglobin molecules that are bound to oxygen.

80. Differentiate

$$h(t) = \sqrt{at}(1-a) + a$$

with respect to t. Assume that a is a positive constant.

81. Differentiate

$$h(t) = \sqrt{at}(t-a) + at$$

with respect to t. Assume that a is a positive constant.

82. Suppose that $f(2) = -4$, and $f'(2) = 1$. Let $y = 1/f(x)$; find $\frac{dy}{dx}$ when $x = 2$.

83. Suppose that $f(2) = -4$, $g(2) = 1$, $f'(2) = 0$, and $g'(2) = -2$. Let $y = f(x)/(2g(x))$; find $\frac{dy}{dx}$ when $x = 2$.

In Problems 84–87, assume that $f(x)$ is differentiable. Find an expression for the derivative of y at $x = 2$, assuming that $f(2) = -1$ and $f'(2) = 1$.

84. $y = \dfrac{f(x)}{x^2+1}$

85. $y = \dfrac{x^2 f(x)}{x^2 + f(x)}$

86. $y = [f(x)]^2 - \dfrac{x}{f(x)}$

87. $y = \dfrac{f(x)+1}{f(x)+x}$

In Problems 88–91, assume that $f(x)$ and $g(x)$ are differentiable at x. Find an expression for the derivative of y in terms of $f(x)$, $g(x)$, $f'(x)$, and $g'(x)$.

88. $y = \dfrac{2f(x)+x}{3g(x)}$

89. $y = \dfrac{f(x)}{[g(x)]^2}$

90. $y = \dfrac{x^2}{f(x)+g(x)}$

91. $y = \sqrt{x} f(x) g(x)$

92. Assume that $f(x)$ is a differentiable function. Find the derivative of the reciprocal function $g(x) = 1/f(x)$ at those points x where $f(x) \neq 0$.

93. Find the tangent line to the hyperbola $yx = c$, where c is a positive constant, at the point (x_1, y_1) with $x_1 > 0$. Show that the tangent line intersects the x-axis at a point that does not depend on c.

94. **Density-dependent Population Growth** ***Smith (1963)*** proposed a model for the growth of a population of microorganisms whose reproductive rate decreases as the number of microorganisms increases.

According to Smith's model, if $N(t)$ is the number of microorganisms, then $dN/dt = R(N)N$ where

$$R(N) = \frac{r(1 - N/K)}{1 + N/a}.$$

where a, K, and r are all positive constants.

(a) Show that $dN/dt \geq 0$ if $N < K$ (remember that since a population may not contain a negative number of microorganisms, $N \geq 0$).

(b) Show that $dN/dt < 0$ if $N > K$.

(a) and (b) imply that the population grows if $N < K$ and that it decreases if $N > K$.

(c) Show that $\frac{d}{dN}(R(N) \cdot N) > 0$ at $N = 0$ and $\frac{d}{dN}(R(N) \cdot N) < 0$ if $N = K$. We will use these results in Chapter 8 to predict the population size $N(t)$ as $t \to \infty$.

(d) Show that $R'(N) < 0$ for all $N \geq 0$. This result means that reproductive rate decreases as the number of microorganisms increases.

95. In this problem we will prove the quotient rule using an argument similar to the one used to prove the product rule in Section 4.4.1.

Let $u(x)$ and $v(x)$ be differentiable functions, and define a quotient function $f(x) = \frac{u(x)}{v(x)}$. The derivative $f'(x)$, if it exists, is equal to:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

(a) Assuming $v(x) \neq 0$, show that the quotient on the right-hand side can be written as:

$$f'(x) = \lim_{h \to 0} \frac{(u(x+h)v(x) - u(x)v(x)) - (u(x)v(x+h) - u(x)v(x))}{hv(x)v(x+h)}$$

and then rearranged into:

$$f'(x) = \lim_{h \to 0} \frac{\left(\frac{u(x+h)-u(x)}{h}\right) v(x) - u(x)\left(\frac{v(x+h)-v(x)}{h}\right)}{v(x)v(x+h)}$$

(b) Using the limit laws, show that the equation from (a) can be rewritten as

$$f'(x) = \frac{\left(\lim_{h \to 0} \frac{u(x+h)-u(x)}{h}\right) v(x) - u(x)\left(\lim_{h \to 0} \frac{v(x+h)-v(x)}{h}\right)}{v(x) \lim_{h \to 0}(v(x+h))}$$

provided all of the limits exist, and provided $\lim_{h \to 0} v(x+h) \neq 0$.

(c) Using the formal definition of a derivative write $f'(x)$ in terms of $u(x)$, $v(x)$, $u'(x)$, and $v'(x)$. Hence prove the quotient law.

4.5 The Chain Rule

4.5.1 The Chain Rule

In Section 1.3, we defined the composition of functions. Composition of functions is often helpful when combining mathematical models. For example, suppose we know how hyena skull size, S, varies as a function of age, t; that is, we have a function $S(t)$. Suppose we also have a model for how bite strength, F, varies with skull size; that is, we also have a function $F(S)$. Then we may compose the two functions to obtain a model for how bite strength varies with age, t:

$$(F \circ S)(t) = F(S(t)).$$

To find the derivative of composite functions, we need the **chain rule**. We will state the chain rule and defer its proof to Subsection 4.5.2. Just as with the quotient rule, you can use the chain rule without seeing its proof.

Chain Rule If g is differentiable at x and f is differentiable at $y = g(x)$, then the composite function $(f \circ g)(x) = f[g(x)]$ is differentiable at x, and the derivative is given by

$$(f \circ g)'(x) = f'[g(x)]g'(x)$$

This formula looks complicated. Let's take a moment to see what we need to do to find the derivative of the composite function $(f \circ g)(x)$. The function g is the inner function; the function f is the outer function. The expression $f'[g(x)]g'(x)$ thus means that we need to find the derivative of the outer function, evaluated at $g(x)$, and the derivative of the inner function, evaluated at x, and then multiply the two together.

EXAMPLE 1 **A Polynomial** Find the derivative of $h(x) = (3x^2 - 1)^2$

Solution The inner function is $g(x) = 3x^2 - 1$; the outer function is $f(u) = u^2$. Then

$$g'(x) = 6x \qquad \text{and} \qquad f'(u) = 2u$$

Evaluating $f'(u)$ at $u = g(x)$ yields

$$f'[g(x)] = 2g(x) = 2(3x^2 - 1)$$

Thus,

$$\begin{aligned} h'(x) &= (f \circ g)'(x) = f'[g(x)]g'(x) \\ &= 2(3x^2 - 1) \cdot 6x = 12x(3x^2 - 1) \end{aligned}$$

•

The derivative of $f \circ g$ can be written in Leibniz notation. If we set $u = g(x)$, then

$$\frac{d}{dx}[(f \circ g)(x)] = \frac{df}{du}\frac{du}{dx}$$

This form of the chain rule emphasizes that, in order to differentiate $f \circ g$, we multiply the derivative of the outer function and the derivative of the inner function, the former evaluated at u, the latter at x.

EXAMPLE 2 **A Polynomial** Find the derivative of $h(x) = (2x + 1)^3$.

Solution If we set $u = g(x) = 2x + 1$ and $f(u) = u^3$, then $h(x) = (f \circ g)(x)$. We need to find both $f'[g(x)]$ and $g'(x)$ to compute $h'(x)$. Now,

$$g'(x) = 2 \qquad \text{and} \qquad f'(u) = 3u^2$$

Hence, since $f'[g(x)] = 3(g(x))^2 = 3(2x + 1)^2$, it follows that

$$\begin{aligned} h'(x) = f'[g(x)]g'(x) &= 3(2x + 1)^2 \cdot 2 \\ &= 6(2x + 1)^2. \end{aligned}$$

We can write the same calculation using Leibniz notation:

$$\begin{aligned} h'(x) = \frac{df}{du}\frac{du}{dx} &= 3u^2 \cdot 2 = 3(2x + 1)^2 \cdot 2 \\ &= 6(2x + 1)^2. \end{aligned}$$

•

EXAMPLE 3 **A Radical** Find the derivative of $h(x) = \sqrt{x^2 + 1}$.

Solution If we set $u = g(x) = x^2 + 1$ and $f(u) = \sqrt{u}$, then $h(x) = (f \circ g)(x)$. We find that

$$g'(x) = 2x \qquad \text{and} \qquad f'(u) = \frac{1}{2\sqrt{u}}$$

We need to evaluate f' at $g(x)$—that is,

$$f'[g(x)] = \frac{1}{2\sqrt{g(x)}} = \frac{1}{2\sqrt{x^2 + 1}}$$

Therefore,

$$h'(x) = f'[g(x)]g'(x) = \frac{1}{2\sqrt{x^2 + 1}} \cdot 2x = \frac{x}{\sqrt{x^2 + 1}}.$$

•

EXAMPLE 4 **A Radical** Find the derivative of $h(x) = \sqrt[7]{2x^2 + 3x}$.

Solution We write $h(x) = (2x^2 + 3x)^{1/7}$. The inner function is $g(x) = 2x^2 + 3x$ and the outer function is $f(u) = u^{1/7}$. Thus, we find that

$$\begin{aligned} h'(x) &= \frac{df}{du}\frac{du}{dx} = \frac{1}{7}u^{1/7-1}(4x+3) \\ &= \frac{1}{7}(2x^2+3x)^{-6/7}(4x+3) \qquad \text{Substitute } u = 2x^2+3x \\ &= \frac{4x+3}{7(2x^2+3x)^{6/7}}. \end{aligned}$$

•

EXAMPLE 5 **A Rational Function** Find the derivative of $h(x) = \left(\frac{x}{x+1}\right)^2$.

Solution If we set $u = g(x) = \frac{x}{x+1}$ and $f(u) = u^2$, then $h(x) = (f \circ g)(x)$. We use the quotient rule to compute the derivative of $g(x)$:

$$g'(x) = \frac{1 \cdot (x+1) - x \cdot 1}{(x+1)^2} = \frac{1}{(x+1)^2}$$

Since $f'(u) = 2u$, we obtain

$$h'(x) = f'[g(x)]g'(x) = 2\frac{x}{x+1}\frac{1}{(x+1)^2} = \frac{2x}{(x+1)^3}.$$

•

The Proof of the Quotient Rule We can use the chain rule to prove the quotient rule. Assume that $g(x) \neq 0$ for all x in the domain of g. If we define $h(x) = \frac{1}{x}$, then

$$(h \circ g)(x) = h[g(x)] = \frac{1}{g(x)}$$

We used the formal definition of the derivative in Example 3 in Section 4.1 to show that $h'(x) = -\frac{1}{x^2}$. This, together with the chain rule, yields

$$(h \circ g)'(x) = -\frac{1}{[g(x)]^2}g'(x), \qquad \text{or} \qquad \left(\frac{1}{g}\right)' = -\frac{g'}{g^2}$$

Since $\frac{f}{g} = f \cdot \frac{1}{g}$, we can use the product rule to find the derivative of $\frac{f}{g}$:

$$\begin{aligned} \left(\frac{f}{g}\right)' &= f'\frac{1}{g} + f\left(\frac{1}{g}\right)' = f'\frac{1}{g} + f\left(-\frac{g'}{g^2}\right) \\ &= \frac{f'g - fg'}{g^2} \end{aligned}$$

•

Note that we did not use the power rule for negative integer exponents (Subsection 4.4.2) to compute $h'(x)$, but instead used the formal definition of derivatives to compute the derivative of $1/x$. Using the power rule for negative integer exponents would have been circular reasoning: We used the quotient rule to prove the power rule for negative integer exponents, so we cannot use the power rule for negative integer exponents to prove the quotient rule.

EXAMPLE 6 **A Function with Unknown Parameters** Find the derivative of $h(x) = (ax^2 - 2)^n$ where $a > 0$ and n is a positive integer.

Solution If we set $g(x) = ax^2 - 2$ and $f(u) = u^n$, then $h(x) = (f \circ g)(x)$. Since

$$g'(x) = 2ax \qquad \text{and} \qquad f'(u) = nu^{n-1}$$

it follows that

$$\begin{aligned} h'(x) &= f'[g(x)]g'(x) = nu^{n-1} \cdot 2ax \\ &= n(ax^2-2)^{n-1} \cdot 2ax \qquad \text{Substitute } u = ax^2 - 2 \\ &= 2anx(ax^2-2)^{n-1}. \end{aligned}$$

Looking at $h'(x) = n(ax^2-2)^{n-1} \cdot 2ax$, we see that we first differentiated the outer function f, which yielded $n(ax^2-2)^{n-1}$ via the power rule, and then multiplied the result by the derivative of the inner function g, which is $2ax$. •

EXAMPLE 7 **A Function That Is Not Specified** Suppose $f(x)$ is differentiable. Find $\frac{d}{dx}\left(\frac{1}{\sqrt{f(x)}}\right)$.

Solution We set

$$h(x) = \frac{1}{\sqrt{f(x)}} = [f(x)]^{-1/2}.$$

Now, $f(x)$ is the inner function and $h(u) = u^{-1/2}$ is the outer function; hence,

$$\begin{aligned} \frac{d}{dx}h(x) &= \frac{dh}{du}\frac{du}{dx} = -\frac{1}{2}u^{-3/2}f'(x) \\ &= -\frac{1}{2u^{3/2}}f'(x) = -\frac{f'(x)}{2[f(x)]^{3/2}}. \qquad \text{Substitute } u = f(x) \end{aligned}$$ •

EXAMPLE 8 **Generalized Power Rule** Suppose $f(x)$ is differentiable and r is a real number. Find

$$\frac{d}{dx}[f(x)]^r.$$

Solution Using the general form of the power rule and the chain rule, we find that

$$\frac{d}{dx}[f(x)]^r = r[f(x)]^{r-1}f'(x) \qquad \text{Let } u^r \text{ be the outer function.}$$ •

EXAMPLE 9 **A Function That Is Not Specified** Suppose that $f'(x) = 3x - 1$. Find

$$\frac{d}{dx}f(x^2) \quad \text{at } x = 3.$$

Solution The inner function is $u = x^2$, the outer function is $f(u)$, and we find that

$$\frac{d}{dx}f(x^2) = 2xf'(x^2).$$

If we substitute $x = 3$ into $f'(x^2)$, we obtain $f'(3^2) = f'(9) = (3)(9) - 1 = 26$. Thus,

$$\left.\frac{d}{dx}f(x^2)\right|_{x=3} = (2)(3)f'(9) = (6)(26) = 156$$ •

The chain rule can be applied repeatedly, as shown in the next two examples.

EXAMPLE 10 **Nested Chain Rule** Find the derivative of

$$h(x) = \left(\sqrt{x^2+1}+1\right)^2.$$

Solution We can set $h(x) = (f \circ g)(x)$, with $g(x) = \sqrt{x^2+1} + 1$ and $f(u) = u^2$. We see that $g(x)$ is itself a composition of two functions, with inner function $v = x^2 + 1$ and outer function $\sqrt{v} + 1$. To differentiate $h(x)$, we proceed stepwise. First,

$$h'(x) = \frac{d}{dx}\left(\sqrt{x^2+1}+1\right)^2 = 2\left(\sqrt{x^2+1}+1\right)\frac{d}{dx}\left(\sqrt{x^2+1}+1\right)$$

Then, since

$$\frac{d}{dx}\left(\sqrt{x^2+1}+1\right) = \frac{2x}{2\sqrt{x^2+1}} = \frac{x}{\sqrt{x^2+1}}$$ Use chain rule to differentiate $\sqrt{x^2+1}$

$$h'(x) = 2\left(\sqrt{x^2+1}+1\right)\frac{x}{\sqrt{x^2+1}} = 2x + \frac{2x}{\sqrt{x^2+1}}.$$ Simplify •

EXAMPLE 11 **Nested Chain Rule** Find the derivative of

$$h(x) = \left(2x^3 - \sqrt{3x^4-2}\right)^3.$$

Solution As in the previous example, we proceed stepwise:

$$h'(x) = 3\left(2x^3 - \sqrt{3x^4-2}\right)^2 \frac{d}{dx}\left(2x^3 - \sqrt{3x^4-2}\right)$$

To evaluate $\frac{d}{dx}\sqrt{3x^4-2}$ we again make use of the chain rule. The inner function is $v = 3x^4 - 2$ and the outer function is $f(v) = \sqrt{v}$. So using Leibniz notation:

$$\frac{d}{dx}\sqrt{3x^4-2} = \frac{df}{dx} = \frac{df}{dv}\cdot\frac{dv}{dx} = \frac{1}{2\sqrt{v}}\cdot 12x^3 = \frac{6x^3}{\sqrt{3x^4-4}}.$$

So substituting into our previous expression for $h'(x)$, we obtain:

$$h'(x) = 3\left(2x^3 - \sqrt{3x^4-2}\right)^2\left(6x^2 - \frac{6x^3}{\sqrt{3x^4-2}}\right)$$
$$= 18x^2\left(2x^3 - \sqrt{3x^4-2}\right)^2\left(1 - \frac{x}{\sqrt{3x^4-2}}\right).$$ •

4.5.2 Proof of the Chain Rule

We conclude our discussion of the chain rule by proving it.

Proof of the Chain Rule We will use the definition of the derivative to prove the chain rule. Formally,

$$(f\circ g)'(x) = \lim_{x\to c}\frac{f[g(x)] - f[g(c)]}{x-c}$$

We need to show that the right-hand side is equal to $f'[g(c)]g'(c)$. As long as $g(x) \neq g(c)$, we can write

$$\lim_{x\to c}\frac{f[g(x)] - f[g(c)]}{x-c} = \lim_{x\to c}\frac{\frac{f[g(x)]-f[g(c)]}{g(x)-g(c)}[g(x)-g(c)]}{x-c}$$
$$= \lim_{x\to c}\left(\frac{f[g(x)] - f[g(c)]}{g(x)-g(c)}\right)\left(\frac{g(x)-g(c)}{x-c}\right)$$

Since $g(x)$ is continuous at $x = c$, it follows that $\lim_{x\to c} g(x) = g(c)$, and hence,

$$\lim_{x\to c}\frac{f[g(x)] - f[g(c)]}{g(x)-g(c)} = f'[g(c)]$$

Furthermore,

$$\lim_{x\to c}\frac{g(x)-g(c)}{x-c} = g'(c)$$

Since these limits exist, we can use the fact that the limit of a product is the product of the limits. We find that

$$\lim_{x\to c}\frac{f[g(x)] - f[g(c)]}{x-c} = \lim_{x\to c}\frac{f[g(x)] - f[g(c)]}{g(x)-g(c)}\cdot\lim_{x\to c}\frac{g(x)-g(c)}{x-c}$$
$$= f'[g(c)]g'(c).$$

In the preceding calculation, we needed to assume that $g(x) - g(c) \neq 0$. But, when we take the limit as $x \to c$, there might be x-values such that $g(x) = g(c)$, and for our proof to be rigorous we must deal with this possibility.

We set $u = g(x)$ and $d = g(c)$. The expression

$$f^*(u) = \frac{f(u) - f(d)}{u - d}$$

is defined only for $u \neq d$. Now

$$\lim_{u \to d} f^*(u) = \lim_{u \to d} \frac{f(u) - f(d)}{u - d} = f'(d).$$

Furthermore we can extend $f^*[g(x)]$ to make $f^*[g(x)]$ a continuous function:

$$f^*[g(x)] = \begin{cases} \dfrac{f[g(x)] - f[g(c)]}{g(x) - g(c)} & \text{for } g(x) \neq g(c) \\ f'[g(c)] & \text{for } g(x) = g(c) \end{cases}$$

This means that, for all x,

$$f[g(x)] - f[g(c)] = f^*[g(x)][g(x) - g(c)]$$ Equality holds whether $g(x) \neq g(c)$ or $g(x) = g(c)$.

We can then repeat our calculations to obtain

$$\begin{aligned} \lim_{x \to c} \frac{f[g(x)] - f[g(c)]}{x - c} &= \lim_{x \to c} \frac{f^*[g(x)][g(x) - g(c)]}{x - c} \\ &= \lim_{x \to c} f^*[g(x)] \cdot \lim_{x \to c} \frac{g(x) - g(c)}{x - c} && \text{Use product rule for limits.} \\ &= f'[g(c)] \cdot g'(c) && \lim_{x \to c} \cdot g(x) = g(c) \quad \lim_{d \to c} f^*[d] = f'(d) \end{aligned}$$

•

Section 4.5 Problems

4.5.1

In Problems 1–28, differentiate the functions with respect to the independent variable.

1. $f(x) = (x - 3)^2$

2. $f(x) = (4x + 5)^3$

3. $f(x) = (1 - 3x^2)^4$

4. $f(x) = (x^2 - 3x)^3$

5. $f(x) = \sqrt{x^2 + 3}$

6. $f(x) = \sqrt{2x + 7}$

7. $f(x) = \sqrt{1 - x^3}$

8. $f(x) = \sqrt{5x + 3x^4}$

9. $f(x) = \dfrac{1}{(x^3 - 1)^4}$

10. $f(x) = \dfrac{2}{(1 - 2x^2)^3}$

11. $f(x) = \dfrac{3x + 1}{\sqrt{2x^2 - 1}}$

12. $f(x) = \dfrac{(1 - 3x^2)^3}{(3 - x^2)^2}$

13. $f(x) = \dfrac{\sqrt{2x - 1}}{(x - 1)^2}$

14. $f(x) = \dfrac{\sqrt{x^2 + 1}}{2 + \sqrt{x^2 + 1}}$

15. $f(s) = \sqrt[n]{s} + \sqrt[n]{s}$

16. $g(t) = \sqrt{t + \sqrt{t + 1}}$

17. $g(t) = \left(\dfrac{t}{t - 3}\right)^3$

18. $h(s) = \left(\dfrac{2s^2}{s + 1}\right)^4$

19. $f(r) = (r^2 - r)^3(r + 3r^3)^{-4}$

20. $h(s) = \dfrac{2(3 - s)^2}{s^2 + (7s - 1)^2}$

21. $h(x) = \sqrt[5]{3 - x^4}$

22. $h(x) = \sqrt[3]{1 + 2x}$

23. $f(x) = \sqrt[7]{x^2 - 2x + 1}$

24. $f(x) = \sqrt[4]{2 + 4x^2}$

25. $g(s) = (3s^7 - 7s)^{3/2}$

26. $h(t) = (t^4 - 5t)^{5/2}$

27. $h(t) = \left(3t + \dfrac{3}{t}\right)^{2/5}$

28. $h(t) = \left(4t^4 + \dfrac{4}{t^4}\right)^{1/4}$

29. Differentiate

$$f(x) = (ax + 1)^3$$

with respect to x. Assume that a is a positive constant.

30. Differentiate

$$f(x) = \sqrt{ax^2 - 2}$$

with respect to x. Assume that a is a positive constant.

31. Differentiate

$$g(N) = \frac{bN}{k + N}$$

with respect to N. Assume that b and k are positive constants.

32. Differentiate

$$g(N) = \frac{N}{(k + bN)^2}$$

with respect to N. Assume that b and k are positive constants.

33. Differentiate

$$g(T) = a(T_0 - T)^3 - b$$

with respect to T. Assume that a, b, and T_0 are positive constants.

34. Suppose that $f'(x) = 2x + 1$. Find the following:

(a) $\dfrac{d}{dx} f(x^2)$ at $x = -1$

(b) $\dfrac{d}{dx} f(\sqrt{x})$ at $x = 4$

35. Suppose that $f'(x) = \frac{1}{x}$. Find the following:

(a) $\dfrac{d}{dx} f(x^2 + 3)$

(b) $\dfrac{d}{dx} f(\sqrt{x - 1})$

In Problems 36–39, assume that $f(x)$ and $g(x)$ are differentiable.

36. Find $\frac{d}{dx} f(2x)$.

37. Find $\frac{d}{dx}\left(\frac{f(x)}{g(x+1)}\right)$.

38. Find $\frac{d}{dx} f[g(x)+1]$.

39. Find $\frac{d}{dx}\left(\frac{f(2x)}{g(2x)+2x}\right)$.

In Problems 40–46, find $\frac{dy}{dx}$ by applying the chain rule repeatedly.

40. $y = (\sqrt{1-2x^2}+1)^2$

41. $y = (\sqrt{x^3-3x}+3x)^4$

42. $y = \left(1+2(x+3)^4\right)^2$

43. $y = \left(1+(3x^2-1)^3\right)^2$

44. $y = \left(\frac{x}{2(x^2-1)^2-1}\right)^2$

45. $y = \left(\frac{2x+1}{3(x^3-1)^3-1}\right)^3$

46. $y = \left(\frac{(2x+1)^2-x}{(3x^3+1)^3-x}\right)^2$

47. Chewing in Mammals ***Druzinsky (1993)*** showed that chewing frequency (i.e., the number of mass an animal chews in one minute) is proportional to its body mass raised to the power of -0.128, that is, if M is the body mass and c the chewing frequency. then,

$$c = kM^{-0.128}$$

for some positive constant k.

(a) Assume that the body mass of a particular mammal is given by a formula $M(t) = 1+2\sqrt{t}$. Calculate dc/dt.

(b) Druzinsky also found that the jaw length of the animal, L, is proportional to its body mass raised to the 0.312 power, i.e., $L = rM^{0.312}$ where r is another positive constant. Calculate dc/dL, the rate of change of chewing frequency with jaw length.

48. Wing Beat Frequency in Hummingbirds ***Altshuler and Dudley (2003)*** found that hummingbirds' wing beat frequency, f, decreases with body mass, m, according to

$$f = 40 - \frac{8}{5}m \tag{4.5}$$

where f is measured in beats per second and m in grams. Assume that the amount of thrust that a flying hummingbird can generate depends on its mass and wing beat frequency as follows

$$T = cf^2m^{4/3}$$

for some positive constant c. (This equation is derived from the thrust mechanics of a moving wing.)

(a) Equation (4.5) should only be used if $m < 25$. Why? (The largest hummingbirds have a mass of 22g.)

(b) Calculate dT/dm.

(c) Show by plotting dT/dm that for larger hummingbirds, T decreases with m. That is, show that $dT/dm < 0$, once m exceeds a critical threshold.

(d) For a hummingbird to fly, its thrust must exceed its weight $W = mg$ (where g is the acceleration due to gravity). Explain, using your answer to (d), why if the hummingbird's mass increases, then T will eventually be smaller than W (Hint: $dW/dm = g$ is constant and positive), no matter what the value of c is.

4.6 Implicit Functions and Implicit Differentiation

4.6.1 Implicit Differentiation

So far, we have considered only functions of the form $y = f(x)$, which define y *explicitly* as a function of x. It is also possible to define y *implicitly* as a function of x.

Consider, for example, the equation for a circle of radius r:

$$x^2 + y^2 = r^2$$

y is still given as a function of x (i.e., y can be thought of as the dependent variable). To calculate the derivative dy/dx we could rewrite the equation as:

$$y = \sqrt{r^2 - x^2}$$

and then differentiate using the rules from Section 4.5. But rearranging the formula is not always possible. For example, the equation:

$$5x^2 + 5y^2 - 6\sqrt{2}xy = 8$$

represents an ellipse with semi-axes 1 and 2 that is tilted (see Figure 4.24). In this case there is no way to solve for y as a function of x.

Fortunately, there is a very useful technique, based on the chain rule, that will allow us to find dy/dx for implicitly defined functions. This technique is called **implicit differentiation**. We explain the procedure in the next example.

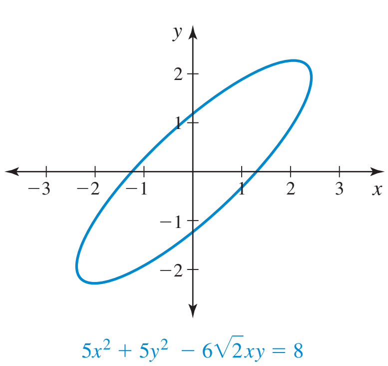


Figure 4.24 The ellipse $5x^2 + 5y^2 - 6\sqrt{2}xy = 8$ is defined implicitly.

EXAMPLE 1 Find $\frac{dy}{dx}$ if $x^2 + y^2 = 1$.

Solution Remembering that y is a function of x, we differentiate both sides of the equation $x^2 + y^2 = 1$ with respect to x:

$$\frac{d}{dx}\left(x^2 + y^2\right) = \frac{d}{dx}(1)$$

$$\frac{d}{dx}(x^2) + \frac{d}{dx}(y^2) = \frac{d}{dx}(1). \qquad \text{Use sum rule from Section 4.3}$$

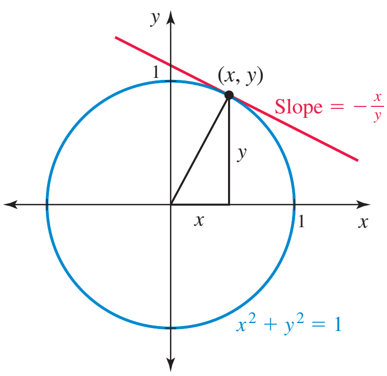


Figure 4.25 The slope of the tangent line at the unit circle $x^2 + y^2 = 1$ at (x, y) is $m = -\frac{x}{y}$.

Starting with the left-hand side and using the power rule, we have $\frac{d}{dx}(x^2) = 2x$. To differentiate y^2 with respect to x, we apply the chain rule to get $\frac{d}{dx}(y^2) = 2y\frac{dy}{dx}$. On the right-hand side $\frac{d}{dx}(1) = 0$. We therefore have

$$2x + 2y\frac{dy}{dx} = 0$$

or:

$$\frac{dy}{dx} = -\frac{2x}{2y} = -\frac{x}{y} \qquad \text{Isolate } dy/dx$$

Since $x^2 + y^2 = 1$ is the equation for the unit circle centered at the origin (Figure 4.25), we can use a geometric argument to convince ourselves that we have indeed obtained the correct derivative. The line that connects $(0, 0)$ and (x, y) has slope y/x and is perpendicular to the tangent line at (x, y). Since the slopes of perpendicular lines are negative reciprocals of each other, the slope of the tangent line at (x, y) must be $-x/y$. •

We summarize the steps we take to find dy/dx when an equation defines y implicitly as a differentiable function of x:

> **STEP 1.** Differentiate both sides of the equation with respect to x, keeping in mind that y is a function of x.
>
> **STEP 2.** Solve the resulting equation for dy/dx.

Note that differentiating terms involving y typically requires the chain rule.

In the previous example we could have rearranged the formula $x^2 + y^2 = 1$ to give y as an explicit function of x, $y = \sqrt{1 - x^2}$, and we can calculate dy/dx from this explicit formula. For the next example, we cannot easily rearrange the formula.

EXAMPLE 2 Find dy/dx when $5x^2 + 5y^2 - 6\sqrt{2}xy = 8$.

Solution This equation defines the ellipse shown in Figure 4.24 implicitly. We follow Steps 1 and 2 above to calculate dy/dx.

$$\frac{d}{dx}(5x^2) + \frac{d}{dx}(5y^2) - \frac{d}{dx}(6\sqrt{2}xy) = \frac{d}{dx}(8). \qquad \text{Differentiate both sides with respect to } x$$

On the left-hand side we have:

$$\frac{d}{dx}(5x^2) = 10x \qquad \text{Use rules from Section 4.3}$$

$$\frac{d}{dx}(5y^2) = 10y\frac{dy}{dx} \qquad \text{Chain rule with } f(y) = 5y^2,\ \tfrac{d}{dx}(5y^2) = \tfrac{df}{dy}\cdot\tfrac{dy}{dx}$$

$$\frac{d}{dx}(6\sqrt{2}xy) = 6\sqrt{2}\left(\left(\frac{d}{dx}x\right)y + x\frac{dy}{dx}\right) \qquad \text{Product rule}$$

$$= 6\sqrt{2}\left(y + x\frac{dy}{dx}\right).$$

Since on the right-hand side $\frac{d}{dx}(8) = 0$, we put the pieces together to obtain:

$$10x + 10y\frac{dy}{dx} - 6\sqrt{2}y - 6\sqrt{2}x\frac{dy}{dx} = 0$$

$$10x - 6\sqrt{2}y + (10y - 6\sqrt{2}x)\frac{dy}{dx} = 0 \qquad \text{Factor } \tfrac{dy}{dx}$$

$$\frac{dy}{dx} = \frac{6\sqrt{2}y - 10x}{10y - 6\sqrt{2}x}. \qquad \text{Solve for } \tfrac{dy}{dx}.$$

The next example prepares us for the power rule for rational exponents.

EXAMPLE 3 Find $\frac{dy}{dx}$ when $y^2 = x^3$. Assume that $x > 0$ and $y > 0$.

Solution We differentiate both sides with respect to x:

$$\frac{d}{dx}(y^2) = \frac{d}{dx}(x^3)$$

$$2y\frac{dy}{dx} = 3x^2. \qquad \text{Chain rule on left hand-side.}$$

Therefore,

$$\frac{dy}{dx} = \frac{3}{2}\frac{x^2}{y}.$$

Since $y = x^{3/2}$, it follows that

$$\frac{dy}{dx} = \frac{3}{2}\frac{x^2}{y} = \frac{3}{2}\frac{x^2}{x^{3/2}} = \frac{3}{2}x^{1/2}.$$

This is the answer we expect from the general version of the power rule:

$$\frac{dy}{dx} = \frac{d}{dx}\left(x^{3/2}\right) = \frac{3}{2}x^{1/2}. \qquad r = 3/2 \text{ in power rule}$$

Proof of the Power Rule for Rational Exponents We can generalize Example 3 to functions of the form $y = x^r$, where r is a rational number. This will provide a proof of the generalized form of the power rule when the exponent is a rational number, something we promised in Section 4.4. We write $r = p/q$, where p and q are integers and are in lowest terms. (If q is even, we require x and y to be positive.) Then

$$y = x^{p/q} \quad \Longleftrightarrow \quad y^q = x^p$$

Differentiating both sides of $y^q = x^p$ with respect to x, we find that

$$qy^{q-1}\frac{dy}{dx} = px^{p-1}$$

Hence,

$$\begin{aligned}\frac{dy}{dx} &= \frac{p}{q}\frac{x^{p-1}}{y^{q-1}} = \frac{p}{q}\frac{x^{p-1}}{(x^{p/q})^{q-1}} \\ &= \frac{p}{q}\frac{x^{p-1}}{x^{p-p/q}} = \frac{p}{q}x^{p-1-p+p/q} \\ &= \frac{p}{q}x^{p/q-1} = rx^{r-1}.\end{aligned}$$

To summarize the preceding result: If r is a rational number, then

$$\frac{d}{dx}(x^r) = rx^{r-1}$$

4.6.2 Related Rates

An important application of implicit differentiation is related-rates problems. We begin with a motivating example.

Consider a parcel of air rising quickly in the atmosphere. As the parcel rises it expands and its temperature changes. Laws of physics tell us that if the parcel does not exchange heat with the surrounding air the volume, V, and the temperature, T, of the parcel of air are related via the formula

$$TV^{\gamma-1} = C \tag{4.6}$$

where γ (lowercase Greek gamma) is approximately 1.4 for sufficiently dry air and C is a constant. The temperature is measured in Kelvin,[1] a scale chosen so that the temperature is always positive. (The Kelvin scale is the absolute temperature scale.) Rising air expands, so the volume, V, of the parcel increases with time, which we denote by t. How does the rate of change of the temperature, dT/dt, depend on the rate of change of volume, dV/dt?

To determine how the temperature of the air parcel changes as it rises, we implicitly differentiate $TV^{\gamma-1} = C$ with respect to t:

$$\frac{dT}{dt}V^{\gamma-1} + T(\gamma-1)V^{\gamma-2}\frac{dV}{dt} = 0 \qquad \text{Product rule and chain rule}$$

or

$$\frac{dT}{dt} = -\frac{T(\gamma-1)V^{\gamma-2}\frac{dV}{dt}}{V^{\gamma-1}} = -T(\gamma-1)\frac{1}{V}\frac{dV}{dt} \qquad \text{Isolate } \tfrac{dT}{dt}.$$

If we use $\gamma = 1.4$, then

$$\frac{dT}{dt} = -0.4 \cdot T\frac{1}{V}\frac{dV}{dt}$$

implying that if air expands (i.e., $dV/dt > 0$), then temperature decreases (i.e., $dT/dt < 0$), since both T and V are positive: The temperature of a parcel of air decreases as the parcel rises, and the temperature of a falling air parcel increases. These phenomena can be observed close to high mountains.

In a typical related-rates problem, one quantity is expressed in terms of another quantity; for example, y is a function of x, $y = f(x)$. Both quantities change with time. If we know how x changes with time (i.e., we know the rate of change $\frac{dx}{dt}$) then we may wish to calculate the rate of change of y, $\frac{dy}{dt}$. The rate of change $\frac{dy}{dt}$ will be related to the rate of change $\frac{dx}{dt}$, which is why these problems are called related-rates problems. We will first discuss an example that shows how to solve related-rates problems mathematically, and then give two applications of related rates to biological problems.

EXAMPLE 4 If $x^2 + y^3 = 1$ find dy/dt at $x = \sqrt{7/8}$ if $dx/dt = 2$.

Solution In this example, both x and y are functions of t. We can imagine that the curve $x^2 + y^3 = 1$ represents a road that a car is traveling along. At time t, the car is located at a point $(x, y) = (x(t), y(t))$. Implicit differentiation with respect to t yields

$$\frac{d}{dt}(x^2 + y^3) = \frac{d}{dt}(1)$$

Hence,

$$2x\frac{dx}{dt} + 3y^2\frac{dy}{dt} = 0$$

So

$$\frac{dy}{dt} = -\frac{2}{3}\frac{x}{y^2}\frac{dx}{dt} \qquad \text{Solve for } dy/dt$$

(1) To compare the Celsius and the Kelvin scales, note that a temperature difference of 1°C is equal to a temperature difference of 1 K, and that 0°C = 273.15 K and 100°C = 373.15 K.

When $x = \sqrt{7/8}$,

$$y^3 = 1 - x^2 = 1 - \frac{7}{8} = \frac{1}{8} \qquad \text{So } y = \tfrac{1}{2}$$

Therefore:

$$\frac{dy}{dt} = -\frac{2}{3}\frac{\sqrt{7/8}}{1/4} \cdot 2 = -\frac{16}{3}\sqrt{\frac{7}{8}} = -\frac{4}{3}\sqrt{14} \qquad \text{Substitute } x = \sqrt{7/8},\ y = 1/2$$

•

EXAMPLE 5 **Yeast Cell Blebbing** When a yeast cell divides a small bud (usually called a bleb) forms on the cell. This bleb grows with time until it eventually separates from the parent cell. To create the bleb the parent cell must make extra cytoplasm, to fill the inside of the bleb, and extra cell wall material, to make its surface. How is the amount of surface related to the amount of volume? Specifically, if the volume, V, increases at a rate dV/dt, how quickly does the surface area, S, increase?

Solution We will approximate the growing bleb as a sphere. If the spherical bleb has radius r, then its volume and surface area are given by formulas:

$$V = \frac{4}{3}\pi r^3 \quad \text{and} \quad S = 4\pi r^2.$$

So S is related to V by an equation:

$$S = 4\pi\left(\frac{3V}{4\pi}\right)^{2/3} \qquad V = \tfrac{4}{3}\pi r^3 \Rightarrow r = \left(\tfrac{3V}{4\pi}\right)^{1/3}, \text{ then substitute for } r \text{ in } S = 4\pi r^2.$$

$$= \left(4\pi(3V)^2\right)^{1/3} = \left(36\pi V^2\right)^{1/3}.$$

Differentiating both sides with respect to t:

$$\frac{dS}{dt} = (36\pi)^{1/3} \cdot \frac{2}{3V^{1/3}}\frac{dV}{dt}$$

Although we are asked to determine $\frac{dS}{dt}$ in terms of $\frac{dV}{dt}$, another way to present this relationship is to calculate the fractional increase in bleb surface in one unit of time, $\frac{1}{S}\frac{dS}{dt}$, as a function of the fractional rate of increase in volume, $\frac{1}{V}\frac{dV}{dt}$. Note that S and V are related by a power law:

$$S = k \cdot V^{2/3} \qquad k = (36\pi)^{1/3}$$

So, differentiating both sides of this equation with respect to time:

$$\frac{dS}{dt} = k \cdot \frac{2}{3}V^{-1/3}\frac{dV}{dt} = \frac{2}{3}\frac{kV}{V} \cdot \frac{dV}{dt}$$

$$= \frac{2}{3}\frac{S}{V}\frac{dV}{dt} \qquad S = kV^{2/3}$$

so

$$\frac{1}{S}\frac{dS}{dt} = \frac{2}{3}\frac{1}{V}\frac{dV}{dt}.$$

So the fractional rate of increase of S is directly proportional to the fraction rate of increase of V.

•

Notice that the coefficient of proportionality is 2/3, which is also the exponent in the power law $S = k \cdot V^{2/3}$. In Problem 26 you will show that if any two quantities y and x are related by a power law, $y = kx^a$, then $\frac{1}{y}\frac{dy}{dt} = \frac{a}{x}\frac{dx}{dt}$. The next example shows how this rule can be applied to the study of allometric equations. (Allometric equations were introduced in Example 7 of Section 1.3.)

EXAMPLE 6 **Allometric Growth** Ichthyosaurs are a group of extract marine reptiles that were fish shaped and comparable in size to dolphins. On the basis of a study of 20 fossil skeletons, ***Benton and Harper (1997)*** found that the skull length (in cm) and backbone length (in cm) of an individual ichthyosaur were related through the allometric equation:

$$[\text{skull length}] = 1.162[\text{backbone length}]^{0.933}$$

How is the growth rate of the backbone related to the growth rate of the skull?

Solution Let x denote the age of the ichthyosaur, and set

$$S = S(x) = \text{skull length at age } x$$
$$B = B(x) = \text{backbone length at age } x$$

so that

$$S(x) = (1.162)[B(x)]^{0.933}.$$

We are interested in the relationship between dS/dx and dB/dx, the growth rates of the skull and the backbone, respectively. Differentiating the equation for $S(x)$ with respect to x, we find that

$$\frac{dS}{dx} = (1.162)(0.933)[B(x)]^{0.933-1}\frac{dB}{dx}$$

Just as in Example 5, we rearrange this equation as in terms of fractional rates of growth.

$$\frac{dS}{dx} = \underbrace{(1.162)[B(x)]^{0.933}}_{S(x)}(0.933)\frac{1}{B(x)}\frac{dB}{dx}$$

Hence,

$$\frac{1}{S(x)}\frac{dS}{dx} = 0.933\frac{1}{B(x)}\frac{dB}{dx}$$

The factor 0.933 is less than 1, which indicates that skulls grow less quickly than backbones. In fact, in most vertebrates the heads of young animals are larger relative to their body size than older animals. •

Section 4.6 Problems

4.6.1

In Problems 1–8, find $\frac{dy}{dx}$ by implicit differentiation.

1. $x^2 + y^2 = 4$

2. $y = x^2 + 3yx$

3. $x^{3/4} + y^{3/4} = 1$

4. $xy - y^3 = 1$

5. $\sqrt{xy} = x^2 + 1$

6. $\frac{1}{2xy} - y^3 = 4$

7. $\frac{x}{y} = \frac{y}{x}$

8. $\frac{x}{xy+1} = 2xy$

In Problems 9–11, find the lines that are (a) tangential and (b) normal to each curve at the given point.

9. $x^2 + y^2 = 25$, $(4, -3)$ (circle)

10. $\frac{x^2}{4} + \frac{y^2}{9} = 1$, $(1, \frac{3}{2}\sqrt{3})$ (ellipse)

11. $\frac{x^2}{25} - \frac{y^2}{9} = 1$, $(\frac{25}{3}, 4)$ (hyperbola)

T **12. Lemniscate**

(a) The curve with equation $y^2 = x^2 - x^4$ is shaped like the numeral eight. Find $\frac{dy}{dx}$ at $(\frac{1}{2}, \frac{1}{4}\sqrt{3})$.

(b) Use a graphing calculator to graph the curve in (a). If the calculator cannot graph implicit functions, graph the upper and the lower halves of the curve separately; that is, graph

$$y_1 = \sqrt{x^2 - x^4}$$
$$y_2 = -\sqrt{x^2 - x^4}$$

Choose the viewing rectangle $-2 \le x \le 2$, $-1 \le y \le 1$.

T **13. Astroid**

(a) Consider the curve with equation $x^{2/3} + y^{2/3} = 4$. Find $\frac{dy}{dx}$ at $(-1, 3\sqrt{3})$.

(b) Use a graphing calculator to graph the curve in (a). If the calculator cannot graph implicit functions, graph the upper and the lower halves of the curve separately. To get the left half of the graph, make sure that your calculator evaluates $x^{2/3}$ in the order $(x^2)^{1/3}$. Choose the viewing rectangle $-10 \le x \le 10$, $-10 \le y \le 10$.

T **14. Kampyle of Eudoxus**

(a) Consider the curve with equation $y^2 = 10x^4 - x^2$. Find $\frac{dy}{dx}$ at $(1, 3)$.

(b) Use a graphing calculator to graph the curve in (a). If the calculator cannot graph implicit functions, graph the upper and the lower halves of the curve separately. Choose the viewing rectangle $-3 \le x \le 3$, $-10 \le y \le 10$.

4.6.2

15. Assume that x and y are differentiable functions of t. Find $\frac{dy}{dt}$ when $x^2 + y^2 = 1$, $\frac{dx}{dt} = 2$ for $x = \frac{1}{2}$, and $y > 0$.

16. Assume that x and y are differentiable functions of t. Find $\frac{dy}{dt}$ when $y^2 + (x+1)^2 = 1$, $\frac{dx}{dt} = 1$ for $x = -\frac{1}{2}$, and $y > 0$.

17. Assume that x and y are differentiable functions of t. Find $\frac{dy}{dt}$ when $x^2 y = 1$ and $\frac{dx}{dt} = 3$ for $x = 2$.

18. Assume that u and v are differentiable functions of t. Find $\frac{du}{dt}$ when $u^2 + v^4 = 12$, $\frac{dv}{dt} = 2$ for $v = 1$, and $u > 0$.

19. Assume that the side length x and the volume $V = x^3$ of a cube are differentiable functions of t. Express dV/dt in terms of dx/dt.

20. Assume that the radius r and the area $A = \pi r^2$ of a circle are differentiable functions of t. Express dA/dt in terms of dr/dt.

21. Assume that the radius r and the surface area $S = 4\pi r^2$ of a sphere are differentiable functions of t. Express dS/dt in terms of dr/dt.

22. Assume that the radius r and the volume $V = \frac{4}{3}\pi r^3$ of a sphere are differentiable functions of t. Express dV/dt in terms of dr/dt.

23. Suppose that water is stored in a cylindrical tank of radius 5 m. If the height of the water in the tank is h, then the volume of the water is $V = \pi r^2 h = (25\text{m}^2)\pi h = 25\pi h\,\text{m}^2$. If we drain the water at a rate of 250 liters per minute, what is the rate at which the water level inside the tank drops? (Note that 1 cubic meter contains 1000 liters.)

24. Suppose that we pump water into an inverted right circular conical tank at the rate of 5 cubic feet per minute (i.e., the tank stands with its point facing downward). The tank has a height of 6 ft and the radius on top is 3 ft. What is the rate at which the water level is rising when the water is 2 ft deep? (Note that the volume of a right circular cone of radius r and height h is $V = \frac{1}{3}\pi r^2 h$.)

25. Two people start biking from the same point. One heads east at 15 mph, the other south at 18 mph. What is the rate at which the distance between the two people is changing after 20 minutes and after 40 minutes?

26. Skull Size Allometric equations describe the scaling relationship between two measurements, such as skull length versus body length. In vertebrates, we typically find that

$$[\text{skull length}] \propto [\text{body length}]^a$$

for $0 < a < 1$. Express the growth rate of the skull length in terms of the growth rate of the body length.

27. Metabolism ***West, Brown, and Enquist (1997)*** argued that because of the distribution of blood vessels through mammalian bodies, the energy needs E of mammals increase with the 3/4 power of their mass, M; i.e.,

$$E = cM^{3/4}$$

for some constant c.

As a mammal grows, M increases. Show how dE/dt is related to dM/dt according to the theory of West, Brown, and Enquist.

28. Tree Growth Rate ***Sperry et al. (2012)*** studied how the growth rates G of trees depend upon their body mass, M. They argued that $G = CM^{0.7}$ for some constant C. As the tree grows, M changes.

(a) Show how dG/dt is related to dM/dt.

(b) Show how the fractional rate of increase of G, $\frac{1}{G}\frac{dG}{dt}$, is related to the fractional rate of growth $\frac{1}{M}\frac{dM}{dt}$.

29. A Model for Bacterial Growth Bacteria, like yeast cells (see Example 5), must create new cell walls as they grow. Consider a rod-like bacterium, which you may model as a cylinder of radius a and length L. Over time the bacterium grows, that is, L increases, but a stays constant. (This is approximately true in many bacteria.) Relate the rate of surface area increase, $\frac{dS}{dt}$, to the rate of volume increase, $\frac{dV}{dt}$.

30. Allometric Equations Suppose that two quantities, y and x are related by a power law:

$$y = kx^a$$

where k and a are both constants. x grows with time at a rate dx/dt.

(a) Explain why $\frac{1}{x}\frac{dx}{dt}$ can be thought of as the relative rate of growth of x.

(b) Show that the relative rates of growth of y and x are related by an equation:

$$\frac{1}{y}\frac{dy}{dt} = \frac{a}{x}\frac{dx}{dt}$$

4.7 Higher Derivatives

The derivative of a function f is itself a function. We refer to this derivative as the **first derivative**, denoted f'. If the first derivative exists, we say that the function is differentiable. Given that the first derivative is a function, we can define its derivative (where it exists). This derivative is called the **second derivative** and is denoted f''. If the second derivative exists, we say that the original function is twice differentiable. This second derivative is again a function; hence, we can define its derivative (where it exists). The result is the **third derivative**, denoted f'''. If the third derivative exists, we say that the original function is three times differentiable. We can continue in this manner; from the fourth derivative on, we denote the derivatives by $f^{(4)}$, $f^{(5)}$, and so on. If the nth derivative exists, we say that the original function is n times differentiable. We denote the nth derivative either by $f^{(n)}$ or by $f'\cdots'$ (f followed by n $'$ symbols).

Polynomials are functions that can be differentiated as many times as desired. The reason is that the first derivative of a polynomial of degree n is a polynomial of degree

$n-1$. Since the derivative is a polynomial as well, we can find its derivative, and so on. Eventually, the derivative will be equal to 0, as is illustrated in the next example.

EXAMPLE 1 Find the nth derivative of $f(x) = x^5$ for $n = 1, 2, \ldots$.

Solution We calculate the nth derivative of $f(x)$ by differentiating n times. We will find that we can use the power rule to form each new derivative. We find the first derivative by differentiating $f(x)$:

$$f'(x) = 5x^4 \qquad \text{Power rule with exponent } r = 5$$

By differentiating $f'(x)$, we find the second derivative:

$$f''(x) = 5(4x^3) = 20x^3$$

By differentiating $f''(x)$, we find the third derivative:

$$f'''(x) = 20(3x^2) = 60x^2$$

By differentiating $f'''(x)$, we find the fourth derivative:

$$f^{(4)}(x) = 60(2x) = 120x$$

By differentiating $f^{(4)}(x)$, we find the fifth derivative:

$$f^{(5)}(x) = 120$$

By differentiating $f^{(5)}(x)$, we find the sixth derivative:

$$f^{(6)}(x) = 0$$

All higher-order derivatives—that is, $f^{(7)}, f^{(8)}, \ldots$—are equal to 0 as well. •

We can also write higher-order derivatives in Leibniz notation: The nth derivative of $f(x)$ is denoted by

$$\frac{d^n f}{dx^n}.$$

EXAMPLE 2 Find the second derivative of $f(x) = \sqrt{x}, x \geq 0$.

Solution First, we find the first derivative:

$$\frac{d}{dx}\left(\sqrt{x}\right) = \frac{d}{dx}x^{1/2} = \frac{1}{2}x^{-1/2} \quad \text{for } x > 0. \qquad \text{Power rule with } r = 1/2$$

To find the second derivative, we differentiate the first derivative

$$\frac{d^2}{dx^2}\left(\sqrt{x}\right) = \frac{d}{dx}\left(\frac{d}{dx}\sqrt{x}\right) = \frac{d}{dx}\left(\frac{1}{2}x^{-1/2}\right) = \frac{1}{2}\left(-\frac{1}{2}\right)x^{(-1/2)-1}$$

$$= -\frac{1}{4}x^{-3/2}. \qquad \text{for } x > 0$$ •

When functions are implicitly defined, we can use the technique of implicit differentiation to find higher derivatives.

EXAMPLE 3 Find $\frac{d^2y}{dx^2}$ when $x^2 + y^2 = 1$.

Solution We found

$$\frac{dy}{dx} = -\frac{x}{y}$$

in Example 1 of Section 4.6. Differentiating both sides of this equation with respect to x, we get

$$\frac{d}{dx}\left[\frac{dy}{dx}\right] = \frac{d}{dx}\left[-\frac{x}{y}\right]$$

The left-hand side can be written as $\frac{d^2y}{dx^2}$. Hence,

$$\begin{aligned}
\frac{d^2y}{dx^2} &= -\frac{1 \cdot y - x \cdot \frac{dy}{dx}}{y^2} && \text{Use quotient rule on right-hand side} \\
&= -\frac{y - x(-\frac{x}{y})}{y^2} && \text{Use } \tfrac{dy}{dx} = -\tfrac{x}{y} \text{ on right-hand side} \\
&= -\frac{y + \frac{x^2}{y}}{y^2} = -\frac{y^2 + x^2}{y^3}.
\end{aligned}$$

Since $x^2 + y^2 = 1$, we can simplify the rightmost expression further and obtain

$$\frac{d^2y}{dx^2} = -\frac{1}{y^3}.$$

An alternate way to derive this equation uses implicit differentiation to avoid having to use the quotient rule.

$$\begin{aligned}
x^2 + y^2 &= 1 \\
\frac{d}{dx}(x^2 + y^2) &= 0 && \text{Differentiate both sides} \\
2x + 2y\frac{dy}{dx} &= 0. && \text{So } \tfrac{dy}{dx} = -\tfrac{x}{y}\text{, as before}
\end{aligned}$$

So:

$$\begin{aligned}
2\frac{d}{dx}\left(x + y\frac{dy}{dx}\right) &= 0. && \text{Differentiate both sides again} \\
1 + \frac{d}{dx}\left(y\frac{dy}{dx}\right) &= 0 && \text{Divide both sides by 2} \\
1 + \frac{dy}{dx} \cdot \frac{dy}{dx} + y\frac{d^2y}{dx^2} &= 0 \\
\Rightarrow \frac{d^2y}{dx^2} &= -\frac{1}{y}\left(1 + \left(\frac{dy}{dx}\right)^2\right) && \text{Solve for } d^2y/dx^2 \\
&= -\frac{1}{y}\left(1 + \frac{x^2}{y^2}\right) && \tfrac{dy}{dx} = -\tfrac{x}{y} \\
&= \frac{-1}{y} \cdot \frac{x^2 + y^2}{y^2} = \frac{-1}{y^3}. && x^2 + y^2 = 1
\end{aligned}$$

•

Why would we need to calculate higher derivatives? In Chapter 5 we will show how the second derivative of a function can help us to sketch its graph. The second derivative is also important for optimization problems (that is, when maximizing or minimizing a function), which will also be discussed in Chapter 5.

For now we will discuss one example: We defined the velocity of an object that moves on a straight line to be the derivative of the object's position. The derivative of the velocity is the **acceleration**. If $s(t)$ denotes the position of an object moving on a straight line, $v(t)$ its velocity, and $a(t)$ its acceleration, then the three quantities are related as follows:

$$v(t) = \frac{ds}{dt} \qquad \text{and} \qquad a(t) = \frac{dv}{dt} = \frac{d^2s}{dt^2}.$$

EXAMPLE 4 **Acceleration** Assume that the position of a car moving along a straight line is given by

$$s(t) = 3t^3 - 2t + 1$$

Find the car's velocity and acceleration.

Solution To find the velocity, we need to differentiate the position:

$$v(t) = \frac{ds}{dt} = 9t^2 - 2.$$

To find the acceleration, we differentiate the velocity:

$$a(t) = \frac{dv}{dt} = \frac{d^2s}{dt^2} = 18t.$$

•

EXAMPLE 5 Neglecting air resistance, we find that the distance (in meters) an object falls when dropped from rest from a height is

$$s(t) = \frac{1}{2}gt^2$$

where $g = 9.81\text{m/s}^2$ is the earth's gravitational constant and t is the time (in seconds) elapsed since the object was released.

(a) Find the object's velocity and acceleration.

(b) If an object is dropped from a height of 30 m, how long will it take until the object hits the ground, and what is its velocity at the time of impact?

Solution **(a)** The velocity is

$$v(t) = \frac{ds}{dt} = gt$$

and the acceleration is

$$a(t) = \frac{dv}{dt} = g.$$

Note that the acceleration is constant.

(b) To find the time it takes the object to hit the ground, we solve for t:

$$s = \frac{1}{2}gt^2$$

$$\Rightarrow \ t = \sqrt{\frac{2s}{g}}$$

Putting $s = 30$ m and $g = 9.81\ \text{ms}^{-2}$;

$$t = \sqrt{\frac{60}{9.81}}\ \text{s} \approx 2.47\ \text{s}$$

(We need consider only the positive solution). The velocity at the time of impact is then

$$v(t) = gt = 9.81\ \text{m/s}^2 \times 2.47\ \text{s} \ \approx 24.3\ \text{m/s}.$$

•

Section 4.7 Problems

4.7

In Problems 1–10, find the first and the second derivatives of each function.

1. $f(x) = x^3 - 3x^2 + 1$

2. $f(x) = (2x + 4)^3$

3. $g(x) = \dfrac{1}{x+1}$

4. $h(s) = \dfrac{1}{s^2 + 2}$

5. $g(t) = \sqrt{3t^3 + 2t}$

6. $f(x) = \dfrac{1}{x^2} + x - x^3$

7. $f(s) = s^{3/2}$

8. $f(x) = \dfrac{2x}{x^2 + 1}$

9. $g(t) = t^{-5/2} - t^{1/2}$

10. $f(x) = x^3 + \dfrac{1}{x^3}$

11. Find the first 10 derivatives of $y = x^6$.

12. Find $f^{(n)}(x)$ and $f^{(n+1)}(x)$ if $f(x) = x^n$.

13. Find a second-degree polynomial $p(x) = ax^2 + bx + c$ with $p(0) = 3$, $p'(0) = 2$, and $p''(0) = 6$.

14. The position at time t of a particle that moves along a straight line is given by the function $s(t)$. The first derivative of $s(t)$ is called the velocity, denoted by $v(t)$; that is, the velocity is the rate of change of the position. The rate of change of the velocity is called **acceleration**, denoted by $a(t)$; that is,

$$\frac{d}{dt}v(t) = a(t)$$

Given that $v(t) = s'(t)$, it follows that

$$\frac{d^2}{dt^2}s(t) = a(t)$$

Find the velocity and the acceleration at time $t = 1$ for the following position functions:

(a) $s(t) = t^2 - 3t$ **(b)** $s(t) = \sqrt{t^2 + 1}$ **(c)** $s(t) = t^4 - 2t$.

15. Neglecting air resistance, the height h (in meters) of an object thrown vertically from the ground with initial velocity v_0 is given by

$$h(t) = v_0 t - \frac{1}{2} g t^2$$

where $g = 9.81\ \mathrm{m/s^2}$ is the earth's gravitational constant and t is the time (in seconds) elapsed since the object was released.

(a) Find the velocity and the acceleration of the object.

(b) Find the time when the velocity is equal to 0. In which direction is the object traveling right before this time? in which direction right after this time?

16. Oxygen Binding by Hemoglobin The fraction of hemoglobin, f, that is bound to oxygen depends on the concentration of oxygen P. This relationship is often modeled by Hill's equation,

$$f(P) = \frac{P^m}{k^m + P^m}$$

where k and m are positive constants that vary from species to species and on whether the animal lives at sea level or at a higher altitude.

(a) Show that $f''(0) > 0$ if $m = 2$

(b) Show that $f''(0) < 0$ if $m = 1$

In Chapter 5 we will show how these observations can be related to the different shapes that the graph of $f(P)$ against P has for different values of m.

17. Interatomic Forces Two atoms are modeled as interacting via a Lennard–Jones 6–12 potential. That is, the energy of interaction, V, depends on their spacing, r, according to a formula.

$$V(r) = \frac{a}{r^{12}} - \frac{b}{r^6},\ r > 0,$$

where a and b are positive constants.

(a) The force between the atoms can be shown to be given by $F(r) = -\frac{dV}{dr}$. When the atoms are in equilibrium, $F(r) = 0$. Find the equilibrium spacing of the pair of atoms, r.

(b) Generally the equilibrium spacing of a pair of atoms interacting with energy is stable if $V''(r) > 0$, and unstable if $V''(r) < 0$. For the Lennard–Jones 6–12 potential, show that the equilibrium spacing that you calculated in (a) is stable.

4.8 Derivatives of Trigonometric Functions

We will need the trigonometric limits from Section 3.4 to compute the derivatives of the sine and cosine functions. Note that all angles are measured in radians.

Theorem The functions $\sin x$ and $\cos x$ are differentiable for all x, and

$$\frac{d}{dx}\sin x = \cos x \qquad \text{and} \qquad \frac{d}{dx}\cos x = -\sin x$$

Graphs of the derivatives of each of the trigonometric functions, based on the geometric interpretation of a derivative as the slope of the tangent line, help to explain these rules. (See Figures 4.26 and 4.27.) Pay particular attention to the points on the graph of $f(x)$ with horizontal tangent lines. These correspond to the points where

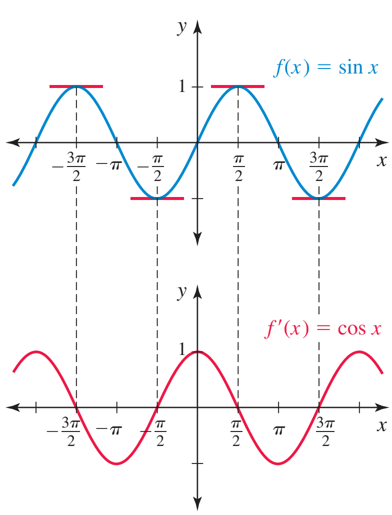


Figure 4.26 The function $f(x) = \sin x$ and its derivative $f'(x) = \cos x$. Red line segments show where $f(x)$ has a horizontal tangent line.

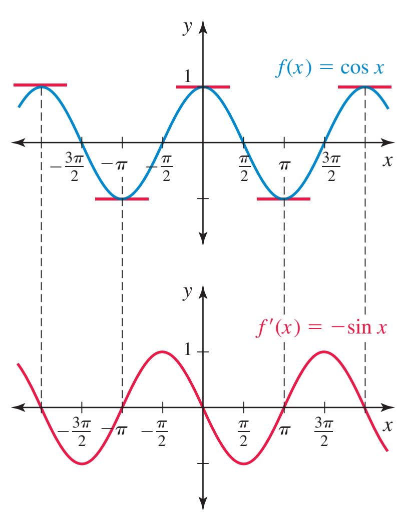


Figure 4.27 The function $f(x) = \cos x$ and its derivative $f'(x) = -\sin x$. Red line segments show where $f(x)$ has a horizontal tangent line.

$f'(x) = 0$. Between these points the tangent slope is either positive ($f'(x) > 0$) or negative ($f'(x) < 0$).

Proof We prove the first formula; a similar proof of the second formula is discussed in Problem 61. We need the trigonometric addition formula

$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta.$$

Using the formal definition of derivatives, we find that

$$\begin{aligned}
\frac{d}{dx}\sin x &= \lim_{h\to 0}\frac{\sin(x+h) - \sin x}{h} \\
&= \lim_{h\to 0}\frac{\sin x\cos h + \cos x\sin h - \sin x}{h} && \text{Use the addition formula on } \sin(x+h) \\
&= \lim_{h\to 0}\left[\sin x \cdot \frac{\cos h - 1}{h} + \cos x\frac{\sin h}{h}\right]
\end{aligned}$$

In Section 3.4, we showed that

$$\lim_{h\to 0}\frac{\cos h - 1}{h} = 0 \qquad \text{and} \qquad \lim_{h\to 0}\frac{\sin h}{h} = 1.$$

We can therefore apply the rules for limits to obtain

$$\begin{aligned}
\frac{d}{dx}\sin x &= \sin x\left(\lim_{h\to 0}\frac{\cos h - 1}{h}\right) + \cos x\left(\lim_{h\to 0}\frac{\sin h}{h}\right) \\
&= (\sin x)(0) + (\cos x)(1) = \cos x
\end{aligned}$$

EXAMPLE 1 Find the derivative of $f(x) = -4\sin x + \cos\frac{\pi}{6}$.

Solution

$$\begin{aligned}
f'(x) &= \frac{d}{dx}(-4\sin x + \cos\frac{\pi}{6}) \\
&= -4\frac{d}{dx}\sin x + \frac{d}{dx}\cos\frac{\pi}{6} \\
f'(x) &= -4(\cos x) + 0 = -4\cos x && \frac{d}{dx}\cos\tfrac{\pi}{6} = 0 \text{ because } \cos\tfrac{\pi}{6} \text{ is a constant.}
\end{aligned}$$

EXAMPLE 2 Find the derivative of $y = \cos(x^2 + 1)$.

Solution We set $u = g(x) = x^2 + 1$ and $f(u) = \cos u$; then $y = f[g(x)]$. Using the chain rule, we then obtain

$$\begin{aligned}
y' &= \frac{df}{du}\frac{du}{dx} = \frac{d}{du}(\cos u)\frac{d}{dx}(x^2 + 1) = (-\sin u)(2x) \\
&= -[\sin(x^2 + 1)]2x = -2x\sin(x^2 + 1)
\end{aligned}$$

EXAMPLE 3 Find the derivative of $y = x^2\sin(3x) - \cos(5x)$.

Solution We will use the product rule for the first term; in addition, we will need the chain rule for both $\sin(3x)$ and $\cos(5x)$:

$$\begin{aligned}
y' &= \frac{d}{dx}[x^2 \cdot \sin(3x) - \cos(5x)] \\
&= \frac{d}{dx}[\overbrace{x^2}^{f(x)} \cdot \overbrace{\sin(3x)}^{g(x)}] - \frac{d}{dx}\cos(5x) \\
&= \overbrace{\left(\frac{d}{dx}x^2\right)\sin(3x)}^{f'(x)g(x)} + \overbrace{x^2\frac{d}{dx}\sin(3x)}^{f(x)g'(x)} - \frac{d}{dx}\cos(5x) \\
&= 2x\sin(3x) + x^2 \cdot 3\cos(3x) - 5(-\sin(5x)) \\
&= 2x\sin(3x) + 3x^2\cos(3x) + 5\sin(5x)
\end{aligned}$$

The derivatives of the other trigonometric functions can be found using the following identities:

$$\tan x = \frac{\sin x}{\cos x} \qquad \cot x = \frac{\cos x}{\sin x}$$

$$\sec x = \frac{1}{\cos x} \qquad \csc x = \frac{1}{\sin x}.$$

For instance,

$$\begin{aligned}\frac{d}{dx}\tan x &= \frac{d}{dx}\left(\frac{\sin x}{\cos x}\right)\\ &= \frac{(\frac{d}{dx}\sin x)\cos x - \sin x(\frac{d}{dx}\cos x)}{\cos^2 x} && \text{Quotient rule}\\ &= \frac{(\cos x)(\cos x) - (\sin x)(-\sin x)}{\cos^2 x}\\ &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x. && \cos^2 x + \sin^2 x = 1\end{aligned}$$

The other derivatives can be found in a similar fashion, as explained in Problems 62–64.

We summarize the derivatives of the six fundamental trigonometric functions in the following box:

$$\frac{d}{dx}\sin x = \cos x \qquad \frac{d}{dx}\cos x = -\sin x$$

$$\frac{d}{dx}\tan x = \sec^2 x \qquad \frac{d}{dx}\cot x = -\csc^2 x$$

$$\frac{d}{dx}\sec x = \sec x \tan x \qquad \frac{d}{dx}\csc x = -\csc x \cot x$$

EXAMPLE 4 Compare the derivatives of

(a) $\tan x^2$ **(b)** $\tan^2 x$

Solution **(a)** If $y = \tan x^2 = \tan(x^2)$, then, using the chain rule, we find that

$$\frac{dy}{dx} = \frac{d}{dx}\tan(x^2) = \underbrace{(\sec^2(x^2))}_{f'(u)}\underbrace{(2x)}_{g'(x)} = 2x\sec^2(x^2) \qquad f(u) = \tan u,\ g(x) = x^2$$

(b) If $y = \tan^2 x = (\tan x)^2$, then, using the chain rule, we obtain

$$\frac{dy}{dx} = \frac{d}{dx}(\tan x)^2 = \underbrace{2(\tan x)}_{f'(u)}\underbrace{\frac{d}{dx}\tan x}_{g'(x)} = 2\tan x\sec^2 x \qquad f(u) = u^2,\ g(x) = \tan x$$

The two derivatives are different, and you should look again at $\tan x^2$ and $\tan^2 x$ to make sure that you understand which is the inner and which the outer function. •

EXAMPLE 5 **Repeated Application of the Chain Rule** Find the derivative of $f(x) = \sec\sqrt{x^2+1}$.

Solution This is a composite function; the inner function is $\sqrt{x^2+1}$ and the outer function is $\sec x$. Applying the chain rule once, we find that

$$\frac{df}{dx} = \frac{d}{dx}\sec\sqrt{x^2+1} = \underbrace{\sec\sqrt{x^2+1}\cdot\tan\sqrt{x^2+1}}_{f'(u)=\sec u\cdot\tan u}\cdot\underbrace{\frac{d}{dx}\sqrt{x^2+1}}_{g'(x)}. \qquad f(u)=\sec u,\ g(x)=\sqrt{x^2+1}$$

To evaluate $\frac{d}{dx}\sqrt{x^2+1}$, we need to apply the chain rule a second time:

$$\frac{d}{dx}\sqrt{x^2+1} = \frac{1}{2\sqrt{x^2+1}}\cdot 2x = \frac{x}{\sqrt{x^2+1}}.$$

Combining the two steps, we obtain

$$\frac{df}{dx} = \left(\sec\sqrt{x^2+1}\right)\left(\tan\sqrt{x^2+1}\right)\frac{x}{\sqrt{x^2+1}}$$

The function $f(x)$ can be thought of as a composition of three functions. The innermost function is $u = x^2+1$, the middle function is $v = \sqrt{u}$, and the outermost function is $f(v) = \sec v$. When we computed the derivative, we applied the chain rule twice in the form

$$\frac{df}{dx} = \frac{df}{dv}\frac{dv}{du}\frac{du}{dx}.$$

Section 4.8 Problems

In Problems 1–58, find the derivative with respect to the independent variable.

1. $f(x) = 2\sin x - \cos x$

2. $f(x) = 3\cos x - 2\sin x$

3. $f(x) = 3\sin x + 5\cos x$

4. $f(x) = -\sin x + \cos x$

5. $f(x) = \cos(x+1)$

6. $f(x) = \sin(2-x)$

7. $f(x) = \sin(3x)$

8. $f(x) = \cos(-5x)$

9. $f(x) = 2\sin(3x+1)$

10. $f(x) = -3\cos(1-2x)$

11. $f(x) = \tan(4x)$

12. $f(x) = \cot(2-3x)$

13. $f(x) = 2\sec(1+2x)$

14. $f(x) = -3\csc(3-5x)$

15. $f(x) = 3\sin(x^2)$

16. $f(x) = 2\cos(x^3-3x)$

17. $f(x) = \sin^2(x^2-3)$

18. $f(x) = \cos^2(x^2-1)$

19. $f(x) = 3\sin^2 x^2$

20. $f(x) = -\sin^2(2x^3-1)$

21. $f(x) = 4\cos x^2 - 2\cos^2 x$

22. $f(x) = -5\cos(2-x^3) + 2\cos^3(x-4)$

23. $f(x) = 4\cos^2 x + 2\cos x^4$

24. $f(x) = -3\cos^2(3x^2-4)$

25. $f(x) = 2\tan(1-x^2)$

26. $f(x) = -\cos(3x^3-4x)$

27. $f(x) = \sin\sqrt{x}$

28. $f(x) = \sqrt{\sin x}$

29. $f(x) = \sqrt{\sin(2x^2-1)}$

30. $g(s) = (\cos^2 s - 3s^2)^2$

31. $g(s) = \sqrt{\cos s} - \cos\sqrt{x}$

32. $g(t) = \dfrac{\sin(3t)}{\cos(5t)}$

33. $g(t) = \dfrac{\sin(2t)+1}{\cos(6t)-1}$

34. $f(x) = \dfrac{\cos(2x)}{\tan(4x)}$

35. $f(x) = \dfrac{\sec(x^2-1)}{\csc(x^2+1)}$

36. $f(x) = \sin x\cos x$

37. $f(x) = \sin(2x-1)\cos(3x+1)$

38. $f(x) = \tan x\cot x$

39. $f(x) = \tan(3x^2-1)\cot(3x^2+1)$

40. $f(x) = \sec x\cos x$

41. $f(x) = \sin x\sec x$

42. $f(x) = \dfrac{1}{\sin^2 x + \cos^2 x}$

43. $f(x) = \dfrac{1}{\tan^2 x - \sec^2 x}$

44. $g(x) = \dfrac{1}{\sin(3x)}$

45. $g(x) = \dfrac{1}{\sin(3x^2-1)}$

46. $g(x) = \dfrac{1}{\csc^2(5x)}$

47. $g(x) = \dfrac{1}{\sin^2(1-5x^2)}$

48. $h(x) = \cot(3x)\csc(3x)$

49. $h(x) = \dfrac{3}{\tan(2x)-x}$

50. $g(t) = \left(\dfrac{1}{\sin t^2}\right)$

51. $h(s) = \sin^3 s + \cos^3 s$

52. $f(x) = (2x^3-x)\cos^2 x$

53. $f(x) = \dfrac{\sin(2x)}{1+x^2}$

54. $f(x) = \dfrac{1+\cos(3x)}{2x^3-x}$

55. $f(x) = \tan\dfrac{1}{x}$

56. $f(x) = \sec\left(\dfrac{1}{1+x^2}\right)$

57. $f(x) = \dfrac{\cos x^2}{\cos^2 x}$

58. $f(x) = \dfrac{\csc(3-x^2)}{1-x^2}$

59. Find the points on the curve $y = \sin(\frac{\pi}{3}x)$ that have a horizontal tangent.

60. Find the points on the curve $y = \cos^2 x$ that have a horizontal tangent.

61. Use the identity

$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$

and the definition of the derivative to show that

$$\frac{d}{dx}\cos x = -\sin x$$

62. Use the quotient rule to show that

$$\frac{d}{dx}\cot x = -\csc^2 x$$

(*Hint*: Write $\cot x = \frac{\cos x}{\sin x}$.)

63. Use the quotient rule to show that

$$\frac{d}{dx}\sec x = \sec x \tan x$$

[*Hint*: Write $\sec x = (\cos x)^{-1}$.]

64. Use the quotient rule to show that

$$\frac{d}{dx}\csc x = -\csc x \cot x$$

[*Hint*: Write $\csc x = (\sin x)^{-1}$.]

In Problems 65 to 72 find the derivatives of the following functions:

65. $f(x) = \sin\sqrt{x^2+1}$

66. $f(x) = \cos\sqrt{x^2+1}$

67. $f(x) = \sin\sqrt{3x^2+3x}$

68. $f(x) = \cos x - \sin 2x$

69. $f(x) = \sin^2(x^2-1)$

70. $f(x) = \cos^2(2x^2+3)$

71. $f(x) = \sin 2x + \sin^2 x$

72. $f(x) = \sec^2(2x^2-1)$

73. **Lake Nitrogen** Suppose that the concentration of nitrogen in a lake exhibits periodic behavior. That is, if we denote the concentration of nitrogen at time t by $c(t)$, then we assume that

$$c(t) = 2 + \sin\left(\frac{\pi}{2}t\right)$$

(a) Find $\frac{dc}{dt}$.

(b) Use a graphing calculator to graph both $c(t)$ and $\frac{dc}{dt}$ in the same coordinate system.

(c) By inspecting the graph in (b), answer the following questions:

(i) When $c(t)$ reaches a maximum, what is the value of dc/dt?

(ii) When dc/dt is positive, is $c(t)$ increasing or decreasing?

(iii) What can you say about $c(t)$ when $dc/dt = 0$?

74. **Circadian Growth** The growth rate of a fungus varies over the course of one day. You find that the size of the fungus is given as a function of time by:

$$L(t) = 3.6t + 1.2\cos(2\pi t/24)$$

where t is the time in hours, and $L(t)$ is the size in millimeters.

(a) Calculate the growth rate dL/dt

(b) What is the largest growth rate of the fungus? What is the smallest growth rate?

4.9 Derivatives of Exponential Functions

In Section 4.2 we discussed how one possible use of the derivative is to calculate the rate of growth of a population whose size, $N(t)$, changes with time. Under many conditions the fractional rate of growth of a population will be constant. That is:

$$\frac{1}{N}\frac{dN}{dt} = c$$

for some constant c. We may interpret this equation as saying that in each unit of time the population increases by a fixed fraction, c, of its current size. Or, alternatively:

$$\frac{dN}{dt} = cN. \tag{4.7}$$

That is, the rate of growth increases in proportion to the population size. Why might a population behave in this way? The rate of population size change $\frac{dN}{dt}$ represents the rate of births (minus the rate of deaths). If two populations of the same type of organism are compared, one twice the size of the other, then we expect the population that has twice the members to have twice the number of births in the same amount of time (and twice the number of deaths); that is, birth rate and death rate are proportional to the population size.

We may regard Equation (4.7) as a **differential equation**; solving the differential equation will give us the population size $N(t)$. When we differentiate $N(t)$, the function we get back is just a constant, c, times $N(t)$. In this section we will show that exponential functions are the solution to this problem.

Recall from Section 1.3 that the function f is an exponential function with base a if

$$f(x) = a^x, \quad x \in \mathbf{R}$$

x is the exponent, a is the base.

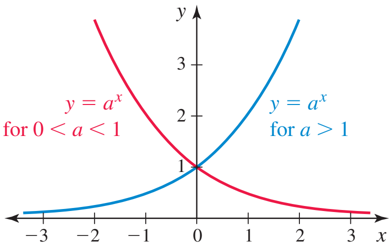


Figure 4.28 The function $y = a^x$.

where a is a positive constant other than 1. (See Figure 4.28.) We can use the formal definition of the derivative to compute $f'(x)$:

$$f'(x) = \frac{d}{dx}a^x = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \lim_{h\to 0}\frac{a^{x+h}-a^x}{h}$$
$$= \lim_{h\to 0}\frac{a^x(a^h-1)}{h} = a^x\lim_{h\to 0}\frac{a^h-1}{h} \tag{4.8}$$

So if $\lim_{h\to 0}\frac{a^h-1}{h}$ exists then we have shown that $f'(x) = cf(x)$ where $c = \lim_{h\to 0}\frac{a^h-1}{h}$ is a constant that depends on the base, a, but does not depend on x. Thus the function $y = a^x$ solves a form of the population growth differential equation $\frac{dy}{dx} = cy$. But does $c = \lim_{h\to 0}\frac{a^h-1}{h}$ exist and how does the limit depend on a? Although it is possible to rigorously prove that the limit exists, the tools needed to do so are outside of the scope of this course. Instead we will use tables to investigate the limit for a few different values of a.

For $a = 2$:

h	0.1	0.01	0.001	0.0001
$\frac{a^h-1}{h}$	0.7177	0.6956	0.6934	0.6932

From the table we see that the limit exists and is equal to approximately 0.693. Similarly for $a = 4$:

h	0.1	0.01	0.001	0.0001
$\frac{a^h-1}{h}$	1.4870	1.3959	1.3873	1.3864

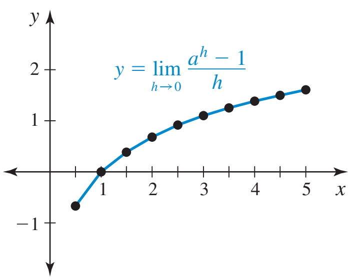


Figure 4.29 The numerically computed value of $\lim_{h\to 0}\frac{a^h-1}{h}$ for $a = 0.5, 1.0, 1.5, \ldots, 5.0$.

Again the limit exists and is equal to approximately 1.386. In fact the limit exists for all bases $a > 0$. We can make a plot showing the numerically computed limits for many different values of a (Figure 4.29).

When $a = 1$, $\lim_{h\to 0}\frac{a^h-1}{h} = 0$, because $a^h = 1^h = 1$ for all values of h. From the plot we see that the value of the limit increases as a increases. In fact it can be shown (though, again, the proof is beyond the scope of this course) that the limit depends continuously on the value of a. From the plot we see that when $a = 2.5$, $\lim_{h\to 0}\frac{a^h-1}{h} = 0.916$, whereas when $a = 3.0$, $\lim_{h\to 0}\frac{a^h-1}{h} = 1.099$. Thus, by the intermediate value theorem there must be a value of a somewhere between $a = 2.5$ and $a = 3.0$ for which $\lim_{h\to 0}\frac{a^h-1}{h} = 1$. We denote this base by e. The number e is therefore defined by:

> **Definition of base e**: e is the only base for which
> $$\lim_{h\to 0}\frac{e^h-1}{h} = 1 \tag{4.9}$$

Because $c = 1$, Equation (4.8) then yields:

$$\frac{d}{dx}e^x = e^x \tag{4.10}$$

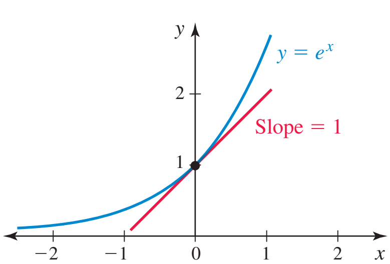


Figure 4.30 The function $y = e^x$. When $x = 0$, $\frac{dy}{dx} = \lim_{n\to 0}\frac{e^h-1}{h} = 1$.

That is, differentiating the function $f(x) = e^x$ returns the same function: $f'(x) = e^x$.

A graph of $f(x) = e^x$ is shown in Figure 4.30. The domain of this function is $\mathbf{R}$ and its range is the open interval $(0, \infty)$. (In particular, $e^x > 0$ for all $x \in \mathbf{R}$.) Denoting by e the base of the exponential function for which (4.9) and (4.10) hold is no accident; it is indeed the natural exponential base that we introduced in Section 1.3. Although we cannot prove this here, a table should convince you: With $e = 2.71828\ldots$, we find that

h	0.1	0.01	0.001	0.0001
$\frac{e^h-1}{h}$	1.0517	1.0050	1.00050	1.000050

Now recall that there is an alternative notation for e^x, namely, $\exp[x]$. Using the identity

$$a^x = \exp[\ln a^x]$$

and the fact that $\ln a^x = x \ln a$, we can find the derivative of a^x with the help of the chain rule:

$$\frac{d}{dx}a^x = \frac{d}{dx}\exp[\ln a^x] = \frac{d}{dx}\exp[x \ln a] \qquad f(u) = e^u,\ g(x) = x \ln a$$

$$= \underbrace{\exp[x \ln a]}_{f'(u)}\ \underbrace{\ln a}_{g'(x)} = (\ln a)a^x$$

That is, we have

$$\boxed{\frac{d}{dx}a^x = (\ln a)a^x} \tag{4.11}$$

which allows us to obtain the following identity:

$$\lim_{h \to 0} \frac{a^h - 1}{h} = \ln a \tag{4.12}$$

EXAMPLE 1 Find the derivative of $f(x) = e^{-x^2/2}$.

Solution This rather odd looking function will turn out to be essential in models of random processes, and we will explore it further in Chapter 12. In particular we will explain how processes that are created by the addition of many small random events, ranging from genes being turned on and off to the diffusion of small molecules, have probabilities given by $f(x)$.

We use the chain rule:

$$f'(x) = e^{-x^2/2}\left(-\frac{2x}{2}\right) = -xe^{-x^2/2} \qquad f(u) = e^u, \quad g(x) = -x^2/2$$

•

Before discussing some uses of exponential functions in models of biological processes, we will work through three examples in which we practice differentiating functions in which exponentials appear, by combining (4.10) and (4.11) with the chain rule, the product rule, and so on.

EXAMPLE 2 Find the derivative of $f(x) = 3^{\sqrt{x}}$.

Solution We can use (4.11) and the chain rule to get

$$\frac{d}{dx}3^{\sqrt{x}} = (\ln 3)3^{\sqrt{x}} \cdot \frac{1}{2\sqrt{x}}. \qquad f(u) = 3^u \quad g(x) = \sqrt{x}$$

However, since every exponential function can be written in terms of the base e, and the differentiation rule for e^x is particularly simple ($\frac{d}{dx}e^x = e^x$), it is often easier to rewrite the exponential function in terms of e and then differentiate. That is, we write

$$3^{\sqrt{x}} = \exp[\ln 3^{\sqrt{x}}] = \exp[\sqrt{x} \ln 3]$$

Then, using the chain rule, we obtain

$$\frac{d}{dx}\exp[\sqrt{x}\ln 3] = \frac{\ln 3}{2\sqrt{x}}\exp[\sqrt{x}\ln 3] = \frac{\ln 3}{2\sqrt{x}}3^{\sqrt{x}} \qquad f(u) = e^u, \quad g(x) = \sqrt{x}\ln 3$$

Since we must frequently differentiate functions of the form $y = e^{g(x)}$, we state this differentiation in a separate rule. Using the chain rule, we have

$$\boxed{\frac{d}{dx}e^{g(x)} = g'(x)e^{g(x)} \qquad (4.13)}$$

EXAMPLE 3 Find the derivative of $f(x) = \exp[\sin\sqrt{x}]$.

Solution We set $g(x) = \sin\sqrt{x}$. To differentiate $g(x)$, we must apply the chain rule:

$$\frac{d}{dx}g(x) = (\cos\sqrt{x})\frac{1}{2\sqrt{x}}$$

Using Equation (4.13), we can now differentiate $f(x)$:

$$\frac{d}{dx}f(x) = (\cos\sqrt{x})\frac{1}{2\sqrt{x}}\exp[\sin\sqrt{x}]$$

Here is an example that shows how (4.13) is used:

EXAMPLE 4 Find the derivative of $f(x) = \frac{e^{-x}}{x}$.

Solution We use the quotient rule to calculate $f'(x)$:

$$f(x) = \frac{\overbrace{e^{-x}}^{u(x)}}{\underbrace{x}_{v(x)}} \Rightarrow f'(x) = \frac{\frac{d}{dx}(e^{-x})x - e^{-x}\frac{d}{dx}(x)}{x^2} \qquad f'(x) = \frac{u'v - uv'}{v^2}$$

Since

$$\frac{d}{dx}(e^{-x}) = (-1)e^{-x} \qquad (4.13) \text{ with } g(x) = -x$$

we obtain:

$$\begin{aligned} f'(x) &= \frac{-e^{-x}\cdot x - e^{-x}\cdot 1}{x^2} \\ &= -\frac{(1+x)e^{-x}}{x^2}. \end{aligned}$$

EXAMPLE 5 **Population Growth** Populations that grow optimally, that is, without any shortage of space or food, typically have constant reproductive rate. That is, in a unit of time, the population size increases by a fixed fraction, r, of its current size. Hence, if the population size is $N(t)$, the rate of change of $N(t)$ obeys a differential equation:

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

where r is a constant.

Show that the exponential function $N(t) = N_o e^{rt}$, where N_o is any constant, solves this differential equation.

Solution To show that $N(t) = N_o e^{rt}$ solves the differential equation we substitute it into both the left and right-hand sides. On the right-hand side we have:

$$rN = rN_o e^{rt}$$

whereas on the left-hand side we have:

$$\frac{dN}{dt} = N_o \frac{d}{dt}(e^{rt}) = N_o r e^{rt} \qquad \text{(4.13) with } g(t) = rt$$

The left-hand and right-hand sides evaluate to the same function, $N_o r e^{rt}$, so $N(t) = N_o e^{rt}$ does indeed solve the differential equation. •

Notice that the function satisfies the differential equation for any value of N_o. We will revisit this issue in more depth in Chapter 8, but in general, differential equations like the one in Example 5 don't completely specify the size of a population, $N(t)$. Two populations with different starting sizes can both grow according to the differential equation and will always have different sizes. So to completely know the population size at time, t, we need to know both the reproductive rate, r, and the starting population size (i.e., the value of N_o).

Differential equations like the one in Example 5 show up in many areas of life science, not just to describe the growth of populations, but also for applications ranging from the filtering of light with depth in lakes, to the decay of isotopes.

EXAMPLE 6 Radioactive isotopes decay (that is, decompose into lighter elements) spontaneously. Knowing how the quantity of a particular isotope shifts over time can be used to date samples, such as fossils. Suppose that at time t, a particular sample contains a total amount $W(t)$ of isotope. In a unit of time a fixed fraction, λ, of the isotope decays, reducing $W(t)$. So the rate of change $\frac{dW}{dt}$ must obey a differential equation:

$$\frac{dW}{dt} = -\lambda W$$

Show that the function $W(t) = W_o e^{-\lambda t}$ solves this differential equation for any value of the constant W_o.

Solution Just as in Example 5, we show that $W(t)$ solves the differential equation by substituting it into both sides. On the right-hand side we have:

$$-\lambda W = -\lambda W_o e^{-\lambda t}$$

While on the left-hand side, we have:

$$\frac{dW}{dt} = W_o \frac{d}{dt}(e^{-\lambda t}) = -W_o \lambda e^{-\lambda t}$$

Since both sides evaluate to the same expression, $W(t)$ indeed solves the differential equation. •

In both Examples 5 and 6, we have made use of the result that:

> If $N(t) = ae^{ct}$ for any constants a and c, then $\frac{dN}{dt} = cae^{ct} = cN$.

It is worth committing this particular rule (which is a special case of (4.13) to memory.

Section 4.9 Problems

Differentiate the functions in Problems 1–52 with respect to the independent variable.

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

2. $f(x) = e^{-2x}$

3. $f(x) = 4e^{1-3x}$

4. $f(x) = 3e^{2-5x}$

5. $f(x) = e^{-2x^2+3x-1}$

6. $f(x) = e^{4x^2-2x+1}$

7. $f(x) = e^{7(x^2+1)^2}$

8. $f(x) = e^{-3(x^3-1)^4}$

9. $f(x) = xe^x$

10. $f(x) = 2xe^{-3x}$

11. $f(x) = x^2e^{-x}$

12. $f(x) = (3x^2-1)e^{1-x^2}$

13. $f(x) = \dfrac{1+e^x}{1+x^2}$

14. $f(x) = \dfrac{x-e^{-x}}{1+xe^{-x}}$

15. $f(x) = \dfrac{e^x+e^{-x}}{2+e^x}$

16. $f(x) = \dfrac{x}{e^x+e^{-x}}$

17. $f(x) = \exp[\sin(3x)]$

18. $f(x) = \exp[\cos(4x)]$

19. $f(x) = \exp[\sin(x^2-1)]$

20. $f(x) = \exp[\cos(1-2x^3)]$

21. $f(x) = \sin(e^x)$

22. $f(x) = \cos(e^x)$

23. $f(x) = \sin(e^{2x}+x)$

24. $f(x) = \cos(3x - e^{x^2-1})$

25. $f(x) = \exp[x - \sin x]$

26. $f(x) = \exp[x^2 - 2\cos x]$

27. $g(s) = \exp[\sec s^2]$

28. $g(s) = \exp[\tan s^3]$

29. $f(x) = e^{x\sin x}$

30. $f(x) = e^{1-x\cos x}$

31. $f(x) = -3e^{x^2+\tan x}$

32. $f(x) = 2e^{-x\sec(3x)}$

33. $f(x) = 2^x$

34. $f(x) = 3^x$

35. $f(x) = 2^{x+1}$

36. $f(x) = 3^{x-1}$

37. $f(x) = 5^{\sqrt{2x-1}}$

38. $f(x) = 3^{\sqrt{1-3x}}$

39. $f(x) = 2^{x^2+1}$

40. $f(x) = 3^{x^3-1}$

41. $h(t) = 2^{t^2-1}$

42. $h(t) = 4^{2t^3-t}$

43. $f(x) = 2^{\sqrt{x}}$

44. $f(x) = 3^{\sqrt{x+1}}$

45. $f(x) = 2^{\sqrt{x^2-1}}$

46. $f(x) = 4^{\sqrt{1-2x^3}}$

47. $h(t) = 5^{\sqrt{t}}$

48. $h(t) = 6^{\sqrt{6t^6-6}}$

49. $g(x) = 2^{2\cos x}$

50. $g(r) = 2^{-3\sin r}$

51. $g(r) = 3^{r^{1/5}}$

52. $g(r) = 4^{r^{1/4}}$

Compute the limits in Problems 53–56.

53. $\displaystyle\lim_{h\to 0} \frac{e^{2h}-1}{h}$

54. $\displaystyle\lim_{h\to 0} \frac{e^{5h}-1}{3h}$

55. $\displaystyle\lim_{h\to 0} \frac{e^{h}-1}{\sqrt{h}}$

56. $\displaystyle\lim_{h\to 0} \frac{2^{h}-1}{h}$

57. Find the equation for the tangent to the curve $y = 2^x$ at the point $(1, 2)$.

58. Find the equation for the tangent to the curve $y = \exp[x^2]$ at the point $(2, e^4)$.

59. **Population Growth** Suppose that the population size at time t is

$$N(t) = e^{2t}, \quad t \geq 0.$$

(a) What is the population size at time 0?

(b) Show that

$$\frac{dN}{dt} = 2N$$

60. **Population Growth** Suppose that the population size at time t is

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

where N_0 is a positive constant and r is a real number.

(a) What is the population size at time 0?

(b) Show that

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

61. **Bacterial Growth** Suppose that a bacterial colony grows in such a way that at time t the population size is

$$N(t) = N_0 2^t$$

where N_0 is the population size at time 0. Find the rate of growth dN/dt. Express your solution in terms of $N(t)$. Show that the growth rate of the population is proportional to the population size.

62. **Bacterial Growth** Suppose that a bacterial colony grows in such a way that at time t the population size is

$$N(t) = N_0 2^t$$

where N_0 is the population size at time 0. Find the per capita growth rate.

$$\frac{1}{N}\frac{dN}{dt}.$$

63. **Logistic Growth** A population of organisms grows according to a logistic growth model:

$$N(t) = \frac{K}{1 + (K-1)e^{-rt}}.$$

where r and K are positive constants.

(a) Find $\frac{dN}{dt}$.

(b) Show that $N(t)$ satisfies the equation

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

[*Hint*: Use the function $N(t)$ for the right-hand side, and simplify until you obtain the derivative of $N(t)$ that you computed in (a).]

(c) Plot the per capita rate of growth $\frac{1}{N}\frac{dN}{dt}$ as a function of N, and note that it decreases with increasing population size.

64. **Fish Recruitment Model** The following model is used in the fisheries literature to describe the recruitment of fish as a function of the size of the parent stock: If we denote the number of recruits by R and the size of the parent stock by P, then

$$R(P) = aPe^{-bP}, \quad P \geq 0$$

where a and b are positive constants.

(a) Sketch the graph of the function $R(P)$ when $b = 1$ and $a = 2$.

(b) Differentiate $R(P)$ with respect to P.

(c) Find all the points on the curve that have a horizontal tangent.

65. **Von Bertalanffy Growth Model** The growth of fish can be described by the von Bertalanffy growth function

$$L(x) = L_\infty - (L_\infty - L_0)e^{-kx}$$

where x denotes the age of the fish and k, L_∞, and L_0 are positive constants.

(a) Set $L_0 = 1$ and $L_\infty = 10$. Graph $L(x)$ for $k = 1.0$ and $k = 0.1$.

(b) Interpret L_∞ and L_0.

(c) Compare the graphs for $k = 0.1$ and $k = 1.0$. According to your graphs which fish reach $L = 5$ more quickly?

(d) Show that

$$\frac{d}{dx}L(x) = k(L_\infty - L(x))$$

That is, $dL/dx \propto L_\infty - L$. What does this proportionality say about how the rate of growth changes with age?

(e) The constant k is the proportionality constant in (d). What does the value of k tell you about how quickly a fish grows?

66. **Radioactive Decay** Suppose $W(t)$ denotes the amount of a radioactive material left after time t (measured in days). Assume that the radioactive decay rate of the material is 0.2/day. Find the differential equation for the radioactive decay function $W(t)$.

67. **Radioactive Decay** Suppose $W(t)$ denotes the amount of a radioactive material left after time t (measured in days). Assume that the radioactive decay rate of the material is 4/day. Find the differential equation for the radioactive decay function $W(t)$.

68. **Radioactive Decay** Suppose $W(t)$ denotes the amount of a radioactive material left after time t (measured in days). Assume that the half-life of the material is 3 days. Find the differential equation for the radioactive decay function $W(t)$.

69. **Radioactive Decay** Suppose $W(t)$ denotes the amount of a radioactive material left after time t (measured in days). Assume that the half-life of the material is 5 days. Find the differential equation for the radioactive decay function $W(t)$.

70. **Radioactive Decay** Suppose $W(t)$ denotes the amount of a radioactive material left after time t. Assume that $W(0) = 15$ and that

$$\frac{dW}{dt} = -2W(t).$$

(a) How much material is left at time $t = 2$?

(b) What is the half-life of this material?

71. **Radioactive Decay** Suppose $W(t)$ denotes the amount of a radioactive material left after time t. Assume that $W(0) = 6$ and that

$$\frac{dW}{dt} = -3W(t).$$

(a) How much material is left at time $t = 4$?

(b) What is the half-life of the material?

72. **Radioactive Decay** Suppose $W(t)$ denotes the amount of a radioactive material left after time t. Assume that $W(0) = 10$ and $W(1) = 8$.

(a) Find the differential equation that describes this situation.

(b) How much material is left at time $t = 5$?

(c) What is the half-life of the material?

73. **Radioactive Decay** Suppose $W(t)$ denotes the amount of a radioactive material left after time t. Assume that $W(0) = 5$ and $W(1) = 2$.

(a) Find the differential equation that describes this situation.

(b) How much material is left at time $t = 3$?

(c) What is the half-life of the material?

4.10 Derivatives of Inverse Functions, Logarithmic Functions, and the Inverse Tangent Function

Recall that the logarithmic function is the inverse of the exponential function. To find the derivative of the logarithmic function, we must therefore learn how to compute the derivative of an inverse function.

4.10.1 Derivatives of Inverse Functions

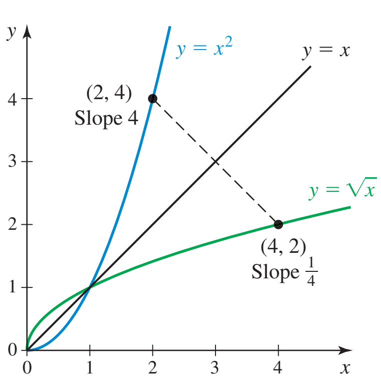


Figure 4.31 The function $y = x^2$, $x \geq 0$, and its inverse function $y = \sqrt{x}$, $x \geq 0$.

We begin with an example (Figure 4.31). Let $f(x) = x^2$, $x \geq 0$. We computed the inverse function of f in Subsection 1.3.6. First note that $f(x) = x^2, x \geq 0$, is one to one (use the horizontal line test from Subsection 1.3.6); hence, we can define its inverse. We repeat the steps from Subsection 1.3.6 to find an inverse function. [Recall that we obtain the graph of the inverse function by reflecting $y = f(x)$ about the line $y = x$.]

1. Write $y = f(x)$:
$$y = x^2$$
2. Solve for x:
$$x = \sqrt{y}$$
3. Interchange x and y:
$$y = \sqrt{x}.$$

Since the range of $f(x)$, which is the interval $[0, \infty)$, becomes the domain for the inverse function, it follows that

$$f^{-1}(x) = \sqrt{x} \quad \text{for } x \geq 0.$$

We already know the derivative of $\sqrt{x}$, namely, $1/(2\sqrt{x})$. But we will try to find the derivative in a different way that we can generalize to get a formula for finding the derivative of any inverse function.

f^{-1} undoes f and vice versa, so applying f^{-1} and then f to x returns x again, i.e., $(f \circ f^{-1})(x) = x$ for any $x \geq 0$.

Suppose we wanted to differentiate both sides of this equation with respect to x. On the right-hand side we obtain $\frac{d}{dx}(x) = 1$. On the left-hand side we could use the chain rule with inner function $g(x) = f^{-1}(x) = \sqrt{x}$ and outer function $f(u) = u^2$:

$$\left(\frac{d}{du}u^2\right) \cdot \left(\frac{d}{dx}\sqrt{x}\right) = 1$$

$$2u\frac{d}{dx}\sqrt{x} =$$

$$2\sqrt{x}\frac{d}{dx}\sqrt{x} = \qquad u = g(x) = \sqrt{x}$$

So the derivative of $f^{-1}(x) = \sqrt{x}$ is:

$$\frac{d}{dx}(\sqrt{x}) = \frac{1}{2\sqrt{x}}$$

To prepare for how the derivatives of a function and its inverse function are related geometrically, look at Figure 4.31, where the slope at the point $(x, y) = (2, 4)$ of $f(x) = x^2$ is $m = 4$ and the slope at the point $(x, y) = (4, 2)$ of $f^{-1}(x) = \sqrt{x}$ is $m = 1/4$. We will see that it is not an accident that the slope of $f^{-1}(x)$ is the reciprocal of the slope of $f(x)$.

We could have computed $\frac{d}{dx}(\sqrt{x})$ directly using the power rule, but we introduce this alternate method here because the steps that led us to the derivative of $\sqrt{x}$ can be used to find a general formula for the derivatives of inverse functions. We assume that $f(x)$ is one to one in its domain. If $g(x)$ is the inverse function of $f(x)$, then $f[g(x)] = x$. Applying the chain rule, we find that

$$\frac{d}{dx}f[g(x)] = f'[g(x)]g'(x)$$

Since $\frac{d}{dx}x = 1$, we obtain

$$f'[g(x)]g'(x) = 1$$

If $f'[g(x)] \neq 0$, we can divide by $f'[g(x)]$ to get

$$g'(x) = \frac{1}{f'[g(x)]}$$

Because $g(x) = f^{-1}(x)$ and $g'(x) = \frac{d}{dx}g(x) = \frac{d}{dx}f^{-1}(x)$, we obtain the following rule:

> **Derivative of an Inverse Function** If $f(x)$ is one to one and differentiable with inverse function $f^{-1}(x)$ and $f'[f^{-1}(x)] \neq 0$, then $f^{-1}(x)$ is differentiable and
>
> $$\frac{d}{dx}f^{-1}(x) = \frac{1}{f'[f^{-1}(x)]} \tag{4.14}$$

This reciprocal relationship is illustrated in Figure 4.32.

We return to the example of $f(x) = x^2$, $x \geq 0$, where $f^{-1}(x) = \sqrt{x}$, $x \geq 0$, to illustrate how to use (4.14). Now, $f'(x) = 2x$ and we need to evaluate

$$f'[f^{-1}(x)] = f'[\sqrt{x}] = 2\sqrt{x}$$

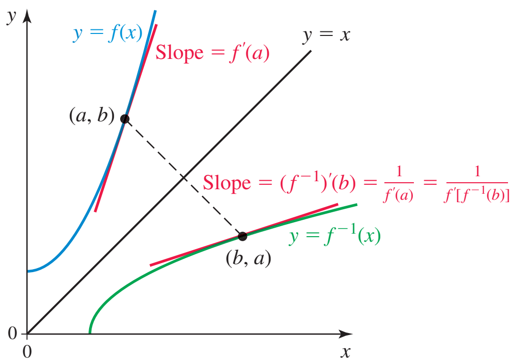


Figure 4.32 The graphs of $y = f(x)$ and its inverse function $y = f^{-1}(x)$ have reciprocal slopes at the points (a, b) and (b, a).

To apply the formula, we assume that $f'[f^{-1}(x)] \neq 0$. Then

$$\frac{d}{dx}\sqrt{x} = \frac{1}{2\sqrt{x}} \quad \text{for } x > 0.$$

Looking at the graphs of $y = x^2$ and $y = \sqrt{x}$ (Figure 4.31) we can see why $f^{-1}(x)$ is not differentiable at $x = 0$. If we draw the tangent line to the curve $y = x^2$ at $x = 0$, we find that the tangent line is horizontal; that is, its slope is 0. Reflecting a horizontal line about $y = x$ results in a vertical line, for which the slope is not defined.

The formula for finding derivatives of inverse functions is easier to remember when we use Leibniz notation. To see this, note that (without interchanging x and y)

$$y = f(x) \quad \Longleftrightarrow \quad x = f^{-1}(y)$$

and hence

$$\boxed{\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}}$$

This formula again emphasizes the reciprocal relationship. We illustrate the formula with the example

$$y = x^2 \quad \Longleftrightarrow \quad x = \sqrt{y}$$

for $x > 0$. Since $\frac{dy}{dx} = 2x$, we have

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{2x} = \frac{1}{2\sqrt{y}}$$

That is,

$$\frac{d}{dy}\sqrt{y} = \frac{1}{2\sqrt{y}}$$

The answer is now in terms of y, because we did not interchange x and y when we computed the inverse function. If we now do so, we again find that

$$\frac{d}{dx}\sqrt{x} = \frac{1}{2\sqrt{x}}.$$

EXAMPLE 1 Let

$$f(x) = \frac{x}{1+x} \quad \text{for } x \geq 0.$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=\frac{1}{3}}$.

Solution To show that $f^{-1}(x)$ exists, we use the horizontal line test and conclude that $f(x)$ is one to one on its domain, since each horizontal line intersects the graph of $f(x)$ at most once. (See Figure 4.33.)

We can actually compute the inverse of $f(x)$. This will give us two different ways to compute the derivative of the inverse of $f(x)$: We can compute the inverse function explicitly and then differentiate the result, or we can use the formula for finding derivatives of inverse functions. We begin with the latter way.

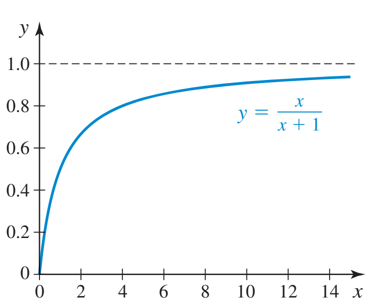


Figure 4.33 The graph of $f(x) = \frac{x}{x+1}$ for $x \geq 0$.

1. To use Equation (4.14), we need to find $f^{-1}(\frac{1}{3})$; this means that we need to find x so that $f(x) = 1/3$. Now,

$$\frac{x}{1+x} = \frac{1}{3} \quad \text{implies that} \quad 2x = 1, \quad \text{or} \quad x = \frac{1}{2}$$

Therefore, $f^{-1}(\frac{1}{3}) = \frac{1}{2}$. Formula (4.14) thus becomes

$$\frac{d}{dx} f^{-1}(x)\bigg|_{x=\frac{1}{3}} = \frac{1}{f'[f^{-1}(\frac{1}{3})]} = \frac{1}{f'(\frac{1}{2})}$$

We use the quotient rule to find the derivative of $f(x)$:

$$f'(x) = \frac{(1)(1+x) - (x)(1)}{(1+x)^2} = \frac{1}{(1+x)^2} \qquad \begin{array}{l} u(x) = x, \Rightarrow u'(x) = 1, \\ v(x) = x + 1 \Rightarrow v'(x) = 1 \end{array}$$

At $x = 1/2$, $f'(\frac{1}{2}) = \frac{1}{(1+1/2)^2} = \frac{4}{9}$. Therefore,

$$\frac{d}{dx} f^{-1}(x)\bigg|_{x=\frac{1}{3}} = \frac{1}{\frac{4}{9}} = \frac{9}{4}.$$

2. We can compute the inverse function: Set $y = \frac{x}{1+x}$. Then, solving for x yields

$$x = \frac{y}{1-y} \qquad y(1+x) = x \Rightarrow y = x - xy$$

Interchanging x and y, we find that

$$y = \frac{x}{1-x}$$

Since the domain of $f(x)$ is $[0, \infty)$, the range of f is $[0, 1)$ (see Figure 4.33). Now, the range of f becomes the domain of the inverse; therefore, the inverse function is

$$f^{-1}(x) = \frac{x}{1-x} \quad \text{for } 0 \le x < 1$$

We use the quotient rule to find $\frac{d}{dx} f^{-1}(x)$:

$$\frac{d}{dx} f^{-1}(x) = \frac{(1)(1-x) - (x)(-1)}{(1-x)^2} = \frac{1}{(1-x)^2}$$

Therefore,

$$\frac{d}{dx} f^{-1}(x)\bigg|_{x=\frac{1}{3}} = \frac{1}{(1-1/3)^2} = \frac{9}{4}$$

which agrees with the answer in part (1). •

The inverse cannot always be computed explicitly, as the next example shows.

EXAMPLE 2 Let

$$f(x) = 2x + e^x \quad \text{for } x \in \mathbf{R}$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=1}$.

Solution In this case, it is not possible to solve $y = 2x + e^x$ for x. Therefore, we must use (4.14) if we wish to compute the derivative of the inverse function at a particular point. Equation (4.14) becomes

$$\frac{d}{dx} f^{-1}(x)\bigg|_{x=1} = \frac{1}{f'[f^{-1}(1)]}$$

We need to find $f'(x)$:

$$f'(x) = 2 + e^x.$$

To use Equation (4.14) we also need to know $f^{-1}(1)$. Since we do not have a closed form expression for $f^{-1}(x)$ we need to guess $f^{-1}(1)$ or obtain it from a graph of $y = f(x)$. From the graph of $y = 2x + e^x$ (Figure 4.34), we see that when $x = 0$, $y = 1$, i.e., $f(0) = 1$, or $f^{-1}(1) = 0$.

So:

$$\frac{d}{dx} f^{-1}(x)\bigg|_{x=1} = \frac{1}{f'[f^{-1}(1)]} = \frac{1}{f'(0)} = \frac{1}{2+1} = \frac{1}{3}$$ •

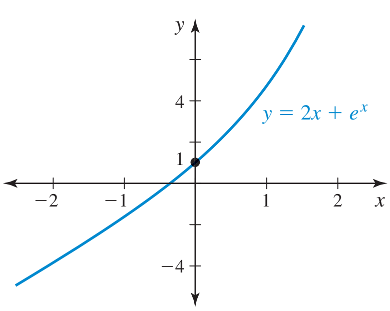


Figure 4.34 The function $f(x) = 2x + e^x$ studied in Example 2.

The next example, in which we again need to use (4.14), involves finding the derivative of the inverse of a trigonometric function.

EXAMPLE 3 Let $f(x) = \tan x$, $-\frac{\pi}{2} < x < \frac{\pi}{2}$. Find $\frac{d}{dx} f^{-1}(x)\big|_{x=1}$.

Solution Recall that $f'(x) = \sec^2 x$. We therefore have

$$\frac{d}{dx} f^{-1}(x)\bigg|_{x=1} = \frac{1}{f'[f^{-1}(1)]} = \frac{1}{f'(\frac{\pi}{4})} = \frac{1}{\sec^2(\frac{\pi}{4})} \qquad f^{-1}(1) = \tan^{-1} 1 = \pi/4$$

$$= \cos^2\left(\frac{\pi}{4}\right) = \left(\frac{1}{2}\sqrt{2}\right)^2 = \frac{1}{2}$$

If we define $f(x) = \tan x$ on the domain $(-\frac{\pi}{2}, \frac{\pi}{2})$, then $f(x)$ is one to one, as can be seen from Figure 4.35. (Use the horizontal line test.) The range of $f(x)$ is $(-\infty, \infty)$. The inverse of the tangent function gets its own name. It is called $y = \arctan x$ (or $y = \tan^{-1} x$) and its domain is $(-\infty, \infty)$. In the next example, we will find the derivative of $y = \arctan x$, for any value of x.

EXAMPLE 4 Let $f(x) = \tan x$, $-\frac{\pi}{2} < x < \frac{\pi}{2}$. Find $\frac{d}{dx} f^{-1}(x)$.

Solution As mentioned previously, $f^{-1}(x)$ exists, since $f(x)$ is one to one on its domain 4. Recall that

$$\frac{d}{dx} \tan x = \sec^2 x$$

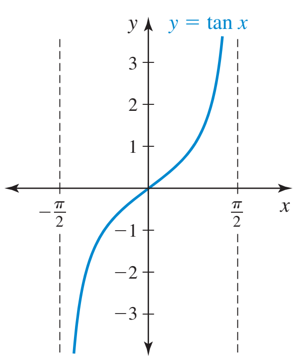


Figure 4.35 The function $f(x) = \tan x$, $(-\frac{\pi}{2}, \frac{\pi}{2})$, is one to one on its domain.

The inverse of the tangent function is denoted by $\tan^{-1} x$ or $\arctan x$. (Note that $\tan^{-1} x$ is different from $\frac{1}{\tan x}$. The superscript "-1" refers to the function being an inverse function.) We set $y = \arctan x$ (and hence $x = \tan y$). Then

$$\frac{dy}{dx} = \frac{d}{dx} \arctan x = \frac{1}{\frac{dx}{dy}} = \frac{1}{\frac{d}{dy} \tan y} = \frac{1}{\sec^2 y} = \frac{1}{1 + \tan^2 y}$$

where we used the trigonometric identity $\sec^2 y = 1 + \tan^2 y$ to get the denominator in the rightmost term. Since $x = \tan y$, it follows that $x^2 = \tan^2 y$, and, hence,

$$\frac{1}{1 + \tan^2 y} = \frac{1}{1 + x^2}$$

Therefore,

$$\frac{d}{dx} \arctan x = \frac{1}{1 + x^2}$$

The result in the preceding example will turn out to be very important, so you should commit it to memory:

$$\boxed{\frac{d}{dx} \arctan x = \frac{d}{dx} \tan^{-1} x = \frac{1}{1 + x^2}}$$

The derivative of the inverse sine function, $y = \arcsin x$, will be derived in Problem 22. We list it here:

$$\boxed{\frac{d}{dx} \arcsin x = \frac{d}{dx} \sin^{-1} x = \frac{1}{\sqrt{1 - x^2}}}$$

The derivatives of the remaining inverse trigonometric functions are listed in the table of derivatives on the inside back cover of the book.

4.10.2 The Derivative of the Logarithmic Function

We introduced the logarithmic function to the base a, $\log_a x$, as the inverse function of the exponential function a^x (Figures 4.36 and 4.37). We can therefore use the formula for derivatives of inverse functions to find the derivative of $y = \log_a x$. Since

$$\log_a x = \frac{\ln x}{\ln a}$$

and $\ln a$ is a constant, it is enough to find the derivative of $\ln x$ (Figure 4.38). We set $f(x) = e^x$; then $f'(x) = e^x$ and $f^{-1}(x) = \ln x$. Therefore,

$$\frac{d}{dx}\ln x = \frac{d}{dx}[f^{-1}(x)] = \frac{1}{f'[f^{-1}(x)]} = \frac{1}{\exp[\ln x]} = \frac{1}{x}$$

Additionally, since $\log_a x = \frac{\ln x}{\ln a}$: we also find $\frac{d}{dx}(\log_a x) = \frac{1}{\ln a} \cdot \frac{d}{dx}\ln x = \frac{1}{x \ln a}$.

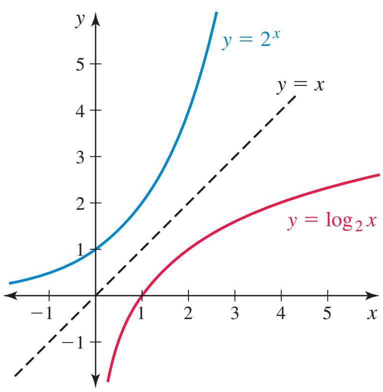


Figure 4.36 The function $y = \log_2 x$ as the inverse function of $y = 2^x$.

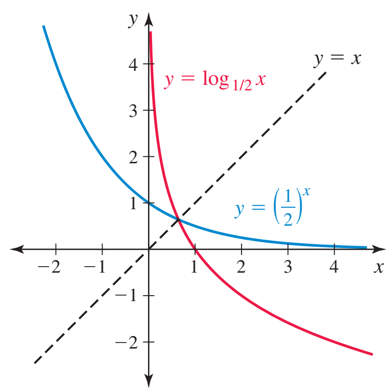


Figure 4.37 The function $y = \log_{1/2} x$ as the inverse function of $y = \left(\frac{1}{2}\right)^x$.

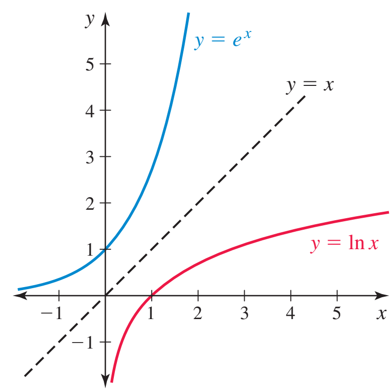


Figure 4.38 The natural logarithm $y = \ln x$ as the inverse function of the natural exponential function $y = e^x$.

We summarize these differentiation rules in the following box:

$$\frac{d}{dx}\ln x = \frac{1}{x}$$

$$\frac{d}{dx}\log_a x = \frac{1}{(\ln a)x}$$

Just as we did in the previous section we will start by practicing combining differentiation of logarithms with the chain rule, product rule, and so on, before discussing some biological examples.

EXAMPLE 5 Find the derivative of $y = \ln(3x)$.

Solution We use the chain rule with $u = g(x) = 3x$ and $f(u) = \ln u$:

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \frac{1}{u}3 = \frac{3}{3x} = \frac{1}{x}.$$

If you are surprised that the factor 3 disappeared, note that

$$y = \ln(3x) = \ln 3 + \ln x.$$

Since $\ln 3$ is a constant its derivative is 0. Hence,

$$\frac{d}{dx}(\ln 3 + \ln x) = \frac{d}{dx}\ln 3 + \frac{d}{dx}\ln x = 0 + \frac{1}{x} = \frac{1}{x}.$$

EXAMPLE 6 Find the derivative of $y = \ln(x^2 + 1)$.

Solution We can use the chain rule with $u = x^2 + 1$ and $f(u) = \ln u$. We obtain

$$y' = \frac{df}{du}\frac{du}{dx} = \frac{1}{u}2x = \frac{1}{x^2+1}2x = \frac{2x}{x^2+1}$$

The preceding example is of the form $y = \ln f(x)$. We will frequently encounter such functions; to find their derivatives, we need to use the chain rule, as shown in the following box:

$$\boxed{\frac{d}{dx}\ln f(x) = \frac{f'(x)}{f(x)}}$$

EXAMPLE 7 Differentiate $y = \ln(\sin x)$.

Solution

$$\frac{dy}{dx} = \frac{\cos x}{\sin x} = \cot x. \qquad y = \ln f(x) \text{ with } f(x) = \sin x,\ f'(x) = \cos x$$

EXAMPLE 8 Differentiate $y = x \ln x - x$.

Solution

$$\begin{aligned}\frac{dy}{dx} &= \frac{d}{dx}(x\ln x) - \frac{d}{dx}(x)\\ &= \frac{d}{dx}(x)\cdot \ln x + x\frac{d}{dx}(\ln x) - 1 \qquad \text{Product rule on first term}\\ &= 1\ln x + x\cdot\frac{1}{x} - 1 = \ln x\end{aligned}$$

EXAMPLE 9 Differentiate $y = \log(2x^3 - 1)$.

Solution

$$y' = \frac{1}{\ln 10}\frac{6x^2}{2x^3 - 1}. \qquad \text{Chain rule with } f(u) = \log u,\ g(x) = 2x^3 - 1,\ f'(u) = \tfrac{1}{(\ln 10)u}\ g'(x) = 6x$$

EXAMPLE 10 **Diatom Sinking** Diatoms are common marine algae that account for most of the photosynthesis that occurs in the ocean. As ***Vogel (2004)*** discusses, although diatoms live in the top layers of the ocean, they are denser than the water around them, so they sink slowly in still water. Only the constant mixing of water in the ocean prevents them from sinking into the deeper layers of the ocean. Diatoms are also apparently adapted to sink as slowly as possible, by evolving shapes that have very high fluid dynamic resistance. One such shape is a long filament. For a diatom whose shape is a filament of length L and radius a, the sinking speed is:

$$U = \frac{W(\ln(2L/a) + \frac{1}{2})}{4\pi\mu L}$$

where W is the weight difference between the diatom and the surrounding water and μ is the viscosity ("stickiness") of the water. Calculate dU/dL.

Solution We need to use the quotient rules to calculate dU/dL:

$$\frac{dU}{dL} = \frac{W}{4\pi\mu} \frac{\overbrace{\left(\ln(2L/a) + \frac{1}{2}\right)}^{u(L)}}{\underbrace{L}_{v(L)}} \qquad \frac{d}{dL}\left(\frac{u(L)}{v(L)}\right) = \frac{u'v - uv'}{v^2}$$

$$= \frac{W}{4\pi\mu} \frac{\left(\frac{1}{L} \cdot L - \left(\ln\left(\frac{2L}{a}\right) + \frac{1}{2}\right) \cdot L\right)}{L^2} \qquad u(L) = \ln\left(\frac{2}{a}\right) + \ln L + \frac{1}{2} \Rightarrow u'(L) = \frac{1}{L}$$

$$= \frac{W}{4\pi\mu L^2}\left(\frac{1}{2} - \ln\frac{2L}{a}\right)$$

For a long filament $\ln(2L/a)$ may be large, so typically $dU/dL < 0$, meaning that U decreases as L increases, i.e., longer filaments sink more slowly. •

EXAMPLE 11 **Population Growth** In Section 4.9 we discussed a model for the growth of a population that has a constant per capita reproductive rate, r. We argued that if the population size is $N(t)$, then $\frac{dN}{dt}$ obeys a differential equation:

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

We showed that this solution to this equation is $N(t) = N_o e^{rt}$. But the solution was pulled out of the air (although we were able to check that it indeed solved the differential equation). Using logarithms we can see where this solution might come from. A full discussion of solutions to differential equations must wait until Chapter 8. Rewrite the differential equation as

$$\frac{1}{N}\frac{dN}{dt} = r$$

We recognize the left-hand side as the implicit derivative with respect to time, t, of $\ln N$. The right-hand side is the derivative of rt. So

$$\frac{d}{dt}(\ln N) = \frac{d}{dt}(rt) \tag{4.15}$$

So the rate of change of $\ln N$ is the same as the rate of change of rt. One possible way this could occur is if:

$$\ln N = rt$$

We will find all possible solutions to (4.15) in Chapter 8, but we have shown directly that

$$\ln N = rt, \text{ i.e., } N(t) = e^{rt}$$

is one possible solution of the differential equation. •

4.10.3 Logarithmic Differentiation

Although we now have the techniques to differentiate complicated functions, sometimes the process of differentiating a function is aided by taking logarithms beforehand and using implicit differentiation. This idea dates back to the earliest days of calculus, and was invented by Leibniz and Bernoulli. Suppose, for example, that you are asked to differentiate:

$$y = \frac{e^x x^{3/2}\sqrt{1+x}}{(x^2+3)^4(3x-2)^3}$$

You could differentiate the expression using the quotient rule, product rule, and chain rule repeatedly. But the calculation is easier if we take logarithms of both sides first (see Example 13).

EXAMPLE 12 Find $\frac{dy}{dx}$ when $y = x^x$. This rather strange-looking function shows up a lot in probability, and knowing how to differentiate it is vital for fitting the parameters in probability models to real data.

Solution We take logarithms on both sides of the equation $y = x^x$:

$$\ln y = \ln x^x$$

Applying properties of the logarithm, we can simplify the right-hand side to $\ln x^x = x \ln x$. We can now differentiate both sides with respect to x.

$$\begin{aligned}
\frac{d}{dx}[\ln y] &= \frac{d}{dx}[x \ln x] \\
\frac{1}{y}\frac{dy}{dx} &= 1 \cdot \ln x + x \cdot \frac{1}{x} && \text{Use implicit differentiation on left-hand side} \\
\frac{dy}{dx} &= y[\ln x + 1] && \text{Solve for } dy/dx \\
\frac{dy}{dx} &= (\ln x + 1)x^x. && y = x^x
\end{aligned}$$

Although the function $y = x^x$ looks strange, we can write it as

$$y = x^x = \exp[\ln x^x] = \exp[x \ln x]$$

That is, $y = e^{x \ln x}$. We can differentiate this function without using logarithmic differentiation;

$$\begin{aligned}
\frac{dy}{dx} &= e^{x \ln x}\frac{d}{dx}(x \ln x) && \text{Using (4.13)} \\
&= e^{x \ln x}\left(1 \cdot \ln x + x\frac{1}{x}\right) \\
&= e^{x \ln x}(\ln x + 1).
\end{aligned}$$

Either approach will give you the correct answer. •

The next example should convince you that logarithmic differentiation can simplify finding the derivatives of complicated expressions.

EXAMPLE 13 Differentiate

$$y = \frac{e^x x^{3/2}\sqrt{1+x}}{(x^2+3)^4(3x-2)^3}.$$

Solution Without logarithmic differentiation, differentiating y would be rather difficult. Taking logarithms on both sides, however, we can simplify the right-hand side. Note that it is very important that we apply the properties of the logarithm before differentiating, as this will simplify the expressions that we must differentiate. We have

$$\begin{aligned}
\ln y &= \ln\left[\frac{e^x x^{3/2}\sqrt{1+x}}{(x^2+3)^4(3x-2)^3}\right] \\
&= \ln e^x + \ln x^{3/2} + \ln\sqrt{1+x} - \ln(x^2+3)^4 - \ln(3x-2)^3 \\
&= x + \frac{3}{2}\ln x + \frac{1}{2}\ln(1+x) - 4\ln(x^2+3) - 3\ln(3x-2) && \text{Simplify logarithms}
\end{aligned}$$

This no longer looks so daunting, and we can differentiate both sides:

$$\begin{aligned}
\frac{d}{dx}[\ln y] &= \frac{d}{dx}\left[x + \frac{3}{2}\ln x + \frac{1}{2}\ln(1+x) - 4\ln(x^2+3) - 3\ln(3x-2)\right] \\
\frac{1}{y}\frac{dy}{dx} &= 1 + \frac{3}{2}\cdot\frac{1}{x} + \frac{1}{2}\cdot\frac{1}{1+x} - 4\cdot\frac{2x}{x^2+3} - 3\cdot\frac{3}{3x-2}
\end{aligned}$$

Finally, solving for dy/dx yields

$$\frac{dy}{dx} = \left(1 + \frac{3}{2x} + \frac{1}{2(1+x)} - \frac{8x}{x^2+3} - \frac{9}{3x-2}\right) \frac{e^x x^{3/2}\sqrt{1+x}}{(x^2+3)^4(3x-2)^3}$$

We can also use this method to prove the general power rule (as promised in Section 4.3).

Power Rule (General Form) Let $f(x) = x^r$, where r is any real number. Then

$$\frac{d}{dx}(x^r) = rx^{r-1}$$

Proof We set $y = x^r$ and use logarithmic differentiation to obtain

$$\frac{d}{dx}[\ln y] = \frac{d}{dx}[\ln x^r]$$

$$\frac{1}{y}\frac{dy}{dx} = \frac{d}{dx}[r \ln x]$$

$$\frac{1}{y}\frac{dy}{dx} = r \cdot \frac{1}{x}.$$

Solving for dy/dx yields

$$\frac{dy}{dx} = r \cdot \frac{y}{x} = r \cdot \frac{x^r}{x} = rx^{r-1}.$$

Section 4.10 Problems

4.10.1

In Problems 1–6, find the inverse of each function and differentiate each inverse in two ways: (i) Differentiate the inverse function directly, and (ii) use (4.14) to find the derivative of the inverse.

1. $f(x) = \sqrt{2x+1}, x \geq -\frac{1}{2}$ **2.** $f(x) = \sqrt{x+1}, x \geq -1$

3. $f(x) = 2x^2 + 2, x \geq 0$ **4.** $f(x) = 3x^2 + 2, x \geq 0$

5. $f(x) = 3 - 2x^3, x \geq 0$ **6.** $f(x) = x^2 + 1, x \geq 0$

In Problems 7–22, use (4.14) to find the derivative of the inverse at the indicated point.

7. Let

$$f(x) = 2x^2 - 2, \quad x \geq 0$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=0}$. [Note that $f(1) = 0$.]

8. Let

$$f(x) = -x^3 + 7, \quad x \geq 0$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=-1}$. [Note that $f(2) = -1$.]

9. Let

$$f(x) = \sqrt{x+1}, \quad x \geq 0$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=2}$. [Note that $f(3) = 2$.]

10. Let

$$f(x) = \sqrt{2+x^2}, \quad x \geq 0$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=\sqrt{3}}$. [Note that $f(1) = \sqrt{3}$.]

11. Let

$$f(x) = x + 2e^x, \quad x \in \mathbf{R}$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=2}$. [Note that $f(0) = 2$.]

12. Let

$$f(x) = 2x + \ln(x+1), \quad x > -1$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=0}$. [Note that $f(0) = 0$.]

13. Let

$$f(x) = x + \sin x, \quad x \in \mathbf{R}$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=\pi}$. [Note that $f(\pi) = \pi$.]

14. Let

$$f(x) = x + e^x \cos x, \quad x \in \mathbf{R}$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=1}$. [Note that $f(0) = 1$.]

15. Let

$$f(x) = \frac{x}{2} + \tan x, \quad x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=0}$. [Note that $f(0) = 0$.]

16. Let

$$f(x) = \frac{x^2}{\pi^2} + \tan x, \quad x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

Find $\frac{d}{dx} f^{-1}(x)\big|_{x=17/16}$. [Note that $f(\frac{\pi}{4}) = \frac{17}{16}$.]

17. Let $f(x) = \ln(\sin x)$, $0 < x < \pi/2$. Find $\frac{d}{dx} f^{-1}(x)$ at $x = -\ln 2$.

18. Let $f(x) = \ln(\tan x)$, $0 < x < \pi/2$. Find $\frac{d}{dx} f^{-1}(x)$ at $x = \frac{\ln 3}{2}$.

19. Let $f(x) = x^5 + x + 1$, $-1 < x < 1$. Find $\frac{d}{dx} f^{-1}(x)$ at $x = 1$.

20. Let $f(x) = e^{-x^2} + x$. Find $\frac{d}{dx} f^{-1}(x)$ at $x = 1$.

21. Let $f(x) = e^{-x^2/2} + 2x$. Find $\frac{d}{dx} f^{-1}(x)$ at $x = 1$.

22. Denote the inverse of $y = \sin x$, $-\frac{\pi}{2} \le x \le \frac{\pi}{2}$, by $y = \arcsin x$, $-1 \le x \le 1$. Show that

$$\frac{d}{dx} \arcsin x = \frac{1}{\sqrt{1-x^2}}, \qquad -1 < x < 1$$

4.10.2

In Problems 23–60, differentiate the functions with respect to the independent variable. (Note that* log *denotes the logarithm to base* 10.*)

23. $f(x) = \ln(x+1)$

24. $f(x) = \ln(3x+4)$

25. $f(x) = \ln(1-2x)$

26. $f(x) = \ln(4-3x)$

27. $f(x) = \ln x^2$

28. $f(x) = \ln(1-x^2)$

29. $f(x) = \ln(2x^3 - x)$

30. $f(x) = \ln(1-x^3)$

31. $f(x) = (\ln x)^2$

32. $f(x) = (\ln x)^3$

33. $f(x) = (\ln x^2)^2$

34. $f(x) = (\ln(1-x^2))^3$

35. $f(x) = \ln\sqrt{x^2+1}$

36. $f(x) = \ln\sqrt{2x^2 - x}$

37. $f(x) = \ln \dfrac{x}{x+1}$

38. $f(x) = \ln \dfrac{2x}{1+x^2}$

39. $f(x) = \ln \dfrac{1-x}{1+2x}$

40. $f(x) = \ln \dfrac{x^2-1}{x^3-1}$

41. $f(x) = \exp[x - \ln x]$

42. $g(s) = \exp[s^2 + \ln s]$

43. $f(x) = \ln(\sin x)$

44. $f(x) = \ln(\cos(1-x))$

45. $f(x) = \ln(\tan x^2)$

46. $g(s) = \ln(\sin^2(3s))$

47. $f(x) = x \ln x$

48. $f(x) = x^2 \ln x^2$

49. $f(x) = \dfrac{\ln x}{x}$

50. $h(t) = \dfrac{\ln t}{1+t^2}$

51. $h(t) = \sin(\ln(3t))$

52. $h(s) = \ln(\ln s)$

53. $f(x) = \ln|x^2 - 3|$

54. $f(x) = \log(2x^2 - 1)$

55. $f(x) = \log(1-x^2)$

56. $f(x) = \log(3x^3 - x + 2)$

57. $f(x) = \log(x^3 - 3x)$

58. $f(x) = \log(\sqrt[3]{\tan x^2})$

59. $f(u) = \log_3(3 + u^4)$

60. $g(s) = \log_5(3^s - 2)$

61. Let $f(x) = \ln x$. We know that $f'(x) = \frac{1}{x}$. We will use this fact and the definition of derivatives to show that

$$\lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n = e$$

(a) Use the definition of the derivative to show that

$$f'(1) = \lim_{h\to 0} \frac{\ln(1+h)}{h}$$

(b) Show that (a) implies that

$$\ln[\lim_{h\to 0}(1+h)^{1/h}] = 1$$

(c) Set $h = \frac{1}{n}$ in (b) and let $n \to \infty$. Show that this implies that

$$\lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n = e$$

62. Assume that $f(x)$ is differentiable with respect to x. Show that

$$\frac{d}{dx} \ln\left[\frac{f(x)}{x}\right] = \frac{f'(x)}{f(x)} - \frac{1}{x}$$

63. Fetal Growth ***Royston and Wright (1998)*** studied how the size of an unborn fetus depends on its age. They fitted data for head circumference (H) as a function of age (t) in weeks using the formula

$$H = -29.32 + 1.705t^2 - 0.3981t^2 \log t$$

(a) Calculate the rate of fetal growth dH/dt.

(b) Is dH/dt larger early in development (say at $t = 8$ weeks) or late (say at $t = 36$ weeks)?

(c) Repeat part (b) but for fractional rate of growth $\frac{1}{H}\frac{dH}{dt}$.

64. Propulsive Force of a Swimming Worm Some worms swim by passing an undulatory wave along their bodies. The force that small worms apply to the water by passing this wave can be modeled using a formula derived by ***Lamb (1911)***

$$F = \frac{4\pi L\mu U}{\left(-0.077 - \ln\left(\frac{\rho U a}{4\mu}\right)\right)}$$

where U is the velocity of undulation, L is the length of the worm, a is the radius of the worm's body, and μ and ρ respectively the viscosity (or "stickiness") and density of the water through which the worm swims.

Calculate dF/dU, the rate of change of the force with increasing undulation velocity.

4.10.3

In Problems 63–74, use logarithmic differentiation to find the first derivative of the given functions.

65. $f(x) = 2x^x$

66. $f(x) = (2x)^{2x}$

67. $f(x) = (\ln x)^x$

68. $f(x) = (\ln x)^{3x}$

69. $f(x) = x^{\ln x}$

70. $f(x) = x^{2\ln x}$

71. $f(x) = x^{1/x}$

72. $f(x) = x^{3/x}$

73. $y = x^{x^x}$

74. $y = (x^x)^x$

75. $y = x^{\cos x}$

76. $y = (\cos x)^x$

77. Differentiate

$$y = \frac{e^{2x}(9x-2)^3}{\sqrt[4]{(x^2+1)(3x^3-7)}}.$$

78. Differentiate

$$y = \frac{e^{x-1}\sin^2 x}{(x^2+5)^{2x}}.$$

4.11 Linear Approximation and Error Propagation

Suppose we want to find an approximation to $\ln(1.05)$ without using a calculator. The method for solving this problem will be useful in many other applications. Let's look at the graph of $f(x) = \ln x$ (Figure 4.39). We know that $\ln 1 = 0$, and we see that 1.05 is quite close to 1—so close, in fact, that the curve connecting $(1, 0)$ to $(1.05, \ln 1.05)$ is close to a straight line. This suggests that we should approximate the curve by a

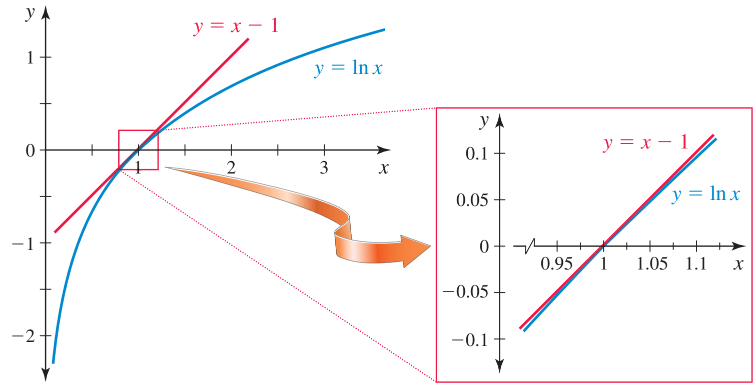


Figure 4.39 The tangent line approximation for $\ln x$ at $x = 1$ to approximate $\ln(1.05)$. When x is close to 1, the tangent line and the graph of $y = \ln x$ are close (see inset).

straight line—but not just any straight line: We choose the tangent line to the graph of $f(x) = \ln x$ at $x = 1$ (Figure 4.39). We can find the equation of the tangent line without a calculator. We note that the slope of $f(x) = \ln x$ at $x = 1$ is $f'(1) = \frac{1}{x}\Big|_{x=1} = 1$. This, together with the point $(1, 0)$, allows us to find the tangent line at $x = 1$:

$$f(x) = L(x) = f(1) + f'(1)(x - 1) = 0 + (1)(x - 1) = x - 1$$

We call $L(x)$ the **tangent line approximation**, or the **linearization**, of $f(x)$ at $x = 1$. If we evaluate $L(x)$ at $x = 1.05$, we find that $L(1.05) = 1.05 - 1 = 0.05$, which is a good approximation to $\ln 1.05 = 0.048790\ldots$. (Here, we used the calculator to see how close the approximation is to the exact value.)

The Tangent Line Approximation

Assume that $y = f(x)$ is differentiable at $x = a$; then

$$L(x) = f(a) + f'(a)(x - a)$$

is the **tangent line approximation**, or the **linearization**, of f at $x = a$.

Geometrically, the linearization of f at $x = a$, $L(x) = f(a) + f'(a)(x - a)$, is the equation of the tangent line to the graph of $f(x)$ at the point $(a, f(a))$. (See Figure 4.40.)

If $|x - a|$ is sufficiently small, then $f(x)$ can be linearly approximated by $L(x)$; that is,

$$f(x) \approx f(a) + f'(a)(x - a)$$

This approximation is illustrated in Figure 4.41.

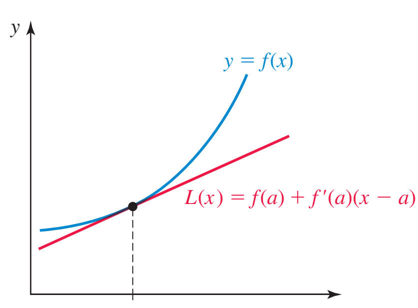


Figure 4.40 The tangent line approximation of $y = f(x)$ at $x = a$.

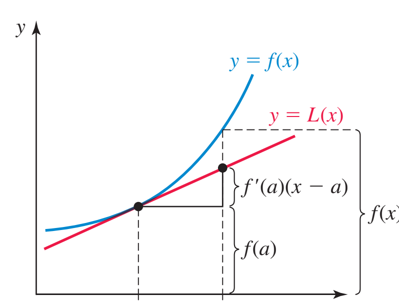


Figure 4.41 The linearization of f at $x = a$ can be used to approximate $f(x)$ if x is close to a.

EXAMPLE 1 **(a)** Find the linear approximation of $f(x) = \sqrt{x}$ at $x = a$, and
(b) use your answer in (a) to find an approximate value of $\sqrt{50}$.

Solution **(a)** The linear approximation at $x = a$ is

$$L(x) = f(a) + f'(a)(x - a)$$
$$= \sqrt{a} + \frac{1}{2\sqrt{a}}(x - a) \qquad f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

(See Figure 4.42.)

(b) To find the approximate value of $f(50) = \sqrt{50}$, we need to choose a value for a close to 50 and for which we know $\sqrt{a}$ exactly. Our choice is $a = 49$. We thus approximate $f(50)$ by $L(50)$ with $a = 49$ and find that

$$\sqrt{50} \approx \sqrt{49} + \frac{50 - 49}{2\sqrt{49}} = 7 + \frac{1}{14} \approx 7.0714$$

Using a calculator to compute $\sqrt{50} = 7.0711\ldots$, we see that the error in the linear approximation is quite small. •

EXAMPLE 2 Find the linear approximation of $f(x) = \sin x$ at $x = 0$.

Solution Since $f'(x) = \cos x$, it follows that

$$L(x) = f(0) + f'(0)(x - 0)$$
$$= \sin 0 + (\cos 0)x = x$$

(Figure 4.43). That is, for small values of x, we can approximate $\sin x$ by x. This approximation is often used in physics. (Note that x is measured in radians.) •

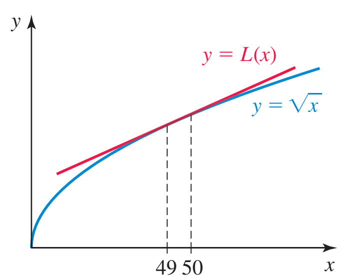


Figure 4.42 The linear approximation of $f(x) = \sqrt{x}$ at $x = 49$ is the line $y = L(x)$.

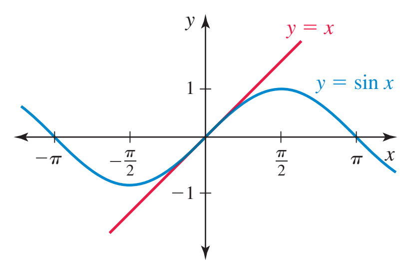


Figure 4.43 The linear approximation of $y = \sin x$ at $x = 0$ is the line $y = x$.

EXAMPLE 3 **Population Growth** Let $N(t)$ be the size of a population at time t, and assume that the growth rate dN/dt of the population is given by

$$\frac{dN}{dt} = f(N)$$

where $f(N)$ is a differentiable function with $f(0) = 0$. Find the linearization of the growth rate at $N = 0$.

Solution We need to find the tangent line approximation of $f(N)$ at $N = 0$. If we denote the linearization of $f(N)$ by $L(N)$, we obtain

$$L(N) = f(0) + f'(0)N$$

For biological reasons: we assume $f(0) = 0$. When the population size is 0, the growth rate should be 0; otherwise, we would have spontaneous creation if $f(0) > 0$, or the population size would become negative if $f(0) < 0$. If we set $r = f'(0)$, we find that, for N close to 0,

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

The preceding formula shows that the population changes approximately exponentially when its size is small. This behavior is observed, for instance, when bacteria are grown in a nutrient-rich environment at a low population density. •

EXAMPLE 4 Let $N(t)$ be the size of a bacterial population at time t, measured in millions, and assume that the per capita growth rate is equal to 2%. We can express this statement in a differential equation, namely,

$$\frac{dN}{dt} = 0.02N$$

Suppose we know that at time $t = 10$ the size of the population is 250,000,000; that is, $N(10) = 250$ (since we measure the population size in millions). Use a linear approximation to predict the approximate population size at time $t = 10.1$.

Solution To predict the population size at time 10.1, we use the following linearization of $N(t)$ at $t = 10$:

$$L(t) = N(10) + N'(10)(t - 10)$$

To evaluate $L(t)$ at $t = 10.1$, we need to find $N'(10)$.

Using the differential equation, we find that

$$N'(t) = (0.02)N(t)$$

When $t = 10$, we obtain

$$N'(10) = (0.02)N(10) = (0.02)(250) = 5$$

Hence,

$$\begin{aligned} L(10.1) &= N(10) + N'(10)(10.1 - 10) \\ &= 250 + (5)(0.1) = 250.5 \end{aligned}$$

Thus, we predict that the population size at time 10.1 is approximately 250,500,000. Note that this approximation is good only if we want to predict the population growth over very short time intervals. •

Error Propagation Linear approximations are used in problems of **error propagation**. Suppose that you wish to determine the surface area of a spherical cell. Since the surface area S of a sphere with radius r is given by

$$S = 4\pi r^2$$

it suffices to measure the radius r. If your measurement of the radius is accurate within 3%, how accurately can you measure the cell's surface area?

First we must discuss what it means for a measurement to be accurate within a certain percentage. Suppose that x_0 is the true value of a variable and x is the measured value. Then $|\Delta x| = |x - x_0|$ is the **absolute error**, or tolerance, in measurement. The **relative error** is defined as $|\Delta x/x_0|$ and the **percentage error** as $100|\Delta x/x_0|$.

Returning to our example, let's find the error that arises in computing the surface area. We start with the absolute error of the surface area,

$$|\Delta S| = |S(r_0 + \Delta r) - S(r_0)|$$

where r_0 is the true radius and $|\Delta r|$ is the absolute error in the measurement of the radius. We approximate $S(r_0 + \Delta r) - S(r_0)$ by its linear approximation $S'(r_0)\Delta r$; that is,

$$\Delta S \approx S'(r_0)\Delta r$$

Since $S'(r) = 8\pi r$, the percentage error in the measurement of the surface area is

$$100\left|\frac{\Delta S}{S(r_0)}\right| \approx 100\left|\frac{S'(r_0)\Delta r}{S(r_0)}\right| = 100\left|\frac{\Delta r}{r_0}\right|\left|\frac{S'(r_0)r_0}{S(r_0)}\right|$$

$$= \underbrace{100\left|\frac{\Delta r}{r_0}\right|}_{=3}\underbrace{\left|\frac{8\pi r_0^2}{4\pi r_0^2}\right|}_{=2} = 6$$

because $100|\Delta r/r_0| = 3$. In other words, the surface area is (approximately) accurate within 6% if the radius is accurate within 3%. The doubling of the percentage error will be explained in the next example.

EXAMPLE 5 Suppose that you wish to determine $f(x)$ from a measurement of x. If $f(x)$ is given by a power function, namely, $f(x) = cx^s$, how does an error in the measurement of x propagate; i.e., what will the error in $f(x)$ be?

Solution Since $f'(x) = csx^{s-1}$, we have

$$\Delta f \approx f'(x)\Delta x = csx^{s-1}\Delta x$$

The percentage error $100\left|\frac{\Delta f}{f}\right|$ is therefore related to the percentage error $100\left|\frac{\Delta x}{x}\right|$ as follows:

$$100\left|\frac{\Delta f}{f}\right| \approx 100\left|\frac{f'(x)\Delta x}{f(x)}\right| = 100\left|\frac{\Delta x}{x}\right|\left|\frac{f'(x)x}{f(x)}\right|$$

$$= 100\left|\frac{\Delta x}{x}\right|\left|\frac{csx^s}{cx^s}\right| = \left(100\left|\frac{\Delta x}{x}\right|\right)|s| \tag{4.16}$$

In our previous example, $s = 2$; hence, the percentage error in the surface area measurement is twice the percentage error in the radius measurement. •

EXAMPLE 6 **Allometric Growth** Suppose that you wish to estimate the total leaf area of a tree. Experimental data collected by ***Niklas, 1994*** indicates that

$$[\text{leaf area}] \propto [\text{stem diameter}]^{1.84}$$

Instead of trying to measure the total leaf area directly, you measure the stem diameter and then use the scaling relationship to estimate the total leaf area. How accurately must you measure the stem diameter if you want to estimate the leaf area within an error of 10%?

Solution We denote the leaf area by A and stem diameter by d. Then

$$A(d) = cd^{1.84}$$

where c is the constant of proportionality. An error in measurement of d is propagated as

$$\Delta A \approx A'(d)\Delta d = c(1.84)d^{0.84}\Delta d$$

The percentage error $100\left|\frac{\Delta A}{A}\right|$ is related to the percentage error $100\left|\frac{\Delta d}{d}\right|$ by:

$$100\left|\frac{\Delta A}{A}\right| \approx 100\left|\frac{A'(d)\Delta d}{A(d)}\right| = 100\left|\frac{\Delta d}{d}\right|\left|\frac{A'(d)d}{A(d)}\right|$$

$$= 100\left|\frac{\Delta d}{d}\right|\left|\frac{c(1.84)d^{0.84}d}{cd^{1.84}}\right| = \left(100\left|\frac{\Delta d}{d}\right|\right)(1.84).$$

We require that $100\left|\frac{\Delta A}{A}\right| = 10$. Hence,

$$10 = (1.84)\left(100\left|\frac{\Delta d}{d}\right|\right)$$

or

$$100\left|\frac{\Delta d}{d}\right| = \frac{10}{1.84} = 5.4.$$

That is, we must measure the stem diameter to within an error of 5.4%.

Using the result of Example 5, we could have found the same error immediately. Since

$$A(d) = cd^{1.84}$$

we get $s = 1.84$, where s is the notation for the exponent that was used in Example 5. Using

$$100\left|\frac{\Delta A}{A}\right| = |s|\left(100\left|\frac{\Delta d}{d}\right|\right)$$

we obtain

$$100\left|\frac{\Delta d}{d}\right| = \frac{1}{|s|}\left(100\left|\frac{\Delta A}{A}\right|\right) = \frac{10}{1.84} = 5.4$$

as before. •

Although the linearization method is particularly helpful for analyzing how measurement errors can be related between different variables, it can also be used as a method to understand biological models generally, as we show in the following example.

EXAMPLE 7 **Blood Oxygenation** Hemoglobin in red blood cells binds to oxygen, allowing the cells to carry oxygen from the lungs to tissues throughout the body. One model for the binding of oxygen to hemoglobin is Hill's equation, which gives the fraction of hemoglobin bound to oxygen, $f(P)$, as a function of oxygen concentration, P (measured in mmHg) as $f(P) = \frac{P^m}{k^m + P^m}$. Here m and k are coefficients that take different values for different species, and for individuals living at high altitude versus sea level, and so on.

Assume that $k = 30$ mmHg and $m = 3$. Show that the percentage change in $f(P)$ due to a change in P is larger when $P = 10$ than when $P = 30$.

Solution First we write out the linearization formula. Since $f(P)$ is not a power law, there is no simple rule for relating $\Delta f/f$ with $\Delta P/P$:

$$\begin{aligned}\Delta f &= f'(P)\Delta P\\ &= \left(\frac{mP^{m-1}(k^m + P^m) - P^m(mP^{m-1})}{(k^m + P^m)^2}\right)\Delta P \qquad \text{Quotient rule with } u(P) = P^m,\ v(P) = k^m + P^m\\ &= \frac{mP^{m-1}\cdot k^m}{(k^m + P^m)^2}\Delta P\end{aligned}$$

so

$$100\frac{\Delta f}{f} = 100\frac{\overbrace{\frac{mP^{m-1}k^m}{(k^m+P^m)^2}\Delta P}^{\Delta f}}{\underbrace{\left(\frac{P^m}{k^m+P^m}\right)}_{f(P)}} = \frac{mk^m}{(k^m+P^m)}\cdot 100\cdot\frac{\Delta P}{P}$$

So, if $100\frac{\Delta P}{P} = 5$ then

$$100\frac{\Delta f}{f} = \frac{5mk^m}{k^m + P^m} = \frac{5\times 3\times 30^3}{30^3 + 10^3} = 14.5\% \text{ when } P = 10$$

and $100\frac{\Delta f}{f} = 7.5\%$ when $P = 30$.

In general if $m = 3$ the same percentage increase in P has a smaller and smaller effect on $f(P)$ as P increases. In Chapter 5 we will show that this result gives information about the graph of $f(P)$. •

Section 4.11 Problems

In Problems 1–10, use the formula

$$f(x) \approx f(a) + f'(a)(x - a)$$

to approximate the value of the given function. Then compare your result with the value you get from a calculator.

1. $\sqrt{65}$; let $f(x) = \sqrt{x}$, $a = 64$, and $x = 65$

2. $\sqrt{35}$; let $f(x) = \sqrt{x}$, $a = 36$, and $x = 35$

3. $\sqrt[3]{124}$

4. $(7.9)^3$

5. $(0.99)^{25}$

6. $\tan(0.01)$

7. $\sin\left(\frac{\pi}{2} + 0.02\right)$

8. $\cos\left(\frac{\pi}{4} - 0.01\right)$

9. $\ln(1.01)$

10. $e^{0.1}$

In Problems 11–30, calculate the linear approximation for $f(x)$:

$$f(x) \approx f(a) + f'(a)(x - a)$$

11. $f(x) = \dfrac{1}{1+x}$ at $a = 0$

12. $f(x) = \dfrac{1}{1-x}$ at $a = 0$

13. $f(x) = \dfrac{2}{1+x}$ at $a = 1$

14. $f(x) = \dfrac{1}{3-2x}$ at $a = 2$

15. $f(x) = \dfrac{1}{(1+x)^2}$ at $a = 0$

16. $f(x) = \dfrac{1}{(1-x)^2}$ at $a = 0$

17. $f(x) = \ln(1 + x)$ at $a = 0$

18. $f(x) = \ln(1 + 2x)$ at $a = 0$

19. $f(x) = \log x$ at $a = 1$

20. $f(x) = \log(1 + x^2)$ at $a = 0$

21. $f(x) = e^x$ at $a = 0$

22. $f(x) = e^{2x}$ at $a = 0$

23. $f(x) = e^{-x}$ at $a = 0$

24. $f(x) = e^{-3x}$ at $a = 0$

25. $f(x) = e^{x-1}$ at $a = 1$

26. $f(x) = e^{2x+1}$ at $a = -1/2$

27. $f(x) = (1 + x)^{-n}$ at $a = 0$. (Assume that n is a positive integer.)

28. $f(x) = (1 - x)^{-n}$ at $a = 0$. (Assume that n is a positive integer.)

29. $f(x) = \sqrt{1 + x^2}$ at $a = 0$

30. $f(x) = \left(1 + \dfrac{1}{x}\right)^{1/4}$ at $a = 1$

31. Population Growth Suppose that the per capita growth rate of a population is 3%; that is, if $N(t)$ denotes the population size at time t, then

$$\frac{1}{N}\frac{dN}{dt} = 0.03$$

Suppose also that the population size at time $t = 4$ is equal to 100. Use a linear approximation to compute the population size at time $t = 4.1$.

32. Population Growth Suppose that the per capita growth rate of a population is 2%; that is, if $N(t)$ denotes the population size at time t, then

$$\frac{1}{N}\frac{dN}{dt} = 0.02$$

Suppose also that the population size at time $t = 2$ is equal to 100. Use a linear approximation to compute the population size at time $t = 2.1$.

33. Plant Biomass Suppose that the specific growth rate of a plant is 1%; that is, if $B(t)$ denotes the biomass at time t, then

$$\frac{1}{B(t)}\frac{dB}{dt} = 0.01$$

Suppose that the biomass at time $t = 1$ is equal to 5 grams. Use a linear approximation to compute the biomass at time $t = 1.1$.

34. Plant Biomass Suppose that a certain plant is grown along a gradient ranging from nitrogen-poor to nitrogen-rich soil. Experimental data show that the average mass per plant grown in a soil with a total nitrogen content of 1000 mg nitrogen per kg of soil is 2.7 g and the rate of change of the average mass per plant at this nitrogen level is 1.05×10^{-3} g per mg change in total nitrogen per kg soil. Use a linear approximation to predict the average mass per plant grown in a soil with a total nitrogen content of 1100 mg nitrogen per kg of soil.

In Problems 35–40, a measurement error in x affects the accuracy of the value $f(x)$. In each case, determine an interval of the form

$$[f(x) - \Delta f, f(x) + \Delta f]$$

that reflects the measurement error Δx. In each problem, the quantities given are $f(x)$ and x = true value of $x \pm |\Delta x|$.

35. $f(x) = 2x$, $x = 1 \pm 0.1$

36. $f(x) = 1 - 3x$, $x = -2 \pm 0.3$

37. $f(x) = 3x^2$, $x = 2 \pm 0.1$

38. $f(x) = \sqrt{x}$, $x = 10 \pm 0.5$

39. $f(x) = e^x$, $x = 2 \pm 0.2$

40. $f(x) = \sin x$, $x = -1 \pm 0.05$

In Problems 41–44, assume that the measurement of x is accurate within 2%. In each case, determine the error Δf in the calculation of f and find the percentage error $100\frac{\Delta f}{f}$. The quantities $f(x)$ and the true value of x are given.

41. $f(x) = 4x^3$, $x = 1.5$

42. $f(x) = x^{1/4}$, $x = 10$

43. $f(x) = \ln x$, $x = 20$

44. $f(x) = \dfrac{1}{1+x}$, $x = 4$

45. The volume V of a spherical cell of radius r is given by

$$V(r) = \frac{4}{3}\pi r^3$$

If you can determine the radius to within an accuracy of 3%, how accurate is your calculation of the volume?

46. Blood Flow Rate The speed v of blood flowing along the central axis of an artery of radius R is given by Poiseuille's law,

$$v(R) = cR^2$$

where c is a constant. If you can determine the radius of the artery to within an accuracy of 5%, how accurate is your calculation of the speed?

47. Allometric Growth Suppose that you are studying reproduction in moss. (***Niklas, 1994***) found a scaling relation

$$N \propto L^{2.11}$$

between the number of moss spores (N) and the capsule length (L). To estimate the number of moss spores, you measure the capsule length. If you wish to estimate the number of moss spores within an error of 5%, how accurately must you measure the capsule length?

48. Effect of Nutrients on Growth Suppose that the rate of growth of a plant in a certain habitat depends on a single resource—for instance, nitrogen. The dependence of the growth rate $f(R)$ on the resource level R is modeled using Monod's equation

$$f(R) = a\frac{R}{k+R}$$

where a and k are constants. Express the percentage error of the growth rate, $100\frac{\Delta f}{f}$, as a function of the percentage error of the resource level, $100\frac{\Delta R}{R}$.

49. Chemical Reaction The reaction rate $R(x)$ of the irreversible reaction

$$\mathrm{A} + \mathrm{B} \to \mathrm{AB}$$

is a function of the concentration x of the product AB and is given by

$$R(x) = k(a-x)(b-x)$$

where k is a constant, a is the concentration of A at the beginning of the reaction, and b is the concentration of B at the beginning of the reaction. Express the error of the reaction rate, ΔR, as a function of the error of the concentration Δx.

Chapter 4 Review

Key Terms

Discuss the following definitions and concepts:

1. Derivative, formal definition
2. Difference quotient
3. Secant line and tangent line
4. Instantaneous rate of change
5. Average rate of change
6. Differential equation
7. Differentiability and continuity
8. Power rule
9. Sum and multiplication rules
10. Product rule
11. Quotient rule
12. Chain rule
13. Implicit function
14. Implicit differentiation
15. Related rates
16. Higher derivatives
17. Derivatives of trigonometric functions
18. Derivatives of exponential functions
19. Derivatives of inverse and logarithmic functions
20. Logarithmic differentiation
21. Tangent line approximation
22. Error propagation
23. Absolute error, relative error, percentage error

Review Problems

In Problems 1–8, differentiate with respect to the independent variable.

1. $f(x) = -3x^4 + \dfrac{2}{\sqrt{x}} + 1$ **2.** $g(x) = \dfrac{1}{\sqrt{x^3+4}}$

3. $h(t) = \left(\dfrac{1-t}{1+t}\right)^{1/3}$ **4.** $f(x) = (x^2+1)e^{-x}$

5. $f(x) = e^{2x}\sin\left(\dfrac{\pi}{2}x\right)$ **6.** $g(s) = \dfrac{\sin(3s+1)}{\cos(3s)}$

7. $f(x) = x^2\ln x - x^2$ **8.** $g(x) = e^{-x}\cos(x+1)$

In Problems 9–12, find the first and second derivatives of the given functions.

9. $f(x) = e^{-x^2/2}$ **10.** $g(x) = \tan(x^2+1)$

11. $h(x) = \dfrac{x}{x+1}$ **12.** $f(x) = \dfrac{e^{-x}}{e^{-x}+1}$

In Problems 13–16, find dy/dx.

13. $x^2y - y^2x = \sin x$ **14.** $e^{x^2+y^2} = 1$

15. $\ln(x-y) = x$ **16.** $xy = 2$

In Problems 17–19, find dy/dx and d^2y/dx^2.

17. $x^2 + y^2 = 16$ **18.** $x = \tan y$ **19.** $e^y = \ln x$

20. Assume that x is a function of t. Find $\frac{dy}{dt}$ when $y = \cos x$ and $\frac{dx}{dt} = \sqrt{3}$ for $x = \frac{\pi}{3}$.

21. Velocity A flock of birds passes directly overhead, flying horizontally at an altitude of 100 feet and a speed of 6 feet per second. How quickly is the distance between you and the birds increasing when the distance is 320 feet? (You are on the ground and are not moving.)

22. Find the derivative of

$$y = \ln|\cos x|$$

23. Suppose that $f(x)$ is differentiable. Find an expression for the derivative of each of the following functions:

(a) $y = e^{f(x)}$ **(b)** $y = \ln f(x)$ **(c)** $y = [f(x)]^2$

24. Find the tangent line and the normal line to $y = \ln(x+1)$ at $x = 1$.

T **25.** Suppose that

$$f(x) = \frac{x^2}{1+x^2},\ x \ge 0$$

(a) Use a graphing calculator to graph $f(x)$ for $x \ge 0$. Note that the graph is S shaped.

(b) Find a line through the origin that touches the graph of $f(x)$ at some point $(c, f(c))$ with $c > 0$. This is the tangent line at $(c, f(c))$ that goes through the origin. Graph the tangent line in the same coordinate system that you used in (a).

In Problems 26–29, find an equation for the tangent line to the curve at the specified point.

26. $y = \sin\sqrt{x}$ at $x = \dfrac{\pi}{2}$ **27.** $y = e^{-x^2}\cos x$ at $x = \dfrac{\pi}{3}$

28. $x^2 + (y+1)^2 = 1^2$ at $x = 1$

29. $x^2 - y^2 = 1$ at $x = 2$

30. In Review Problem 19 of Chapter 3, we introduced the following hyperbolic functions:

$$\sinh x = \frac{e^x - e^{-x}}{2}, \quad x \in \mathbf{R}$$

$$\cosh x = \frac{e^x + e^{-x}}{2}, \quad x \in \mathbf{R}$$

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad x \in \mathbf{R}$$

(a) Show that

$$\frac{d}{dx} \sinh x = \cosh x$$

and

$$\frac{d}{dx} \cosh x = \sinh x$$

(b) Use the facts that

$$\tanh x = \frac{\sinh x}{\cosh x}$$

and

$$\cosh^2 x - \sinh^2 x = 1$$

together with your results in (a) to show that

$$\frac{d}{dx} \tanh x = \frac{1}{\cosh^2 x}$$

31. Find a second-degree polynomial

$$p(x) = ax^2 + bx + c$$

with $p(-1) = 6$, $p'(1) = 8$, and $p''(0) = 4$.

32. Use the geometric interpretation of the derivative to find the equations of the tangent lines to the curve

$$x^2 + y^2 = 1$$

at the following points:

(a) $(1, 0)$ **(b)** $(\frac{1}{2}, \frac{1}{2}\sqrt{3})$

(c) $(-\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2})$ **(d)** $(0, -1)$

33. Sinking Diatoms In Section 4.10 we studied the sinking of diatoms. We mentioned then that many diatoms have a filamentous shape. Why do they have this shape? In this problem you will compare sinking of filamentous diatoms and spherical diatoms.

The sinking speed of a spherical diatom of radius r is given by the formula:

$$U = \frac{2}{9} \frac{r^2 \rho g}{\mu}$$

where μ is the viscosity ("stickiness") of seawater, ρ is the buoyancy corrected density of the diatom (that is, the mass of 1 cm^3 of diatom, minus the mass of 1 cm^3 of seawater) and g is the gravitational acceleration constant.

(a) We want to consider the effect of changing the volume of the diatom upon its sinking velocity. Supposing that the diatom is a sphere, its volume is given by $V = \frac{4}{3}\pi r^3$.

If V increases by 5%, predict the corresponding percentage change in U.

(b) In contrast a filamentous diatom with length L and radius a will sink at a speed:

$$U = \frac{\rho a^2 g}{4\pi \mu} \left(\ln(2L/a) + \frac{1}{2} \right)$$

If the volume of the diatom is $V = \pi a^2 L$ and V is varied by varying L (i.e., changing the length of the filament but not its radius) predict the percentage change in U that would occur if V were changed by 5%. Assume that initially $L/a = 20$.

34. Mushroom Spore Dispersal ***Dressaire et al. (2016)*** have shown that mushrooms are capable of dispersing their spores by creating microscopic winds that blow the spores out from under the mushroom. For a mushroom with height h, ***Dressaire et al.***, showed that these winds blow the spores a distance:

$$d = k \cdot h^2$$

where k is a positive constant.

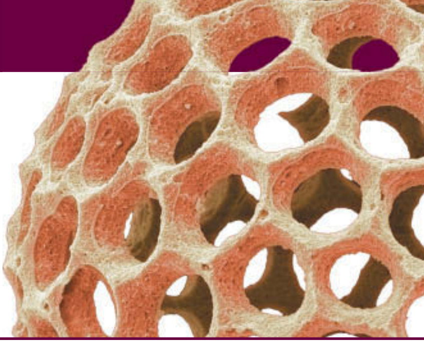
(a) Suppose that the mushroom increases its height by 5%. How much further will its spores blow (i.e., what is the percentage change in d)?

(b) In fact the distance traveled by spores also depends on the size of the spores, a, on the amount of asymmetry of the mushroom, A, and on how much cooler it is compared to its surroundings, ΔT. ***Dressaire et al.*** find:

$$d = \frac{cAh^2 \Delta T}{a^2}$$

where c is another positive constant.

Over the course of evolution the all of the parameters $A, h, \Delta T, a$ may vary. If a 5% change can be made to one of these parameters, which one should be varied (i.e., changing what parameter yields the largest percentage increase in d)?



CHAPTER

5

Applications of Differentiation

Differentiation is an important tool for understanding the behavior of functions, and for building mathematical models to predict real world phenomena. In this chapter, we will learn how to

- deduce the behavior of functions by using differentiation;
- solve optimization problems;
- sketch the graphs of functions on the basis of their behavior;
- investigate the long-term behavior of difference equations;
- find the solutions of equations using a computer and choose the correct parameters to fit a mathematical model to real data;
- derive differential equations to model biological phenomena; and
- solve some types of differential equations.

5.1 Extrema and the Mean-Value Theorem

Calculus can help us to understand the behavior of functions. We will show in Section 5.3 that points where a function is smallest or largest, called extrema, are of particular importance. For example, identifying these points can help to draw the graph of a function, or to calculate an optimal blood vessel radius or the frequency with which genes occur in a population. This section defines extrema, gives conditions that guarantee extrema (via the Extreme-Value Theorem), and provides a characterization of extrema (Fermat's theorem). From Fermat's theorem we will derive the Mean-Value Theorem. We will need the Mean-Value Theorem at many points in this chapter to characterize the behavior of functions.

5.1.1 The Extreme-Value Theorem

Suppose that you measure the depth of a creek along a transect between two points A and B (see Figure 5.1). Looking at the profile of the creek, you see that there is a location of maximum depth and a location of minimum depth. The existence of such locations is guaranteed by the extreme-value theorem. Before we meet this theorem, we must introduce some terminology.

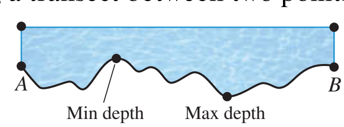


Figure 5.1 A transect of a creek between the points A and B.

Definition Let f be a function defined on the set D that contains the number c. Then

f has a **global** (or **absolute**) **maximum** at $x = c$ if

$$f(c) \geq f(x) \quad \text{for all } x \in D$$

and

f has a **global** (or **absolute**) **minimum** at $x = c$ if

$$f(c) \leq f(x) \quad \text{for all } x \in D$$

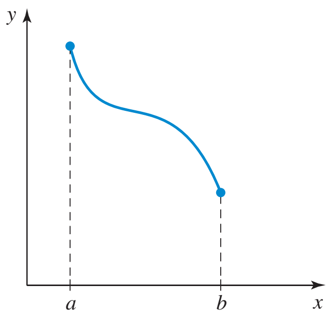


Figure 5.2 Extreme values at endpoints.

The following result gives conditions under which global maxima and global minima, collectively called **global** (or **absolute**) **extrema**, exist:

Theorem The Extreme-Value Theorem If f is continuous on a closed interval $[a, b]$, $-\infty < a < b < \infty$, then f has a global maximum and a global minimum on $[a, b]$.

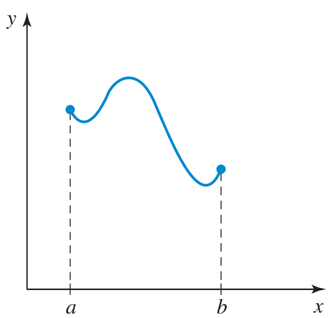


Figure 5.3 Extreme values in the interior.

The proof of the Extreme-Value Theorem is beyond the scope of this text and will be omitted. However, the consequences are quite intuitive, and we illustrate them in Figures 5.2 and 5.3. Figure 5.2 shows that a function may attain its extreme values at the endpoints of the interval $[a, b]$, whereas in Figure 5.3 the extreme values are attained in the interior of the interval $[a, b]$. The function must be continuous and defined on a closed interval in order for it to have global maxima and global minima. Note that the Extreme-Value Theorem tells us only that global extrema exist, not where they are. Furthermore, they need not be unique, meaning that a function can have more than one global maximum or global minimum.

EXAMPLE 1

Bio Info • Optimal Strategy Each year a plant must devote some fraction of its resources to growth and some fraction to reproduction. Evolutionary biologists believe that the plant splits its resources to maximize its fitness, meaning the number of offspring it ultimately produces. Devoting resources in a given year to reproduction may lead to more offspring being produced that year. But devoting resources to growth may mean that the plant is bigger and more competitive, so able to have more offspring in future years.

Suppose that a plant devotes a fraction p of its resources to growth $(0 \leq p \leq 1)$ and the remaining fraction $1 - p$ of its resources to reproduction. Denote by $f(p)$ the plant's fitness as a function of p. Assuming that $f(p)$ is a continuous function, why is there a strategy of resource allocation (called an optimal strategy) that maximizes the plant's fitness?

Solution We don't know what the function $f(p)$ is. Nevertheless, we can show that it has a maximum value for some p in $[0, 1]$. According to the Extreme-Value Theorem, since $f(p)$ is continuous on the closed interval $[0, 1]$, $f(p)$ has a global maximum (and a global minimum) on $[0, 1]$. The global maximum represents the optimal strategy. Note that the theorem guarantees only the existence of an optimal strategy; it does not tell us *which* strategy is optimal. Furthermore, there could be more than one global maximum, meaning that there could be more than one optimal strategy for resource allocation. •

The Extreme-Value Theorem guarantees the existence of global maximum and minimum points only if f is continuous and if the interval is closed. The next two examples show that the theorem cannot be used if either of those assumptions is broken.

EXAMPLE 2 Show that the function $f(x)$ defined by:

$$f(x) = \begin{cases} 2x & \text{if } 0 \le x < 2 \\ 3 - x & \text{if } 2 \le x \le 3 \end{cases}$$

does not have a global maximum point.

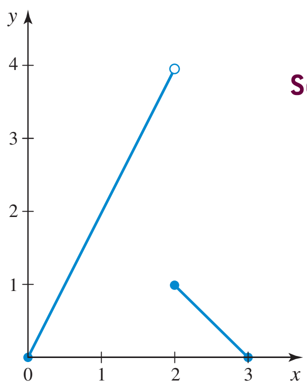


Figure 5.4 The function $f(x)$ defined in Example 2. $f(x)$ is not continuous and has no global maximum.

Solution $f(x)$ is defined on a *closed* interval, namely, $[0, 3]$. However, f is discontinuous at $x = 2$, as can be seen in Figure 5.4. The graph of $f(x)$ shows that there is no value $c \in [0, 3]$ where $f(c)$ attains a global maximum. It looks like there may be a global maximum near $x = 2$: For example if $x = 1.9$, then $f(x) = 2x = 3.8$. However, we can find a point with a larger value of $f(x)$ by choosing x closer to 2. If $x = 1.99$, then $f(x) = 3.98$ which is larger than 3.8. Indeed, for any candidate for a global maximum that you might come up with, you will be able to find a point whose y-coordinate exceeds the y-coordinate of your previous candidate by choosing a point closer to 2. The reason for this is that $\lim_{x\to 2^-} f(x) = 4$ but the function f takes the value 1 at $x = 2$. So the value of $f(x)$ can get arbitrarily close to 4, but it cannot reach 4. We conclude that the function does not have a global maximum. It does however have global minima, at $x = 0$ and $x = 3$, where $f(x) = 0$. (This is a function that has more than one global minimum.) •

EXAMPLE 3 Show that the function defined by $f(x) = x$ for $0 < x < 1$ does not have a global maximum point.

Solution $f(x)$ is continuous on its domain, $(0, 1)$, but is *not* defined on a closed interval (Figure 5.5). The function $f(x)$ attains neither a global maximum nor a global minimum: $\lim_{x\to 0^+} f(x) = 0$ and $\lim_{x\to 1^-} f(x) = 1$ and $0 < f(x) < 1$ for all $x \in (0, 1)$. But the function $f(x)$ has no global minimum because although it can get arbitrarily close to 0, it cannot reach 0. Similarly it has no global maximum because it can get arbitrarily close to 1, but it cannot reach 1. •

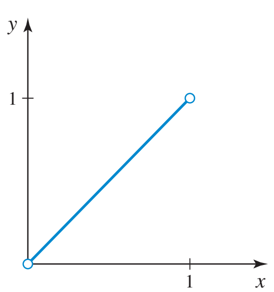


Figure 5.5 The function $f(x)$ defined in Example 3. Although this function is continuous, it is not defined on a closed interval. There is no global minimum nor maximum.

5.1.2 Local Extrema

Local (or relative) extrema, are points where a graph is higher or lower than all *nearby* points. They appear as the peaks and valleys of the graph of a function (see Figure 5.6.) The graph of the function in Figure 5.6 has three peaks—at $x = a$, c, and e—and two valleys—at $x = b$ and d. A peak (or local maximum) has the property that the graph is lower nearby; a valley (or local minimum) has the property that the graph is higher nearby.

The formal definitions follow (see Figures 5.7 and 5.8):

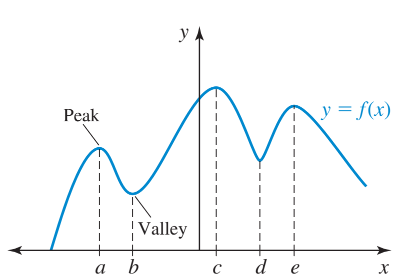


Figure 5.6 The function $y = f(x)$ has valleys (local minima) at $x = b$ and d and peaks (local maxima) at $x = a$, c, and e.

Definition A function f defined on an interval D has a **local** (or **relative**) **maximum** at a point c if there exists a $\delta > 0$ such that

$$f(c) \ge f(x) \qquad \text{for all } x \in (c - \delta, c + \delta) \cap D$$

A function f defined on a set D has a **local** (or **relative**) **minimum** at a point c if there exists a $\delta > 0$ such that

$$f(c) \le f(x) \qquad \text{for all } x \in (c - \delta, c + \delta) \cap D$$

Local maxima and local minima are collectively called **local** (or **relative**) **extrema**. Local extrema need not also be global extrema: for example, $x = c$ is a local maximum

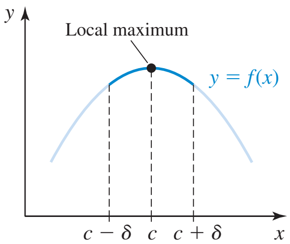


Figure 5.7 The function $y = f(x)$ has a local maximum at $x = c$, since for all x in the interval $c - \delta < x < c + \delta$, $f(x) \le f(c)$.

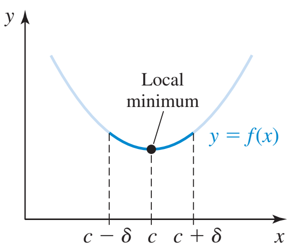


Figure 5.8 The function $y = f(x)$ has a local minimum at $x = c$, since for all x in the interval $c - \delta < x < c + \delta$, $f(x) \ge f(c)$.

if $f(x) \le f(c)$ for all x near to c, but there may be other points in D at which $f(x)$ takes a larger value than $f(c)$. (For example in Figure 5.6, point a is a local maximum but not a global maximum). However, a global maximum point will necessarily also be a local maximum point. For this reason, one way to find global maxima is to first identify all local maxima and then compare them to find the largest value of $f(x)$. A similar method (comparing local minima to find the smallest value of $f(x)$) can be used to find the global minimum of f. We illustrate this method in the next two examples.

How do we show that a point is a local minimum or maximum? If we have a graph of the function, such as Figure 5.6, we can spot local extrema on the graph. (Later we will show in this section how calculus can be used to identify local extrema). To show that a particular point $x = c$ is a local maximum (or minimum) according to the definition we must identify a small interval, containing c, in which $f(x) \le f(c)$ (or $f(x) \ge f(c)$ for a local minimum). Notice that this interval must be chosen to be completely contained in D, the domain of the function. Let's expand a little on what that entails. If the point c lies in the interior of the interval D, then if δ is small enough then the entire of the interval $(c-\delta, c+\delta)$ lies inside of D. So to prove that c is a local maximum, we need to find a value for δ that is small enough so that $c-\delta$ and $c+\delta$ both lie in D, and show also that if $x \in (c-\delta, c+\delta)$ then $f(x) \le f(c)$. If on the other hand c is an end point of the interval D then it is impossible for all points in $(c-\delta, c+\delta)$ to also lie in D. For definiteness, lets assume that c is the left end point of the interval D, meaning that there are no numbers in D smaller than c. Then if δ is small enough, we can ensure that $[c, c+\delta)$ lies in D. So $(c-\delta, c+\delta) \cap D = [c, c+\delta)$. So to prove that c is a local maximum, we need to find a value for δ that is small enough so that $c+\delta$ lies in D, and show also that if $x \in [c, c+\delta)$ then $f(x) \le f(c)$.

We can find local and global extrema by looking at the graphs of functions. For now we will make these graphs using a graphing calculator or a spreadsheet, but in Section 5.6, we will discuss how to make graphs of functions using calculus. In the first example, we consider a function that is defined on a closed interval. In the second example, we consider a function that is defined on a half-open interval; the value of the function can be computed at one endpoint of its domain but not at the other endpoint.

EXAMPLE 4 Let

$$f(x) = (x-1)^2(x+2) \quad \text{for } -2 \le x \le 3$$

(a) Use the graph of $f(x)$ to find all local extrema.

(b) Find the global extrema.

Solution **(a)** The graph of $f(x)$ is illustrated in Figure 5.9. The function f is defined on the closed interval $[-2, 3]$. We begin by identifying all local extrema that occur at interior points of the domain $D = [-2, 3]$; looking at the figure, we see that a local maximum occurs at $x = -1$, as there are no greater values of f nearby. That is, we can find a small interval $(-1-\delta, -1+\delta)$ about $x = -1$ so that $f(-1) \ge f(x)$ for all $x \in (-1-\delta, -1+\delta)$. For instance, we could choose $\delta = 0.1$ (Figure 5.10).

There is a local minimum at $x = 1$, since there are no smaller values of f nearby. To satisfy the definition of local minimum, we need to find a small interval containing $x = 1$. For example we could choose $\delta = 0.1$ or the interval $(0.9, 1.1)$. Then for any $x \in (0.9, 1.1)$, $f(x) \ge f(1)$.

A local minimum also occurs at $x = -2$, one of the endpoints of the domain of f. To show that there is a local minimum at $x = -2$, we must find $\delta > 0$ such that $f(x) \ge f(-2)$ for all $x \in (-2-\delta, -2+\delta) \cap D = [-2, -2+\delta)$. So we need to compare $f(-2)$ with all points that are no more than δ greater than -2, but we do not need to consider any points that are smaller than -2. We can again choose $\delta = 0.1$ and see that $f(x) \ge f(-2)$ for all $x \in [-2, -1.9)$ (Figure 5.11).

Similarly, we see that there is a local maximum at $x = 3$, since $f(x) \le f(3)$ for all $x \in (2.9, 3]$.

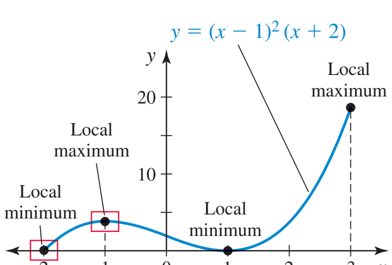


Figure 5.9 The graph of $f(x) = (x-1)^2(x+2)$ used in Example 4. The red-boxed regions are expanded for Figures 5.10 and 5.11.

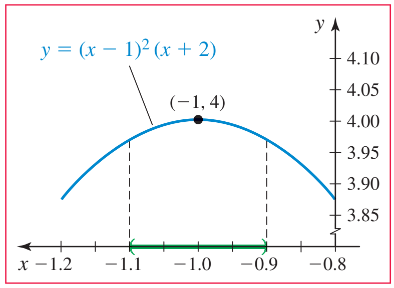


Figure 5.10 Zooming in on the graph of $f(x) = (x-1)^2(x+2)$ near $x = -1$. To apply the definition of local maximum choose $\delta = 0.1$. We see from the graph that for all $x \in (-1.1, -0.9)$, $f(x) \leq f(-1)$.

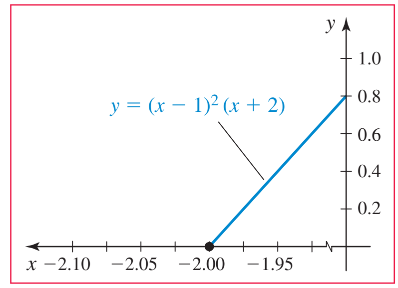


Figure 5.11 Zooming in on the graph of $f(x) = (x-1)^2(x+2)$ near $x = -2$. The point $x = -2$ is a local minimum.

(b) Global maxima and minima are respectively the points at which a function is largest and smallest. Since f is defined on a closed interval, namely $[-2, 3]$, both a global maximum and a global minimum exist because of the Extreme-Value Theorem. These global extrema may occur either in the interior or at the endpoints of the domain $D = [-2, 3]$.

To find the global minimum, we compare the local minima ($x = -2$ and $x = 1$). Since $f(-2) = 0$ and $f(1) = 0$, f takes the same value at both local minima, it follows that the function has two global minimum points at $x = -2$ and $x = 1$ (Figure 5.9). To find the global maximum, we compare the local maxima ($x = -1$ and $x = 3$). Since $f(3) = 20$ and $f(-1) = 4$, f takes a larger value at $x = 3$; therefore, the global maximum occurs at $x = 3$ (Figure 5.9). •

EXAMPLE 5 Let $f(x) = |x^2 - 4|$ for $-2.5 \leq x < 3$. Find all local and global extrema.

Solution The graph of $f(x)$, plotted in Figure 5.12, reveals local minima at $x = -2$ and $x = 2$ and local maxima according to the definition at $x = -2.5$ and $x = 0$. To prove that $x = -2$ is a local minimum, we can compare $f(-2)$ with $f(x)$ for all $x \in (-2.1, -1.9)$. Similar arguments can be made for all of the other local extrema. Since $f(x)$ increases as we approach the right end of the interval, it is tempting to look for a local maximum there; but $x = 3$ does not lie in the interval for which $f(x)$ is defined, so it cannot be a local maximum. Candidates for global extrema are all the local extrema, which must be compared against the value of the function near the boundary $x = 3$. However, since the interval on which the function is defined $[-2.5, 3)$ is not closed, we must also consider the possibility that the global extrema do not exist.

Let's discuss the global maximum first. At the local maxima $f(-2.5) = 2.25$ and $f(0) = 4$. However, $\lim_{x\to 3^-} f(x) = 5$, so we can find larger values of $f(x)$ than both of the above, by choosing x close to 3. For example $f(2.9) = 4.41$ which is greater than $f(-2.5)$ and $f(0)$. But because $f(x)$ is not defined at $x = 3$, the function has no global maximum.

To find the global minimum, we need only compare $f(-2)$ and $f(2)$. We find that $f(-2) = 0$ and $f(2) = 0$, so f takes the same value at both local minima, and from the graph we see that the function never takes a smaller value than 0. Both local minima are therefore global minima. •

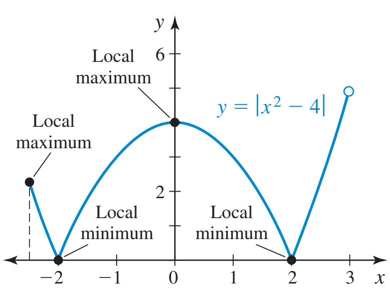


Figure 5.12 The graph of $f(x) = |x^2 - 4|$ for $-2.5 \leq x < 3$ in Example 5. This function has two local minima and two local maxima.

Looking at Figures 5.9 and 5.12, we see that if the function f is differentiable at an interior point where f has a local extremum, then there is a horizontal tangent line at that point. This statement is known as Fermat's theorem (see Figure 5.13), and it provides a method to test whether a point is a local extremum.

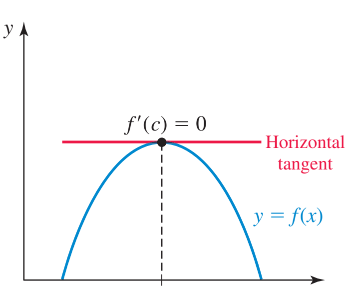


Figure 5.13 Fermat's theorem.

> **Theorem Fermat's Theorem** If f has a local extremum at an interior point c and $f'(c)$ exists, then $f'(c) = 0$.

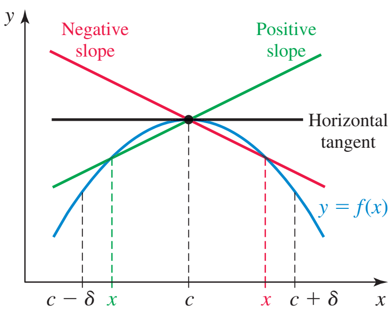


Figure 5.14 To prove Fermat's theorem we calculate the slope of the secant line that joins c to a point x, in the limit as $x \to c$. If $x > c$ the secant slope is negative, while if $x < c$ the secant slope is positive.

Proof We prove Fermat's theorem for a local maximum point (the proof for a local minimum point is similar). We need to show that $f'(c) = 0$. To do so, we use the formal definition of the derivative $f'(c)$ from Chapter 4:

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \tag{5.1}$$

To compute this limit, we will separately compute the left-hand limit ($x \to c^-$) and the right-hand limit ($x \to c^+$). We begin with the following observation (Figure 5.14): Suppose that f has a local maximum at an interior point c. Then there exists $\delta > 0$ such that

$$f(x) \le f(c) \quad \text{for all } x \in (c - \delta, c + \delta)$$

So if $x < c$ then $x - c < 0$ and the local maximum property implies that $f(x) - f(c) \le 0$. The denominator and numerator in Equation (5.1) are therefore both negative, meaning that their ratio (which we can also interpret as the slope of the secant line; see Figure 5.14) is positive or 0. Since this is true for any $x < c$ in the interval it follows that:

$$\lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} \ge 0 \tag{5.2}$$

Conversely, if we consider points x that are right of $x = c$ but still in the interval $(c - \delta, c + \delta)$; $f(x) - f(c) \le 0$, but now $x - c > 0$, so the ratio of these two terms is negative. Since this is true for any $x > c$ that is close to c it follows that:

$$\lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} \le 0 \tag{5.3}$$

Because f is differentiable at c,

$$f'(c) = \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} = \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c}$$

Now by Equation (5.2), $f'(c) \ge 0$ while by Equation (5.3), $f'(c) \le 0$. If you have a number that is simultaneously nonnegative and nonpositive, the number must be 0. Therefore, $f'(c) = 0$.

EXAMPLE 6 Explain, without drawing the graph of the function, why $y = \tan x$ does not have a local extremum at $x = 0$.

Solution $y = \tan x$ is differentiable at $x = 0$, with

$$\frac{d}{dx} \tan x = \sec^2 x = 1 \quad \text{for} \quad x = 0$$

Since the derivative is not equal to 0, Fermat's theorem implies that $x = 0$ is not a local extremum.

Caution!

1. The condition that $f'(c) = 0$ is a necessary, but not sufficient, condition for the existence of local extrema. For instance, $f(x) = x^3$, $x \in \mathbf{R}$, is differentiable at $x = 0$ and $f'(0) = 0$, but there is no local extremum at $x = 0$. The graph of $y = x^3$ is shown in Figure 5.15. Although there is a horizontal tangent at $x = 0$, there is no local extremum at $x = 0$. Fermat's theorem does tell you, however, that if $x = c$ is an interior point and $f'(c) \ne 0$, then $x = c$ cannot be a local extremum (see, e.g., Example 6). Interior points with horizontal tangents are *candidates* for local extrema.
2. The function f may not be differentiable at a local extremum. For instance, in Example 5, the function $f(x)$ is not differentiable at $x = -2$ and $x = 2$, but both points turned out to be local extrema. This means that, in order to identify candidates for local extrema, it will not be enough simply to look at points with horizontal tangents; you also must look at points where the function $f(x)$ is not differentiable.

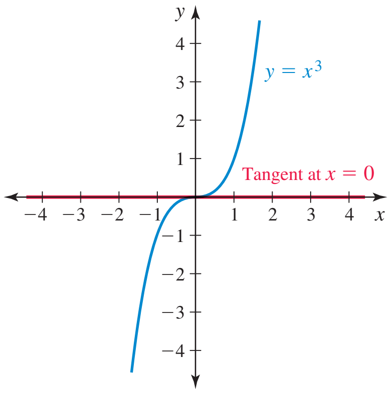


Figure 5.15 The graph of $y = x^3$ has a horizontal tangent at $x = 0$, but this point is not an extremum.

3. Local extrema may occur at endpoints of the domain. Since Fermat's theorem says nothing about what happens at endpoints, you will have to look at endpoints separately.

To summarize our discussion, here are guidelines for finding the candidates for local extrema of a function f:

1. Find all points c where $f'(c) = 0$.
2. Find all points c where $f'(c)$ does not exist.
3. Find the endpoints of the domain of f.

We will return to local extrema in Section 5.3, where we will learn methods for deciding whether candidates for local extrema are indeed local extrema.

5.1.3 The Mean-Value Theorem

Knowing the global extrema of a function $f(x)$ defined on an interval D gives us the total amount that $f(x)$ varies over the interval. The derivative $f'(x)$ tells us the rate at which $f(x)$ varies near the point x. These two pieces of information are partly connected through the **Mean-Value Theorem** (abbreviated as MVT).

Here is an example that explains the MVT: Consider the function

$$f(x) = x^2 \quad \text{for } 0 \leq x \leq 1$$

The secant line connecting the two endpoints $(x, y) = (0, 0)$ and $(1, 1)$ of the graph of $f(x)$ has slope

$$m = \frac{f(1) - f(0)}{1 - 0} = \frac{1 - 0}{1 - 0} = 1$$

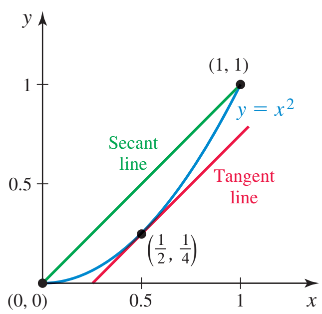


Figure 5.16 For the function $f(x) = x^2$ the secant line through $(x, y) = (0, 0)$ and $(1, 1)$ has slope 1. The MVT guarantees that there is a point, $0 < c < 1$ for which $f'(c) = 1$.

The graph of $f(x)$ and the secant line are shown in Figure 5.16. Note that $f(x)$ is differentiable in $(0, 1)$; that is, you can draw a tangent line at every point of the graph. If you look at the graph of $f(x)$, you see that there exists a number $c \in (0, 1)$ such that the slope of the tangent line at $(x = c)$ is the same as the slope of the secant line through the points $(0, 0)$ and $(1, 1)$. That is, there exists a point $c \in (0, 1)$ for which:

$$f'(c) = \frac{f(1) - f(0)}{1 - 0}$$

We can compute the value of c in this example. Since $f'(x) = 2x$ and the slope of the secant line is $m = 1$, we must solve $2c = 1$, which implies $c = \frac{1}{2}$. (This tangent line is shown in Figure 5.16.)

The Mean-Value Theorem guarantees that an interior point can be found at which the gradient is equal to 1.

Theorem The Mean-Value Theorem (MVT) If f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), then there exists at least one number $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

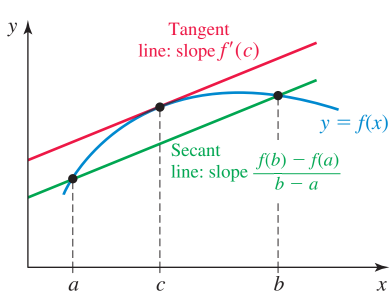


Figure 5.17 The mean-value theorem guarantees the existence of a number $c \in (a, b)$ such that the tangent line at $(c, f(c))$ has the same slope as the secant line through $(a, f(a))$ and $(b, f(b))$.

The fraction on the right-hand side of the equation in the theorem is the slope of the secant line connecting the points $(x, y) = (a, f(a))$ and $(b, f(b))$, which we can identify as the average slope of f over the entire interval $[a, b]$. The quantity on the left-hand side is the slope of the tangent line at $x = c$ (see Figure 5.17).

The MVT can be interpreted geometrically: It states that there exists a point on the graph between $(a, f(a))$ and $(b, f(b))$, (i.e., the point c in the statement of the MVT) where the tangent line is parallel to the secant line through $(a, f(a))$ and $(b, f(b))$. The MVT is an "existence" result: It tells us neither how many such points there are nor where they are.

Going back to the example $f(x) = x^2$, $0 \le x \le 1$, we see that $f(x)$ satisfies the assumptions of the MVT, namely $f(x)$ is continuous on the closed interval $[0, 1]$ and differentiable on the open interval $(0, 1)$. The MVT then guarantees the existence of at least one point $c \in (0, 1)$ such that the slope of the secant line through the points $(0, 0)$ and $(1, 1)$ is equal to the slope of the tangent line at $x = c$.

The next example provides a physical illustration of the theorem.

EXAMPLE 7 **Velocity** A car moves in a straight line. At time t (measured in seconds), its position (measured in meters) is

$$s(t) = \frac{1}{25}t^3, \quad 0 \le t \le 10$$

Show that there is a time $t \in (0, 10)$ when the velocity is equal to the average velocity between $t = 0$ and $t = 10$.

Solution The average velocity between $t = 0$ and $t = 10$ is

$$\frac{s(10) - s(0)}{10 - 0} = \frac{\frac{1}{25} \cdot 1000\,\text{m}}{10\,\text{s}} = 4\,\text{m/s}$$

This is the slope of the secant line connecting the points $(0, 0)$ and $(10, 40)$. Since $s(t)$ is continuous on $[0, 10]$ and differentiable on $(0, 10)$, the MVT tells us that there must exist a number $c \in (0, 10)$ such that $s'(c) = 4\,\text{m/s}$. Now, $s'(t)$ is the (instantaneous) velocity. So, at some point during this short trip, the speedometer must have read $4\,\text{m/s}$.

In the rest of the section we will prove the Mean-Value Theorem and show some of its applications. More applications will be given in Section 5.2.

To prove the MVT we start with a special case called Rolle's theorem:

> **Theorem** **Rolle's Theorem** If f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), and if $f(a) = f(b)$, then there exists a number $c \in (a, b)$ such that $f'(c) = 0$.

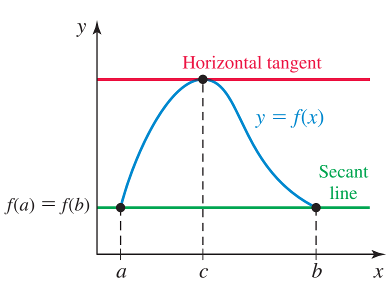


Figure 5.18 An illustration of Rolle's theorem.

Figure 5.18 illustrates Rolle's theorem. The function in the graph is defined on the closed interval $[a, b]$ and takes the same value at the two endpoints of $[a, b]$ (that is, $f(a) = f(b)$). Thus, the secant line connecting the two endpoints has slope 0. We see, then, that there is a point in (a, b) at which the tangent has slope 0.

Before we prove Rolle's theorem, we check why it is a special case of the MVT. If we compare the assumptions in the two theorems, we find that Rolle's theorem has an additional requirement, $f(a) = f(b)$; that is, the function values must agree at the endpoints of the interval on which f is defined. If we apply the MVT to such a function, it says that there exists a number $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a} = 0 \qquad f(a) = f(b)$$

which is the conclusion of Rolle's theorem.

Proof of Rolle's Theorem If f is the constant function, then $f'(x) = 0$ for all $x \in (a, b)$ and the theorem is true in this particular case. For the more general case, we assume that f is not constant. Since $f(x)$ is continuous on the closed interval $[a, b]$, the Extreme-Value Theorem tells us that the function has a global maximum and a global minimum in that interval. At least one of these extrema must lie in the open interval (a, b), that is, will not coincide with $x = a$ or $x = b$. Observe that if f is not constant, then there exists an $x_0 \in (a, b)$ such that either $f(x_0) > f(a) = f(b)$ or $f(x_0) < f(a) = f(b)$. In the first case $x = a$ and $x = b$ cannot be the global maximum. In the second case $x = a$ and $x = b$ cannot be the global minimum. So in both cases there must be a global extremum inside the interval. This global extremum is also a local extremum. Suppose that the local extremum is at $c \in (a, b)$; then it follows from Fermat's theorem that $f'(c) = 0$. •

The MVT follows from Rolle's theorem and can be thought of as a "tilted" version of that theorem. Although the secant and tangent lines are no longer horizontal, they are still parallel.

Proof of the MVT We define the following function:

$$F(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

Since f is continuous on $[a, b]$ and differentiable on (a, b) so is the function F. Furthermore:

$$F(a) = f(a) - \frac{f(b) - f(a)}{b - a}(a - a) = f(a)$$

$$F(b) = f(b) - \frac{f(b) - f(a)}{b - a}(b - a) = f(a).$$

Therefore, $F(a) = F(b)$. We can apply Rolle's theorem to the function $F(x)$: Rolle's theorem implies there exists a point $c \in (a, b)$ with $F'(c) = 0$. Since

$$F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}$$

it follows that, for this value of c,

$$0 = F'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}$$

and hence

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$ •

We next discuss two consequences of the MVT.

Corollary 1 If f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b) such that

$$m \le f'(x) \le M \qquad \text{for all } x \in (a, b)$$

then

$$m(b - a) \le f(b) - f(a) \le M(b - a)$$

This corollary is useful in obtaining information about a function on the basis of its derivative.

EXAMPLE 8 **Population Growth** Denote the population size at time t by $N(t)$, and assume that $N(t)$ is continuous on the interval $[0, 10]$ and differentiable on the interval $(0, 10)$ with $N(0) = 100$ and $|dN/dt| \le 3$ for all $t \in (0, 10)$. What are the maximum and minimum possible values for $N(10)$?

Solution Since $|dN/dt| \le 3$ implies that $-3 \le dN/dt \le 3$, we can set $m = -3$ and $M = 3$ in Corollary 1. So, Corollary 1 yields the following estimate:

$$(-3)(10 - 0) \le N(10) - N(0) \le (3)(10 - 0) \qquad a = 0, b = 10$$

Simplifying and solving for $N(10)$ gives

$$-30 + N(0) \le N(10) \le 30 + N(0)$$

Since $N(0) = 100$, we have
$$70 \le N(10) \le 130.$$
That is, the population size at time $t = 10$ is bounded between 70 and 130.

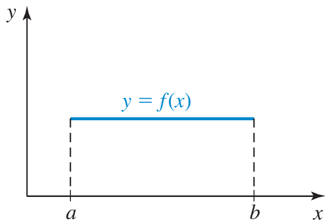


Figure 5.19 An illustration of Corollary 2: if the derivative of a function is 0 everywhere in a closed interval, then the function must be constant.

We will make use of the next corollary in Section 5.10.

> **Corollary 2** If f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), with $f'(x) = 0$ for all $x \in (a, b)$, then f is constant on $[a, b]$.

Figure 5.19 illustrates Corollary 2: Each point on the graph has a horizontal tangent so the function must be constant.

EXAMPLE 9 Assume that f is continuous on $[-1, 1]$ and differentiable on $(-1, 1)$, with $f(0) = 2$ and $f'(x) = 0$ for all $x \in (-1, 1)$. Find $f(x)$.

Solution Corollary 2 tells us that $f(x)$ is a constant. Since we know that $f(0) = 2$, we have $f(x) = 2$ for all $x \in [-1, 1]$.

Proof of Corollary 2 Define any two points $x_1, x_2 \in (a, b)$, $x_1 < x_2$. Then f satisfies the assumptions of the MVT on the closed interval $[x_1, x_2]$. Therefore, there exists a number $c \in (x_1, x_2)$ such that
$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c)$$
Since $f'(c) = 0$, it follows that $f(x_2) = f(x_1)$. Finally, because x_1, x_2 are arbitrary numbers from the interval (a, b), we conclude that f is constant.

EXAMPLE 10 Show that
$$\sin^2 x + \cos^2 x = 1 \quad \text{for all } x \in [0, 2\pi]$$

Solution This identity can be shown using pre-calculus methods, but let's see what we get if we use Corollary 2. We define $f(x) = \sin^2 x + \cos^2 x$, $0 \le x \le 2\pi$. Then $f(x)$ is continuous on $[0, 2\pi]$ and differentiable on $(0, 2\pi)$, with
$$f'(x) = 2 \sin x \cos x - 2 \cos x \sin x = 0 \qquad \text{Use chain rule}$$
Using Corollary 2 now, we conclude that $f(x)$ is equal to a constant on $[0, 2\pi]$. To find the constant, we need only evaluate $f(x)$ at one point in the interval, say, $x = 0$. We find that
$$f(0) = \sin^2 0 + \cos^2 0 = 1 \qquad \sin 0 = 0, \cos 0 = 1$$
This proves the identity.

Section 5.1 Problems

5.1.1

T ***In Problems 1–8, each function is continuous and defined on a closed interval. It therefore satisfies the assumptions of the extreme-value theorem. With the help of a graphing calculator or spreadsheet, graph each function and locate its global extrema. (Note that a function may have more than one global minimum or maximum point.)***

1. $f(x) = 2x$, $0 \le x \le 1$

2. $f(x) = -x^2 + 1$, $-1 \le x \le 1$

3. $f(x) = \sin(x - 2)$, $0 \le x \le \pi$

4. $f(x) = \sin \frac{x}{2}$, $0 \le x \le 2\pi$

5. $f(x) = |x|$, $-1 \le x \le 1$

6. $f(x) = (x - 1)^2(x + 1)$, $-2 \le x \le 2$

7. $f(x) = e^{-|x|}$, $-1 \le x \le 1$

8. $f(x) = x \ln x$, $1 \le x \le 2$

9. Sketch the graph of a function that is continuous on the closed interval $[0, 3]$ and has global maxima at both the right and left end points of the interval.

10. Sketch the graph of a function that is continuous on the closed interval $[-2, 1]$ and has a global maximum and a global minimum in the interior of the interval.

11. Sketch the graph of a function that is continuous on the open interval $(0, 1)$ and has a global maximum but does not have a global minimum.

12. Sketch the graph of a function that is continuous on the closed interval $[0, 4]$, except at $x = 2$, and has neither a global maximum nor a global minimum in its domain.

5.1.2

T ***In Problems 13–18, use a graphing calculator or spreadsheet to plot the function and determine all local and global extrema.***

13. $f(x) = 4 - x, x \in [-1, 4)$

14. $f(x) = 3x - 5, x \in (-2, 1)$

15. $f(x) = x^2 - 2, x \in [-1, 1]$

16. $f(x) = (x - 2)^2, x \in [0, 3]$

17. $f(x) = x \ln x, x \in [1, 5]$

18. $f(x) = x^2 - x, x \in [0, 1]$

In Problems 19–26, find c such that $f'(c) = 0$ and determine whether $f(x)$ has a local extremum at $x = c$.

19. $f(x) = x^2$

20. $f(x) = (x - 4)^2$

21. $f(x) = -x^2$

22. $f(x) = e^{-x^2}$

23. $f(x) = x^3$

24. $f(x) = e^{x^3}$

25. $f(x) = (x + 1)^2$

26. $f(x) = (x + 1)^3$

27. Show that $f(x) = |x|$ has a local minimum at $x = 0$ but $f(x)$ is not differentiable at $x = 0$.

28. Show that $f(x) = |x - 1|$ has a local minimum at $x = 1$ but $f(x)$ is not differentiable at $x = 1$.

29. Show that $f(x) = -|x^2 - 4|$ has local maxima at $x = 2$ and $x = -2$ but $f(x)$ is not differentiable at $x = 2$ or $x = -2$.

30. Show that $f(x) = |x^2 - 1|$ has local minima at $x = 1$ and $x = -1$ but $f(x)$ is not differentiable at $x = 1$ or $x = -1$.

31. Graph

$$f(x) = (1 - |x|)^2, \quad -1 \leq x \leq 2$$

and determine all local and global extrema on $[-1, 2]$.

32. Graph

$$f(x) = (|x| - 2)^3, \quad -3 \leq x \leq 3$$

and determine all local and global extrema on $[-3, 3]$.

33. **Population Growth** Suppose the size of a population at time t is $N(t)$ and its growth rate is given by the logistic growth model

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right), \quad t \geq 0$$

where r and K are positive constants.

(a) Graph the growth rate of the population $\frac{dN}{dt}$ as a function of population size, N, assuming that $r = 2$ and $K = 100$, and find the population size for which the growth rate is maximal.

(b) Show that whatever the value of the parameters N and K, $f(N) = rN(1 - N/K)$, $N \geq 0$, is differentiable for $N > 0$, and compute $f'(N)$.

(c) If $r = 2$ and $K = 100$, show that $f'(N) = 0$ if N is equal to the fastest growing population size, (you calculated the size in part (a)).

34. **Population Growth** Suppose that the size of a population at time t is $N(t)$ and its growth rate is given by the logistic growth model

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right), \quad t \geq 0$$

where r and K are positive constants. For a population whose size changes with time, we can define a per capita reproductive rate $R(N)$ by: $R(N) = \frac{1}{N}\frac{dN}{dt}$

(a) Show that

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

(b) Graph $R(N)$ as a function of N for $N \geq 0$ when $r = 2$ and $K = 100$, and find the population size for which the reproductive rate is maximal.

5.1.3

35. Suppose $f(x) = x^3, x \in [0, 1]$.

(a) Find the slope of the secant line connecting the points $(x, y) = (0, 0)$ and $(1, 1)$.

(b) Find a number $c \in (0, 1)$ such that $f'(c)$ is equal to the slope of the secant line you computed in (a), and explain why such a number must exist in $(0, 1)$.

36. Suppose $f(x) = e^x, x \in [0, 1]$.

(a) Find the slope of the secant line connecting the points $(x, y) = (0, 1)$ and $(1, e)$.

(b) Find a number $c \in (0, 1)$ such that $f'(c)$ is equal to the slope of the secant line you computed in (a), and explain why such a number must exist in $(0, 1)$.

37. Suppose that $f(x) = x^2, x \in [-1, 1]$.

(a) Find the slope of the secant line connecting the points $(x, y) = (-1, 1)$ and $(1, 1)$.

(b) Find a number $c \in (-1, 1)$ such that $f'(c)$ is equal to the slope of the secant line you computed in (a), and explain why such a number must exist in $(-1, 1)$.

38. Suppose that $f(x) = \ln x, x \in [1, e]$.

(a) Find the slope of the secant line connecting the points $(x, y) = (1, 0)$ and $(e, 1)$.

(b) Find a number $c \in (1, e)$ such that $f'(c)$ is equal to the slope of the secant line you computed in (a), and explain why such a number must exist in $(1, e)$.

39. Suppose that $f(x) = -x^2 + 2$. Explain why there exists a point c in the interval $(-1, 2)$ such that $f'(c) = -1$.

40. Suppose that $f(x) = x^3$. Explain why there exists a point c in the interval $(-1, 1)$ such that $f'(c) = 1$.

41. Suppose that $f(x) = x(2 - x)$. Explain why there exists a point c in the interval $(0, 2)$ such that $f'(c) = 0$.

42. Suppose that $f(x) = x^4(5 - x)$. Explain why there exists a point c in the interval $(0, 5)$ such that $f'(c) = 0$.

43. Suppose that $f(x) = x^2, x \in [a, b]$.

(a) Compute the slope of the secant line through the points $(a, f(a))$ and $(b, f(b))$.

(b) Find the point $c \in (a, b)$ such that the slope of the tangent line to the graph of f at $(c, f(c))$ is equal to the slope of the secant line determined in (a). How do you know that such a point exists? Show that c is the midpoint of the interval (a, b); that is, show that $c = (a + b)/2$.

44. Assume that f is continuous on $[a, b]$ and differentiable on (a, b). Show that if $f(a) < f(b)$, then f' is positive at some point between a and b.

45. Assume that f is continuous on $[a, b]$ and differentiable on (a, b). Show that if $f'(x) > 0$ for all $x \in (a, b)$, then $f(a) < f(b)$.

46. Assume that f is continuous on $[0, 1]$ and differentiable on $(0, 1)$. Assume that $f'(1/2) = 0$, show by sketching the graph of a function $f(x)$ that satisfies all of these conditions (you do not need to write down the equation of the function) that it is not necessary that $f(0) = f(1)$.

47. A car moves in a straight line. At time t (measured in seconds), its position (measured in meters) is

$$s(t) = \frac{1}{10}t^2, \; 0 \le t \le 10$$

(a) Find its average velocity between $t = 0$ and $t = 10$.

(b) Find its instantaneous velocity for $t \in (0, 10)$.

(c) At what time is the instantaneous velocity of the car equal to its average velocity?

48. A car moves in a straight line. At time t (measured in seconds), its position (measured in meters) is

$$s(t) = \frac{1}{100}t^3, \; 0 \le t \le 10$$

(a) Find its average velocity between $t = 0$ and $t = 10$.

(b) Find its instantaneous velocity for $t \in (0, 10)$.

(c) At what time is the instantaneous velocity of the car equal to its average velocity?

49. Prof. Roper drives to work in stop-and-go traffic. His speed measured in miles per hour (mph) is given by the following function of time, t, measured in minutes

$$v(t) = 30 + 20\sin(t/5).$$

The total journey time is 1 hour. Explain using the MVT why the total distance that he travels in this hour is somewhere between 10 and 50 miles.

50. Prof. Roper runs with his dog Molly. Their speed varies a lot over the course of the run, following the graph shown in Figure 5.20. Their total run time is 90 minutes. Explain using the MVT why the total distance that Prof. Roper and Molly travel is somewhere between 1.5 miles and 12 miles.

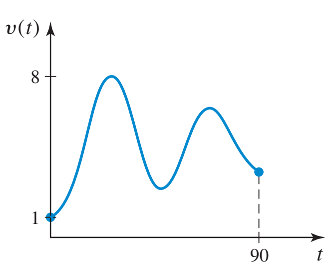


Figure 5.20 Speed of Prof. Roper and Molly on their run (Problem 50). $v(t)$ is measured in mph and t is measured in minutes.

51. Denote the size of a population at time t by $N(t)$, and assume that $N(0) = 50$ and $|dN/dt| \le 20$ for all $t \in [0, 5]$. What can you say about $N(5)$? [*Hint*: Remember also that it is impossible for the number of organisms to become negative].

52. Denote the total biomass in a given area of soil at time t by $B(t)$, and assume that $B(0) = 3$ and $|dB/dt| \le 1$ for all $t \in [0, 3]$. What can you say about $B(3)$?

53. Suppose that f is differentiable for all $x \in \mathbf{R}$ and, furthermore, that f satisfies $f(0) = 0$ and $1 \le f'(x) \le 2$ for all $x > 0$.

(a) Use Corollary 1 of the MVT to show that

$$x \le f(x) \le 2x$$

for all $x \ge 0$.

(b) Use your result in (a) to explain why $f(1)$ cannot be equal to 3.

(c) Find an upper and a lower bound for the value of $f(1)$.

54. Suppose that f is differentiable for all $x \in \mathbf{R}$ with $f(2) = 3$ and $f'(x) = 0$ for all $x \in \mathbf{R}$. Find $f(x)$.

55. Suppose that $f(x) = e^{-|x|}$, $x \in [-2, 2]$.

(a) Show that $f(-2) = f(2)$.

(b) Compute $f'(x)$, where defined.

(c) Show that there is no number $c \in (-2, 2)$ such that $f'(c) = 0$.

(d) Explain why your results in (a) and (c) do not contradict Rolle's theorem.

(e) Use a graphing calculator to sketch the graph of $f(x)$.

56. **Population Growth** In Chapter 4 we learned that

$$f(x) = e^{rx}$$

satisfies the differential equation

$$\frac{df}{dx} = rf(x)$$

with $f(0) = 1$. This exercise will show that $f(x)$ is in fact the only solution. Suppose that r is a constant and f is a differentiable function satisfying

$$\frac{df}{dx} = rf(x) \tag{5.4}$$

for all $x \in \mathbf{R}$, and $f(0) = 1$. The following steps will show that $f(x) = e^{rx}$, $x \in \mathbf{R}$, is the only solution of (5.4).

(a) Define the function

$$F(x) = f(x)e^{-rx}, \quad x \in \mathbf{R}$$

Use the product rule to show that

$$F'(x) = e^{-rx}[f'(x) - rf(x)]$$

(b) Use (a) and (5.4) to show that $F'(x) = 0$ for all $x \in \mathbf{R}$.

(c) Use Corollary 2 to show that $F(x)$ is a constant and, hence, $F(x) = F(0) = 1$.

(d) Show that (c) implies that

$$1 = f(x)e^{-rx},$$

and therefore,

$$f(x) = e^{rx}$$

57. Medication in the Human Body. We will solve a differential equation model for the concentration of drug in a patient's blood. Under a particular dosing regime, the amount $M(t)$ in the patient's blood changes in time according to the following differential equation: $\frac{dM}{dt} = a - k_1 M$, where a and k_1 are positive constants, representing the rate of absorption and the rate of elimination respectively, and the initial concentration is $M(0) = 0$. In Section 5.9 we will demonstrate that a solution to this equation is $M(t) = \frac{a}{k_1}\left(1 - e^{-k_1 t}\right)$. This question will show that this function is, in fact, the only solution of the equation. First let's assume that $a = k_1 = 1$ (this assumption simplifies the algebra, but the case of general parameters follows exactly the same lines).

Assume that $M(t)$ satisfies the differential equation:

$$\frac{dM}{dt} = 1 - M, \quad M(0) = 0 \tag{5.5}$$

(a) Define the function

$$C(t) = M(t)e^t - e^t, \quad t \in \mathbf{R}$$

Show that

$$C'(t) = M'(t)e^t + M(t)e^t - e^t$$

(b) Use (a) and (5.5) to show that $C'(t) = 0$ for all $t \in \mathbf{R}$.

(c) Use Corollary 2 to show that $C(t)$ is a constant and, hence, $M(t) = C(0) = -1$.

(d) Show that (c) implies that

$$M(t)e^t - e^t = -1,$$

and therefore:

$$M(t) = 1 - e^{-t}.$$

5.2 Monotonicity and Concavity

Fish are indeterminate growers; they increase in body size throughout their life. However, as they become older, their growth slows. Fish growth is often described mathematically by the von Bertalanffy equation, which fits a large number of both freshwater and marine fishes. This equation is given by

$$L(t) = L_\infty - (L_\infty - L_0)e^{-Kt}$$

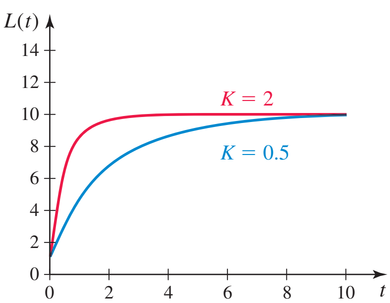


Figure 5.21 Length of a fish, $L(t)$, given by Equation (5.6), plotted against time t for $K = 0.5$ and $K = 2$.

where $L(t)$ denotes the length of the fish at time t. This equation features three different constants: L_0, L_∞, and K. For different fish species and for fish growing in different environments, different values for the constants must be used. In Section 5.6 we will learn how to describe the behavior of functions like $L(t)$ without needing to know the values of all of these constants. For now let's assume that we know that $L_0 = 1$ and $L_\infty = 10$; that is, we know all of the constants but K:

$$L(t) = 10 - 9e^{-Kt} \tag{5.6}$$

The constant K might depend on how much food is available, or how much space the fish have to live in. We want to determine how changing K affects fish growth; namely, what features of the plot of $L(t)$ as a function of t change, and which features stay the same. We show two such plots in Figure 5.21. We see from the graph that in both cases $L(0) = 1$ and as $t \to \infty$, $L(t) \to 10$. Both of these features can be seen from Equation (5.6), because if $t = 0$, $L(t) = L(0) = 10 - 9e^{-0K} = 10 - 9 = 1$, while as $t \to \infty$, $e^{-Kt} \to 0$ for any positive constant K, so $L(t) \to 10 - 9 \times 0 = 10$. So, independently of the value of K, the fish length starts the same, and also has the same long time limit. However, changing the value of K seems to change the shape of the graph. Larger values of K seem to make the fish grow faster.

Looking at Figure 5.21 we see that for either value of K, $L(t)$ is an increasing function of t. So the rate of growth of the fish, which is given by the slope of the tangent line at each point of the graph, is always positive, meaning that fish increase their body size throughout their life. We can reach this same conclusion without having to plot the graph by using calculus methods. Specifically we recall that the rate of growth is also given by the first derivative of $L(t)$:

$$L'(t) = 9Ke^{-Kt}$$

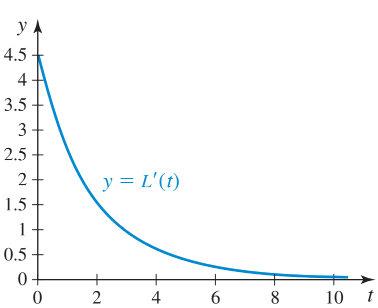


Figure 5.22 Rate of growth for a fish, $L'(t)$, plotted against time, t, for $K = 0.5$.

Since $e^{-Kt} > 0$ (this holds for all t, regardless of K), we see that indeed $L'(t) > 0$. This method has the advantage that we can see that the fish grows with positive growth rate throughout its life for *any* value of K. Figure 5.21 only shows that fact for two specific values of K.

The graph of $L'(t)$ is shown in Figure 5.22. This plot shows that $L'(t)$ is a decreasing function of t: Although fish continue to grow throughout their life, their rate of growth, $L'(t)$, decreases as they get older. This observation can also be turned into a calculus statement: at any time the tangent line to the curve in Figure 5.22 has negative slope.

The slope of the tangent line is given by the derivative of $L'(t)$ with respect to t, which is $L''(t)$ (the second derivative of the function $L(t)$). So, the slope of $L'(t)$ is negative if and only if $L''(t) < 0$. We can calculate $L''(t)$ without having to make any plots:

$$L''(t) = -9K^2 e^{-Kt}$$

Since $K^2 > 0$ and $e^{-Kt} > 0$ for any values of K and t, we see that $L''(t) < 0$. Again arguing from the derivative, rather than from the plot, shows that $L'(t)$ is decreasing (the rate of growth is slowing) for *all* values of K and t. Figure 5.22 only shows that to be true for a specific value of K.

The fact that the rate of growth decreases with age can also be seen directly from the graph of $L(t)$ in Figure 5.21. For both values of K shown in Figure 5.21 the curve bends downward. As we will see in this section, the second derivative gives us information about which way the graph of $L(t)$ bends.

This section discusses the concepts of **monotonicity**—whether a function is decreasing or increasing—and **concavity**—whether a function bends upward or downward. These concepts give us a way to talk about the behavior of functions, like the von Bertalanffy model for fish growth.

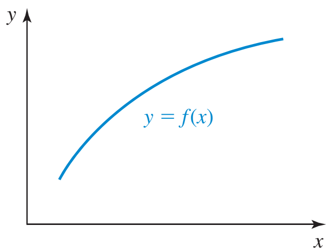


Figure 5.23 An example of an increasing function.

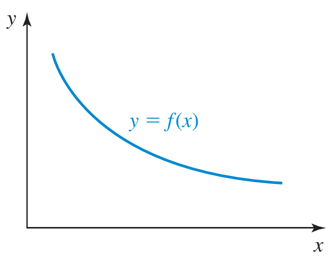


Figure 5.24 An example of a decreasing function.

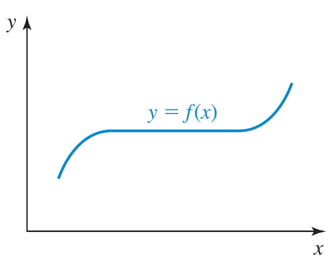


Figure 5.25 A nondecreasing function may have regions where the function is constant.

5.2.1 Monotonicity

Definition A function f defined on an interval I is called **increasing** on I if

$$f(x_1) < f(x_2) \qquad \text{whenever } x_1 < x_2 \text{ in } I$$

and is called **decreasing** on I if

$$f(x_1) > f(x_2) \qquad \text{whenever } x_1 < x_2 \text{ in } I$$

Examples of increasing and decreasing function are shown in Figures 5.23 and 5.24. We call a function **monotonic** if it is either an increasing function or a decreasing function. If, instead of the strict inequality $f(x_1) < f(x_2)$, we have the inequality $f(x_1) \leq f(x_2)$, whenever $x_1 < x_2$ in I, we call f *nondecreasing*. If $f(x_1) \geq f(x_2)$ whenever $x_1 < x_2$ in I, then f is called *nonincreasing*. (See Figures 5.25 and 5.26 for examples.) The inequality $f(x_1) \leq f(x_2)$ whenever $x_1 < x_2$ implies that either $f(x_1) < f(x_2)$ (the function increases in going from x_1 to x_2) or $f(x_1) = f(x_2)$, meaning that the function is constant between the two values. A nondecreasing function, unlike an increasing function, may therefore be constant over some intervals.

Almost all functions that show up in biological models are differentiable. If a function f is differentiable then there is a useful test to determine whether it is increasing or decreasing. We will first state the test, and then prove it using the Mean-Value Theorem.

First-Derivative Test for Monotonicity Suppose f is continuous on $[a, b]$ and differentiable on (a, b).

(a) If $f'(x) > 0$ for all $x \in (a, b)$, then f is increasing on $[a, b]$.

(b) If $f'(x) < 0$ for all $x \in (a, b)$, then f is decreasing on $[a, b]$.

Proof (See Figure 5.27.) We choose two numbers x_1 and x_2 in $[a, b]$, $x_1 < x_2$. Then f is continuous on $[x_1, x_2]$ and differentiable on (x_1, x_2). We can therefore apply the MVT to f defined on $[x_1, x_2]$: There exists a number $c \in (x_1, x_2)$ such that

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c)$$

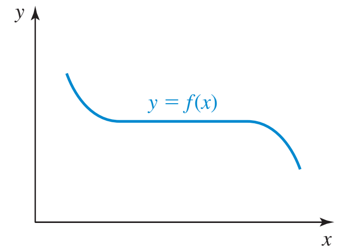


Figure 5.26 A nonincreasing function may have regions where the function is constant.

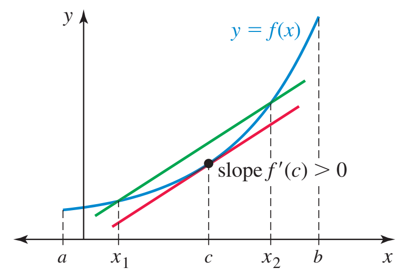


Figure 5.27 Given any pair of points $x_1 < x_2$, the MVT theorem allows us to relate the difference between $f(x_2)$ and $f(x_1)$ to some $f'(c)$ for some point $c \in (x_1, x_2)$. If $f'(c) > 0$ then $f(x_1) < f(x_2)$.

In part (a) of the theorem, we assume that $f'(x) > 0$ for all $x \in (a, b)$. The point $c \in (x_1, x_2) \subset (a, b)$, so $f'(c) > 0$. Since $x_2 > x_1$ it follows that

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) > 0$$

which implies that $f(x_2) > f(x_1)$. Because x_1 and x_2 are arbitrary numbers in $[a, b]$ satisfying $x_1 < x_2$, it follows that f is increasing. The proof of part (b) is similar, and you can work through it in Problem 27.

EXAMPLE 1 Determine where the function

$$f(x) = x^3 - \frac{3}{2}x^2 - 6x + 3, \quad x \in \mathbf{R}$$

is increasing and where it is decreasing.

Solution Since $f(x)$ is continuous and differentiable for all $x \in \mathbf{R}$, we can use the first-derivative test to find the intervals on which it is increasing or decreasing. The derivative of $f(x)$ is

$$f'(x) = 3x^2 - 3x - 6 = 3(x - 2)(x + 1), \quad x \in \mathbf{R}$$ Factorize $f'(x)$

$f'(x)$ is a quadratic polynomial, so the graph of $f'(x)$ is a parabola. By factorizing the polynomial we find that this parabola intersects the x-axis at $x = -1$ and $x = 2$, meaning that the function $f'(x)$ changes sign at $x = -1$ and $x = 2$. $f'(x)$ must take the same sign over any interval that does not cross these points. To find the sign in the interval $x < -1$, evaluate $f'(x)$ for a single point in this interval, say for $f'(-2) = 3(-4)(-1) = 12 > 0$. So $f'(x) > 0$ for $x < -1$. $f'(x)$ changes sign at $x = -1$ and again at $x = 2$, so:

$$f'(x) \begin{cases} > 0 & \text{if } x < -1 \text{ or } x > 2 \\ < 0 & \text{if } -1 < x < 2 \end{cases}$$

You can also evaluate $f'(x)$ at points in each interval, e.g., since $f'(0) = -6 < 0$, $f'(x) < 0$ for $-1 < x < 2$. Thus, $f(x)$ is increasing for $x < -1$ or $x > 2$ and decreasing for $-1 < x < 2$. A look at the graph of $f(x)$ in Figure 5.28 confirms this conclusion.

EXAMPLE 2 **Bio Info • Hemoglobin Saturation Curve** Hemoglobin is an oxygen binding chemical that allows red blood cells to carry oxygen molecules from the lungs to the organs and cells where they are needed. Each hemoglobin molecule can bind up to four oxygen molecules. Each oxygen molecule that is added to the hemoglobin molecule changes the shape of the hemoglobin molecule and makes it bind more readily to further oxygen molecules. When each hemoglobin molecule has bound to four oxygen molecules it becomes saturated, and can accept no more oxygen molecules. The total oxygen content of blood varies depending on how much oxygen there is in the cells and tissues surrounding the blood. Let Y be the oxygen

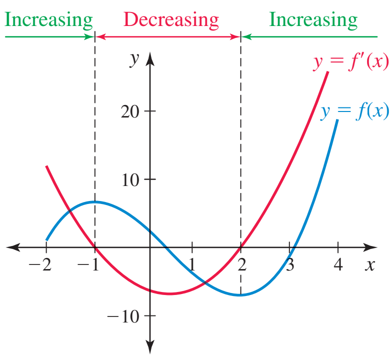


Figure 5.28 From the graph of $f'(x) = 3x^2 - 3x - 6$, we can find the regions on which the function $f(x) = x^3 - \frac{3}{2}x^2 - 6x + 3$ is increasing or decreasing.

content of the blood. Y varies from 0, meaning no hemoglobin molecules have bound oxygen, to 1, meaning all hemoglobin molecules have bound their full complement of four oxygen molecules. The oxygen concentration of the surrounding tissues is measured in terms of the partial oxygen pressure P, which is measured in units of mm Hg. You may assume that $P > 0$. The dependence of blood oxygen content upon oxygen levels in the surrounding tissues is often modeled using Hill's equation:

$$Y(P) = \frac{P^n}{P^n + 30^n}$$

The function $Y(P)$ contains a single unknown positive constant, n, usually called Hill's coefficient. Hemoglobin molecules from different animals have different Hill's coefficients. According to ***Milo et al. (2007)***, n ranges from 2.3 for camels to 3.6 for antelope (for humans $n \approx 2.95$); Figure 5.29 shows $Y(P)$ for different values of n.

Show that the oxygen saturation level in the blood (Y) increases with the oxygen partial pressure (P).

Solution Unlike Example 1, there is an unknown constant in this equation. So we cannot plot the function without knowing the value of n. Figure 5.29 shows the saturation curve for several different values of n, but we cannot be sure that the plots in this figure capture all possible behaviors of the function. We can use the first derivative test to explore the behavior of $Y(P)$ without plotting it. To calculate the derivative it is helpful to use long-division to simplify the way the equation is written so that P only appears in one place:

$$Y(P) = 1 - \frac{30^n}{P^n + 30^n}$$

$$\Rightarrow \qquad Y'(P) = \frac{n \cdot 30^n P^{n-1}}{(P^n + 30^n)^2} \qquad \text{Use chain rule, letting } u = P^n.$$

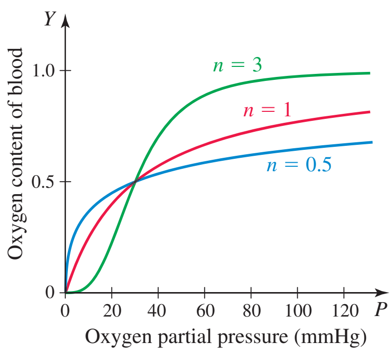


Figure 5.29 The amount of oxygen bound to hemoglobin varies according to Hill's equation. Different values of the Hill coefficient n produce curves with different shapes.

Since we don't know what n is we can't say much about $Y'(P)$, but we do know that the denominator is positive. Similarly since $P > 0$ and $n > 0$, the numerator $nP^{n-1} > 0$ also. So $Y'(P) > 0$ so, by the first derivative test, $Y(P)$ increases as P increases.

5.2.2 Concavity

We saw in the fish growth example that the second derivative tells us something about whether a function bends upward or downward.

A function is called **concave up** if it bends upward, and **concave down** if it bends downward. In other words:

> **Definition** A differentiable function $f(x)$ is **concave up** on an interval I if the first derivative $f'(x)$ is an increasing function on I. $f(x)$ is **concave down** on an interval I if the first derivative $f'(x)$ is a decreasing function on I.

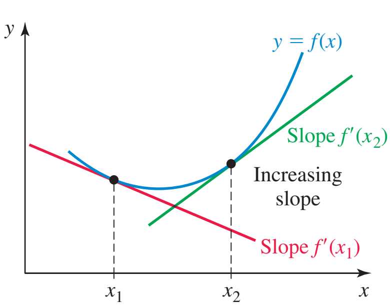


Figure 5.30 A function is concave up if its derivative is increasing. Here $f'(x_1) < f'(x_2)$ if $x_1 < x_2$. The function bends upward.

We can see how this definition works for a concave up function in Figure 5.30 and for a concave down function in Figure 5.31. Let's check whether this test correctly diagnoses whether the functions $f(x) = x^2$ and $g(x) = -x^2 + 4$, shown in Figure 5.32, are concave up or concave down. First look at $f(x) = x^2$, which bends up. $f'(x) = 2x$, which increases with x, so the function is concave up. On the other hand $g(x) = -x^2 + 4$ bends down. For this function $g'(x) = -2x$ decreases with x, so the function is concave down.

Note that the definition assumes that $f(x)$ is differentiable. There is a more general definition that does not require differentiability. (After all, not all functions are differentiable.) The more general definition is more difficult to use, however. The definition given here suffices for our purposes. We can combine it with the results of Section 5.2.1

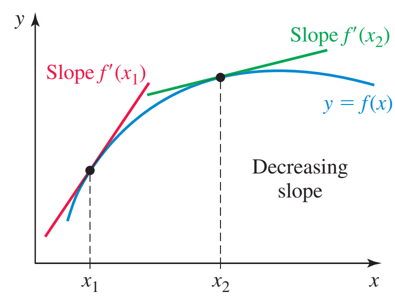


Figure 5.31 A function is concave down if its derivative is decreasing. Here $f'(x_1) > f'(x_2)$ if $x_1 < x_2$. The function bends downward.

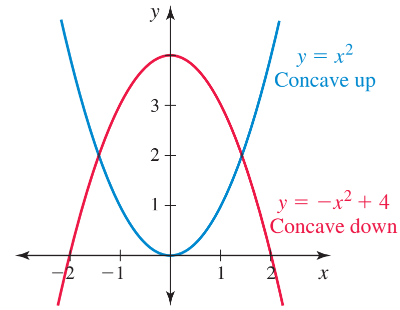


Figure 5.32 The function $f(x) = x^2$ is concave up, and the function $g(x) = -x^2 + 4$ is concave down.

to obtain a test that can be used to determine whether a twice-differentiable function is concave up or concave down:

Second-Derivative Test for Concavity Suppose that f is twice differentiable on an open interval I.

(a) If $f''(x) > 0$ for all $x \in I$, then f is concave up on I.

(b) If $f''(x) < 0$ for all $x \in I$, then f is concave down on I.

Proof Since f is twice differentiable, we can apply the first-derivative criterion to the function $f'(x)$. To prove part (a) argue as follows: If $f''(x) > 0$ on I, then $f'(x)$ is an increasing function on I. From the definition of *concave up*, it follows that f is concave up on I. The proof of part (b) is similar; you will prove part (b) in Problem 28. •

You can use the function $y = x^2$ to remember which functions are concave up: The "u" in "concave up" should remind you of the U-shaped form of the graph of $y = x^2$. You can also use the function $y = x^2$ to help you to memorize the second-derivative criterion. The graph of $y = x^2$ is concave up, and you can easily compute the second derivative of $y = x^2$, namely, $y'' = 2 > 0$. So a concave up function has positive second derivative.

EXAMPLE 3 Determine where the function

$$f(x) = x^3 - \frac{3}{2}x^2 - 6x + 3, \qquad x \in \mathbf{R}$$

is concave up and where it is concave down.

Solution This is the same function as in Example 1 (redrawn in Figure 5.33). Since $f(x)$ is a polynomial, it is twice differentiable. In Example 1, we found that $f'(x) = 3x^2 - 3x - 6$; differentiating $f'(x)$, we get the second derivative of f:

$$f''(x) = 6x - 3$$

We find that

$$f''(x) \begin{cases} > 0 & \text{if } x > \frac{1}{2} \qquad 6x - 3 > 0 \Rightarrow 6x > 3 \\ < 0 & \text{if } x < \frac{1}{2} \qquad 6x - 3 < 0 \Rightarrow 6x < 3 \end{cases}$$

So $f(x)$ is concave up for $x > 1/2$ and concave down for $x < 1/2$. A look at Figure 5.33 confirms this result. •

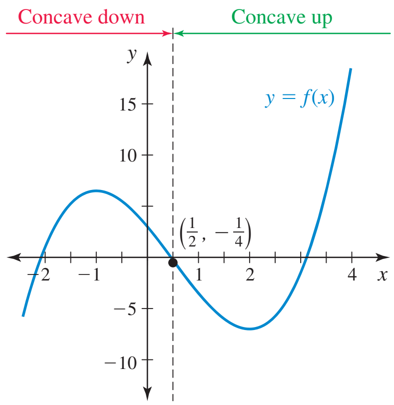


Figure 5.33 (Example 3) The graph of $f(x) = x^3 - \frac{3}{2}x^2 - 6x + 3$, showing concave down and concave up intervals.

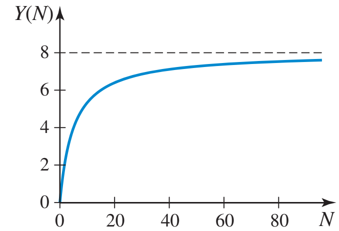


Figure 5.34 (Example 4) Crop yield, $Y(N)$, as a function of amount of nitrogen added to soil, N.

A very common mistake is to mistake monotonicity and concavity. One has nothing to do with the other. For instance, an increasing function can bend downward or upward (see Problem 21.)

There are many biological examples of increasing functions that have a decreasing derivative and are therefore concave down.

EXAMPLE 4 **Crop Yield** Adding nitrogen fertilizer to soil typically increases crop yields. The dependence of crop yield Y upon the soil nitrogen level N is sometimes described by a function of the form:

$$Y(N) = Y_{\max}\frac{N}{K+N}, \qquad N \geq 0$$

This function has two unknown coefficients: $Y_{\max}$ and K. You may assume that $Y_{\max} > 0$ and $K > 0$. Show that whatever the values of $Y_{\max}$ and K are, $Y(N)$ is an increasing function of N, and that it is concave down. How would you interpret these facts for a farmer trying to maximize her crop yield?

Solution A graph of $Y(N)$ against N is shown for $K = 5$ and $Y_{\max} = 8$ in Figure 5.34. We see from the graph that for this particular pair of coefficients $Y(N)$ is an increasing function of N, and that the graph bends downward and hence is concave down. Are these observations true for all choices of coefficients? We use the first derivative test to determine whether the function is increasing. Before calculating the derivative it is sensible to simplify $Y(N)$ using long-division:

$$Y(N) = Y_{\max}\left(1 - \frac{K}{K+N}\right)$$

so:

$$Y'(N) = Y_{\max}\frac{K}{(K+N)^2}$$

Alternatively use the quotient rule; $Y(N) = \frac{p(N)}{q(N)}$, $p(N) = Y_{\max}N$, $q(N) = K + N$

and:

$$Y''(N) = \frac{-2Y_{\max}K}{(K+N)^3}$$

Chain rule; $u = K + N$

Since $Y_{\max}$ and K are positive constants and $N \geq 0$ we see that $Y'(N) > 0$, meaning that the yield function is increasing with N, and $Y''(N) < 0$ meaning that the yield function is concave down.

Taken together, these facts mean that $Y(N)$ is an increasing function, but the rate of increase is decreasing. We say that Y is *increasing at a decelerating rate*. What does this mean? It means that as we increase fertilizer levels, the yield will increase, but

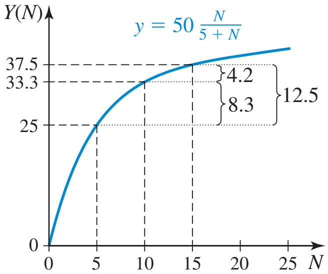


Figure 5.35 The graph of $Y(N)$ in Example 4 for $Y_{max} = 50$ and $K = 5$.

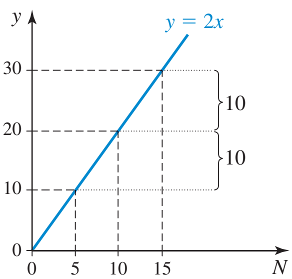


Figure 5.36 The graph of a linear function: Increases are proportional.

slower and slower. This behavior is called **diminishing returns**. To understand what diminishing returns means it is helpful to consider the behavior of the function for specific values of Y_{max} and K. Let's set $Y_{max} = 50$ and $K = 5$, so that:

$$Y(N) = 50 \cdot \frac{N}{5+N}, \quad N \geq 0$$

The graph of this function is shown in Figure 5.35.

Suppose that initially $N = 5$. If we increase N by 5 (i.e., from 5 to 10), then $Y(N)$ changes from $Y(5) = 25$ to $Y(10) = 33.3$, an increase of 8.3. If we double the increase in N (i.e., increase by 10, from 5 to 15), then $Y(15) = 37.5$ so the increase in yield is only 12.5, less than twice 8.3. That is we get less than twice as much **return** from adding twice as much fertilizer. An alternate way of thinking about diminishing returns is that if we start with $N = 5$ and increase by 5 to $N = 10$, then the change in Y (the Y-increment) is 8.3. Increasing N by the same amount, but starting at 10, we see that the Y-increment is $Y(15) - Y(10) = 4.2$. In general, increasing N has less of an effect on Y for larger values of N; thus, we say that the return is diminishing.

This is important information for a farmer trying to maximize her yield. Adding more nitrogen will always increase the yield from the field. But adding more and more nitrogen has progressively smaller and smaller effect upon yields.

You should compare a function representing a diminishing return with a linear function, say, $f(x) = 2x$, which is neither concave up nor concave down. (See Figure 5.36.) With a linear function, if we increase x from 5 to 10, $f(x)$ changes from 10 to 20. That is, $f(x)$ increases by 10. Then, if we increase x from 10 to 15, $f(x)$ changes from 20 to 30, again an increase of 10. That is, for linear functions, the increase in y is always proportional to the increase in x. •

Section 5.2 Problems

5.2.1 and 5.2.2

In Problems 1–20, use the first derivative test and the second derivative test to determine where each function is increasing, decreasing, concave up, and concave down. You do not need to use a graphing calculator for these exercises.

1. $y = 2x - x^2, x \in \mathbf{R}$

2. $y = x^2 + 5x, x \in \mathbf{R}$

3. $y = x^2 - x - 4, x \in \mathbf{R}$

4. $y = 2x^2 - x + 3, x \in \mathbf{R}$

5. $y = \frac{1}{3}x^3 - 2x^2 + 3x + 4, x \in \mathbf{R}$

6. $y = x^3 - 5x^2 + 8x + 2, x \in \mathbf{R}$

7. $y = \sqrt{2x+1}, x \geq -1/2$

8. $y = (3x-1)^{1/3}, x \in \mathbf{R}$

9. $y = \dfrac{1}{x}, x \neq 0$

10. $y = \dfrac{x+1}{x}, x \neq 0$

11. $(x^2+1)^{1/3}, x \in \mathbf{R}$

12. $y = \dfrac{5}{x-2}, x \neq 2$

13. $y = \dfrac{1}{(1+x)^2}, x \neq -1$

14. $y = \dfrac{x^2}{x^2+1}, x \geq 0$

15. $y = \sin x, 0 \leq x \leq 2\pi$

16. $y = \sin(\pi x^2), 0 \leq x \leq 1$

17. $y = e^{-x}, x \in \mathbf{R}$

18. $y = xe^{-x}, x > 0$

19. $y = e^{-x^2}, x \in \mathbf{R}$

20. $y = \dfrac{1}{1+e^{-x}}, x \in \mathbf{R}$

21. Sketch the graph of

(a) a function that is increasing at an accelerating rate; and

(b) a function that is increasing at a decelerating rate.

(c) Assume that your functions in (a) and (b) are twice differentiable. Explain in each case how you could check the respective properties by using the first and the second derivatives. Which of the functions is concave up, and which is concave down?

22. Show that if $f(x)$ is the linear function $y = mx + b$ where m and b are constants, then increases in $f(x)$ are proportional to increases in x. That is, suppose initially that $x = x_0$, and $y = y_0 = mx_0 + b$. Then we increase x by Δx to $x = x_0 + \Delta x$. Calculate the increase in y. Show that the increase in y depends on Δx but does not depend on x_0. This means that the same increment in x always produces the same increment in y, independently of the starting value of x. Contrast this behavior with a concave down function.

You must solve Problem 23 before attempting any of Problems 24–26.

23. We frequently must solve equations of the form $f(x) = 0$. When f is a continuous function on $[a, b]$ and $f(a)$ and $f(b)$ have opposite signs, the intermediate-value theorem guarantees that there exists at least one solution of the equation $f(x) = 0$ in $[a, b]$. Explain in words why there exists exactly one solution in (a, b) if, in addition, f is differentiable in (a, b) and $f'(x)$ is either strictly positive or strictly negative throughout (a, b).

24. Use the result from Problem 23 to show that

$$x^3 - 4x + 1 = 0$$

has exactly one solution in $[-1, 1]$.

25. Use the result from Problem 23 to show that

$$x^3 - 2x^2 + x + 1 = 0$$

has exactly one solution in $[-1, 0]$.

26. Use the result from Problem 23 to show that

$$x \ln x - 2 = 0$$

has exactly one solution in $[1, 5]$.

27. First-Derivative Test for Monotonicity Suppose that f is continuous on $[a, b]$ and differentiable on (a, b). Show that if $f'(x) < 0$ for all $x \in (a, b)$, then f is decreasing on $[a, b]$.

28. Second-Derivative Test for Concavity Suppose that f is twice differentiable on an open interval I. Show that if $f''(x) < 0$, then f is concave down.

One model for the growth of a population is the logistic growth model. This model says that the rate of growth of the population is given as function of the population size, N, by the equation:

$$f(N) = rN\left(1 - \frac{N}{K}\right), \quad N \geq 0$$

where r and K are positive coefficients that are different for different populations. In Problems 29 and 30 you will study the reproductive rate and per capita reproductive rates predicted by the logistic model.

29. (a) Plot the function $f(N)$ for $r = 3$ and $K = 10$.

(b) For the parameters $r = 3$ and $K = 10$, use calculus to find $f'(N)$, and determine where the function $f(N)$ is increasing and where it is decreasing.

(c) Now suppose that $r = 3$, but the value of K is not given to you (You may assume $k > 0$.) Show that the growth rate $f(N)$ is an increasing function of N for $N < K/2$.

30. For a population growing according to the logistic model we can calculate a per capita reproductive rate, which is defined to be equal to:

$$g(N) = \frac{f(N)}{N} = r\left(1 - \frac{N}{K}\right), \quad N \geq 0$$

(a) Plot the function $g(N)$ for $r = 3$ and $K = 10$.

(b) For the parameters $r = 3$ and $K = 10$, use calculus to find $g'(N)$, and determine where the function $g(N)$ is increasing and where it is decreasing.

(c) Now suppose that $K = 10$, but the value of r is not given to you (You may assume $r > 0$.) Show that the reproduction rate $g(N)$ is a decreasing function of N for all $N > 0$.

31. Resource-Dependent Growth Growth rates for many microbes and plants depend on the amount of nutrients that are available to them. ***Monod (1949)*** introduced a model, now widely adopted, for how the rate of growth of *E. coli* bacteria depends on the level of glucose in the medium in which the bacteria are grown. Specifically, Monod observed that the reproduction rate of the bacteria (number of cell divisions in one hour) is given as a function of the glucose concentration (C, measured in units of mM) by an equation:

$$r(C) = 1.35\frac{C}{C + 0.022}, \qquad C > 0$$

Is $r(C)$ an increasing function of C?

32. Monod Growth Monod's equation describes the rate of growth of microorganisms or plants as a function of the amount, C, of nutrients available to the organisms. The most general form of the equation includes two coefficients, a and K:

$$r(C) = \frac{aC}{C + K}, \qquad C > 0$$

You may assume that $a > 0$ and $K > 0$. Using different values of these coefficients the equation can be used to model different species and different types of nutrients. Show that for any value of a and any value of K the reproductive rate $r(C)$ is an increasing function of C.

33. Oxygen Binding by Hemoglobin In Example 2 we met Hill's equation as a model for the oxygen saturation of blood. In the form given in Example 2, Hill's equation had a single unknown coefficient, n, and modeled the oxygen saturation of blood as a function of the oxygen concentration, P, by a function:

$$f(P) = \frac{P^n}{P^n + 30^n}, \qquad P > 0$$

We showed that $f(P)$ is an increasing function of P, no matter what the value of n.

(a) Suppose that $n = 1$. Show that $f(P)$ is concave down. (So $f(P)$ increases with P but at a decreasing rate).

(b) In most mammals, n is close to 3. Assuming that $n = 3$, show that:

$$f''(P) = -6 \cdot \frac{30^3 P(2P^3 - 30^3)}{(P^3 + 30^3)^3}, \qquad P > 0$$

Explain why for $2P^3 < 30^3$, (or equivalently $P < 30/2^{1/3} = 23.8$), $f(P)$ increases with P at an increasing rate.

34. Host–Parasitoid Interactions Parasitoids are insects that lay their eggs in, on, or close to other (host) insects. Parasitoid larvae then devour the host insect. The likelihood of the host insect escaping from being eaten depends on the number of parasitoids in her vicinity. One model for this dependence is that the probability of escaping parasitism is equal to

$$f(P) = e^{-aP}$$

where P is the number of parasitoids in the host insect's vicinity and a is a positive constant. Determine whether the probability of the host insect escaping being eaten increases or decreases with the number of parasitoids nearby.

35. Drug Elimination Model Suppose that a patient is receiving a particular drug at a constant rate by intravenous line (a needle that delivers the blood directly into one of the patient's veins). In Section 5.9 we will show that one model for how the amount of drug in the patient's blood varies with time, t, is:

$$M(t) = a - ae^{-kt}, \qquad t > 0$$

This model contains two coefficients; $a > 0$ depends on rate at which the drug is introduced through the intravenous line, and $k > 0$ represents the rate at which it is broken down within the body. Assume that for one particular drug $a = 2$, but the value of k is not known.

(a) Show that whatever the value of k is, the amount of drug $M(t)$ is an increasing function of time.

(b) Show that whatever the value of k is, the amount of drug $M(t)$ increases at a decreasing rate with time, meaning that $M(t)$ is concave down.

36. Drug Elimination Model If instead of receiving a drug by intravenous line, a patient takes the drug in pill form then the model from Problem 35 must be modified. The amount of drug in a patient's blood is often modeled by the following equation:

$$M(t) = ae^{-k_1 t} - ae^{-k_2 t}, \qquad t > 0.$$

This model contains three coefficients: $a > 0$ is a measure of the total amount of drug taken, $k_1 > 0$ is the rate at which the drug is absorbed into the blood from the patient's gut, and $k_2 > 0$ is the rate at which the drug is broken down by the body.

(a) Assuming initially that you know that $k_1 = 1$ and $k_2 = 2$, show that there is an interval containing $t = 0$ over which the amount of drug increases with time, whatever the value of a is.

(b) Suppose instead that you know that $a = 1$, and that $k_1 = 1$ but you do not know k_2. However, you do know that $k_2 > k_1$, meaning that the drug is broken down by the body more rapidly than it is absorbed from the gut. Show that provided $t < \frac{\ln k_2}{k_2 - 1}$, $M(t)$ is an increasing function of t. In other words, there is an initial phase after taking the pill where the amount of drug in the patient's blood increases with time.

(c) Under the assumptions of part (b) what happens to the amount of drug in the patient's blood if $t > \frac{\ln k_2}{k_2 - 1}$?

37. Plant Reproduction Strategies Plants employ two basic reproductive strategies: *polycarpy*, in which reproduction occurs repeatedly during the lifetime of the organism, and *monocarpy*, in which the plant flowers and produces seeds only once before dying. (Bamboo, for instance, is a monocarpic plant.) ***Iwasa et al. (1995)*** argued that the best strategy for a plant depends on how reproductive success (that is, number of progeny that the plant produces) varies with the investment (that is amount of resource that the plant uses up to reproduce)

> The optimal strategy is polycarpy if reproductive success increases with the investment at a decreasing rate, [or] monocarpy if the reproductive success increases at an increasing rate.

(a) Sketch the graph of reproductive success as a function of reproductive investment for the cases of **(i)** polycarpy and **(ii)** monocarpy.

(b) Given that the second derivative describes whether a curve bends upward or downward, explain the preceding quote in terms of the second derivative of the reproductive success function.

38. Pollinator Visits ***Iwasa et al. (1995)*** argued that the number of times that a plant can expect to be visited by pollinating insects will depend on the number, F, of flowers that the plant makes. They assumed a power law dependence; namely that the number of pollinator visits is given by:

$$X(F) = cF^{\gamma}$$

where c and γ are positive constants.

(a) Show that if $\gamma = 1/2$ then, for all values of c, the average number of pollinator visits to a plant increases with the number of flowers, F, but the rate of increase decreases with F.

(b) Show that if $\gamma = 3/2$ then, for all values of c, the average number of pollinator visits to a plant increases with the number of flowers, F, and the rate of increase increases with F.

39. Clutch Size ***Lloyd (1987)*** studied how the likelihood of a chick surviving to adulthood depends on the amount of resource that the chick's parents invested in it. He proposed the following model for the relationship between the amount of resource R invested in the chick and the likelihood $p(R)$ that it survives to adulthood:

$$p(R) = \frac{R^2}{k^2 + R^2}, \qquad R > 0$$

In the model, $k > 0$ is a coefficient that varies between different species and different environments.

(a) Show that for all values of k, $p(R)$ is an increasing function of R.

(b) Assume now that $k = 1$. Show that if $R > 1/\sqrt{3}$, then the $p(R)$ is concave down. Explain why this means that for R in this interval, there are diminishing returns from increasing the investment in the chick.

(c) Assuming that $k = 1$, what can you say about $p(R)$ if $R < 1/\sqrt{3}$?

40. Pulse-Chase Experiments A pulse-chase experiment can be used to see how a particular chemical is processed by the cell; for example, how pancreatic cells convert amino acids into insulin. The experiment starts with a pulse phase, in which the cells are fed a radioactively labelled form of the amino acid. Following this there is a chase phase, in which they are fed the same amino acid, but without the radioactive label. Any insulin that the cells produce is removed and tested for the presence of the radioactive label.

You perform this experiment for pancreatic cells that have been treated with different drugs before the pulse-chase experiment. You measure the amount of radioactive labelled insulin $c(t)$ produced as a function of time, t:

$$c(t) = 12.7te^{-kt}, \qquad t > 0$$

where $k > 0$ is a coefficient that depends on which drugs the cells have been treated with before the pulse chase experiment. You expect the level of radioactive label in the insulin to start increasing (during the pulse phase), and then decrease (during the chase phase), as the radioactive amino acid works its way through the cell and is replaced by the non-radioactive amino acid. Show that the function $c(t)$ has this behavior for all values of the coefficient k.

We are not always given the function of interest in explicit form. In each of Problems 41–44 y is related to x by an implicit equation. Determine using implicit differentiation and the first derivative test whether y is an increasing or a decreasing function of x.

41. $x^2 + y^2 = 1, \qquad 0 < x < 1,\ y > 0$

42. $x^2 - y^2 = 1, \qquad x > 1,\ y > 0$

43. $\ln y = 1 - \frac{y}{x}, \qquad x > 0, y > 0$

44. $xy = e^{-y}, \qquad x > 0$

45. Fish Schooling Many fish join with others of their species to form schools, large groups that swim together in a coordinated way. It is thought that schooling helps the fish evade predators (predators are unable to pick out a single individual in the school to prey upon), and may also allow the fish to swim more efficiently by slip-streaming off each other. There is a lot of interest in how individual fish within the school interact to produce the complex swimming patterns seen in real schools.

(a) Fish tend to prefer not to be too close or too far from their neighbors. ***D'Orsogna et al. (2006)*** propose that interactions between neighbors can be modeled by incorporating an energy of interaction $U(r)$, that depends on the distance r between the fish. The force between the two fish can be derived from this energy from the formula: $F(r) = -\frac{dU}{dr}$. A positive force means that the fish repel each other, and a negative force means that they attract each other. D'Orsogna et al. assume the following form for the energy of interaction:

$$U(r) = a_1 e^{-k_1 r} - a_2 e^{-k_2 r}$$

Where a_1, a_2, k_1, and k_2 are all positive constants. Let's assume that for one particular species of fish $a_1 = 3$, $a_2 = 2$, $k_1 = 2$, $k_2 = 1$. Show that the fish repel each other when $r < \ln 3$ and that they attract each other with $r > \ln 3$.

(b) Based on your answer to (a) why do you think that $r = \ln 3$ is referred to as the equilibrium spacing of the fish?

(c) In addition to maintaining their distance from each other, fish must also follow the movements of their neighbors.

Strandburg-Peshkin et al. (2013) proposed a model in which a fish watches its neighbors, and turns only if enough of its neighbors turn. If a fraction f of a particular fish's neighbors turn, then Strandburg-Peshkin et al. propose that the fish will also turn with probability:

$$P(f) = \frac{1}{1+e^{-f}}$$

Show that $P(f)$ is an increasing function of f, and that it is concave down.

(d) Based on your answer to (c) explain why although having more neighbors turn increases the likelihood that a fish will turn, the likelihood of the fish turning increases less and less with each of its neighboring fish that turns.

46. Distance Between Two Randomly Chosen Points When trying to understand the processes by which proteins are organized through a cell, it is helpful to compare where the proteins are located in the cell to what would be expected if they were just placed at random (see for example ***Cameron, Roper and Zambryski, 2012***). One way to make this comparison is to measure the real distances between each protein and its nearest neighbor. For randomly placed proteins, the likelihood that two proteins are within distance d of each other is given approximately by a function:

$$P(d) = 1 - e^{-d/\mu}, \qquad d \geq 0$$

where $\mu > 0$ is a coefficient that depends on the size and geometry of the cell, and on how many proteins it contains.

(a) Show that no matter what the value of μ is, $P(d)$ is an increasing function of d.

(b) Is $P(d)$ concave up or concave down? Explain in words what $P(d)$ being concave up or down means for the distance between proteins.

47. *(Adapted from **Reiss, 1989**)* Suppose that the rate at which body weight W changes with age t is

$$\frac{dW}{dt} \propto W^a \tag{5.7}$$

where $a > 0$ is a coefficient that takes different values for different species of animal.

(a) The relative growth rate (percentage weight gained per unit of time) is defined as

$$G(W) = \frac{1}{W}\frac{dW}{dt}$$

Write down a formula for $G(W)$. For which values of a is the relative growth rate increasing, and for which values is it decreasing?

(b) As fish grow larger, their weight increases each day but the relative growth rate decreases. If the rate of growth is described by (5.7) explain what constraints must be imposed on a.

48. pH The pH value of a solution measures the concentration of hydrogen ions, denoted by $[H^+]$, and is defined as

$$\mathrm{pH} = -\log[H^+]$$

Use calculus to decide whether the pH value of a solution increases or decreases as the concentration of H^+ increases.

49. Allometric Growth Allometric equations describe the scaling relationship between two measurements, such as tree height versus tree diameter or skull length versus backbone length. These equations are often of the form

$$Y = bX^a \tag{5.8}$$

where b is some positive constant and a is a constant that can be positive, negative, or zero.

(a) Assume that X and Y are body measurements (and therefore positive) and that their relationship is described by an allometric equation of the form (5.8). For what values of a is Y an increasing function of X?

(b) For what values of a is Y an increasing function of X but Y/X is a decreasing function of X? Is Y concave up or concave down in this case?

(c) In vertebrates, we typically find

$$[\text{skull length}] \propto [\text{body length}]^a \tag{5.9}$$

for some $a \in (0, 1)$. One measure of the animals' proportions is to calculate the ratio of skull length to total body length. Use your answer in (b) to explain what (5.9) means for the ratio of skull length to body length in juveniles versus adults. It may help to draw a picture!

5.3 Extrema and Inflection Points

5.3.1 Extrema

In Section 5.1.2 we showed that it is possible to find the local extrema of a differentiable function $f(x)$ that is defined on a closed interval $[a, b]$ by following a 3-step procedure:

Finding candidate local extrema

1. Find all points c where $f'(c) = 0$.
2. Find all points c where $f'(c)$ does not exist.
3. Find the endpoints of the domain of f.

Although this method gives us a way to identify local extrema, it does not tell us whether an extremum point is a local maximum point or a local minimum point. We will show how knowledge about the second derivative $f''(x)$ can be used to identify

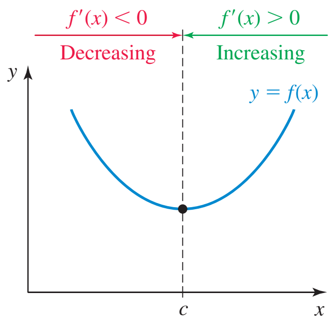


Figure 5.37 The function $y = f(x)$ has a local minimum at $x = c$.

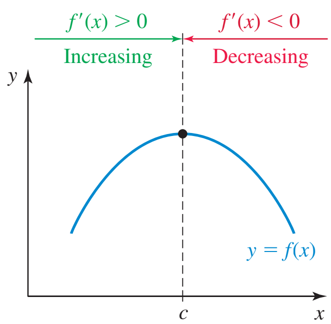


Figure 5.38 The function $y = f(x)$ has a local maximum at $x = c$.

what type of local extremum a point is. First note that near a local extremum we expect a continuous function to have one of the following behaviors (see Figures 5.37 and 5.38):

> **Local behavior of a function near a local extremum** A continuous function has a local minimum at c if the function is decreasing to the left of c and increasing to the right of c. A continuous function has a local maximum at c if the function is increasing to the left of c and decreasing to the right of c.

If the function is differentiable, then we can use the derivative to identify regions where the function is increasing and regions where it is decreasing. Suppose that $a < c < b$, meaning that c is not one of the end points of the interval. If $x = c$ is a local minimum then $f(x)$ is decreasing to the left of c means that $f'(x) < 0$ while $f(x)$ is increasing to the right of $x = c$ means that $f'(x) > 0$ (Figure 5.37). Whereas if $x = c$ is a local maximum then $f(x)$ is increasing to the left of $x = c$, which means that $f'(x) > 0$, and $f(x)$ is decreasing to the right of $x = c$, which means that $f'(x) < 0$ (Figure 5.38).

We summarize these observations in Figure 5.39, in which we draw $y = f'(x)$ along with $y = f(x)$ for a local minimum point and for a local maximum point. For any local extremum that is not an endpoint, Fermat's rule requires that $f'(c) = 0$, so $f'(x)$ crosses through 0 as x crosses through $x = c$. If $x = c$ is a local minimum then $f'(x)$ increases through 0, while if $x = c$ is a local maximum then $f'(x)$ decreases through 0 as x crosses through the value $x = c$.

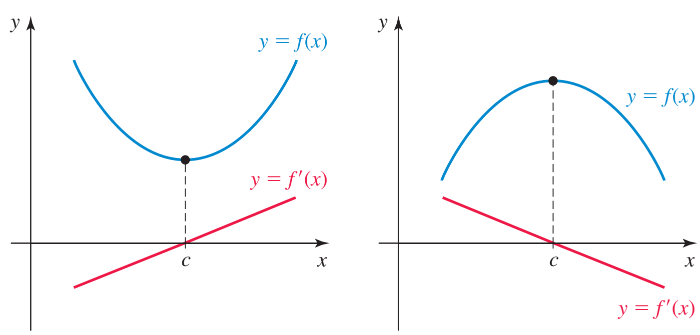


Figure 5.39 A function $f(x)$ has a local extremum at a point $x = c$ that is not an endpoint of the function's domain. If c is a local minimum point $f'(x)$ increases through 0, and if c is a local maximum point $f'(x)$ decreases through 0.

What if c coincides with the end point of the interval; meaning that $c = a$ or $c = b$? In most cases then $f'(c) \neq 0$, meaning that Fermat's rule does not apply. But the description of the local behavior of the function is still valid. The main difference between boundary points $c = a$ or $c = b$ and interior points, $a < c < b$, is that we can only examine the behavior of $f(x)$ on one side of a boundary point. For example a is a local maximum if $f(x)$ is decreasing right of a or a local minimum if $f(x)$ is increasing right of a. The behavior of $f(x)$ left of a doesn't matter, because all points left of a are outside of the interval on which we are defining the function. Now if $f(x)$ is differentiable at a, we can determine where $f(x)$ slopes downward or upward from the sign of $f'(a)$, as in Figure 5.40. (Strictly speaking we need to add the condition that $f'(x)$ is continuous on an interval containing $x = a$; this condition is met for essentially all of the functions that appear in biological models.) If $f'(a) > 0$ then the function slopes upward at $x = a$, so a is a local minimum. If $f'(a) < 0$ then the function slopes downward at $x = a$ so a is a local maximum. We can make similar arguments if the local extremum point coincides with the right end point of the interval (see Problem 19). To summarize, the first derivative can be

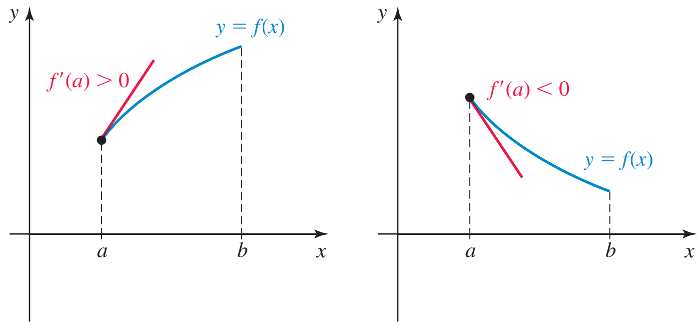


Figure 5.40 A function $f(x)$ defined on an interval $[a, b]$ has a local extremum at the point $x = a$ (the left end point of the interval). $f'(a) > 0$ if $x = a$ is a local minimum, since this corresponds to an upward slope, and $f'(a) < 0$ if $x = a$ is a local maximum, since this corresponds to a downward slope.

used to determine whether the endpoints of the domain of f are local minima or local maxima:

> **The First-Derivative Test for Local Extrema at Endpoints** Suppose that f is differentiable on the interval $[a, b]$ and its derivative is continuous at $x = a$ and $x = b$. Then:
>
> If $f'(a) > 0$ then f has a local minimum at $x = a$.
> If $f'(a) < 0$ then f has a local maximum at $x = a$.
> If $f'(b) > 0$ then f has a local maximum at $x = b$.
> If $f'(b) < 0$ then f has a local minimum at $x = b$.

If the local extremum, c, is not an endpoint and if f is twice differentiable at c, then there is a shortcut for determining whether c is a local maximum or a local minimum. Figure 5.39 shows for a local minimum point, $f'(x)$ is increasing through 0, which means that the slope of the curve $y = f'(x)$ is positive. The slope of the curve $y = f'(x)$ is given by the second derivative, $f''(x)$. For a local minimum, $f'(x)$ is increasing at $x = c$, which implies that $f''(c) > 0$. Similarly, for a local maximum, $f'(x)$ is decreasing at $x = c$, which implies that $f''(c) < 0$.

> **The Second-Derivative Test for Local Extrema** Suppose that f is twice differentiable on an open interval containing c.
>
> If $f'(c) = 0$ and $f''(c) < 0$, then f has a local maximum at $x = c$.
> If $f'(c) = 0$ and $f''(c) > 0$, then f has a local minimum at $x = c$.

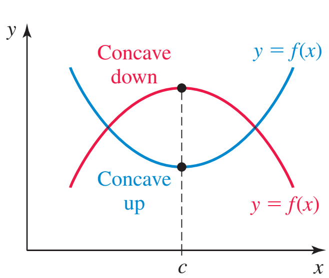


Figure 5.41 The function $y = f(x)$ has a local extremum at $x = c$. If $x = c$ is a local minimum, then $f(x)$ is concave up at that point, and if $x = c$ is a local maximum, then $f(x)$ is concave down.

Remembering that the second derivative tells us whether the graph of the function curves upward or downward provides another way to understand the second-derivative test. Looking at Figure 5.41 we see that if the function has a local minimum at $x = c$, its graph will curve upward there, meaning that the function is concave upward, so $f''(c) > 0$. If, instead the function has a local maximum at $x = c$, it will curve downward there, meaning that the function is concave downward, so $f''(c) < 0$.

Finding the point c where $f'(c) = 0$ gives us a candidate for a local extremum. If the second derivative $f''(c) \neq 0$, then this point is truly a local extremum, and the sign of $f''(c)$ tells whether it is a local maximum or minimum.

The next two examples make use of the second-derivative test to identify extrema.

EXAMPLE 1 Find all local and global extrema of

$$f(x) = \frac{3}{2}x^4 - 2x^3 - 6x^2 + 2, \quad x \in \mathbf{R}.$$

Solution The second derivative test requires us to know both the first derivative, to find the points at which $f'(x) = 0$, (the candidate local extrema), and the second derivative, to determine whether these points are local minima or local maxima. Since $f(x)$ is twice differentiable for all $x \in \mathbf{R}$, we begin by finding the first two derivatives of f. The first derivative is

$$f'(x) = 6x^3 - 6x^2 - 12x.$$

The second derivative is

$$f''(x) = 18x^2 - 12x - 12.$$

Since $f'(x)$ exists for all $x \in \mathbf{R}$ and the domain has no endpoints, the only candidates for local extrema are points where $f'(x) = 0$. To find these points we factorize $f'(x)$:

$$f'(x) = 6x^3 - 6x^2 - 12x = 6x(x-2)(x+1)$$

x is a factor, so $f'(x) = x(6x^2 - 6x - 12)$. Then factorize the quadratic.

We thus find that $x = 0$, $x = 2$, and $x = -1$ are all possible local extrema. Since $f''(x)$ exists, we can use the second-derivative test to determine which of these points are local extrema and, if so, of what type they are. We need to evaluate the second derivative at each x-coordinate:

$$f''(0) = -12 < 0 \implies \text{local maximum at } x = 0$$
$$f''(2) = 36 > 0 \implies \text{local minimum at } x = 2$$
$$f''(-1) = 18 > 0 \implies \text{local minimum at } x = -1$$

The function $f(x)$ is defined on $\mathbf{R}$. As mentioned at the beginning of this section, in order to find global extrema, we must check the local extrema and compare their values against each other and against the function values as $x \to \infty$ and $x \to -\infty$ (which serve the role of the endpoints of the interval). At our local extrema we have

$$f(0) = 2 \qquad f(2) = -14 \qquad f(-1) = -\frac{1}{2}$$

and

$$\lim_{x\to\infty} f(x) = \infty \qquad \lim_{x\to-\infty} f(x) = \infty$$

See Section 3.3

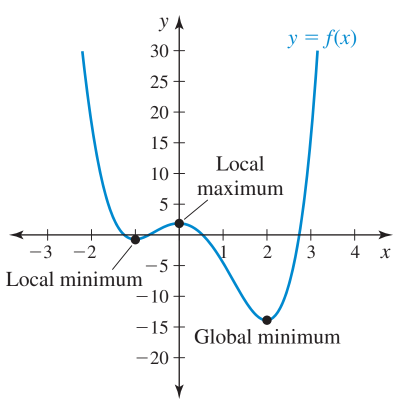


Figure 5.42 The graph of $f(x) = \frac{3}{2}x^4 - 2x^3 - 6x^2 + 2$ in Example 1. $f(x)$ has a global minimum at $x = 2$, but no global maximum.

The local minimum at $x = 2$ is the global minimum. Although $x = 0$ is a local maximum, it is not a global maximum: Since the function goes to ∞ as $x \to \pm\infty$, it certainly exceeds the value 2. In fact, there is no global maximum. Figure 5.42 shows the graph of $f(x)$.

EXAMPLE 2 Find all local and global extrema of

$$f(x) = x(1-x)^{2/3}, \quad x \in \mathbf{R}$$

Solution Since we need to know $f'(x)$ to find local extrema and $f''(x)$ to diagnose whether they are local minima or local maxima, we start by calculating these derivatives.

$$f(x) = \overbrace{x}^{u(x)} \cdot \overbrace{(1-x)^{2/3}}^{v(x)}$$

$$f'(x) = (1-x)^{2/3} + x \cdot \frac{2}{3}(1-x)^{-1/3}(-1) \qquad \text{Product rule}$$

$$= (1-x)^{2/3} - \frac{2x}{3(1-x)^{1/3}}.$$

$$f''(x) = \frac{2}{3}(1-x)^{-1/3}(-1) - \frac{2}{3}\left[(1-x)^{-1/3} + x\left(-\frac{1}{3}\right)(1-x)^{-4/3}(-1)\right] \quad \text{Product rule for second term}$$

$$= -\frac{2}{3(1-x)^{1/3}} - \frac{2}{3(1-x)^{1/3}} - \frac{2x}{9(1-x)^{4/3}}$$

$$= -\frac{4}{3(1-x)^{1/3}} - \frac{2x}{9(1-x)^{4/3}}.$$

Both of these functions are defined except when $x = 1$. It follows that $f(x)$ is twice differentiable provided $x \neq 1$. We will have to separately consider $x = 1$, after finding any local maxima or minima at which f is twice differentiable. First we find points where $f'(x) = 0$, which requires:

$$(1-x)^{2/3} = \frac{2x}{3(1-x)^{1/3}}$$

$$3(1-x) = 2x \quad \times 3(1-x)^{1/3} \text{ on both sides}$$

$$3 = 5x \implies x = \frac{3}{5} \quad \text{Solve for } x$$

$x = \frac{3}{5}$ is a possible local extremum; to diagnose what kind of extremum it is, we check the second derivative:

$$f''\left(\frac{3}{5}\right) = -\frac{4}{3(\frac{2}{5})^{1/3}} - \frac{2(\frac{3}{5})}{9(\frac{2}{5})^{4/3}} = \left(\frac{5}{2}\right)^{1/3}\left(-\frac{4}{3} - \frac{2(\frac{3}{5})}{9(\frac{2}{5})}\right)$$

$$= \left(\frac{5}{2}\right)^{1/3}\left(-\frac{4}{3} - \frac{1}{3}\right) = -\frac{5}{3}\left(\frac{5}{2}\right)^{1/3} < 0$$

We have simplified the expression, but you could also use a calculator to evaluate it directly. Since $f''(\frac{3}{5}) < 0$, by the second derivative test $f(x)$ has a local maximum at $x = \frac{3}{5}$.

We are not finished finding local extrema yet; recall that points at which $f'(x)$ is not defined may also be local extrema. The first derivative is not defined at $x = 1$. We therefore must investigate the function in the neighborhood of $x = 1$. Note that $f(x)$ is continuous at $x = 1$ with $f(1) = 0$. Since we cannot compute $f'(1)$, we need to rely on our description of the local behavior of a function near a local maximum or minimum point; namely whether the function is increasing or decreasing on either side of $x = 1$. We can determine whether it is increasing or decreasing from the sign of $f'(x)$ on either side of $x = 1$. Any continuous function $g(x)$ can only change sign at points where $g(x) = 0$ or at points where $g(x)$ is not defined, so $f'(x)$ can only change sign at $x = \frac{3}{5}$ and at $x = 1$. Thus, $f'(x)$ must have the same sign for all $x < \frac{3}{5}$, the same sign for all $\frac{3}{5} < x < 1$, and the same sign for all $x > 1$. To determine what the sign of $f'(x)$ is for each of these intervals, we can evaluate it at a single point in each interval.

$$x < \frac{3}{5}: f'(0) = 1 > 0 \implies f'(x) > 0$$

$$\frac{3}{5} < x < 1: f'\left(\frac{4}{5}\right) = \left(\frac{1}{5}\right)^{2/3} - \frac{2(\frac{4}{5})}{3(\frac{1}{5})^{1/3}} = \frac{1}{5^{2/3}}\left(1 - \frac{8}{3}\right) = \frac{-5^{1/3}}{3} < 0$$

$$\implies f'(x) < 0. \quad \text{You can also use a calculator to evaluate: } f'(4/5) = -0.57$$

$$x > 1: f'(2) = (-1)^{2/3} - \frac{2 \times 2}{3(-1)^{1/3}} = 1 + \frac{4}{3} = \frac{7}{3} > 0$$

$$\implies f'(x) > 0$$

The sign of $f'(x)$ is summarized in the following number line:

We conclude that $f(x)$ has a local minimum at $x = 1$ since $f'(x)$ is negative left of $x = 1$, and positive right of $x = 1$, just like Figure 5.37. We can say a bit more about what is happening at $x = 1$:

$$\lim_{x\to 1^-} f'(x) = \lim_{x\to 1^-}\left[(1-x)^{2/3} - \frac{2x}{3(1-x)^{1/3}}\right] = -\infty$$
$$\lim_{x\to 1^+} f'(x) = \lim_{x\to 1^+}\left[(1-x)^{2/3} - \frac{2x}{3(1-x)^{1/3}}\right] = \infty$$

So the tangent is vertical at $x = 1$; such points are called **cusps**.

To determine which (if any) of the local extrema are also global extrema, we must compare the value of the function there with its value at the endpoints of the interval. For a function defined on **R**, this means comparing with the values that the function takes in the limits as $x \to \infty$ and $x \to -\infty$: As $x \to +\infty$, $f(x) \to +\infty$, since both x and $(1-x)^{2/3}$ are large and positive. As $x \to -\infty$, x becomes large and negative while $(1-x)^{2/3}$ is large and positive, so $f(x) = x(1-x)^{2/3} \to -\infty$. So neither of the two local extrema is the global minimum or maximum of the function.

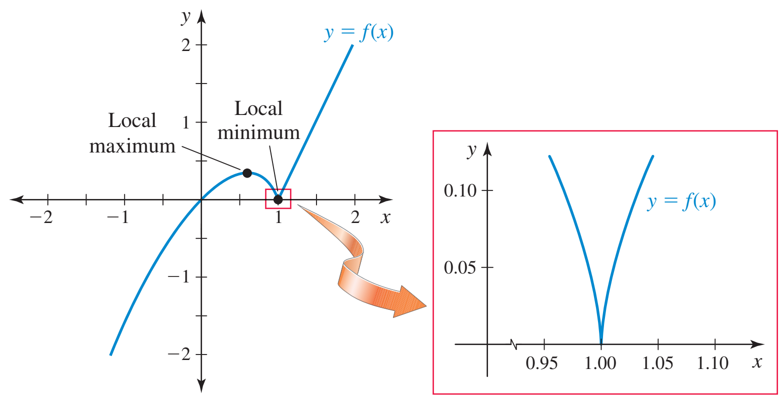


Figure 5.43 The function $f(x) = x(1-x)^{2/3}$ has a local maximum at $x = 3/5$ and a local minimum at $x = 1$. At $x = 1$ the function has a cusp, and is not differentiable. There is no global maximum or minimum.

The function $f(x)$ is plotted in Figure 5.43. The inset to this figure shows $f(x)$ near the cusp at $x = 1$, where the slope of $f(x)$ becomes infinite. •

Example 2 shows that the second derivative rule cannot be used to diagnose whether an extremum point is a local maximum or minimum if $f'(x)$ or $f''(x)$ is not defined at the extremum. The next example shows it also cannot be used if $f''(x) = 0$ at the extremum point.

EXAMPLE 3 Find the local and global extrema of $f(x) = x^4(1-x)$ on the interval $-1 \le x \le 1$.

Solution We multiply out $f(x)$ to make calculating the first and second derivatives easier.

$$f(x) = x^4 - x^5$$
$$f'(x) = 4x^3 - 5x^4 = x^3(4-5x) \qquad \text{Factorize}$$
$$f''(x) = 12x^2 - 20x^3 = 4x^2(3-5x).$$

Factorizing $f'(x)$ allows us to find points at which $f'(x) = 0$ more easily: they are $x = 0$ and $x = \frac{4}{5}$. Since $f'(x)$ is defined for all $x \in \mathbf{R}$, these and the endpoints of the interval ($x = -1$ and $x = 1$) are the only possible local extrema.

We start with the internal local extrema: $x = 0$ and $x = \frac{4}{5}$. $f''(0) = 0$ so the second derivative test cannot be used for $x = 0$, while for $x = \frac{4}{5}$:

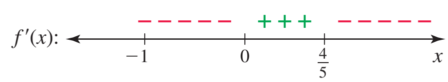
$$f''\left(\frac{4}{5}\right) = 4 \cdot \left(\frac{4}{5}\right)^2 \cdot (3 - 4) = -4 \cdot \left(\frac{4}{5}\right)^2 < 0,$$

so $x = \frac{4}{5}$ is a local maximum. To determine the nature of the extremum at $x = 0$, we need to examine the sign of $f'(x)$ to the left and the right of $x = 0$, to determine where $f(x)$ is increasing or decreasing. Since $f'(x)$ is a continuous function, it must take the same sign for all $x < 0$, the same sign for all $0 < x < \frac{4}{5}$, and the same sign for all $x > \frac{4}{5}$.

Since $x = \frac{4}{5}$ is a local maximum $f'(x)$ changes from positive to negative there. ($f(x)$ is increasing left of $x = 4/5$ and decreasing right of $x = 4/5$, as in Figure 5.38.) Therefore $f'(x) > 0$ for all $0 < x < \frac{4}{5}$ and $f'(x) < 0$ for all $x > \frac{4}{5}$. To get the sign of $f'(x)$ for $x < 0$, we just need to evaluate $f'(x)$ for a single value of x in this interval, e.g., at $x = -1$,

$$f'(-1) = (-1)^3(4 - 5(-1)) = -9 < 0 \implies f'(x) < 0 \quad \text{for all} \quad x < 0.$$

The sign of $f'(x)$ is summarized in the following number line.



Since $f'(x) < 0$ just left of $x = 0$, and $f'(x) > 0$ just right of $x = 0$, $x = 0$ must be a local minimum. Since $f(x)$ is differentiable over the entire of the interval $-1 \leq x \leq 1$, the only other places local extrema can occur are at the endpoints of the interval. From the number line we see that $f''(-1) < 0$ so by the first derivative criterion for endpoints, $x = -1$ is a local maximum. Also from the number line we see that $f'(1) < 0$, so by the first derivative criterion, $x = 1$ is a local minimum point.

To find the global extrema we must compare the values between all local extrema (since $f(x)$ is defined on a closed interval, the Extreme-Value Theorem guarantees us that global extrema exist). To find the global maximum we compare the value of $f(x)$ at its two local maxima: $x = -1$ and $x = \frac{4}{5}$:

$$f(-1) = (-1)^4(1 - (-1)) = 2,$$

$$f\left(\frac{4}{5}\right) = \left(\frac{4}{5}\right)^4\left(1 - \frac{4}{5}\right) = \left(\frac{4}{5}\right)^4 \cdot \frac{1}{5} = \frac{4^4}{5^5} \approx 0.082 < 2.$$

$x = -1$ is the global maximum point, and $f(-1) = 2$ is the global maximum value. To find the global minimum we compare the value of $f(x)$ at its two local minima: $x = 0$ and $x = 1$:

$$f(0) = 0^4(1 - 0) = 0, \qquad f(1) = 1^4(1 - 1) = 0.$$

So both $x = 0$ and $x = 1$ are global minima, and the minimum value of the function is 0.

$f(x)$ is plotted in Figure 5.44.

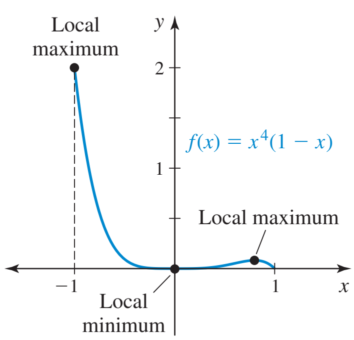


Figure 5.44 The function $f(x) = x^4(1 - x)$ has a local minimum at $x = 0$, but $f''(0) = 0$.

5.3.2 Inflection Points

We saw in Sections 5.2 and 5.3.1 that knowing whether a function $f(x)$ is concave up or concave down is an important piece of qualitative information. It can tell us whether a local extremum is a maximum or minimum point, and whether there are increasing or diminishing returns as the variable x increases. For this reason it is often helpful to identify the points in which the function shifts from concave up to concave down, or vice versa. These points are called **inflection points.** We start by defining these points, and then give a method for locating them. (See Figure 5.45.)

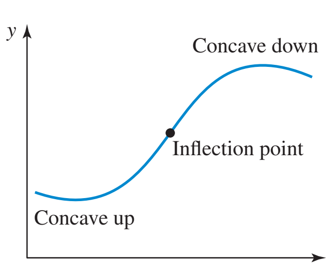


Figure 5.45 Inflection points are points where the concavity of a function changes.

> Inflection points are points where the concavity of a function changes—that is, where the function changes from concave up to concave down or from concave down to concave up.

EXAMPLE 4 Show that the function

$$f(x) = \frac{1}{2}x^3 - \frac{3}{2}x^2 + 2x + 1, \quad x \in \mathbf{R}$$

has an inflection point at $x = 1$.

Solution The graph of $f(x)$ is shown in Figure 5.46. We compute the first two derivatives:

$$f'(x) = \frac{3}{2}x^2 - 3x + 2$$

$$f''(x) = 3x - 3 = 3(x - 1)$$

Since $f''(x)$ is positive for $x > 1$ and negative for $x < 1$, $f''(x)$ changes sign at $x = 1$. We therefore conclude that $f(x)$ has an inflection point at $x = 1$. •

If a function f is twice differentiable, it is concave up if $f'' > 0$ and concave down if $f'' < 0$. At an inflection point, f'' must therefore change sign; that is, the second derivative must be 0 at an inflection point. More formally,

> **Second derivative test for inflection points.** If $f(x)$ is twice differentiable and has an inflection point at $x = c$, then $f''(c) = 0$.

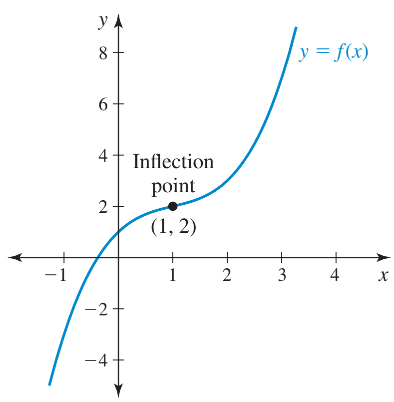


Figure 5.46 The function $f(x) = \frac{1}{2}x^3 - \frac{3}{2}x^2 + 2x + 1$ has an inflection point at $x = 1$.

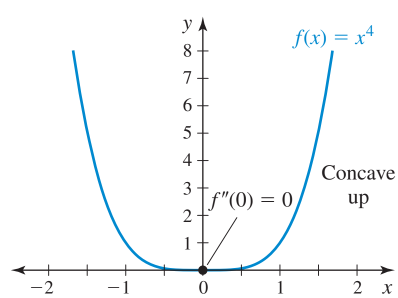


Figure 5.47 If $f(x) = x^4$ then $f''(0) = 0$ but $x = 0$ is not an inflection point.

Note that $f''(c) = 0$ is a necessary, but not sufficient, condition for the existence of an inflection point of a twice-differentiable function. It is similar to Fermat's Theorem (from Section 5.1) for finding candidate local extremum points. The second derivative test can only be used to find candidate inflection points; to determine whether a candidate is an inflection point, we must check whether the second derivative changes sign. For instance, $f(x) = x^4$ has $f''(0) = 0$, but the function $f(x) = x^4$ is concave up (see Figure 5.47), so there is no inflection point at $x = 0$.

Section 5.3 Problems

5.3.1

Find the local maxima and minima of each of the functions in Problems 1–17. Determine whether each function has local maxima and minima and find their coordinates. For each function, find the intervals on which it is increasing and the intervals on which it is decreasing.

1. $y = (x - 1)^2, -2 \le x \le 3$

2. $y = \sqrt{x + 1}, 1 \le x \le 2$

3. $y = \sqrt{(2x - 1)}, 1/2 \le x \le 2$

4. $y = \ln\left(\frac{x}{x+1}\right), x > 0$

5. $y = xe^{-x}, 0 \le x \le 1$

6. $y = |x^2 - 25|, -5 \le x \le 8$

7. $y = x^3 - 3x^2 + 3x + 3, x \in \mathbf{R}$

8. $y = x^3 - 3x + 1, x \in \mathbf{R}$

9. $y = e^{-x^2}, -1 \le x \le 1$

10. $y = \cos(\pi x^2), -1 \le x \le 1$

11. $y = e^{-|x|}, x \in \mathbf{R}$

12. $y = \sin 2\pi x, 0 \le x \le 1$

13. $y = \frac{1}{3}x^3 + \frac{1}{2}x^2 + 2, x \in \mathbf{R}$

14. $y = \frac{1}{3}x^3 - x^2 + x + 1, x \in \mathbf{R}$

15. $y = x^2(1 - x), x \in \mathbf{R}$

16. $y = (x - 1)^{1/3}, x \ge 1$

17. $y = \sqrt{1 + x^2}, x \in \mathbf{R}$

18. [This problem illustrates the fact that $f'(c) = 0$ is not a sufficient condition for the existence of a local extremum of a differentiable function.] Show that the function $f(x) = x^3$ has a horizontal tangent at $x = 0$; that is, show that $f'(0) = 0$, but $f'(x)$ does not change sign at $x = 0$ and, hence, $f(x)$ does not have a local extremum at $x = 0$.

19. Explain the first-derivative criterion for local extrema located at the right endpoint of the domain of a function. That is, if $f(x)$ is defined on a domain $[a, b]$, and is differentiable on some interval containing $x = b$, show that $x = b$ is a local maximum point if $f'(b) > 0$ and a local minimum point if $f'(b) < 0$.

20. Suppose that $f(x)$ is differentiable on $\mathbf{R}$, with $f(x) > 0$ for $x \in \mathbf{R}$. Show that if $f(x)$ has a local maximum at $x = c$, then $g(x) = \ln f(x)$ also has a local maximum at $x = c$.

21. Suppose that $f(x)$ is differentiable on $\mathbf{R}$. Show that if $f(x)$ has a local maximum at $x = c$, then $g(x) = e^{f(x)}$ also has a local maximum at $x = c$.

5.3.2

In Problems 22–27, determine all inflection points.

22. $f(x) = x^3 - 2, x \in \mathbf{R}$

23. $f(x) = (x - 3)^5, x \in \mathbf{R}$

24. $f(x) = e^{-x^2}, x \ge 0$

25. $f(x) = xe^{-x}, x \ge 0$

26. $f(x) = \frac{x}{x+1}, x \ge 0$

27. $f(x) = \frac{x^2}{x^2+1}$

28. $f(x) = \tan x, -\dfrac{\pi}{2} < x < \dfrac{\pi}{2}$

29. $f(x) = x \ln x, x > 0$

30. [This problem illustrates the fact that $f''(c) = 0$ is not a sufficient condition for an inflection point of a twice-differentiable function.] Show that the function $f(x) = x^4$ has $f''(0) = 0$ but that $f''(x)$ does not change sign at $x = 0$ and, hence, $f(x)$ does not have an inflection point at $x = 0$.

31. **Logistic Equation** Suppose that the size of a population at time t is denoted by $N(t)$ and satisfies

$$N(t) = \frac{100}{1 + 3e^{-2t}}$$

for $t \ge 0$.

(a) Show that $N(0) = 25$.

(b) Show that $N(t)$ is strictly increasing.

(c) Show that

$$\lim_{t\to\infty} N(t) = 100$$

(d) Show that $N(t)$ has an inflection point when $N(t) = 50$—that is, when the size of the population is at half its maximum value.

32. **Hill's Function for Hemoglobin Binding** Hill's equation for the oxygen saturation of blood states that the level of oxygen saturation (fraction of hemoglobin molecules that are bound to oxygen) in blood can be represented by a function:

$$f(P) = \frac{P^n}{P^n + 30^n}$$

where P is the oxygen concentration around the blood ($P \ge 0$) and n is a parameter that varies between different species.

(a) Assume that $n = 1$. Show that $f(P)$ is an increasing function of P and that $f(P) \to 1$ as $P \to \infty$.

(b) Assuming that $n = 1$ show that $f(P)$ has no inflection points. Is it concave up or concave down everywhere?

(c) Knowing that $f(P)$ has no inflection points, could you deduce which way the curve bends (whether it is concave up or concave down) without calculating $f''(P)$?

(d) For most mammals n is close to 3. Assuming that $n = 3$ show that $f(P)$ is an increasing function of P and that $f(P) \to 1$ as $P \to \infty$.

(e) Assuming that $n = 3$, show that $f(P)$ has an inflection point, and that it goes from concave up to concave down at this inflection point.

(f) Using a graphing calculator plot $f(P)$ for $n = 1$ and $n = 3$. How do the two curves look different?

33. **Drug Absorption** A two-compartment model of how drugs are absorbed into the body predicts that the amount of drug in the blood will vary with time according to the following function:

$$M(t) = a(e^{-kt} - e^{-3kt}), \qquad t \ge 0$$

where $a > 0$ and $k > 0$ are parameters that vary depending on the patient and the type of drug being administered. For parts (a)–(c) of this question you should assume that $a = 1$.

(a) Show that $M(t) \to 0$ as $t \to \infty$.

(b) Show that the function $M(t)$ has a single local maximum, and find the maximum concentration of drug in the patient's blood.

(c) Show that the $M(t)$ has a single inflection point (which you should find). Does the function go from concave up to concave down at this inflection point or vice versa?

(d) Would any of your answers to (a)–(c) be changed if a were not equal to 1. Which answers?

5.4 Optimization

In Sections 5.1 and 5.3 we showed how calculus can be used to find the extrema of functions. Calculating extrema is often useful: For example if the function represents how the crop yield in a field depends on the amount of fertilizer that is added, then the global maximum of the function tells us how much fertilizer needs to be added to maximize yields. These kinds of problems are called **optimization problems**: by maximizing the yield function we find the optimal amount of fertilizer to add. Optimization can be used to design experiments of treatments, for example, by predicting the optimal amount of drug to give to a patient. Optimization can also be used to understand organisms better: for example, organisms as diverse as fungal spores and fish are shaped to travel as fast as possible by minimizing their air or water resistance. Blood vessels are evolved to transport blood around the body at the smallest cost to the organism of materials and energy lost to flow. Organisms also optimize their life histories; the age at which they start reproduction and how much to invest in each offspring. By building a mathematical model for the drag an organism experiences, the cost of building and maintaining networks, or for the fitness of an organism and then optimizing this model to find the optimum shape, transport networks, or life history, we may be able to show that the model for our organism shape or behavior is correct. In each case, once you obtain a function that you wish to optimize, the results from Sections 5.1 and 5.3 will help you to find the global extremum. It is important to state the domain of the function, since global extrema may be found at the endpoints of the function's domain as well as within it.

EXAMPLE 1 **Crop Yield** The yield of crop in a field depends on the amount of nitrogen fertilizer that is added to the field. In Section 5.2 we met one formula that relates the yield of crops (measured, for example, in kilograms per m^2 of field) to the amount of nitrogen added (measured, for example, in grams per m^2 of field). Suppose that the yield is given by a function:

$$Y(N) = \frac{N}{N+1}$$

$Y(N)$ increases monotonically with N; we can check this by calculating the derivative:

$$Y'(N) = \frac{1 \times (N+1) - N \times 1}{(N+1)^2} = \frac{1}{(N+1)^2} > 0 \qquad \text{Quotient rule}$$

So adding more nitrogen fertilizer always increases the yield. But we also need to consider the cost of the fertilizer. Instead of maximizing yield, most farmers are, in practice, most interested in maximizing their return (that is, their net profit, once the cost of materials and so on has been tallied). Suppose that the revenue from selling one kilo of crop is 1, and the cost of one gram of fertilizer is C; then the return per m^2 of crop is:

$$\begin{aligned} r(N) &= \text{revenue} - \text{cost of fertilizer} = Y(N) - CN \\ &= \frac{N}{N+1} - CN \end{aligned}$$

What value of N maximizes the farmer's total return, $r(N)$?

Solution First we try to understand why there might be an optimum fertilizer level, N. We start by plotting the function $r(N)$ for the specific value $C = 1/2$ (see Figure 5.48). We see from the plot that the return initially increases when more nitrogen is added, but then starts to decrease again. We can understand this behavior in terms of a trade-off between increased yield and the cost of the fertilizer. For small values of N the yield increases with increasing N, leading to increased returns. However as N increases $Y(N) \to 1$ meaning that the yield levels stop increasing. However, the cost of adding extra nitrogen continues to increase proportionally to the amount of nitrogen added; eventually the cost of adding more nitrogen fertilizer will exceed the benefit.

The plot in Figure 5.48 strongly suggests that for this particular value of C, there is a specific value for N that maximizes the farmer's return. We can use the methods of Sections 5.1 and 5.3 to find this value and how it depends on the cost of fertilizer, C.

To find any local extrema, we differentiate $r(N)$:

$$r'(N) = Y'(N) - C = \frac{1}{(N+1)^2} - C \qquad Y'(N) \text{ was calculated above.}$$

A point $N > 0$ is a candidate for a local extremum only if $r'(N) = 0$ meaning that

$$\frac{1}{(N+1)^2} = C \implies (N+1)^2 = \frac{1}{C} \implies N = -1 \pm \sqrt{\frac{1}{C}}$$

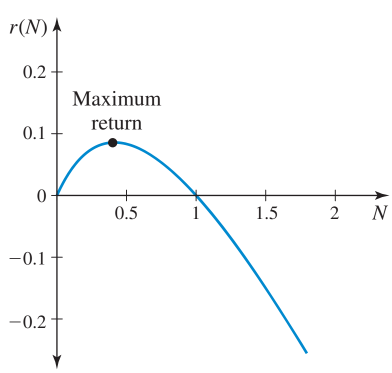


Figure 5.48 Starting at $N = 0$, the return increases initially when the amount of nitrogen added is increased. At high nitrogen levels the return starts to decrease.

$N = -1 - \sqrt{\frac{1}{C}}$ is certainly negative, so it cannot lie in the domain on which we are trying to maximize $r(N)$. The other extremum point will be positive if and only if $\sqrt{\frac{1}{C}} > 1$; i.e., if $C < 1$.

Since $r(N)$ is differentiable over the entire interval on which it is defined, the only other candidate points for the global extremum occur at endpoints of the interval on which $r(N)$ is defined, meaning at $N = 0$ or as $N \to +\infty$. At $N = 0$, $r(0) = 0$ and $r'(0) = 1 - C$. So by the first derivative test $N = 0$ is a local minimum if $C < 1$ (since then $r'(0) > 0$) and a local maximum if $C > 1$ (since then $r'(0) < 0$). As $N \to +\infty$, $Y(N) \to 1$, so $r(N) \to -\infty$, meaning that there is no global minimum return.

To link what we know about local extrema and to find the global maximum value of $r(N)$, let's first assume that $C < 1$. In this case $N = 0$ is a local minimum, with $r'(0) > 0$. There is a local extremum point at $N = -1 + \sqrt{\frac{1}{C}}$. Since $r'(N) \to -C < 0$ as $N \to \infty$ it follows that $r'(N)$ must go from positive to negative at the local extremum point, so $N = -1 + \sqrt{\frac{1}{C}}$ is a local maximum. Since it is the only local maximum point, it must also be the global maximum: $N = -1 + \sqrt{\frac{1}{C}}$ therefore gives the maximum return.

If, on the other hand, $C > 1$, then $r'(0) < 0$, so $N = 0$ is a local maximum. There are no points in $N > 0$ with $r'(N) = 0$, and $r'(N) \to -C < 0$ as $N \to \infty$. So $r'(N) < 0$ for all $N > 0$, meaning that the return decreases with N. $N = 0$ (adding no nitrogen) gives the global maximum. Put another way, if the cost per gram, C, of fertilizer is too high, then the best return is achieved by not adding fertilizer, even though there is no crop yield at $N = 0$. •

EXAMPLE 2 **Maximizing Area** An ecologist wants to enclose a rectangular study plot. She has 1600 ft of fencing. Using this fencing, determine the dimensions of the study plot that will have the largest area.

Solution Figure 5.49 illustrates the situation. The area A of this study plot is given by

$$A = xy \tag{5.10}$$

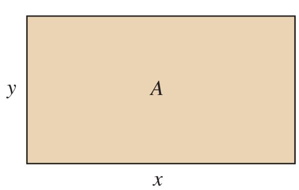


Figure 5.49 The rectangular study plot in Example 2.

and the perimeter of the study plot is given by

$$1600 = 2x + 2y \tag{5.11}$$

From two equations in two variables, we must reduce to one function with one independent variable. To do this, solve (5.11) for y:

$$y = 800 - x$$

and substitute for y in (5.10):

$$A(x) = x(800 - x) = 800x - x^2 \quad \text{for } 0 \le x \le 800$$

It is important to state the domain of the function. Clearly, the smallest value of x is 0, in which case the enclosed area is also 0, since $A(0) = 0$. The largest possible value for x is 800, which will also produce a rectangle with two sides of length 0; the corresponding area is $A(800) = 0$.

We wish to maximize the enclosed area $A(x)$. The function $A(x)$ is differentiable for $x \in (0, 800)$ with

$$A'(x) = 800 - 2x \quad \text{for } 0 \le x \le 800 \qquad \text{Need } A'(x) \text{ to find local extrema}$$

$$A''(x) = -2 \quad \text{for } 0 \le x \le 800 \qquad \text{Need } A''(x) \text{ to diagnose maxima/minima}$$

At local extrema $A'(x) = 0$ so:

$$800 - 2x = 0, \qquad \text{or} \qquad x = 400$$

Since $A''(400) < 0$, $x = 400$ is a local maximum. To find the global maximum, we also need to check the function $A(x)$ at the endpoints of the interval $[0, 800]$. We have

$$A'(0) = 800 > 0 \quad \text{and} \quad A'(800) = -800 < 0$$

so the endpoints $x = 0$ and $x = 800$ are both local minima by the first derivative test. The area is maximized when $x = 400$, which implies that the study plot is a square. This relationship is true in general: For a rectangle with fixed perimeter, the maximum area occurs when the rectangle is a square (see Problem 2). •

EXAMPLE 3

Bio Info • The life history of an organism may include many variables: the age at which an organism first reproduces, the age at which it stops reproducing, the number of offspring that the organism has each year, and how much of its resources it invests in raising each of them. Compare, for example, king penguins, which may produce only 10 chicks over the course of their 20-year life-span, with mushrooms, which can release billions of propagules per day. Why do different organisms differ so widely in their reproductive strategies? In this example we will study how the number of offspring that an organism produces may be optimized to ensure that the largest number of offspring survive to adulthood. The number of offspring is determined by the amount of resources the parents can provide to each offspring. On one hand, the parent has limited resources, and the more offspring they have, the fewer resources they can devote to each individual offspring. On the other hand, over-investing in a few offspring may not be a good strategy, because it limits the number of progeny that the parent may have, and because something outside of the parent's control (such as changing environment) may kill its offspring. This trade-off between having too few offspring and having too many suggests that an intermediate number of offspring might be optimal.

Lloyd (1987) (summarized in ***Roff, 1992***) proposed the following model to determine the optimal number of offspring for an organism. Suppose the total amount of resources that the organism can provide for its offspring is R, and it has N offspring. R varies between organisms and also depends on the organism's environment. The constant amount of resources invested in each offspring is $x = R/N$. The chance $f(x)$ that an offspring survives is an S-shaped, or sigmoidal, function of x; that is, survival chances are very low when the investment is low, survival chances increase as the investment increases, but show diminishing returns as investments get larger and larger. If the organism produces N offspring, the mean number that will survive (usually called **fitness** by biologists) is,

$$\begin{aligned} w(x) &= (\text{number of offspring}) \times (\text{probability of survival of offspring}) \\ &= N f(x) = \frac{R}{x} f(x) \quad \text{for } x > 0 \qquad x = \tfrac{R}{N} \Rightarrow N = \tfrac{R}{x} \end{aligned} \tag{5.12}$$

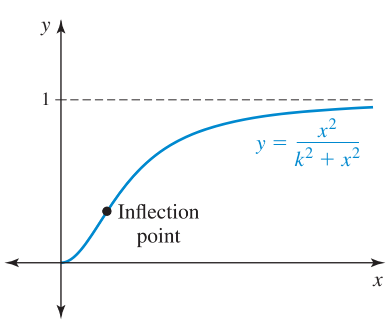


Figure 5.50 Probability of offspring survival in Example 3.

The optimal value of x, from the organism's point of view, is the one that maximizes the fitness $w(x)$.

A common choice to model the function $f(x)$ is

$$f(x) = \frac{x^2}{k^2 + x^2} \quad \text{for } x \geq 0$$

where k is a positive constant: the value of the constant k will depend on what species of organism is being modeled.

The graph of $f(x)$ is shown in Figure 5.50. The curve is sigmoidal: initially the curve is concave upward, but for large x, $f(x)$ asymptotes to a constant and the curve becomes concave downward. So, for small k the organism has increasing returns for increasing its investment in its offspring, but once x crosses the inflection point of $f(x)$, returns start to diminish. According to (5.12), the fitness of the organism is:

$$w(x) = R\frac{f(x)}{x} = \frac{Rx}{k^2 + x^2}$$

and the domain of the function is $x \geq 0$ since $R \geq 0$ and $N > 0$. To calculate the optimal reproduction strategy for the organism we need to calculate $w'(x)$

$$w'(x) = \frac{R(k^2 + x^2) - Rx(2x)}{(k^2 + x^2)^2} = \frac{R(k^2 - x^2)}{(k^2 + x^2)^2} \qquad \text{Quotient rule.}$$

Since $(k^2 + x^2) > 0$ and $R > 0$, the sign of $w'(x)$ is the same as the sign of the numerator: $(k^2 - x^2)$, meaning that $w'(x) > 0$ if $x < k$ and $w'(x) < 0$ if $x > k$. That is, $w(x)$ switches from increasing to decreasing at $x = k$. It follows that that $x = k$ is a local maximum. The other candidate extrema are at the endpoints of the interval. $w(x)$ is increasing away from $x = 0$, so this point is a local minimum. Moreover, $w(x)$ decreases as $x \to +\infty$, so the global maximum cannot be attained as $x \to +\infty$. It follows that $x = k$ is both a local maximum and a global maximum.

The optimal number of offspring under this model, $N = R/k$, and therefore depends both on the amount of resource available to the parent, R, and on the parameter k. •

Bio Info • The blood vessels of animals transport oxygen-rich blood around the body and return the de-oxygenated blood to the lungs. Blood vessels form vascular networks in which a large vessel carries blood from the heart, and then branches into two and then four and so on. Vascular networks contain on the order of ten billion vessels, and differ greatly from one species to another. However, all vascular networks have the same fundamental principles in common. If vessels are too narrow then the cost of transporting blood through the network is too high; on the other hand if the vessels are too wide then the organism must invest a lot of resource in building and maintaining the network.

In the next two examples, which are adapted from ***Sherman (1981)***, we will show that the tradeoff between the cost of transport and the cost of building the network leads to optimal radii for vessels.

EXAMPLE 4 A single blood vessel has a radius r and carries a total blood flow f (f gives the volume of blood that passes through the vessel in a second). What is the optimal value for the vessel radius, r? For a blood vessel of length ℓ, the total cost of transporting blood in the vessel is given by a function:

$$T(r) = 0.071\frac{f^2\ell}{r^4}$$

To derive this formula, we would need to study the physics of how blood flows through the vessel. This particular formula is true only for smaller vessels. Both ℓ and r are measured in centimeters, and f is measured in milliliters/s. The cost of transport $T(r)$ decreases as r increases. So if the only cost that the organism cared about were associated with the cost of transport, then any increase in blood vessel radius would benefit the organism. However, the organism must pay another cost, often called the metabolic cost, to build the artery. This cost increases in proportion to the volume of the artery. Suppose that the cost of building 1 cm^3 of artery is b; then the cost of building an artery of radius r and length ℓ is:

$$M(r) = b\pi r^2\ell \qquad b \times \text{ volume of a cylinder of radius } r \text{ and length } \ell$$

The total cost of building and transporting blood through a blood vessel of radius r is:

$$C(r) = T(r) + M(r) = 0.071\frac{f^2\ell}{r^4} + b\pi r^2\ell, \quad r > 0$$

To recap, ℓ is the length of the vessel (in cm), r is its radius (also in cm), f is flow rate, measured in milliliters/s and b is the cost of building and maintaining 1 cm^3 of blood vessel. $C(r) \to \infty$ if $r \to 0$, since the cost of transport becomes extremely large for small vessels. Also $C(r) \to \infty$ if $r \to \infty$, since the metabolic cost becomes extremely large for large vessels. So we expect that there is a value of r that minimizes $C(r)$. What is this optimal radius?

Solution Since C is differentiable for all $r > 0$, any interior local extremum will need to be a point at which $C'(r) = 0$. Now:

$$C'(r) = -0.284\frac{f^2\ell}{r^5} + 2b\pi r\ell$$

so

$$C'(r) = 0 \implies 0.284 f^2 \ell = 2b\pi r^6 \ell$$
$$\implies r = \frac{(0.142)^{1/6} f^{1/3}}{\pi^{1/6} b^{1/6}} \approx 0.597 \frac{f^{1/3}}{b^{1/6}} \tag{5.13}$$

Now since $C(r) \to +\infty$ as $r \to 0$ we must have $C'(r) < 0$ for small r. Similarly, since $C(r) \to +\infty$ as $r \to +\infty$, we must have $C'(r) > 0$ for large r. Since it is a continuous function, $C'(r)$ can only change sign at points where $C'(r) = 0$. At the local extremum, given in (5.13), $C'(r)$ must change sign from negative to positive, meaning that $C(r)$ goes from decreasing to increasing, i.e., $r = \frac{0.597 f^{1/3}}{b^{1/6}}$ is a local minimum point. Since it is the only local minimum, and $C(r) \to \infty$ at both ends of the interval, it must furthermore be the global minimum point, or equivalently the radius that minimizes the total vessel cost.

There is no independent way of measuring the metabolic cost per volume, b, but our derivation above shows that there is a power law relationship between the blood vessel radius r and the flow f within the vessel. ***Zamir (1977)*** used data on the blood vessels in bat wings collected by ***Wiedeman (1963)*** to test the power law relationship predicted by (5.13). Zamir's data is shown in Figure 5.51, plotted on log-log axes to show the power law.

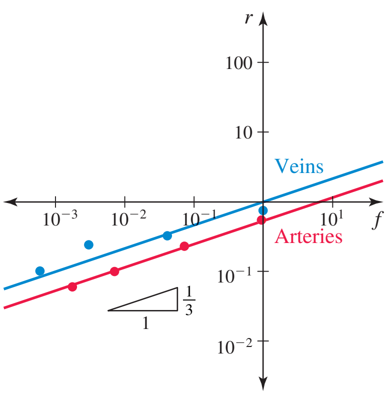


Figure 5.51 ***Zamir (1977)***'s data on the correlation between blood vessel radius and flow rate. We plot the data on log-log axes to show the power law dependence. Both arteries and veins follow the power law $r \propto f^{1/3}$.

We also notice that the length of the vessel does not show up in the power law for the vessel radius. It makes sense that the optimal radius should not depend on the length of the vessel, because both metabolic cost and transport cost increase proportional to the length of a vessel. The relative sizes of the metabolic cost and transport costs are not affected by changing ℓ.

Although Example 4 shows that there is a scaling relationship between the radius of each vessel and the flux of blood within the vessel, directly measuring blood flow in each vessel of a living animal is very difficult. However, it is easier to study how larger vessels branch into smaller vessels. As the next example shows, the coefficients f and b, will disappear from this calculation, allowing vessel radius alone to be used to show that the vascular network minimizes total transport costs.

EXAMPLE 5

A blood vessel has radius R and carries a total blood flow, F (F measures the volume of blood passing through the vessel in one second. It is usually measured in milliliters/s). The blood vessel branches into two daughter vessels whose radii are r_1 and r_2. For the network to minimize transport and metabolic costs, how should r_1 and r_2 be related to R?

Solution

We draw a diagram of the branching vessels in Figure 5.52. We will call the vessel that splits in two the parent vessel. Notice that the problem statement does not include any of the lengths either of the parent vessel or of its daughters because these lengths do not affect the optimal radii of any of these vessels. However, the blood flow rates in the two daughter vessels do affect their optimal radii, so we identify the flow rates in the two daughter vessels as f_1 and f_2 respectively.

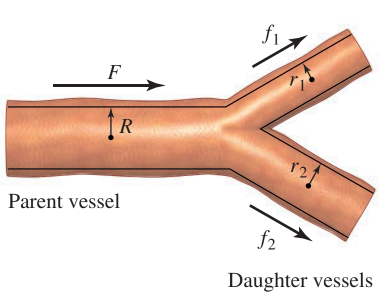


Figure 5.52 A single vessel of radius R, which we call the parent, branches into two daughter vessels of radius r_1 and r_2 respectively. We denote the flow rate of blood in the parent vessel by F, and in the two daughter vessels by f_1 and f_2 respectively.

We use Equation (5.13) from Example 4 to calculate the optimal radii for the parent and daughter vessels:

$$R = 0.597 \frac{F^{1/3}}{b^{1/6}} \quad \text{and} \quad r_i = 0.597 \frac{f_i^{1/3}}{b^{1/6}}$$

where the second equation can be read as a formula for r_1 in terms of f_1 or for r_2 in terms of f_2. However, we do not know what the flow rates of the two vessels are. Calculating these flow rates would require a physical model that included the pressures within the different vessels, the branching angle, and so on. But we do know that any blood that enters through the parent vessel must leave through one or the other of the two daughter vessels. In other words, the sum of the flow in the two daughter vessels must be equal to the flow in the parent

$$F = f_1 + f_2 \tag{5.14}$$

Now, we can rewrite our equation for R in terms of F as an equation for F in terms of R:

$$F = \frac{b^{1/2}R^3}{0.597} \quad \text{and} \quad f_i = \frac{b^{1/2}r_i^3}{0.597}$$

Using these formulas to substitute for F, f_1, and f_2 in Equation (5.14) we obtain:

$$\frac{b^{1/2}R^3}{0.597} = \frac{b^{1/2}}{0.597}\left(r_1^3 + r_2^3\right) \implies R^3 = r_1^3 + r_2^3 \quad \text{Cancel common factors} \tag{5.15}$$

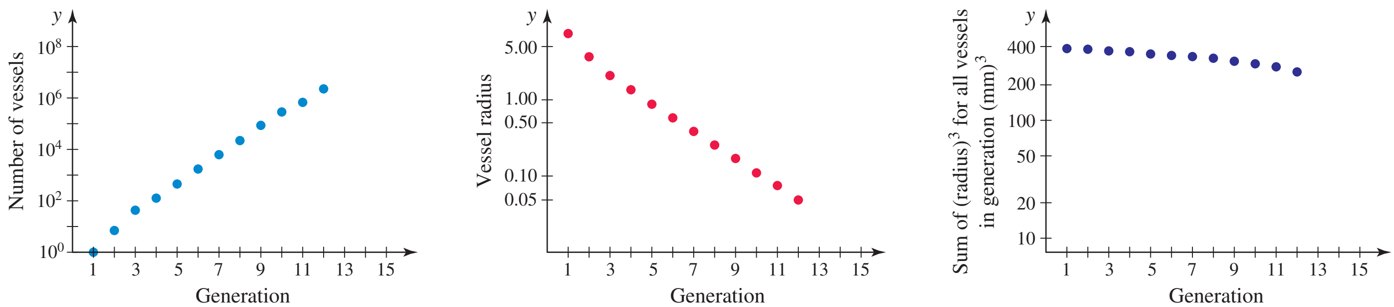


Figure 5.53 Blood vessels in the human lung are divided into generations, according to how many times the largest vessel branches to reach each vessel. (1) The number of vessels increases from each generation to the next. (2) The vessel radius decreases from each generation to the next. (3) The sum of the radii raised to the third power in each generation is approximately constant, consistent with Murray's law.

So we can relate the radii of the two daughter vessels to the radius of the parent vessel without needing to know b, or any of the flow rates, or how the flow of blood is split between the daughter vessels. Equation (5.16) is known as **Murray's law**. Murray's law has been carefully investigated in many different animals. Many, but not all blood vessel networks obey Murray's law. We show some measurements of the radii of blood vessels within human lungs in Figure 5.53. This data is taken from ***Huang et al. (1996)*** Since the blood vessel network contains billions of branching arteries all with different radii, one way to test Murray's law is to divide all of the vessels into generations. If the largest vessel entering the lung is the first generation, then all daughters that branch off it will form the second generation and so on. Murray's law says that if the radii of each vessel in the same generation are cubed and the cubes are summed, we should get the same answer for each generation. As Figure 5.53 shows, when this calculation is performed the vessels do obey this law. •

EXAMPLE 6

Bio Info • One of the most important roles for mathematics in biology is the theory of population genetics; that is, building models to understand the frequency of different genes within a population. One problem of particular interest is how harmful mutations can persist in a population. One such harmful mutation is sickle cell anemia. Sickle cell anemia is caused by a change in a single gene. The gene comes in two forms, s (non-mutant) and S (mutant). The s gene makes part of the hemoglobin molecule that enables red blood cells to carry oxygen. The mutant S gene makes a variant form of the protein that interferes with the function of hemoglobin. Red blood cells containing the mutant form of hemoglobin have a curved (sickle-like shape) that gives the disease its name. They are removed from circulation at a much higher rate than the non-mutant red blood cells, leading to anemia (low red blood cell count), as well as pain and clotting problems. However, humans are diploid, which means that we have two copies of each gene. Their **genotype** is a list of which genes they have. There are three possible sickle cell genotypes. An individual may have two copies of the non-mutant gene, we say these individuals have ss genotype. Or they may have two copies of the mutant gene (we say these individuals have the SS genotype). Or they may have one copy of the mutant gene and one copy of the non-mutant gene (we say these individuals have Ss genotype). Individuals with two copies of the mutant gene suffer from severe sickle cell anemia. However, Ss genotype individuals have only a mild form of sickle cell anemia. Additionally the changes to the red blood cells seem to protect them from malaria. Because of this protection, in areas with high incidence of malaria, individuals with the *Ss* genotype tend to be healthier than individuals with the *ss* genotype. The goal of this example is to try to predict the number of individuals of each genotype within the population. The derivation of the model for gene frequency uses some ideas about probability. We will cover these ideas at more length in Chapter 12. You should not worry if the details of the derivation are a little unclear right now.

To find the frequency of the different genotypes, we need a model for how genes are transmitted from one generation to the next. One commonly used model is to assume that genotypes are created at random from the pool of all available genes. So if the fraction of all genes that are of type S is p, and the fraction that are of type s is $1 - p$, then the probability that a randomly chosen individual has genotype SS is p^2 (since their first copy of the gene must be S-type, with probability p, and then the second copy of the gene must also be S type, also with probability p, so $p \times p$ probability overall). Similarly, the probability that an individual has genotype ss is $(1 - p)^2$. The probability that an individual has genotype Ss is $2p(1 - p)$ because either the first copy of the gene is S and the second gene copy is s with probability $p \times (1 - p)$, or the first copy is s and the second copy is S, with probability $(1 - p) \times p$. Summing the two probabilities gives $2p(1 - p)$.

The frequency of each genotype is determined by the fitness of individuals carrying the genotype. The fitness of an individual is usually defined to be the number of offspring they have – the main factor in this is the likelihood that they survive to adulthood (and that they are not killed or disabled by childhood diseases). We will define the fitnesses of SS-type, Ss-type, and ss-type individuals to be w_{11}, w_{12}, and w_{22} respectively. The average fitness for the population is then given by the formula:

$$\begin{aligned}\text{average fitness} = {} & \text{probability randomly chosen individual is } SS\text{-type} \\ & \qquad \times \text{ fitness of } SS\text{-type} \\ & + \text{probability randomly chosen individual is } Ss\text{-type} \\ & \qquad \times \text{ fitness of } Ss\text{-type} \\ & + \text{probability randomly chosen individual is } ss\text{-type} \\ & \qquad \times \text{ fitness of } ss\text{-type}\end{aligned}$$

or, if we use the formulas derived above to write out each term in the above word equation, the average fitness of an individual from a population in which a proportion p of genes are sickle cell mutants is:

$$w(p) = p^2 w_{11} + 2p(1 - p)w_{12} + (1 - p)^2 w_{22} \tag{5.16}$$

Since p represents the proportion of genes that are S-type, the domain of this function is $0 \le p \le 1$.

One theory for how proportions of genes evolve with time suggests that p will evolve to maximize the function $w(p)$. For what value of p is this maximum attained?

Solution The answer depends on the relative values of w_{11}, w_{12}, and w_{22}. In an area with many malaria infections, individuals with the Ss-genotype tend to be fitter than individuals with either the SS-genotype or the ss-genotype, because the sickle cell gene protects them from malaria. Therefore, in Equation (5.16), $w_{12} > w_{11}$ and $w_{12} > w_{22}$. In the language of population genetics, this kind of trait (in which the healthiest individuals have two different copies of the gene) is known as an **over-dominant** trait. To find local extrema we calculate $w'(p)$:

$$w'(p) = 2pw_{11} + 2(1 - 2p)w_{12} - 2(1 - p)w_{22} = 2\left(p(w_{11} + w_{22} - 2w_{12}) + w_{12} - w_{22}\right)$$

Points at which $w'(p) = 0$ are candidates for local extrema. Are there any of these points in the domain $0 \le p \le 1$? We know that $w(p)$ is a quadratic polynomial in p, and $w'(p)$ is a linear function of p. So $w'(p) = 0$ has exactly one root. Although we can calculate this root from the formula for $w'(p)$ it is easier for checking whether it lies in [0, 1] to evaluate $w'(p)$ at the two ends of the interval:

$$w'(0) = 2(w_{12} - w_{22}) > 0$$
$$w'(1) = 2(w_{11} - w_{12}) < 0$$

so $w(p)$ is increasing at $p = 0$ and decreasing at $p = 1$. So there must be a local maximum somewhere between the two points where the function goes from increasing to decreasing. This point will also be the global maximum for the interval.

We can solve $w'(p) = 0$ to compute the local maximum

$$p = \frac{w_{22} - w_{12}}{w_{11} + w_{22} - 2w_{12}} = \frac{w_{12} - w_{22}}{2w_{12} - w_{11} - w_{22}}$$

Since $w_{12} > w_{11}$ and $w_{12} > w_{22}$, denominator and numerator are positive.

According to population genetics theory, this represents the proportion of genes that are of S-type. It is difficult to measure w_{11}, w_{12}, and w_{22} directly but, according to ***Durrett (2008)***, in West Africa, where malaria remains common, $w_{11} = 0.22w_{22}$, and $w_{12} = 1.11w_{22}$, so the model predicts that:

$$p = \frac{(1.11 - 1)w_{22}}{(2.22 - 0.22 - 1)w_{22}} = 0.11$$

Equivalently, a proportion $p^2 = 0.01$ of individuals have sickle cell anemia and a proportion $2p(1 - p) = 0.20$ carry both sickle cell and non-sickle cell genes. These numbers agree well with the measured frequencies of the gene in West Africa.

The power of this theory is that it explains why a gene that causes the people who carry it a lot of suffering is not removed by natural selection. Although SS-type individuals are very seriously ill, Ss-type individuals are typically healthier than individuals who have no copies of the sickle cell gene, and this fitness advantage keeps the gene in the population at a low level. •

Section 5.4 Problems

Problems 1–22 are not inspired by biology, but will enable you to practice setting up optimization problems and solving them.

1. Find the smallest perimeter possible for a rectangle whose area is 25 in.^{2}.

2. Show that, among all rectangles with a given perimeter, the square has the largest area.

3. A rectangle has its base on the x-axis and its upper two vertices on the parabola $y = 3 - x^2$, as shown in Figure 5.54. What is the largest area the rectangle can have?

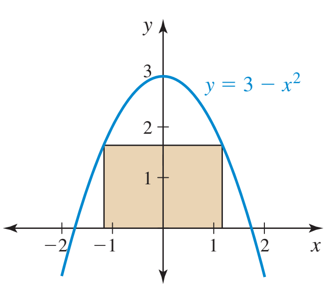


Figure 5.54 The graph of $y = 3 - x^2$ together with the inscribed rectangle in Problem 3.

4. A rectangular field is bounded on one side by a river and on the other three sides by a fence. Find the dimensions of the field that will maximize the enclosed area if the fence has a total length of 320 ft.

5. Find the largest possible area of a right triangle whose hypotenuse is 4 cm long.

6. Suppose that a and b are the side lengths in a right triangle whose hypotenuse is 5 cm long. What is the largest perimeter possible?

7. Suppose that a and b are the side lengths in a right triangle whose hypotenuse is 10 cm long. Show that the area of the triangle is largest when $a = b$.

8. A rectangle has its base on the x-axis, its lower left corner at (0, 0), and its upper right corner on the curve $y = 1/x$. What is the smallest perimeter the rectangle can have?

9. Denote by (x, y) a point on the straight line $y = 4 - 3x$. (See Figure 5.55.)

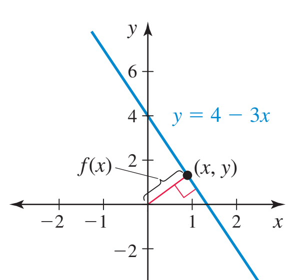


Figure 5.55 The graph of $y = 4 - 3x$ in Problem 9.

(a) Show that the distance from (x, y) to the origin is given by

$$f(x) = \sqrt{x^2 + (4 - 3x)^2}.$$

(b) Give the coordinates of the point on the line $y = 4 - 3x$ that is closest to the origin. (*Hint*: Find x so that the distance you computed in (a) is minimized.)

(c) Show that the *square* of the distance between the point (x, y) on the line and the origin is given by

$$g(x) = [f(x)]^2 = x^2 + (4 - 3x)^2$$

and find the minimum of $g(x)$. Show that this minimum agrees with your answer in (b).

10. How close does the line $y = 1 + 2x$ come to the origin?

11. How close does the curve $y = 1/x$ come to the origin? (*Hint*: Find the point on the curve that minimizes the *square* of the distance between the origin and the point on the curve. If you use the square of the distance instead of the distance, you avoid dealing with square roots.)

12. How close does the curve $y = 1/x^2$ come to the origin? (*Hint*: Find the point on the curve that minimizes the *square* of the distance between the origin and the point on the curve. If you use the square of the distance instead of the distance, you avoid dealing with square roots.)

13. Show that if $f(x)$ is a positive twice-differentiable function that has a local minimum at $x = c$, then $g(x) = [f(x)]^2$ has a local minimum at $x = c$ as well.

14. Show that if $f(x)$ is a differentiable function with $f(x) < 0$ for all $x \in \mathbf{R}$ and with a local maximum at $x = c$, then $g(x) = [f(x)]^2$ has a local minimum at $x = c$.

15. Show that if $f(x)$ is a differentiable function for all $x \in \mathbf{R}$ and with a local minimum at $x = c$, then $g(x) = \exp(-f(x))$ has a local maximum at $x = c$.

16. Optimizing Crop Yield A farmer is trying to optimize the amount of nitrogen fertilizer to add to a field. She finds that her yield per square meter increases with the amount, N, of nitrogen added according to the formula:

$$Y(N) = \frac{1}{e^{-N} + 1} \quad N \geq 0$$

(a) Show that $Y(N)$ increases monotonically with N. If the cost of fertilizer is not important, this result suggests that she should add as much fertilizer as she can to the field.

(b) Show either by using calculus, or by making a plot on a graphing calculator, that $Y(N)$ is concave downward, so there are diminishing returns from using more fertilizer.

(c) Suppose that the farmer includes the cost of fertilizer when determining the optimal amount to use. If the cost of one unit of fertilizer is C, then her return, N, becomes:

$$r(N) = Y(N) - CN = \frac{1}{e^{-N} + 1} - CN \quad N \geq 0$$

You may assume that C is a positive constant. Explain in words why, whatever the value of C is, we would expect there to be an optimal value of N that maximizes the return $r(N)$.

(d) Calculate the optimal value of N if $C = 1/8$ (*Hint*: To find when $r'(N) = 0$, make the substitution $u = e^{-N}$ and solve for u. You will need to use the quadratic formula.)

(e) If $C = 1$ show that the optimal amount of fertilizer for the farmer to add is $N = 0$.

17. Ticket Price Optimization Dalmatian Airlines flies a daily flight from Los Angeles to Minneapolis. Currently they sell each ticket for \$300, and on average 100 people take the flight, so their revenue per flight is 100 tickets × \$300/ticket = \$30,000. They are interested in seeing whether they can increase their revenue by changing the price of a ticket. Based on market research they discover that for every \$1 increase in ticket price, one fewer person will buy a ticket. Similarly for every \$1 decrease in ticket price, one more person will buy a ticket.

(a) What ticket price would maximize Dalmatian Airlines' revenue? (*Hint*: Denote the number of extra people flying on the route due to a price change by x, and the cost of a ticket by \$300-x. Then explain why the revenue to be maximized is $R(x) = (300 - x)(100 + x)$. You should also explain what the domain of this function is.

(b) The plane can seat a maximum of 150 people. How does this information change the domain of $R(x)$? What is the new optimal ticket price?

18. Ticket Price Optimization Dalmatian Airlines also flies a daily flight from Los Angeles to Sacramento. Currently they sell each ticket for \$100, and on average 200 people take the flight, so their revenue per flight is 200 tickets × \$100/ticket = \$20,000. They are interested in seeing whether they can increase their revenue by changing the price of a ticket. Based on market research they discover that for every \$2 increase in ticket price, one fewer person will buy a ticket. Similarly for every \$2 decrease in ticket price, one more person will buy a ticket.

(a) What ticket price would maximize Dalmatian Airlines' revenue? (*Hint:* Denote the number of extra people flying on the route due to a price change by x, and the cost of a ticket by \$100−2x. Then explain why the revenue to be maximized is $R(x) = (100 - 2x)(200 + x)$. You should also explain what the domain of this function is.)

(b) The plane can seat a maximum of 250 people. How does this information change the domain of $R(x)$? Does this constraint affect your answer to part (a)?

19. Optimal Soda Can A soda can manufacturer wants to minimize the cost of the aluminum used to make their can. The can has to hold a volume V of soda. Assuming that the thickness of the can is the same everywhere, the amount of aluminum used to make the can will be proportional to its surface area. That is, suppose the height of the can is h and the radius of the can is r, as in Figure 5.56. Then the manufacturer wants to minimize:

$$S = 2\pi rh + 2\pi r^2 \tag{5.17}$$

subject to the constraint that $\pi r^2 h = V$. Here we have used the formulas for the total surface area and volume of a cylinder.

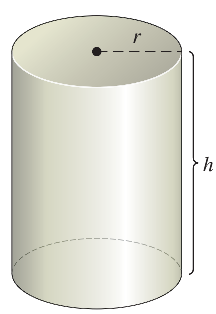


Figure 5.56 A soda can is a circular cylinder with radius r and height h. The curved surface area is $2\pi rh$ and the area of each end cap is πr^2.

(a) A real soda can has volume $V = 355\text{ cm}^3$ (or 12 fl. oz.). By substituting for h in Equation (5.17), write S as a function of r only.

(b) Describe the behavior of $S(r)$ as $r \to \infty$

(c) Describe the behavior of $S(r)$ as $r \to 0$.

(d) Based on your answers to (b) and (c), explain why you expect there to be a value of r that minimizes $S(r)$. Calculate this optimum radius r.

20. Redo Problem 19, but with the additional information that the two ends of the can have twice the thickness as the walls of the can. *Hint*: Explain why minimizing the amount of material needed to make the can is equivalent to minimizing:

$$S = 2\pi rh + 4\pi r^2$$

21. A circular sector with radius r and angle θ has area A. Find r and θ so that the perimeter is smallest when (a) $A = 2$ and (b) $A = 10$. (*Note*: $A = \frac{1}{2}r^2\theta$, and the length of the arc $s = r\theta$, when θ is measured in radians; see Figure 5.57.)

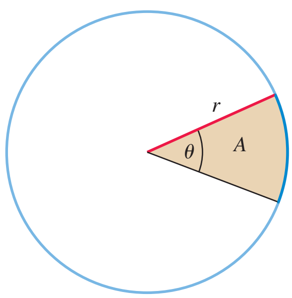


Figure 5.57 The circular sector in Problems 21 and 22.

22. A circular sector with radius r and angle θ has area A. Find r and θ so that the perimeter is smallest for a given area A. (*Note*: $A = \frac{1}{2}r^2\theta$, and the length of the arc $s = r\theta$, when θ is measured in radians; see Figure 5.57.)

23. Molecular Dynamics One popular model for the interactions between two molecules is the Leonard-Jones 6-3 potential. According to this model, the energy of interaction between two molecules that are distance r apart is given by a function:

$$V(r) = \frac{1}{r^6} - \frac{A}{r^3} \quad r > 0$$

Molecules will attract or repel each other until they reach a distance that minimizes the function $V(r)$. The coefficient A is a positive constant.

(a) What is the behavior of $V(r)$ as $r \to 0$? What is the behavior of $V(r)$ as $r \to +\infty$?

(b) Explain why you expect there to be a value of r that minimizes $V(r)$, and then calculate that value of r (it may help for your argument to determine the sign of $V'(r)$ for large r).

(c) Would you still expect there to be a spacing that minimizes $V(r)$ if A were a negative number? Justify your answer.

24. Fish Schooling One model that is used for the interactions between animals, including fish in a school, is that the fish have an energy of interaction that is given by a Morse potential:

$$V(r) = e^{-r} - Ae^{-ar} \quad r > 0$$

The fish will attract or repel each other until they reach a distance that minimizes the function $V(r)$. The coefficients A and a are positive numbers.

(a) Assume initially that $a = 1/2$ and $A = 1$, what is the behavior of $V(r)$ as $r \to 0$. What is the behavior of $V(r)$ as $r \to +\infty$?

(b) Find the value of r that minimizes $V(r)$.

(c) Explain what happens to the spacing that minimizes the energy of interaction if $a = 1/2$ and $A = 4$?

Population Genetics ***In Example 6 we discussed a model for the frequency of a particular gene. We imagined that the gene comes in two types S and s, and argued that the proportion, p, of copies of the gene that are of S-type will be the value of p that maximizes the function w(p) in Equation (5.16). We reproduce the equation for that function here:***

$$w(p) = p^2 w_{11} + 2p(1-p)w_{12} + (1-p)^2 w_{22}$$

w_{11}, w_{12}, and w_{22} are the fitnesses (mean number of offspring) for individuals with SS-genotype, Ss genotype, and ss genotype, respectively. In Example 6 we showed that if $w_{12} > w_{11}$ and $w_{12} > w_{22}$, the optimal value of p is somewhere in the interval $(0, 1)$. In Problems 25 and 26, you will explore other possibilities for the location of this optimum, depending on the relative sizes of w_{11}, w_{12}, and w_{22}.

25. Directional selection occurs if the fitness of individuals with two different genes is somewhere between the fitness of individuals who have identical copies of the gene.

(a) Suppose that: $w_{11} > w_{12} > w_{22}$, that is individuals with two copies of the gene are fitter than individuals with one copy, and both are fitter than individuals with no copies. By optimizing $w(p)$, show that the theory predicts that the frequency of this gene will be $p = 1$. In this case, the S-gene is said to be dominant over the s-gene.

(b) If, instead, $w_{11} < w_{12} < w_{22}$, show by optimizing $w(p)$, show that the theory predicts that the frequency of this gene will be $p = 0$. In this case, the s-gene is said to be dominant over the S-gene.

(c) Interpret in words your answers from (a) and (b).

26. The trait is said to be **under-dominant** if the fitness of individuals with two different genes is less than the fitness of individuals who have identical copies of the gene. Show that if $w_{12} < w_{11}$ and $w_{12} < w_{22}$, then the value of p that maximizes $w(p)$ is either $p = 0$ or $p = 1$ (that is, there is no local maximum for $w(p)$ with $p \in (0, 1)$).

27. In Example 3, we showed how tradeoffs between having too few and too many offspring led to the existence of an optimal number of offspring. Our calculation was based on maximizing the fitness $w(x)$ of the organism, where x is the amount of resource that it invests in each of its offspring. We showed that the fitness could be written as a function:

$$w(x) = \frac{R}{x} f(x) \quad x > 0$$

where R is the total resource that the organism could split between all of its offspring. In Example 3, we assumed a specific form for the function $f(x)$. In this question you will explore the conditions under which an optimal number of offspring might exist, making as few assumptions as possible about the function $f(x)$.

(a) Assume that $f(x)$ is twice differentiable. Show that $w'(x) = 0$ if and only if $f'(x) = \frac{f(x)}{x}$.

(b) The condition from part (a) can be interpreted geometrically. If $\hat{x}$ is the solution of $w'(x) = 0$, then $f'(\hat{x}) = \frac{f(\hat{x})}{\hat{x}}$. Based on this equation, draw a sketch that shows how the secant line between the origin and $\hat{x}$ is related to the slope of the curve $y = f(x)$ at $x = \hat{x}$.

(c) To determine whether $\hat{x}$ is a local maximum or minimum, we use the second-derivative test. Show that:

$$\frac{d^2w}{dx^2} = -\frac{R}{x^2}\left(f'(x) - \frac{f(x)}{x}\right) + \frac{R}{x}\left(f''(x) - \frac{d}{dx}\frac{f(x)}{x}\right) \quad (5.18)$$

(d) Show that when $x = \hat{x}$, the right-hand side of Equation (5.18) can be simplified so that:

$$\left.\frac{d^2w}{dx^2}\right|_{x=\hat{x}} = R\frac{f''(\hat{x})}{\hat{x}}$$

So for any function $f(x)$ that is concave down at $\hat{x}$, $w''(\hat{x}) < 0$, which means that $w(x)$ has a local maximum at $\hat{x}$.

28. Murray's Law for Plants This problem is based on ***McCulloh et al. (2003)***. The plant xylem is a transport network within plants that forms a network like the blood vessels of animals. The xylem transports water from the roots, up the plant stem to its leaves. Unlike blood vessels, in some plants the xylem vessels are not single tubes, but are made up of bundles of smaller tubes. Larger xylem vessels contain more tubes, smaller vessels contain fewer tubes. Because vessels are made of smaller tubes, the way that transport costs depend on vessel radius is different for the xylem than for the blood vessels of an animal. Specifically, it can be shown that the cost of transporting water at a flow rate f (measured in milliliters/s) in a xylem vessel of radius r and length ℓ (both measured in cm) is given by the function

$$T(r) = 0.071\frac{f^2\ell}{r_T^2 r^2}$$

where r_T is the radius of one of the tubes within the xylem vessel (you may assume that $r_T = 5 \times 10^{-2}$ cm).

(a) Assume that the cost of building the xylem vessel is still proportional to its volume:

$$M(r) = b\pi r^2\ell$$

where b is the metabolic cost of building and maintaining 1 cm^3 of xylem vessel. If the plant controls xylem vessel radius to minimize the total cost $T(r) + M(r)$, derive a formula relating xylem radius r to flow rate f. Your formula will include b as an unknown coefficient.

(b) If a xylem vessel of radius R branches into two smaller vessels of radii r_1 and r_2, and all vessels minimize the total cost of transport and maintenance, show that the xylem vessel radii are related by Murray's law for plants:

$$R^2 = r_1^2 + r_2^2$$

29. Evaluating Different Maintenance Cost Functions in Murray's Law When we derived Murray's law in Examples 4 and 5, we assumed that the cost of building and maintaining a blood vessel is proportional to the volume of the vessel (that is, we defined a quantity b, to be the cost of building and maintaining 1 cm^3 of blood vessel). But this is not the only possibility. Another possibility is that the cost of building a vessel will be proportional to its surface area, and not to its volume. The idea here is that a wall needs to be built around the vessel, and the cost of building the vessel wall is much higher than the cost of filling it with blood. Under this assumption, the cost of building and maintaining a vessel of radius r and length ℓ is:

$$M(r) = 2\pi c r\ell$$

where c is the cost of maintaining a 1 cm^2 area of vessel wall. (*Hint*: $2\pi r\ell$ is the surface area, not including end-caps for a cylinder of length ℓ and radius r.)

(a) Using the same function for the cost of transport within the vessel as was given in Example 4:

$$T(r) = 0.071\frac{f^2\ell}{r^4}$$

and assuming that the organism controls vessel radius to minimize the total cost of transport and maintenance $T(r) + M(r)$, derive a formula relating blood vessel radius r to flow rate f. Your formula will include c as an unknown coefficient.

(b) If a blood vessel of radius R branches into two smaller vessels of radii r_1 and r_2, and all vessels minimize the total cost of transport and maintenance, show that the blood vessel radii are related by a modified form of Murray's law:

$$R^{5/2} = r_1^{5/2} + r_2^{5/2}$$

5.5 L'Hôpital's Rule

L'Hôpital's rule gives us a way to calculate the limits of fractions in which both the numerator and the denominator tend to 0. (l'Hôpital is pronounced lop-it-al.) The rule also works when both the numerator and the denominator tend to infinity. Sometimes we can use algebraic or trigonometric manipulations to find the limit—for instance:

$$\lim_{x\to 3}\frac{x^2-9}{x-3} = \lim_{x\to 3}(x+3) = 6 \qquad \text{Factorize } x^2-9=(x+3)(x-3) \text{ and cancel } (x-3)$$

$$\lim_{x\to\infty}\frac{x}{1+x} = \lim_{x\to\infty}\frac{1}{\frac{1}{x}+1} = 1 \qquad \text{Divide numerator and denominator by } x$$

Using algebraic manipulations, however, is not always possible, as in the following example:

$$\lim_{x\to 0}\frac{e^x-1}{x} \tag{5.19}$$

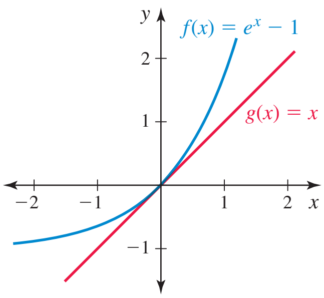


Figure 5.58 Both the numerator and denominator in Equation (5.19) go to 0 as $x \to 0$.

Here, both numerator and denominator tend to 0 as $x \to 0$ (see Figure 5.58.). There is no way of algebraically simplifying the ratio. Instead, we linearize both numerator and denominator. Recall from Section 4.11 that the linear approximation of a function $f(x)$ at $x = a$ is defined as

$$f(x) \approx f(a) + f'(a)(x-a)$$

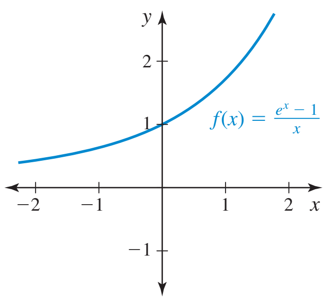


Figure 5.59 Plotting $f(x) = \frac{e^x-1}{x}$ shows that the function tends to 1 as $x \to 0$.

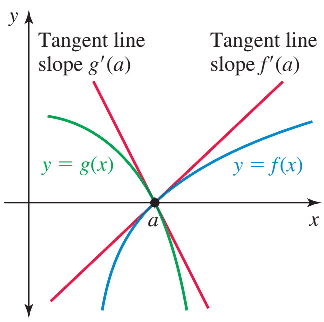


Figure 5.60 Illustration of l'Hôpital's rule applied to two functions that vanish at $x = a$.

If $f(x) = e^x - 1$, then $f'(x) = e^x$, and so $f(0) = 0$ and $f'(0) = 1$. That is, the linear approximation of the numerator at $x = 0$ is

$$f(x) \approx 0 + (1)(x - 0) = x$$

The denominator is already a linear function, namely, $g(x) = x$. So when we linearize both numerator and denominator functions we get $\frac{e^x-1}{x} \approx \frac{x}{x} = 1$. So based on the linearization, we predict that the limiting value of $\frac{e^x-1}{x}$ as $x \to 0$ is 1. This value is supported by a graph of the function (Figure 5.59).

To see clearly what we have just done, we look at the general case

$$\lim_{x\to a} \frac{f(x)}{g(x)}$$

and assume that both

$$\lim_{x\to a} f(x) = 0 \qquad \text{and} \qquad \lim_{x\to a} g(x) = 0$$

(See Figure 5.60.) Using a linear approximation as before, we find that, for x close to a,

$$\frac{f(x)}{g(x)} \approx \frac{f(a) + f'(a)(x-a)}{g(a) + g'(a)(x-a)}$$

Since $f(a) = g(a) = 0$ and $x \neq a$, the right-hand side is equal to

$$\frac{f'(a)(x-a)}{g'(a)(x-a)} = \frac{f'(a)}{g'(a)}$$

provided that $\frac{f'(a)}{g'(a)}$ is defined. This calculation is one case of l'Hôpital's rule, which allows the limit of $\frac{f(x)}{g(x)}$ to be calculated from $\frac{f'(x)}{g'(x)}$. We will only state the rule here; its proof is beyond the scope of this book.

> **L'Hôpital's Rule** Suppose that f and g are differentiable functions and that
>
> $$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$$
>
> or
>
> $$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \infty$$
>
> Then
>
> $$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}$$
>
> provided this limit exists.

L'Hôpital's rule works for $a = +\infty$ or $-\infty$ as well, and it also applies to one-sided limits. Using l'Hôpital's rule, we can redo the three examples we just presented. In each case, the ratio $\frac{f(a)}{g(a)}$ is an **indeterminate expression**: either it is $\frac{0}{0}$ or $\frac{\infty}{\infty}$. We apply l'Hôpital's rule in each case: We differentiate both numerator and denominator and then take the limit of $\frac{f'(x)}{g'(x)}$ as $x \to a$. L'Hôpital's rule then gives

$$\lim_{x\to 3} \frac{x^2 - 9}{x - 3} = \lim_{x\to 3} \frac{2x}{1} = 6$$

$$\lim_{x\to\infty} \frac{x}{1 + x} = \lim_{x\to\infty} \frac{1}{1} = 1$$

$$\lim_{x\to 0} \frac{e^x - 1}{x} = \lim_{x\to 0} \frac{e^x}{1} = 1$$

In fact L'Hôpital's rule can be applied to calculate the true limits for seven kinds of indeterminate expressions:

$$\frac{0}{0}, \quad \frac{\infty}{\infty}, \quad 0 \cdot \infty, \quad \infty - \infty, \quad 0^0, \quad 1^\infty, \quad \text{or} \quad \infty^0.$$

We present examples for each kind of limit. Sometimes (see Example 3) l'Hôpital's rule needs to be applied more than once to calculate a particular limit.

EXAMPLE 1 **0/0** Evaluate $\lim_{x\to 2} \frac{x^6 - 64}{x^2 - 4}$.

Solution This limit is of the form $\frac{0}{0}$, since $2^6 - 64 = 0$ and $2^2 - 4 = 0$. Applying l'Hôpital's rule yields

$$\lim_{x\to 2} \frac{x^6 - 64}{x^2 - 4} = \lim_{x\to 2} \frac{6x^5}{2x} = \frac{6 \times 2^5}{2 \times 2} = (6)(2^3) = 48.$$

•

EXAMPLE 2 **0/0** Evaluate $\lim_{x\to 0} \frac{1 - \cos^2 x}{\sin x}$.

Solution This limit is of the form $\frac{0}{0}$, since $1 - \cos^2 0 = 0$ and $\sin 0 = 0$. Applying l'Hôpital's rule yields

$$\lim_{x\to 0} \frac{1 - \cos^2 x}{\sin x} = \lim_{x\to 0} \frac{2\cos x \sin x}{\cos x} = \lim_{x\to 0} (2\sin x) = 0.$$

In fact this is another limit that can be calculated without l'Hôpital's rule, since $1 - \cos^2 x = \sin^2 x$, so $\frac{1-\cos^2 x}{\sin x} = \sin x$. •

EXAMPLE 3 **0/0** Evaluate $\lim_{x\to 0} \frac{1 - \cos x}{x^2}$.

Solution As this example shows we may need to apply l'Hôpital's rule more than once to find a particular limit. In this example, the limit is of the form $\frac{0}{0}$ since $1 - \cos 0 = 0$ and $0^2 = 0$. Applying l'Hôpital's rule yields $\lim_{x\to 0} \frac{1-\cos x}{x^2} = \lim_{x\to 0} \frac{\sin x}{2x}$ which is still of the form $\frac{0}{0}$. You may remember from Section 3.4 that $\lim_{x\to 0} \frac{\sin x}{x} = 1$. However, even if we do not remember this result we may rederive if by applying l'Hôpital's rule again. Differentiating the numerator and denominator one more time we obtain:

$$\lim_{x\to 0} \frac{1 - \cos x}{x^2} = \lim_{x\to 0} \frac{\sin x}{2x} = \lim_{x\to 0} \frac{\cos x}{2} = \frac{1}{2}.$$

•

EXAMPLE 4 **∞/∞** Evaluate $\lim_{x\to\infty} \frac{\ln x}{x}$.

Solution This limit is of the form $\frac{\infty}{\infty}$. We can again apply l'Hôpital's rule:

$$\lim_{x\to\infty} \frac{\ln x}{x} = \lim_{x\to\infty} \frac{\frac{1}{x}}{1} = \lim_{x\to\infty} \frac{1}{x} = 0$$

•

EXAMPLE 5 **∞/∞** Evaluate $\lim_{x\to\infty} \frac{x^3 - 3x + 1}{3x^3 - 2x^2}$.

Solution This limit that we previously encountered in Section 3.3 is of the form $\frac{\infty}{\infty}$ and can be calculated by applying l'Hôpital's rule:

$$\lim_{x\to\infty} \frac{x^3 - 3x + 1}{3x^3 - 2x^2} = \lim_{x\to\infty} \frac{3x^2 - 3}{9x^2 - 4x}$$

which is still of the form $\frac{\infty}{\infty}$ so we apply l'Hôpital's rule to the new ratio, to obtain:

$$\lim_{x\to\infty} \frac{3x^2 - 3}{9x^2 - 4x} = \lim_{x\to\infty} \frac{6x}{18x - 4}$$

Since this is still of the form $\frac{\infty}{\infty}$, we can apply l'Hôpital's rule yet again. We now find that

$$\lim_{x\to\infty} \frac{6x}{18x-4} = \lim_{x\to\infty} \frac{6}{18} = \frac{1}{3}$$

In fact we could have found the answer without using l'Hôpital's rule by following the rules in Section 3.3. Divide both numerator and denominator by x^3 to obtain:

$$\lim_{x\to\infty} \frac{x^3-3x+1}{3x^3-2x^2} = \lim_{x\to\infty} \frac{(1-\frac{3}{x^2}+\frac{1}{x^3})}{(3-\frac{2}{x})} = \frac{1}{3},$$

because terms $\frac{3}{x^2}$, $\frac{1}{x^3}$, etc., all $\to 0$. •

EXAMPLE 6 ∞/∞ Evaluate $\lim_{x\to\infty} \frac{e^x}{x}$.

Solution This limit is of the form $\frac{\infty}{\infty}$. The example will show that l'Hôpital's rule has no problems detecting when a limit is infinite.

$$\lim_{x\to\infty} \frac{e^x}{x} = \lim_{x\to\infty} \frac{e^x}{1} = \infty \qquad \text{(limit does not exist)}$$ •

$0\cdot\infty$ L'Hôpital's rule can sometimes be applied to limits of the form

$$\lim_{x\to a} f(x)g(x)$$

where

$$\lim_{x\to a} f(x) = 0 \qquad \text{and} \qquad \lim_{x\to a} g(x) = \infty.$$

To apply l'Hôpital's rule to these limits write them in one of the two forms:

$$\lim_{x\to a} f(x)g(x) = \lim_{x\to a} \frac{f(x)}{\frac{1}{g(x)}} = \lim_{x\to a} \frac{g(x)}{\frac{1}{f(x)}}.$$

In the first case, the ratio is $\frac{0}{0}$, while in the second case the ratio is $\frac{\infty}{\infty}$. Usually only one of the two methods for rewriting the expression actually makes the limit easier to evaluate, as the following example shows:

EXAMPLE 7 $0\cdot\infty$ Evaluate $\lim_{x\to 0^+} x\ln x$.

Solution This is a one-sided limit, but l'Hôpital's rule still applies. $\ln x$ is not defined for $x \le 0$ so we must approach $x = 0$ from the right. The limit is of the form $(0)(-\infty)$. We apply l'Hôpital's rule after rewriting it in the form $\frac{\infty}{\infty}$:

$$\lim_{x\to 0^+} x\ln x = \lim_{x\to 0^+} \frac{\ln x}{\frac{1}{x}} = \lim_{x\to 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}}$$

$$= \lim_{x\to 0^+} \frac{1}{x}\left(-\frac{x^2}{1}\right) = \lim_{x\to 0^+}(-x) = 0$$

If we had written the limit in the form

$$\lim_{x\to 0^+} \frac{x}{\frac{1}{\ln x}}$$

and then applied l'Hôpital's rule, we would have obtained

$$\lim_{x\to 0^+} \frac{x}{\frac{1}{\ln x}} = \lim_{x\to 0^+} \frac{1}{(-1)(\ln x)^{-2}\frac{1}{x}} = \lim_{x\to 0^+} \left[-x(\ln x)^2\right]$$

which is more complicated than the initial expression. In general you may need to try both ways of rewriting expressions of the form $0\cdot\infty$ to find an expression you can evaluate. •

EXAMPLE 8 **$0 \cdot \infty$** Evaluate $\lim_{x\to\infty} xe^{-x^2/2}$.

Solution This limit is of the form $0 \cdot \infty$. We have two choices. We can rewrite the limit as

$$\lim_{x\to\infty} xe^{-x^2/2} = \lim_{x\to\infty} \frac{x}{e^{x^2/2}} = \lim_{x\to\infty} \frac{1}{xe^{x^2/2}} = 0$$

applying l'Hôpital's rule at the second step. Or we could have written:

$$\lim_{x\to\infty} xe^{-x^2/2} = \lim_{x\to\infty} \frac{e^{-x^2/2}}{\frac{1}{x}} = \lim_{x\to\infty} \frac{-xe^{-x^2/2}}{-\frac{1}{x^2}}$$

In the second case applying l'Hôpital's rule produces a more complicated limit than the one we were initially given to calculate. •

$\infty - \infty$ Suppose we want to calculate:

$$\lim_{x\to a} (f(x) - g(x))$$

where $f(a) = \infty$ and $g(a) = \infty$, so that our limit looks like $\infty - \infty$. L'Hôpital's rule can sometimes be applied in this case if we factorize the expression either as:

$$\lim_{x\to a} f(x)\left(1 - \frac{g(x)}{f(x)}\right) \quad \text{or} \quad \lim_{x\to a} g(x)\left(\frac{f(x)}{g(x)} - 1\right) \tag{5.20}$$

If the limit is then of the form $\infty \cdot 0$ it can then be evaluated using l'Hôpital's rule. Just as in the previous case, you may need to try both ways of rewriting $f(x) - g(x)$ to find one to which l'Hôpital's rule may be more readily applied.

EXAMPLE 9 **$\infty - \infty$** Evaluate

$$\lim_{x\to\infty} (x - \sqrt{x^2 + x}).$$

Solution This limit is of the form $\infty - \infty$. We can write the limit as the difference between two functions $f(x) = x$, $g(x) = \sqrt{x^2 + x}$. Since $f(x)$ is the simpler function, we factor it out to get

$$\lim_{x\to\infty} (x - \sqrt{x^2 + x}) = \lim_{x\to\infty} \left[x\left(1 - \sqrt{1 + \frac{1}{x}}\right)\right]$$

This is of the form $\infty \cdot 0$. We can transform it to the form $\frac{0}{0}$ and then apply l'Hôpital's rule:

$$\begin{aligned}
\lim_{x\to\infty} x\left(1 - \sqrt{1 + \frac{1}{x}}\right) &= \lim_{x\to\infty} \frac{1 - \sqrt{1 + \frac{1}{x}}}{\frac{1}{x}} \\
&= \lim_{x\to\infty} \frac{-\frac{1}{2}(1 + \frac{1}{x})^{-1/2}(-\frac{1}{x^2})}{-\frac{1}{x^2}} \qquad \text{Apply chain rule with } u = 1 + \tfrac{1}{x} \\
&= \lim_{x\to\infty} \frac{-1}{2\sqrt{1 + \frac{1}{x}}} = -\frac{1}{2}
\end{aligned}$$

•

When trying to calculate limits of the form $\infty - \infty$ remember to check that the factorized expression (5.20) actually is of the form $0 \cdot \infty$ before applying l'Hôpital's rule. As the next example shows, it is not guaranteed that the factorized expression will be of this form.

EXAMPLE 10 $\infty - \infty$ Evaluate $\lim_{x\to+\infty}(x - \ln x)$.

Solution Our limit takes the form of the difference between the functions $f(x) = x$ and $g(x) = \ln x$. Since x is the simpler function we factorize it out to obtain:

$$\lim_{x\to+\infty}(x - \ln x) = \lim_{x\to+\infty} x\left(1 - \frac{\ln x}{x}\right)$$

The first factor has limit ∞. However, the limit of the second factor is actually 1, since:

$$\lim_{x\to+\infty}\left(1 - \frac{\ln x}{x}\right) = 1 - \lim_{x\to+\infty}\frac{\ln x}{x} = 1 - 0 = 1$$

$\lim_{x\to\infty}\frac{\ln x}{x} = 0$, from Example 4

So the product of the two factors has a limit of the form $\infty \cdot 1$. The limit is therefore ∞ or undefined.

0^0, 1^∞, ∞^0 Suppose we want to evaluate

$$\lim_{x\to a}[f(x)]^{g(x)}$$

when $(f(a), g(a)) = (0, 0), (1, \infty)$, or $(\infty, 0)$. The key to solving such limits is to rewrite them as exponentials:

$$\begin{aligned}\lim_{x\to a}[f(x)]^{g(x)} &= \lim_{x\to a}\exp\left\{\ln[f(x)]^{g(x)}\right\}\\ &= \lim_{x\to a}\exp[g(x)\cdot\ln f(x)]\\ &= \exp\left[\lim_{x\to a}(g(x)\cdot\ln f(x))\right]\end{aligned}$$

The last step, in which we interchanged lim and exp, uses the fact that the exponential function is continuous. Rewriting the limit in this way transforms

$$\begin{aligned}0^0 &\quad \text{into} \quad \exp[0\cdot(-\infty)]\\ \infty^0 &\quad \text{into} \quad \exp[0\cdot\infty]\\ 1^\infty &\quad \text{into} \quad \exp[\infty\cdot\ln 1] = \exp[\infty\cdot 0]\end{aligned}$$

Since we know how to deal with limits of the form $0\cdot\infty$, we are in good shape again.

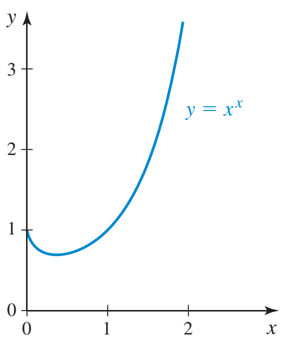


Figure 5.61 The function $y = x^x$ converges to 1 as $x \to 0$.

EXAMPLE 11 0^0 Evaluate $\lim_{x\to0^+} x^x$.

Solution This limit is of the form 0^0; we rewrite the limit first:

$$\begin{aligned}\lim_{x\to0^+} x^x = \lim_{x\to0^+}\exp[\ln x^x] &= \lim_{x\to0^+}\exp[x\ln x]\\ &= \exp\left[\lim_{x\to0^+}(x\ln x)\right]\end{aligned}$$

In Example 7 we showed $\lim_{x\to0^+}(x\ln x) = 0$. Hence,

$$\lim_{x\to0^+} x^x = \exp\left[\lim_{x\to0^+}(x\ln x)\right] = \exp[0] = 1$$

(See Figure 5.61.) •

EXAMPLE 12 1^∞ Evaluate $\lim_{x\to0}(1+x)^{1/x}$.

Solution Since $\lim_{x\to0}(1+x) = 1$ and $\lim_{x\to0}\frac{1}{x} = \infty$, this limit is of the form 1^∞. We rewrite it as

$$\lim_{x\to0}(1+x)^{1/x} = \lim_{x\to0}\exp\left(\frac{1}{x}\ln(1+x)\right) = \exp\left[\lim_{x\to0}\left(\frac{1}{x}\ln(1+x)\right)\right]$$

The exponentiated function is of the form $\infty \cdot 0$ (since $\ln 1 = 0$). We evaluate the limit of this function by writing it in the form $\frac{0}{0}$ and then applying l'Hôpital's rule:

$$\lim_{x\to 0}\left(\frac{\ln(1+x)}{x}\right) = \lim_{x\to 0}\left(\frac{\frac{1}{1+x}}{1}\right) = 1.$$

Therefore, $\lim_{x\to 0}(1+x)^{1/x} = \exp(1) = e$. •

Section 5.5 Problems

Find the limits in Problems 1–60; not all limits require use of l'Hôpital's rule.

1. $\lim_{x\to 5}\frac{x^2-25}{x-5}$

2. $\lim_{x\to 2}\frac{x-2}{x^2-4}$

3. $\lim_{x\to -2}\frac{3x^2+5x-2}{x+2}$

4. $\lim_{x\to -1}\frac{x+1}{x^2-2x-3}$

5. $\lim_{x\to 0}\frac{3-\sqrt{2x+9}}{2x}$

6. $\lim_{x\to 0}\frac{\sqrt{2x+4}-2}{x}$

7. $\lim_{x\to 0} x^2\ln x$

8. $\lim_{x\to\infty}\frac{\ln x}{x^2}$

9. $\lim_{x\to\infty} \ln x - \sqrt{x}$

10. $\lim_{x\to\infty}\sqrt{x}-\sqrt{x+3}$

11. $\lim_{x\to 0^+}\frac{\sqrt{x}}{\ln(x+1)}$

12. $\lim_{x\to\infty}\frac{\ln x}{\sqrt{x}}$

13. $\lim_{x\to\infty}\frac{\ln(\ln x)}{x}$

14. $\lim_{x\to\infty}\frac{\ln(\ln x)}{\ln x}$

15. $\lim_{x\to 0}\frac{5^x-1}{7^x-1}$

16. $\lim_{x\to 0}\frac{2^x-1}{3^x-1}$

17. $\lim_{x\to 0}\frac{3^{-x}-1}{2^x-1}$

18. $\lim_{x\to 0}\frac{2^{-x}-1}{5^x-1}$

19. $\lim_{x\to\infty}\frac{e^x-1-x}{e^x-x^2}$

20. $\lim_{x\to 0}\frac{e^x-1}{3e^{2x}-x}$

21. $\lim_{x\to\infty}\frac{(\ln x)^2}{x^2}$

22. $\lim_{x\to\infty}\frac{x^7}{e^x}$

23. $\lim_{x\to 0}\frac{\sin x}{x}$

24. $\lim_{x\to 0}\frac{\sin x}{x^2}$

25. $\lim_{x\to\infty} xe^{-x}$

26. $\lim_{x\to\infty} x^2e^{-x}$

27. $\lim_{x\to\infty} x^5e^{-x}$

28. $\lim_{x\to\infty} x^n e^{-x}$, $n\in\mathbf{N}$

29. $\lim_{x\to 0^+}\sqrt{x}\ln x$

30. $\lim_{x\to 0^+} x^\alpha\ln x$, $\alpha>0$ is any positive constant

31. $\lim_{x\to+\infty}\frac{\ln x}{x^{1/2}}$

32. $\lim_{x\to+\infty}\frac{\ln x}{x^\alpha}$, $\alpha>0$ is any positive constant

33. $\lim_{x\to+\infty}(e^x-x^3)$

34. $\lim_{x\to+\infty}(e^x-x^n)$, $n\in\mathbf{N}$

35. $\lim_{x\to\infty}\sqrt{x}\sin\frac{1}{x}$

36. $\lim_{x\to\infty} x^2\sin\frac{1}{x^2}$

37. $\lim_{x\to 0^+}(e^x-\frac{1}{x})$

38. $\lim_{x\to\infty}(x-\sqrt{x^2+1})$

39. $\lim_{x\to 0^+}\left(\frac{1}{\sin x}-\frac{1}{x}\right)$

40. $\lim_{x\to 0}\frac{\sqrt{x^2+x^4}-x}{x}$

41. $\lim_{x\to 0^+} x^{2x}$

42. $\lim_{x\to 0^+} x^{x^2}$

43. $\lim_{x\to\infty} x^{1/x}$

44. $\lim_{x\to\infty}(1+e^x)^{1/x}$

45. $\lim_{x\to\infty}\left(1+\frac{3}{x}\right)^x$

46. $\lim_{x\to\infty}\left(1+\frac{5}{x}\right)^x$

47. $\lim_{x\to\infty}\left(1-\frac{2}{x}\right)^x$

48. $\lim_{x\to\infty}\left(1+\frac{3}{x^2}\right)^x$

49. $\lim_{x\to\infty}\left(\frac{x}{1+x}\right)^x$

50. $\lim_{x\to\infty}(1+x)^{1/x}$

51. $\lim_{x\to 0} xe^x$

52. $\lim_{x\to 0^+}\frac{e^x}{x}$

53. $\lim_{x\to 1^-}(\ln(1-x)-\frac{1}{x-1})$

54. $\lim_{x\to(\pi/2)^-}\frac{\sin x}{\cos x}$

55. $\lim_{x\to 1}\frac{x^2-1}{x+1}$

56. $\lim_{x\to 0}\frac{1-\cos x}{x}$

57. $\lim_{x\to-\infty} xe^x$

58. $\lim_{x\to 0^+}\left(\frac{1}{x}-\frac{1}{\sqrt{x}}\right)$

59. $\lim_{x\to+\infty}\frac{\sqrt{x+1}}{\sqrt{x}}$

60. $\lim_{x\to\infty}\left(\frac{x+1}{x+2}\right)^x$

61. Use l'Hôpital's rule to find

$$\lim_{x\to 0}\frac{a^x-1}{b^x-1}$$

where $a, b>0$.

62. Use l'Hôpital's rule to find

$$\lim_{x\to\infty}\left(1+\frac{c}{x}\right)^x$$

where c is a constant.

63. For $p>0$, determine the values of p for which the following limit is either 1 or ∞ or a constant that is neither 1 nor ∞:

$$\lim_{x\to\infty}\left(1+\frac{1}{x^p}\right)^x$$

64. **Cell Division Time** The Gamma distribution is used as a model for the amount of time taken for a cell to undergo a certain number of divisions. According to the Gamma distribution the likelihood that it takes t hours for the cell to complete all of its divisions is proportional to:

$$f(t) = t^p e^{-t}$$

where $p > -1$ is a constant that depends on the number of cells that are dividing, and on the conditions that the cells are growing in.

(a) Show that for all values of p, $f(t) \to 0$ as $t \to \infty$.

(b) For what values of p does $f(t)$ converge to a finite value as $t \to 0$?

65. **Lifespan Modeling** The Weibull distribution is used to model the lifespan of organisms. According to the Weibull distribution, the likelihood that an animal dies at the age of t is proportional to:

$$f(t) = t^{k-1} \exp\left(-t^k\right), \quad t \geq 0$$

where $k > 0$ is a constant that depends on the type of organism being studied, and on the environment that it is living in.

(a) Show that for all values of k, $f(t) \to 0$ as $t \to \infty$.

(b) For what values of k does $f(t)$ converge to a finite value as $t \to 0$?

66. Show that

$$\lim_{x\to\infty} \frac{\ln x}{x^p} = 0$$

for any number $p > 0$. This shows that the logarithmic function grows more slowly than any positive power of x as $x \to \infty$.

67. **Species Diversity** Chapter 3 introduced the Shannon diversity index for the diversity of a habitat. If a population contains two different species of animals, and they are present in the proportions p and $1 - p$ respectively, then the Shannon diversity index for the population is given by the formula:

$$H(p) = -p \ln p - (1-p)\ln(1-p), \quad 0 < p < 1$$

(a) Show that if $H(0) = 0$ and $H(1) = 0$, then the function $H(p)$ is continuous at $p = 0$ and $p = 1$. (*Hint*: This is equivalent to showing that $\lim_{p\to 0+} H(p) = 0$ and $\lim_{p\to 1-} H(p) = 0$.)

(b) Show that the global maximum diversity over the interval $0 \leq p \leq 1$ occurs when $p = 1/2$.

68. The height y in feet of a tree as a function of the tree's age x in years is given by

$$y = 121e^{-17/x} \quad \text{for } x > 0$$

(a) Determine (1) the rate of growth when $x \to 0^+$ and (2) the limit of the height as $x \to \infty$.

(b) Find the age at which the growth rate is maximal.

(c) Show that the height of the tree is an increasing function of age. At what age is the height increasing at an accelerating rate and at what age at a decelerating rate?

5.6 Graphing and Asymptotes

Calculus tools can be used to draw the graph of a function. Being able to graph a function may feel like a redundant skill since graphing calculators or spreadsheets can be used to plot any function. But there are many cases where the function that we are interested in has one (or more) unknown constants. For example, some insects (e.g., locusts and cicadas) undergo explosive population growth and then decay. One possible mathematical model for the population growth and decay is

$$N(t) = Mt^2e^{-mt}$$

where M and m are positive constants. We may not initially know what the values of M and m are. They will typically depend on the species that you are modeling. However, even if you do not know what the values of M and m are, you may wish to know whether the model qualitatively describes the scenario that we have described of a population that rapidly grows and then decays. One way to do this is to plot the function $N(t)$ for a particular choice of values of M and m, using either a spreadsheet or a graphics calculator. We show such a plot in Figure 5.62, using parameters $M = 100$ and $m = 0.1$.

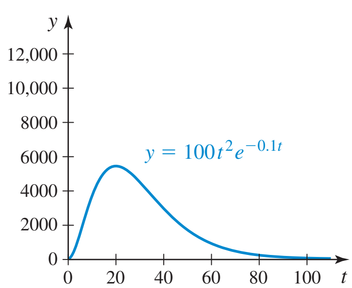


Figure 5.62 Population growth predicted by the model $N(t) = Mt^2e^{-mt}$ with $M = 100$ and $m = 0.1$.

The population appears to grow, reach a maximum, and then decay. But is this behavior a feature of the model only for this particular choice of coefficients? And how do the features of the graph, including the maximum population size and the time taken for the population to grow to its maximum size, depend quantitatively on M and m?

We can make observations about the effect of changing these coefficients upon the way that the population grows and decays by making plots for a few different values of M and m. We see in Figure 5.63 that if M is increased, say to $M = 200$, the maximum population increases but the time that the population takes to reach its maximum size remains the same.

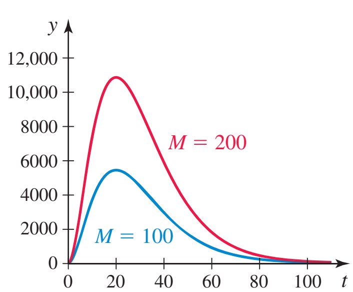


Figure 5.63 Comparing the population growth predicted by the model $N(t) = Mt^2e^{-mt}$ for $M = 200$ with $M = 100$. $m = 0.1$ for both curves.

On the other hand, if m is increased but M remains the same, then the population reaches its maximum size more quickly, and starts to decay earlier. The maximum population size also decreases (see Figure 5.64).

But we don't know from making these plots whether these observations are always true. For example, does increasing m always decrease the maximum population

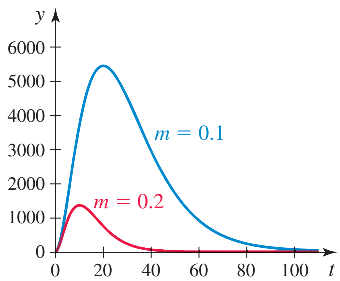


Figure 5.64 Comparing the population growth predicted by the model $N(t) = Mt^2e^{-mt}$ for $m = 0.2$ with $m = 0.1$. $M = 100$ for both curves.

size? Furthermore, if these observations *are* generally true then our plots do not tell us why. To understand these observations we can use calculus-based graph sketching methods that will be introduced in this section. These methods have the added benefit of allowing us to be quantitative about the dependence of the features of the graph upon M and m, allowing us to answer questions like if M is increased then *how much* will the maximum population size change?

The first examples we will consider in this section will not have unknown constants, so we can compare our graphs with plots produced by a spreadsheet or graphics calculator. Later we will graph functions with one or more free coefficients.

To graph the function $y = f(x)$ it is helpful to follow the following steps.

Steps for graphing a function

1. Find zeros of $f(x)$.
2. Find intervals on which $f(x)$ is positive and intervals on which it is negative.
3. If $f(x)$ is differentiable, calculate $f'(x)$.
4. Find points where either $f(x)$ or $f'(x)$ are undefined, and find behavior of $f(x)$ near these points.
5. Locate local extrema.
6. Find intervals on which $f(x)$ is increasing and intervals on which $f(x)$ is decreasing.
7. Classify extrema as either maxima or minima.
8. Calculate behavior of $f(x)$ at the end points of its interval, or if $f(x)$ is defined for all x in $\mathbf{R}$ find behavior as $x \to \pm\infty$.
9. Find intervals on which $f(x)$ is concave up and intervals on which $f(x)$ is concave down.
10. Find inflection points.

These ten steps can be performed whether or not the function $f(x)$ has unknown constants in it, as the following examples will show. In many examples we will have enough information to develop a reasonably good sketch of the graph without carrying out all ten steps (in Example 3 for instance the last two steps are omitted). The calculations that are needed to perform most of these steps are covered in Sections 5.1–5.3, and in the following examples we will discuss how the pieces of information can be linked together to draw the graph of a function. The graph of the function that we take for our first example may be familiar to you. We start this example because it allows us to expand on Steps 4 and 8.

EXAMPLE 1 Sketch the graph of the function:

$$f(x) = \frac{1}{x}, \quad x \neq 0.$$

Solution 1: Find zeros of $f(x)$. $f(x) \neq 0$ for all x.
2: Find positive and negative intervals. $\frac{1}{x} > 0$ when $x > 0$, and $\frac{1}{x} < 0$ when $x < 0$.
3: Find $f'(x)$, if $f(x)$ is differentiable. f is differentiable for $x \neq 0$ with $f'(x) = -\frac{1}{x^2}$.
4: Find points where $f(x)$ or $f'(x)$ are not defined. Neither $f(x)$ nor $f'(x)$ are defined at $x = 0$. To determine the behavior of $f(x)$ near these points we need to use the limit laws from Chapter 3.

$$\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-} \frac{1}{x} = -\infty \qquad \text{and} \qquad \lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} \frac{1}{x} = +\infty$$

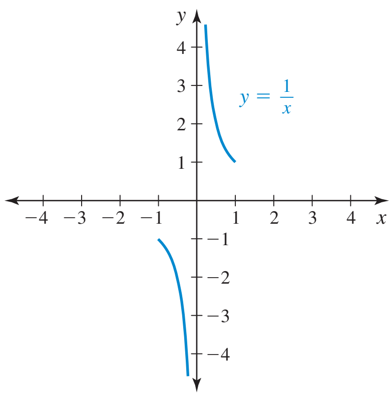


Figure 5.65 Using Steps 1–7 we begin our sketch of $y = 1/x$. Specific features shown on this graph include the vertical asymptote at $x = 0$, and that $f(x)$ is decreasing both for $x > 0$ and $x < 0$. Also there are no extrema or points where $f(x)$ crosses the x-axis.

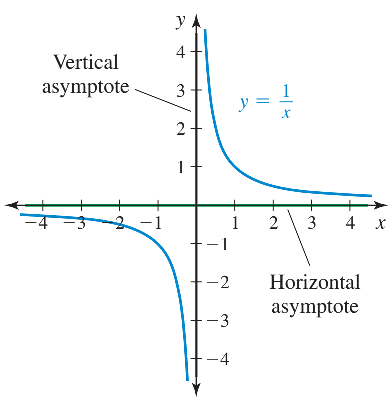


Figure 5.66 The graph of $f(x) = \frac{1}{x}$ incorporating all of the information from Steps 1–9. The two asymptotes are labelled.

That is: $f(x) \to -\infty$ as $x \to 0$ from the left and $f(x) \to +\infty$ as $x \to 0$ from the right. Taken together this means that there is a break in the graph at $x = 0$, and as x approaches this break, the function $f(x)$ approaches a vertical line $x = 0$. We call the line $x = 0$ a **vertical asymptote** of the graph. More generally, points at which functions become infinite, like this one, are known as **singularities**.
5: Find local extrema. $f'(x) = 0$ has no solutions, so by Fermat's criterion, there are no candidate local extremum points. The only such point is $x = 0$, and the behavior of $f(x)$ near $x = 0$ has already been discussed.
6: Find increasing and decreasing intervals. $f'(x) = -\frac{1}{x^2} < 0$ for $x \neq 0$, so $f(x)$ is decreasing for $x < 0$ and for $x > 0$. (Because $f(x)$ is not defined at $x = 0$, although $f'(x) < 0$, the function is not monotonic decreasing.)
7: Classify extrema. $f(x)$ has no local extrema so this step can be skipped.

From Steps 1–7 we already have enough information to begin our sketch of the function (see Figure 5.65).
8: Behavior at end points of domain. We cannot complete the sketch without knowing how the function behaves when x is large, that is, as $x \to \pm\infty$. Using the limit laws:

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \frac{1}{x} = 0 \qquad \text{and} \qquad \lim_{x\to-\infty} f(x) = \lim_{x\to-\infty} \frac{1}{x} = 0$$

So the graph of $f(x)$ approaches the horizontal line $y = 0$ when $x \to \infty$ and also when $x \to -\infty$. Such a line is called a **horizontal asymptote**.
9: Find concave up/concave down regions. f is twice differentiable provided $x \neq 0$, so we can use the second derivative to calculate whether $f(x)$ is concave up or concave down:

$$f''(x) = \frac{2}{x^3} \begin{cases} > 0 & \text{(concave up) if } x > 0 \\ < 0 & \text{(concave down) if } x < 0 \end{cases}$$

We add to our plot the information that $y = f(x)$ bends upward for $x > 0$ and downward for $x < 0$ along with the horizontal asymptote to produce Figure 5.66.

In Steps 4 and 8 we found that the function $f(x) = 1/x$ approached a vertical line as $x \to 0$ and a horizontal line as $x \to \pm\infty$. We call these lines the asymptotes of the curve. It is useful to distinguish between horizontal and vertical asymptotes:

Definition A line $x = c$ is a **vertical asymptote** if

$$\lim_{x\to c^+} f(x) = +\infty \qquad \text{or} \qquad \lim_{x\to c^+} f(x) = -\infty$$

or

$$\lim_{x\to c^-} f(x) = +\infty \qquad \text{or} \qquad \lim_{x\to c^-} f(x) = -\infty$$

A line $y = b$ is a **horizontal asymptote** if either

$$\lim_{x\to-\infty} f(x) = b \qquad \text{or} \qquad \lim_{x\to\infty} f(x) = b$$

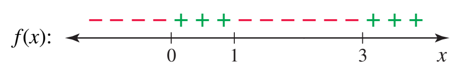
We can now use all of the steps to produce the graph of a given function. Our second example, like the first, will have no unknown constants in it, and so you could also make the plot using a graphing calculator.

EXAMPLE 2 Sketch the graph of the function:

$$f(x) = x(x-1)(x-3), \quad x \in \mathbf{R}.$$

Solution We build up information about the graph of this function by following Steps 1–10.
1: Find zeros of $f(x)$. $f(x)$ is already written in factored form, from which we can read off the roots $x = 0, 1, 3$.

2: Find positive and negative intervals. $f(x)$ is defined for all $x \in \mathbf{R}$, so it must take the same sign over the subintervals $(-\infty, 0), (0, 1), (1, 3), (3, \infty)$. To determine whether $f(x) > 0$ or $f(x) < 0$ on each subinterval, we can evaluate $f(x)$ for one value of x drawn from each subinterval. For example, for all $x \in (-\infty, 0)$, $f(x)$ takes the same sign as $f(-1) = (-1)(-2)(-4) = -8 < 0$ so $f(x) < 0$. Similarly since $f(\frac{1}{2}) = \frac{1}{2}(-\frac{1}{2})(\frac{-5}{2}) = \frac{5}{8} > 0$ so $f(x) > 0$ for $x \in (0, 1)$. By analyzing each subinterval in this way we can summarize the sign of $f(x)$ in the following number line:



To construct this number line we may evaluate $f(x)$ for one value of x in each subinterval, or note that because $f(x)$ has no repeated roots, it must change again at each root. That is, because $f(x) < 0$ for $x < 0$, we must have $f(x) > 0$ for $0 < x < 1$, and then $f(x) < 0$ for $1 < x < 3$, and so on.

3: Find $f'(x)$. $f(x)$ is differentiable for all $x \in \mathbf{R}$, and:

$$f(x) = x^3 - 4x^2 + 3x \qquad \text{Multiply out factors}$$

$$f'(x) = 3x^2 - 8x + 3.$$

4: Find points where $f(x)$ or $f'(x)$ are undefined. Both $f(x)$ and $f'(x)$ are defined for all $x \in \mathbf{R}$.

5: Find local extrema. Set $f'(x) = 0$ to identify candidate local extrema:

$$0 = 3x^2 - 8x + 3$$

$$\implies x = \frac{1}{6}\left(8 \pm \sqrt{8^2 - 4 \times 3 \times 3}\right) = \frac{4}{3} \pm \frac{\sqrt{7}}{3} \qquad \text{Solve for } x \text{ using the quadratic formula}$$

$$x \approx 0.45 \text{ or } x \approx 2.22$$

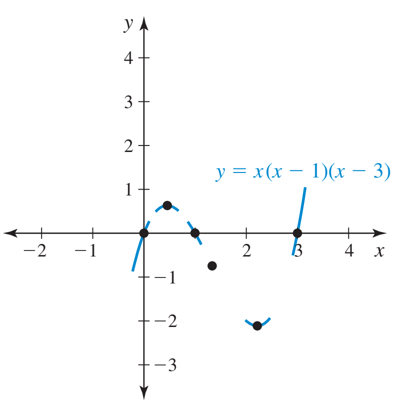


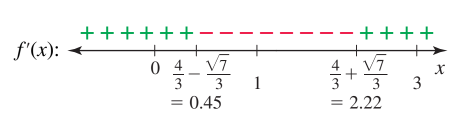
Figure 5.67 Sketch of $y = x(x-1)(x-3)$ showing (x, y) locations of roots and local extrema.

There are no points where $f'(x)$ is not defined, and the interval on which the function is defined has no endpoints, so there are no other candidates for local extrema. From the decimal values of the extrema, we see that there is one local extremum between 0 and 1, and another between 1 and 3. From the first four steps, we can start to sketch the shape of the curve $y = f(x)$. Before we do so we calculate the y-coordinates for each of the extrema: $f(0.45) = 0.63$ and $f(2.22) = -2.11$. Our first sketch therefore shows the (x, y) positions of both roots and extrema (see Figure 5.67).

6: Find intervals where $f(x)$ is increasing or decreasing. To find the intervals on which the function is increasing ($f'(x) > 0$) or decreasing ($f'(x) < 0$) we factorize $f'(x)$.

$$f'(x) = 3(x^2 - \frac{8x}{3} + 1) = 3\left(x - \left(\frac{4}{3} - \frac{\sqrt{7}}{3}\right)\right)\left(x - \left(\frac{4}{3} + \frac{\sqrt{7}}{3}\right)\right)$$

If $x < 4/3 - \sqrt{7}/3$ then both factors are negative, so $f'(x) > 0$, meaning that f is increasing. If $\frac{4}{3} - \frac{\sqrt{7}}{3} < x < \frac{4}{3} + \frac{\sqrt{7}}{3}$ then the first factor is positive and the second factor is negative, so $f'(x) < 0$, and f is decreasing. If $x > \frac{4}{3} + \frac{\sqrt{7}}{3}$ then both factors are positive so $f'(x) > 0$ and f is increasing. We summarize the information about $f'(x)$ on a number line:



7: Classify extrema. From the number line we can see that f goes from increasing to decreasing at $x = \frac{4}{3} - \frac{\sqrt{7}}{3}$ so this point is a local maximum. At $x = \frac{4}{3} + \frac{\sqrt{7}}{3}$ the function goes from decreasing to increasing, so this point is a local minimum.

8: Endpoint behavior of $f(x)$. Because $f(x)$ is defined for all real x, we need to find the behavior of $f(x)$ $x \to \pm\infty$. Since $f(x)$ is a polynomial, we recall from Chapter 3

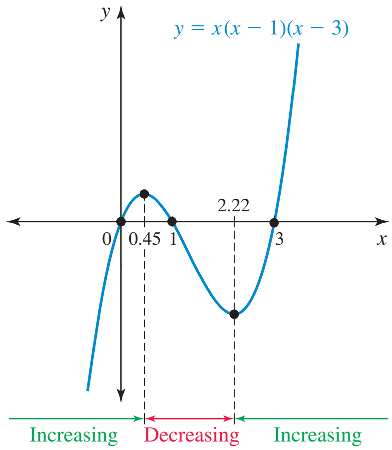


Figure 5.68 Sketch of $y = x(x-1)(x-3)$ including gradient information and behavior as $x \to \pm\infty$.

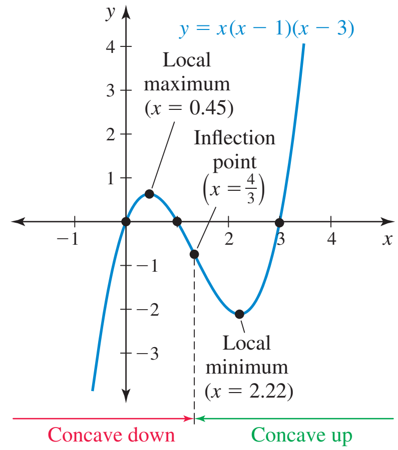


Figure 5.69 Sketch of $y = x(x-1)(x-3)$ including concave down and concave up intervals.

that its behavior for large x is dominated by the highest order term (x^3 in this case). So $f(x) \to -\infty$ as $x \to -\infty$ and $f(x) \to +\infty$ and $x \to +\infty$.

We add the information from Steps 5–8 to our sketch (see Figure 5.68). At this point we have a fairly good sense of what the function looks like, in particular where it increases, and where it decreases. But we do not yet know whether the curve bends upward or downward.

9: Identify concave up and concave down regions. To identify where the function is concave up and where it is concave down we need to calculate $f''(x)$:

$$f''(x) = 6x - 8.$$

The function is concave down if $f''(x) < 0$, i.e., if $6x - 8 < 0$ meaning that $x < \frac{4}{3}$ and it is concave up if $f''(x) > 0$ i.e., if $x > \frac{4}{3}$

10: Find inflection points. There is an inflection point where $f''(x)$ changes sign at $x = \frac{4}{3}$.

Including this information about curvature allows us to complete our sketch of the function (see Figure 5.69). •

In the next example we consider a function that has multiple singularities (points where the function goes to infinity). These kinds of functions can be challenging for a graphing calculator to plot, because y varies over an infinite range.

EXAMPLE 3 Sketch

$$f(x) = \frac{x}{x^2 + 4x + 3}.$$

Solution A function formed as the ratio $f(x) = \frac{p(x)}{q(x)}$ with $p(x)$ and $q(x)$ both polynomials is called a rational function. (Rational functions were introduced in Section 1.3). In this case $p(x) = x$ and $q(x) = x^2 + 4x + 3$. The zeros of the numerator polynomial, $p(x)$, are the roots of $f(x)$, and the zeros of the denominator polynomial are points at which $f(x)$ is undefined.

1: Find zeros. If $f(x) = 0$ then $p(x) = 0$. So $f(x) = 0$ implies that $x = 0$.
Note: $p(0) = 0$ but $q(0) = 3 \neq 0$, so $f(0) = 0/3 = 0$.

2: Find positive and negative intervals. $f(x)$ can change sign (that is change from positive to negative or vice versa) only at points where either $f(x) = 0$ or where $f(x)$ is undefined.

$f(x) = 0$ only at $x = 0$. $f(x)$ is undefined if $q(x) = 0$. Now $q(x) = (x+3)(x+1)$, so $q(x) = 0$ if $x = -1$ or $x = -3$. $f(-1) = \frac{p(-1)}{q(-1)} = \frac{-1}{0}$ and $f(-3) = \frac{p(-3)}{q(-3)} = \frac{-3}{0}$, so f is undefined at $x = -1$ and $x = -3$.

So $f(x)$ must take the same sign on subintervals $(-\infty, -3)$, $(-3, -1)$, $(-1, 0)$, and $(0, \infty)$. To determine what sign it takes on each of these intervals notice that $f(x)$ has three factors: x, $(x+1)$, and $(x+3)$. If $x < -3$, then $x < 0$, $(x+1)<0$ and $(x+3)<0$, so:

$$f(x) = \frac{(-)}{(-)(-)} < 0.$$

We use this notation to mean that the numerator < 0 and the two factors of the denominator are both < 0. For $-3 < x < -1$; $x < 0$, $(x+1) < 0$, and $(x+3) > 0$, so

$$f(x) = \frac{(-)}{(+)(-)} > 0.$$

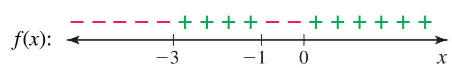
If $-1 < x < 0$, then $x < 0$ while $(x+1) > 0$, and $(x+3) > 0$, so

$$f(x) = \frac{(-)}{(+)(+)} < 0.$$

For $x > 0$; x, $(x+1)$, and $(x+3)$ are all > 0, so

$$f(x) = \frac{(+)}{(+)(+)} > 0.$$

We can summarize the sign of $f(x)$ using a number line:



3: Find $f'(x)$. On intervals where $f(x)$ is differentiable, we can calculate $f'(x)$ using the quotient rule:

$$f'(x) = \frac{(x^2+4x+3)(1) - (2x+4)(x)}{(x^2+4x+3)^2}$$
$$= \frac{-x^2+3}{(x^2+4x+3)^2}.$$

4: Find where $f(x)$ or $f'(x)$ are defined. $f(x)$ is undefined when $q(x) = 0$, i.e., at $x = -1$ or $x = -3$. $f'(x)$ is also undefined at these points. Although $f(x)$ goes to ∞ as $x \to -3$, whether $f(x)$ goes to $-\infty$ or $+\infty$ depends on whether $x = -3$ is approached from the left or right. Since $f(x) < 0$ for $x < -3$, we must have $\lim_{x\to -3-} f(x) = -\infty$. Whereas since $f(x) > 0$ for $-3 < x < -1$:

$$\lim_{x\to -3+} f(x) = +\infty.$$

Similarly $f(-1)$ is undefined. Since $f(x) < 0$ for $-1 < x < 0$, $\lim_{x\to -1-} f(x) = +\infty$ and since $f(x) > 0$ for $x > -1$, $\lim_{x\to -1+} f(x) = -\infty$.

We use the information from Steps 1–4 to begin a sketch of $f(x)$ that includes only its singularities and its root (Figure 5.70).

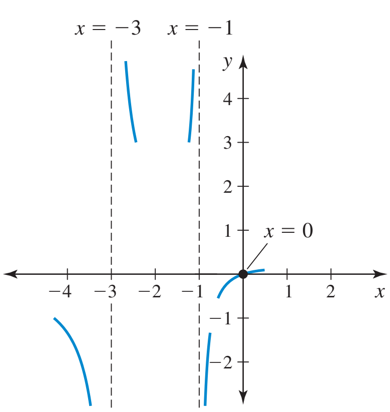


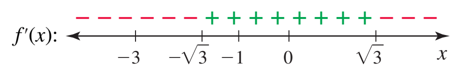
Figure 5.70 Sketch of $y = \frac{x}{x^2+4x+3}$ including zeros and singularities.

5: Find local extrema. $f'(x) = 0$, when $-x^2+3 = 0$, so there are candidate local extrema at $x = \pm\sqrt{3}$. Since $\sqrt{3} \approx 1.73$, there is one candidate local extremum between $x = -3$ and $x = -1$, and one right of $x = 0$. There are no endpoints, and the only points at which $f'(x)$ is undefined are also points where $f(x)$ is undefined, so $x = \pm\sqrt{3}$ are the only candidate local extrema.

6: Find increasing and decreasing intervals. We have shown that

$$f'(x) = \frac{-(x-\sqrt{3})(x+\sqrt{3})}{(x^2+4x+3)^2}.$$

The denominator is always positive so the sign of $f'(x)$ is determined by the sign of the numerator, that is, by the function: $-(x-\sqrt{3})(x+\sqrt{3})$. If $x<-\sqrt{3}$, both factors are negative, so $f'(x)=-\frac{(-)(-)}{(+)}<0$, meaning that $f(x)$ is decreasing. If $-\sqrt{3}<x<\sqrt{3}$ then $(x-\sqrt{3})<0$ and $(x+\sqrt{3})>0$ so $f'(x)=-\frac{(-)(+)}{(+)}>0$, so $f(x)$ is increasing. Finally if $x>\sqrt{3}$ then $(x-\sqrt{3})>0$ and $(x+\sqrt{3})>0$ so $f'(x)=-\frac{(+)(+)}{(+)}<0$, so $f(x)$ is decreasing again. We can summarize this information about $f'(x)$ in another number line.



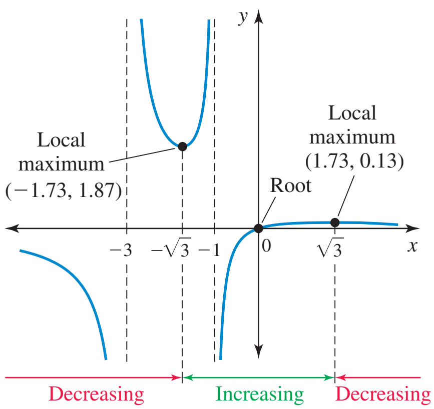


Figure 5.71 Sketch of $y=\frac{x}{x^2+4x+3}$ including zeros, singularities, local extrema and increasing/decreasing intervals, and behavior as $x\to\pm\infty$. We have added y-coordinates for each of the local extrema, $x=\pm\sqrt{3}\approx\pm1.73$.

7: Classify extrema. $f(x)$ goes from decreasing to increasing at $x=-\sqrt{3}$ so this point is a local minimum. $f(x)$ goes from increasing to decreasing at $x=\sqrt{3}$ so this point is a local maximum.

8: Calculate behavior at endpoints. Since $f(x)$ is defined on an infinite interval, we interpret the endpoint behavior to mean the limit of $f(x)$ as x approaches $\pm\infty$. Remembering that the limits of the polynomials $p(x)$ and $q(x)$ are dominated by their highest order terms (respectively x and x^2) we see that as $x\to\pm\infty$, $f(x)\approx\frac{x}{x^2}=\frac{1}{x}$. So as $x\to\pm\infty$, $f(x)\to 0$. $y=0$ is therefore a horizontal asymptote both as $x\to-\infty$ and $x\to+\infty$.

9 and 10: Determine where $f(x)$ is concave up or concave down, and find inflection points. To find where $f(x)$ is concave up or concave down, we would need to calculate $f''(x)$. We could do this in principle, since we can apply the quotient rule to differentiate $f'(x)$. But the calculation is very involved, and with the results of Steps 1–8 we already have enough information to almost completely flesh out a sketch of the function (see Figure 5.71).

The functions graphed in the first three examples had no unknown constants in them. Such functions can be plotted using a graphing calculator (although the singularities in the function may be difficult for the graphing calculator to display), so you can check that you are following the steps for graphing the function correctly by comparing your sketches with plots made by a graphing calculator. In our final two examples, the function will have one or more unknown constants and calculus-based methods are essential to understand how these constants affect the graph of the function.

EXAMPLE 4 **Insect Population Explosion** An insect population grows with time according to the model

$$N(t)=Mt^2e^{-mt};\qquad t\geq 0;$$

where M and m are positive constants. Sketch the population growth curve and explain how its features depend on the unknown coefficients.

Solution This function has two unknown constants, M and m. This function was introduced at the beginning of this section, and we used plots for specific values of M and m to develop some ideas for how these coefficients affect the shape of the curve. Now we will use calculus-based graph sketching to understand the role played by these parameters.

1: Find roots of $N(t)$. e^{-mt} is always positive, so $N(t)=0$ only when $t^2=0$. So the only root occurs at $t=0$.

2: Find positive and negative intervals. $N(t)>0$ for $t>0$, because $e^{-mt}>0$ and $M>0$ for $t>0$.

3: Calculate $N'(t)$. $N(t)$ is differentiable for all $t\geq 0$, with derivative:

$$\begin{aligned}\frac{dN}{dt}&=M(2t)e^{-mt}+Mt^2(-me^{-mt}) \qquad \text{Product rule with } u(t)=Mt^2,\ v(t)=e^{-mt}.\\&=Mt(2-mt)e^{-mt}\end{aligned}$$

4: Find points where $N(t)$ or $N'(t)$ are undefined. $N(t)$and $N'(t)$ are defined and continuous for all t.

5: Find local extrema. Candidate local extrema occur in the interior of the interval where $\frac{dN}{dt} = 0$. Since $e^{-mt} > 0$, the only such points either have $t = 0$ or $2 - mt = 0$ (meaning that $t = 2/m$). Since $t = 0$ is an endpoint of the interval, it would automatically be a local extremum also, even if $N'(0) \neq 0$.
6: Find increasing and decreasing intervals. Since $M > 0$ and $e^{-mt} > 0$, the sign of $\frac{dN}{dt}$ is determined by the sign of $t(2 - mt)$. If $0 < t < \frac{2}{m}$ then $t > 0$ and $(2 - mt) > 0$ so

$$\frac{dN}{dt} = (+)(+) > 0;$$

meaning that $N(t)$ is increasing in this interval. If $t > \frac{2}{m}$ then $t > 0$ and $(2 - mt) < 0$ so

$$\frac{dN}{dt} = (+)(-) < 0;$$

so $N(t)$ is decreasing in this interval.
7: Classify extrema. Since N is increasing to the right of $t = 0$, $t = 0$ is a local minimum. Since N goes from increasing to decreasing at $t = \frac{2}{m}$, it is a local maximum.

Using the information from Steps 1–7, we can make a preliminary sketch of the graph of $N(t)$ showing its local extrema and increasing and decreasing intervals (Figure 5.72). To assist with making this graph we add the y-coordinates for both of the local extrema, $N(0) = 0$ and $N(2/m) = 4\frac{M}{m^2}e^{-2}$.

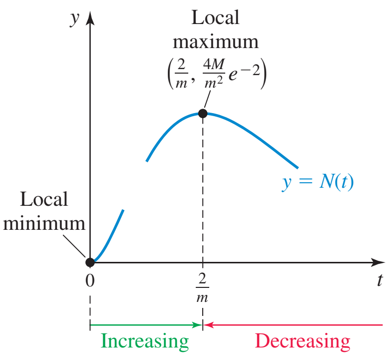


Figure 5.72 Sketch of the graph of $y = Mt^2e^{-mt}$ including increasing and decreasing intervals, root at $t = 0$ and local extrema.

8: Calculate behavior at endpoints. We have already shown that $t = 0$ is a local minimum, and that $N'(0) = 0$ so the tangent is horizontal at $t = 0$ (indeed we drew this behavior in Figure 5.72). For the other endpoint we need to determine the behavior of $N(t)$ as $t \to +\infty$:

$$\begin{aligned} \lim_{t\to\infty} N(t) &= \lim_{t\to\infty} \frac{Mt^2}{e^{mt}} && \tfrac{\infty}{\infty} \text{ limit so use l'Hôpital's rule.} \\ &= \lim_{t\to\infty} \frac{2Mt}{me^{mt}} && \text{Use l'Hôpital's rule again.} \\ &= \lim_{t\to\infty} \frac{2M}{m^2e^{mt}} = \frac{2M}{\infty} = 0 \end{aligned}$$

9: Find concave up and concave down intervals. To find concave up and concave down intervals, we need to calculate $N''(t)$:

$$\begin{aligned} \frac{d^2N}{dt^2} &= M(2 - 2mt)e^{-mt} + M(2t - mt^2)(-me^{-mt}) && \text{Product rule: } u(t) = M(2t - mt^2) \;\; v(t) = e^{-mt} \\ &= M(2 - 4mt + m^2t^2)e^{-mt}. && \text{Simplify} \end{aligned}$$

Since $e^{-mt} > 0$ for all m and t, the sign of $\frac{d^2N}{dt^2}$ is the same as the sign of $(2 - 4mt + m^2t^2)$. The roots of this equation are

$$\begin{aligned} t &= \frac{4m \pm \sqrt{(4m)^2 - 8m^2}}{2m^2} && \text{Quadratic formula} \\ &= \frac{4m \pm \sqrt{8m^2}}{2m^2} = \frac{1}{m}(2 \pm \sqrt{2}), \end{aligned}$$

Having found its roots we can factorize the polynomial part of $N''(t)$:

$$\frac{d^2N}{dt^2} = Mm^2\left(t - \frac{(2-\sqrt{2})}{m}\right)\left(t - \frac{(2+\sqrt{2})}{m}\right)e^{-mt}.$$

The sign of $N''(t)$ is determined by the sign of the two factors $\left(t - \frac{(2-\sqrt{2})}{m}\right)$ and $\left(t - \frac{(2+\sqrt{2})}{m}\right)$. It is easy to get confused by the square roots. What is important to keep in mind is that we are looking at a quadratic equation with two roots: the smaller root is $\frac{(2-\sqrt{2})}{m}$ and the larger root is $\frac{(2+\sqrt{2})}{m}$. If $t < \frac{1}{m}(2 - \sqrt{2})$ then both factors are negative:

$$\frac{d^2N}{dt^2} = (-)(-) > 0$$

so $N(t)$ is concave up. If $\frac{1}{m}(2-\sqrt{2}) < t < \frac{1}{m}(2+\sqrt{2})$ then $\left(t - \frac{(2-\sqrt{2})}{m}\right) > 0$ and $\left(t - \frac{(2+\sqrt{2})}{m}\right) < 0$ so

$$\frac{d^2N}{dt^2} = (+)(-) < 0$$

meaning that $N(t)$ is concave down. If $t > \frac{1}{m}(2+\sqrt{2})$ then both $\left(t - \frac{(2-\sqrt{2})}{m}\right) > 0$ and $\left(t - \frac{(2+\sqrt{2})}{m}\right) > 0$ so

$$\frac{d^2N}{dt^2} = (+)(+) > 0$$

meaning that $N(t)$ is concave up again.

10: Find inflection points. The graph changes curvature at $t = \frac{2\pm\sqrt{2}}{m}$ so both of these points are inflection points.

Assembling the information from Steps 1–10 we can finish our drawing of the graph. To properly place each of the points identified above on the graph, we find their coordinates as decimals. The local maximum is at $t = 2/m$ and $N(t) = 4Me^{-2}/m^2 = 0.54M/m^2$, the first inflection point is at $t = (2-\sqrt{2})/m = 0.59/m$, $N(t) = 0.19M/m^2$, and the second inflection point is at $t = 3.41/m$, $N(t) = 0.38M/m^2$ (see Figure 5.73).

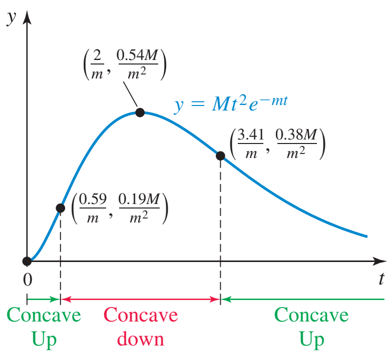


Figure 5.73 Sketch of the graph of $y = Mt^2e^{-mt}$ including all information from Steps 1–10.

We can interpret the graph and describe how its features relate to the parameters M and m. $N(t)$ initially increases and reaches its maximum value at time $t = \frac{2}{m}$. The maximum population size is approximately $0.54\frac{M}{m^2}$. Until time $t = \frac{0.59}{m}$ the growth rate of the population increases with time (the graph curves up). The growth rate starts decreasing at $t = \frac{0.59}{m}$, the decline becomes faster and faster until $t = \frac{3.41}{m}$, whereupon the rate of decline then starts to slow and the population size then tends to 0. In particular, we see that the maximum population size is proportional to M and inversely proportional to m^2. The time at which this maximum size is attained is inversely proportional to m. So if m is increased, the maximum population size becomes smaller, and is reached earlier. All of these relationships are consistent with our numerical explorations of the same function in the beginning of this chapter, but the calculus-based approach shows us quantitatively how the features of the graph depend on the coefficients m and M. •

EXAMPLE 5 Sketch the graph of the function

$$f(x) = e^{-x^2/2\sigma^2}, \quad x \in \mathbf{R}$$

where you may assume that σ is a positive constant.

Solution Although this function may look strange, it is actually one of the most important functions in the entire of mathematics. In particular, it shows up frequently in the theory of probability distributions, which you will meet in Chapter 12. To sketch the graph we will follow all ten steps for graphing a function.

1: Find zeros of $f(x)$. Remember that e raised to any power gives a positive number; so $e^{-x^2/2\sigma^2} > 0$.

2: Find positive and negative intervals. $f(x) > 0$ for all x.

3: Find $f'(x)$. $f(x)$ is differentiable everywhere, and:

$$f'(x) = -\frac{x}{\sigma^2}e^{-x^2/2\sigma^2}, \quad x \in \mathbf{R} \qquad \tfrac{d}{dx}\exp(g(x)) = g'(x)\exp(g(x))$$

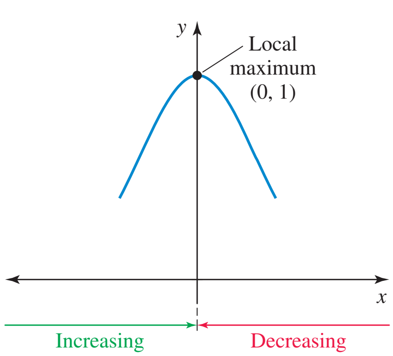


Figure 5.74 Sketch of $y = e^{-x^2/2\sigma^2}$, including increasing and decreasing intervals and local maximum at $x = 0$. Since f has no roots, the curve does not cross the x-axis.

4: Find points where $f(x)$ or $f'(x)$ are undefined. Both $f(x)$ and $f'(x)$ are defined everywhere.

5: Find local extrema. The interval has no endpoints and $f'(x)$ is defined everywhere, so the only candidates for local extrema are points where $f'(x) = 0$. Since $e^{-x^2/2\sigma^2} > 0$ for all x and σ, while $\frac{1}{\sigma^2}$ is an overall positive coefficient, the only way $f'(x)$ can go to zero is if $x = 0$.

6: Find increasing and decreasing intervals. Since x is the only part of $f'(x)$ that can change sign, $f'(x) > 0$ for $x < 0$ and $f'(x) < 0$ for $x > 0$.
7: Classify local extrema. Since $f(x)$ goes from increasing to decreasing at $x = 0$, this point is a local maximum.

From the information in Steps 1–7, we can develop a preliminary sketch of the graph of the function $f(x)$. (See Figure 5.74.)
8: Find behavior at endpoints. For this function endpoint behavior means the behavior of $f(x)$ as $x \to \pm\infty$. Since for all σ, $-x^2/2\sigma^2 \to -\infty$ as $x \to \pm\infty$, $e^{-x^2/2\sigma^2} \to 0$ in this limit. So $f(x) \to 0$ in both directions.
9: Find concave up and concave down intervals.

$$f''(x) = -\frac{1}{\sigma^2}\left(e^{-x^2/2\sigma^2} + \frac{x}{\sigma^2}\left(-\frac{x}{\sigma^2}e^{-x^2/2\sigma^2}\right)\right) \qquad \text{Product rule, } u(x) = x\sigma^2 \;\; v(x) = e^{-x^2/2\sigma^2}$$
$$= \frac{1}{\sigma^2}e^{-x^2/2\sigma^2}\left(\frac{x^2}{\sigma^2} - 1\right)$$

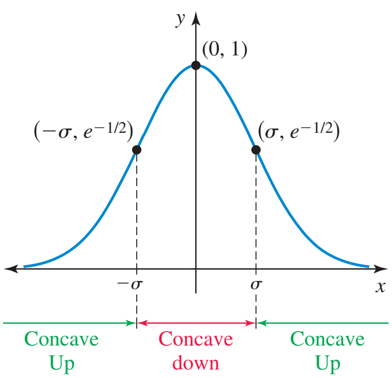


Figure 5.75 Graph of $y = e^{-x^2/2\sigma^2}$, including concave up and concave down intervals, and behavior as $x \to \pm\infty$.

Since $\frac{1}{\sigma^2}e^{-x^2/2\sigma^2} > 0$ for all x and σ, the sign of $f''(x)$ is the same as the sign of $\left(\frac{x^2}{\sigma^2} - 1\right)$, so $f''(x) < 0$ (f is concave down) for $|x| < \sigma$ and $f''(x) > 0$ (f is concave up) for $|x| > \sigma$.
10: Find inflection points. From the information in Step 9, we see that f has inflection points (goes from concave up to concave down, or conversely) at $x = \pm\sigma$.

We can incorporate all of the information from Steps 1–10 into our graph of f (see Figure 5.75).

From the graph in Figure 5.75 we can see how the shape of the function is affected by the changing the value of σ. The maximum height is 1 and is attained at the local maximum point, $x = 0$. But if σ is increased the inflection points move further and further from the origin. The height of f at the inflection points is always $e^{-1/2}$. As σ increases the graph gets wider and wider, though its height remains constant.

Section 5.6 Problems

For each of Problems 1–14, follow Steps 1–10 to find the roots, increasing and decreasing intervals and concave up and concave down intervals of each function along with its behavior at any endpoints (or as $x \pm \infty$). Sketch the function, noting on your graph where any local extrema or inflection points are. Determine whether the functions have global maxima and minima, and, if so, note their location on the graph.

1. $y = \frac{2}{3}x^3 - 2x^2 - 6x$ for $-1 \le x \le 4$

2. $y = x^4 - 4x^2, x \in \mathbf{R}$

3. $y = x^3 - 27, -5 \le x \le 5$

4. $y = x^3 - 3x, x \in \mathbf{R}$

5. $y = |x^3|, x \in \mathbf{R}$

6. $y = |x^2 - 1|, -2 \le x \le 2$

7. $y = x + e^{-2x}, x \ge 0$

8. $y = xe^{-x}, x \ge 0$

9. $y = xe^{-x^2/2}, x \in \mathbf{R}$

10. $y = e^{-(x-1)^2/2}, x \in \mathbf{R}$

11. $y = \dfrac{x^2 - 1}{x^2 + 1}, x \in \mathbf{R}$

12. $y = \dfrac{x^2 + 1}{x^2 - 1}, x \in \mathbf{R}$

13. $y = \ln(x^2 + 1), x \in \mathbf{R}$

14. $y = x \ln x, x \ge 0$

15. Let
$$f(x) = \frac{x}{x-1}, \quad x \ne 1$$

(a) Show that
$$\lim_{x\to-\infty} f(x) = 1$$
and
$$\lim_{x\to+\infty} f(x) = 1$$
That is, show that $y = 1$ is a horizontal asymptote of the curve $y = \frac{x}{x-1}$.

(b) Show that
$$\lim_{x\to1^-} f(x) = -\infty$$
and
$$\lim_{x\to1^+} f(x) = +\infty$$
That is, show that $x = 1$ is a vertical asymptote of the curve $y = \frac{x}{x-1}$.

(c) Determine where $f(x)$ is increasing and where it is decreasing. Does $f(x)$ have local extrema?

(d) Determine where $f(x)$ is concave up and where it is concave down. Does $f(x)$ have inflection points?

(e) Sketch the graph of $f(x)$ together with its asymptotes.

16. Let

$$f(x) = \frac{2}{1 - x^2}, \quad x \neq -1, 1$$

(a) Show that

$$\lim_{x \to +\infty} f(x) = 0$$

and

$$\lim_{x \to -\infty} f(x) = 0.$$

That is, show that $y = 0$ is a horizontal asymptote of $f(x)$.

(b) Show that

$$\lim_{x \to -1^-} f(x) = -\infty$$

and

$$\lim_{x \to -1^+} f(x) = +\infty$$

and that

$$\lim_{x \to 1^-} f(x) = +\infty$$

and

$$\lim_{x \to 1^+} f(x) = -\infty.$$

That is, show that $x = -1$ and $x = 1$ are vertical asymptotes of $f(x)$.

(c) Determine where $f(x)$ is increasing and where it is decreasing. Does $f(x)$ have local extrema?

(d) Determine where $f(x)$ is concave up and where it is concave down. Does $f(x)$ have inflection points?

(e) Sketch the graph of $f(x)$ together with its asymptotes.

17. Let

$$f(x) = \frac{1}{x(x+1)}, \quad x \neq 0, \ -1$$

(a) Show that $x = 0$ and $x = -1$ are vertical asymptotes.

(b) Determine where $f(x)$ is increasing and where it is decreasing. Does $f(x)$ have local extrema?

(c) Determine where $f(x)$ is concave up and where it is concave down. Does $f(x)$ have inflection points?

(d) What is the behavior of the function as $x \to \pm\infty$?

(e) Sketch the graph of $f(x)$ together with its asymptotes.

18. Let

$$f(x) = \frac{1}{x^2 + 1}$$

(a) Show that $y = 0$ is a horizontal asymptote.

(b) Does $f(x)$ have any vertical asymptotes?

(c) Follow the Steps 1–8 for graphing a function, to make a sketch of the graph of $f(x)$.

19. Let

$$f(x) = \frac{x^2}{1 + x^2}, \ x \in \mathbf{R}$$

(a) Determine where $f(x)$ is increasing and where it is decreasing.

(b) Where is the function concave up and where is it concave down? Find all inflection points of $f(x)$.

(c) Find $\lim_{x \to \pm\infty} f(x)$ and decide whether $f(x)$ has a horizontal asymptote.

(d) Sketch the graph of $f(x)$ together with its asymptotes and inflection points (if they exist).

20. Protein Binding We previously met **Hill's function** for describing cooperative binding of proteins to ligands (chemicals that have biological function) when we discussed the binding of hemoglobin to oxygen. If the concentration x of ligand is written in the right units, then Hill's function can be written in the form.

$$f(x) = \frac{x^k}{1 + x^k}, \ x \geq 0$$

where you should assume that k is a positive constant greater than 1.

(a) What are the roots of $f(x)$?

(b) Determine where $f(x)$ is increasing and where it is decreasing.

(c) Where is the function concave up and where is it concave down? Find all inflection points of $f(x)$.

(d) Find $\lim_{x \to \infty} f(x)$ and decide whether $f(x)$ has a horizontal asymptote.

(e) Sketch the graph of $f(x)$ together with its asymptotes and inflection points (if they exist).

(f) Now assume that k is less than 1. For definiteness, let $k = 1/2$. Is $f(x)$ increasing or decreasing? Show that $f(x)$ still has a horizontal asymptote, but that it is concave down for all $x \geq 0$, and use this information to make a sketch of $f(x)$.

21. Chemical Reaction We previously met the **Michaelis-Menten** rate function as a model for the rate at which a reaction occurs as a function of the concentration x of one of the reactants:

$$f(x) = \frac{x}{a + x}, \ x \geq 0$$

where a is a positive constant.

(a) Determine where $f(x)$ is increasing and where it is decreasing.

(b) Where is the function concave up and where is it concave down? Find all inflection points of $f(x)$.

(c) Find $\lim_{x \to \infty} f(x)$ and decide whether $f(x)$ has a horizontal asymptote.

(d) Sketch the graph of $f(x)$ together with its asymptotes and inflection points (if they exist).

(e) Describe in words how the graph of the function $f(x)$ changes if a is increased.

22. Population Growth The **logistic function** defined by:

$$N(t) = \frac{1}{a + e^{-t}}, \ t \geq 0$$

represents the growth of a population. We will derive this function by solving a differential equation model in Chapter 8. Assume that a is a positive constant.

(a) Determine where $N(t)$ is increasing and where it is decreasing.

(b) Find and classify any local extrema that the function has.

(c) Where is the function concave up and where is it concave down? Find all inflection points of $N(t)$.

(d) Find $\lim_{t\to\infty} N(t)$ and decide whether $N(t)$ has a horizontal asymptote.

(e) Sketch the graph of $N(t)$ together with its asymptotes and inflection points (if they exist).

(f) Describe in words how the graph of the function changes if a is increased.

23. Tumor Growth The **Gompertz function** is used in mathematical models for the rate of growth of certain tumors. The mass $M(t)$ of a tumor described by Gompertz's equation changes with time according to:

$$M(t) = \exp(ae^{-t}), \quad t \geq 0$$

where you may assume that $a > 0$ is a positive coefficient.

(a) Determine where $M(t)$ is increasing and where it is decreasing.

(b) Find and classify any local extrema that the function has.

(c) Where is the function concave up and where is it concave down? Find all inflection points of $M(t)$.

(d) Find $\lim_{t\to\infty} M(t)$ and decide whether $M(t)$ has a horizontal asymptote.

(e) Sketch the graph of $M(t)$ together with its asymptotes and inflection points (if they exist).

(f) Describe in words how the graph of the function changes if a is increased.

24. Insect Predation Spruce budworms are a major pest in forests. They are preyed upon by birds. A model for predation rate is given by

$$f(N) = \frac{aN}{k^2 + N^2}, N \geq 0$$

where N denotes the number of spruce budworm and a and k are both positive constants. The model therefore has two coefficients that will vary between different habitats.

(a) Determine where $f(N)$ is increasing and where it is decreasing.

(b) Determine the behavior of $f(N)$ at $N = 0$ and as $N \to +\infty$.

(c) Sketch the graph of $f(N)$.

5.7 Recurrence Equations: Stability

This Section uses material from Sections 2.1 and 2.2, and it should only be studied after those sections.

In Chapter 2, we introduced recurrence equations, and we showed that many biological systems can be represented by first order recurrence equation in which a sequence of values $\{x_t : t = 0, 1, 2, \ldots\}$ is specified by a formula:

$$x_{t+1} = f(x_t), \ t = 0, 1, 2, \ldots \tag{5.21}$$

Here $f(x)$ is a function that enables the value of x_{t+1} to be computed from x_t. To completely determine the sequence $\{x_t\}$ we must also specify an initial condition, that is, the value of x_0. In Chapter 2 we were able to analyze recurrence equations only numerically. We saw that equilibria (or fixed points) played a special role. An **equilibrium point** (also called fixed point), x^*, of (5.21) satisfies the equation

$$x^* = f(x^*) \tag{5.22}$$

and has the property that if $x_0 = x^*$, then $x_t = x^*$ for $t = 1, 2, 3, \ldots$. We also saw in a number of applications that, under certain conditions, x_t converged to the fixed point as $t \to \infty$ even if $x_0 \neq x^*$. However, back then, we were not able to predict when this behavior would occur.

In this Section, we will study these fixed points once again using calculus to come up with a condition that allows us to check whether convergence to a fixed point occurs. We start with the simplest example: exponential growth.

5.7.1 Exponential Growth

Exponential growth in discrete time is given by the recursion

$$N_{t+1} = RN_t, \quad t = 0, 1, 2, \ldots \tag{5.23}$$

where N_t is the population size at time t and $R > 0$ is the growth parameter. We assume throughout that $N_0 \geq 0$, which implies that $N_t \geq 0$.

The fixed point of (5.23) can be found by solving $N = RN$. The only solution of this equation is $N^* = 0$, unless $R = 1$. (If $R = 1$, then $N_0 = N_1 = N_2 = \ldots$, regardless of N_0.) Since 0 is a fixed point of the sequence, if $N_0 = 0$, then $N_1 = 0$, $N_2 = 0$, and so on. But what happens if $N_0 \neq 0$? We showed in Section 2.2 that if the sequence of population sizes converges to a limit N (i.e., $\lim_{t\to\infty} N_t = N$), then N must be a fixed

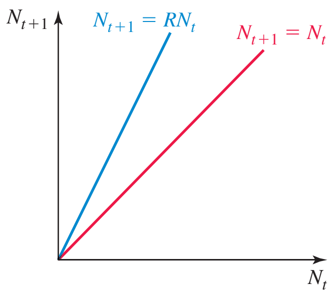


Figure 5.76 The graphs of $N_{t+1} = N_t$ and $N_{t+1} = RN_t$ intersect at $N = 0$ only if $R \neq 1$.

point of the sequence. We therefore ask under what circumstances does $N_t \to 0$ as $t \to \infty$?

In Chapter 2, we found that $N_t = N_0 R^t$ is a solution of (5.23) with initial condition N_0. So, no matter what the value of N_0, if $0 < R < 1$ then $N_t \to 0$ as $t \to \infty$. On the other hand, if $N_0 > 0$ and $R > 1$, then $N_t \to \infty$ as $t \to \infty$. So if $0 < R < 1$ then the sequence converges to the fixed point $N = 0$, while if $R > 1$ the sequence does not converge to any fixed point. We say that $N^* = 0$ is **stable** if $0 < R < 1$ and **unstable** if $R > 1$. If $R = 1$, $N = 0$ is neutral, since the sequence neither converges to 0 nor grows without bound.

Cobwebbing. There is a graphical approach to determining whether a fixed point is stable or unstable. The fixed points of (5.23) are found graphically where the graphs of $N_{t+1} = RN_t$ and $N_{t+1} = N_t$ intersect. We see (Figure 5.76) that provided $R \neq 1$, the two graphs intersect only where $N_t = 0$.

We can use the two graphs in Figure 5.76 to follow population growth ($R > 1$ in the figure). Start at N_0 on the horizontal axis. Since $N_1 = RN_0$, we can read off from the curve $N_{t+1} = RN_t$, the value of N_1 by tracing up from $N_t = N_0$ and then across to $N_{t+1} = N_1$ (shown by a dashed line in Figure 5.77). By then tracing back to the line $N_{t+1} = N_t$ we can locate N_1 on the horizontal axis, (the purple line in Figure 5.77). Then tracing to up to $N_{t+1} = RN_t$ and across to the line $N_{t+1} = N_t$ (green lines in Figure 5.77), we can locate N_2 on the horizontal axis. We can then repeat the preceding steps to find N_3 on the vertical axis, and so on (Figure 5.78). This procedure is called **cobwebbing**.

In Figures 5.77 and 5.78, $R > 1$, and we see that if $N_0 > 0$, then N_t will not converge to the fixed point $N^* = 0$, but instead will move away from 0 (and, in fact, go to infinity as $t \to \infty$).

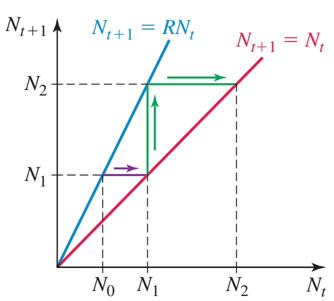


Figure 5.77 Calculating N_2 and N_1 from N_0.

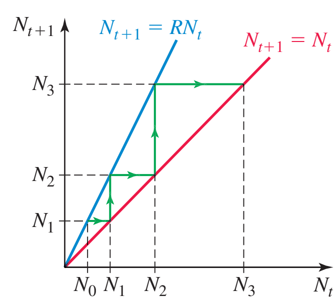


Figure 5.78 The cobwebbing procedure when $R > 1$.

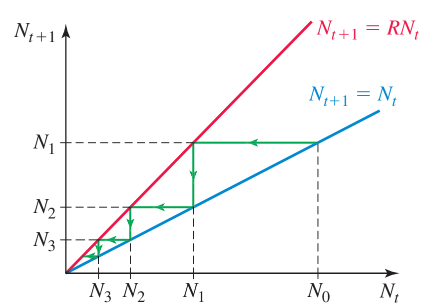


Figure 5.79 The cobwebbing procedure when $0 < R < 1$.

In Figure 5.79, we use the cobwebbing procedure when $0 < R < 1$. We see that if $N_0 > 0$, then N_t will converge to the fixed point $N^* = 0$.

When we discuss the general case, we will find that the slope of the function $N_{t+1} = f(N_t)$ at the fixed point [i.e., $f'(N^*)$] determines whether the solution moves away from the fixed point or converges to it. In the example of exponential growth, $N^* = 0$ and $f'(0) = R$. For $0 < R < 1$, $N^* = 0$ is stable; if $R > 1$, $N^* = 0$ is unstable.

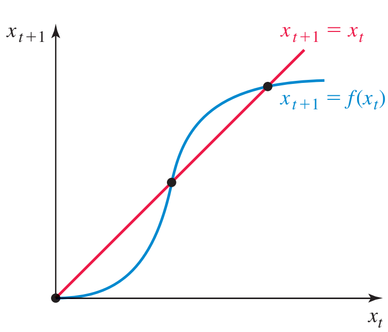


Figure 5.80 Example of a first order recurrence equation with multiple equilibria.

5.7.2 Stability: General Case

The general form of a first-order recurrence equation is

$$x_{t+1} = f(x_t), \quad t = 0, 1, 2, \ldots \tag{5.24}$$

We assume that the function f is differentiable in its domain. To find fixed points algebraically, we solve $x = f(x)$. To find them graphically, we look for points of intersection of the graphs of $x_{t+1} = f(x_t)$ and $x_{t+1} = x_t$ (Figure 5.80). The graphs in Figure 5.80 intersect more than once, which means that there are multiple equilibria or fixed points.

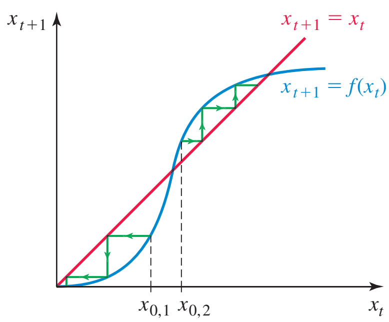


Figure 5.81 Depending on the initial value, the dynamical system converges to different limiting values.

We can use the cobwebbing procedure from the previous subsection to graphically investigate the behavior of the difference equation for different initial values. Two cases are shown in Figure 5.81, one starting at a point $x_{0,1}$ and the other at a point $x_{0,2}$. We see that x_t converges to different values, depending on the initial value. This is important to keep in mind in the discussion that follows.

Stability. To determine the **stability** of an equilibrium—that is, whether it is stable or unstable—we will proceed as in the previous subsection: We will start at a value that is different from the equilibrium and check whether the solution will return to the equilibrium. There is one important difference, however: We will *not* allow just any initial value that is different from the equilibrium; rather, we allow only initial values that are "close" to the equilibrium. We think of starting at a different value as a **perturbation** of the equilibrium, and since the initial value is close to the equilibrium, we call it a **small perturbation**.

Since we are concerned only with small perturbations, we can approximate the function $f(x)$ by its tangent-line approximation at the equilibrium x^* (Figure 5.82). We will therefore first look at graphs in which we replace $f(x)$ by its tangent-line approximation at x^*.

There are four different cases, which can be divided according to whether the slope of the tangent line at x^* is between 0 and 1 (Figure 5.83a), greater than 1 (Figure 5.83b), between -1 and 0 (Figure 5.84a), or less than -1 (Figure 5.84b).

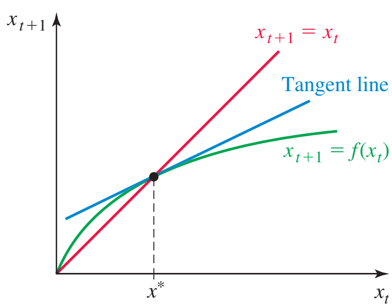


Figure 5.82 Linearizing about the equilibrium.

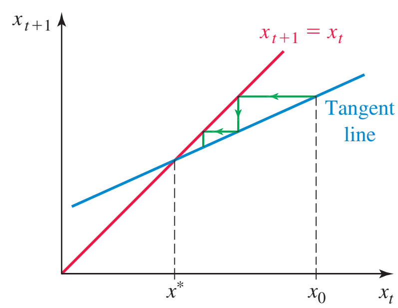


Figure 5.83a If $0 < f'(x^*) < 1$ then $x = x^*$ is locally stable.

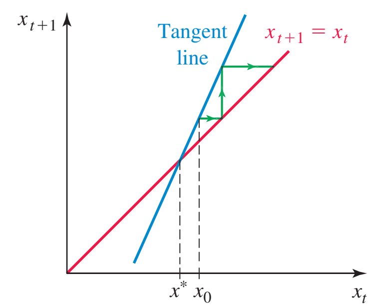


Figure 5.83b If $f'(x^*) > 1$ then $x = x^*$ is unstable.

We see that when the slope of the tangent line is between -1 and 1, x_t converges to the equilibrium (Figures 5.83a and 5.84a). The difference between Figures 5.83a and 5.84a is that in Figure 5.84a the solution x_t approaches the equilibrium in a spiral (thus exhibiting oscillatory behavior), whereas in Figure 5.83a it travels in one direction only as it approaches in one direction (thus exhibiting nonoscillatory behavior). Looking at Figure 5.83b, in which the slope is greater than 1, and Figure 5.84b, in which the slope is less than -1, we see that the solution x_t does not return to the equilibrium. In Figure 5.84b the solution moves away from the equilibrium in a spiral (thus exhibiting oscillatory behavior), whereas in Figure 5.83b it travels in one direction only

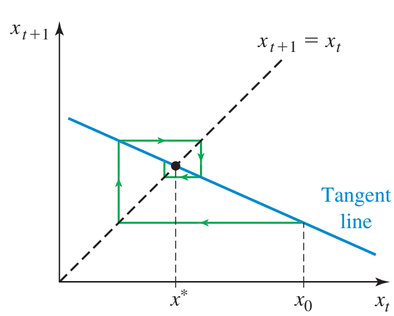


Figure 5.84a If $-1 < f'(x^*) < 0$ then $x = x^*$ is a locally stable spiral.

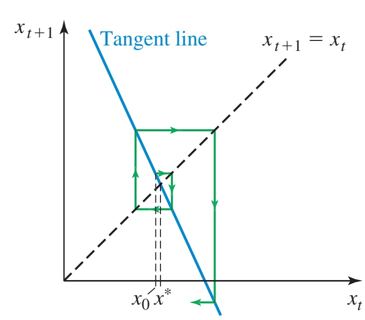


Figure 5.84b If $f'(x^*) < -1$ then $x = x^*$ is an unstable spiral.

(thus exhibiting nonoscillatory behavior). We call the equilibria in Figures 5.83a and 5.84a **locally stable**, and in Figures 5.83b and 5.84b **unstable**. Note that we added the word *locally* to *stable* to emphasize that this is a local property since we consider only perturbations close to the equilibrium.

Since the slope of the tangent-line approximation of $f(x)$ at x^* is given by $f'(x^*)$, we are led to the following criterion, which we will prove by calculus:

> **Stability Criterion for a Fixed Point** A fixed point x^* of $x_{t+1} = f(x_t)$ is locally stable if
> $$|f'(x^*)| < 1$$

Proof In Figures 5.83 and 5.84, we looked at the linearization of $f(x)$ about the equilibrium x^* graphically to investigate how the sequence behaves if it is slightly perturbed from the equilibrium. Suppose that at some time t, the solution to our recurrence equation is close to x^*. We represent the distance of the sequence from x^* by a function X_t. That is, we define:

$$x_t = x^* + X_t$$

For small perturbations from x^*, X_t will be very small.

Then

$$x_{t+1} = f(x_t) = f(x^* + X_t) \tag{5.25}$$

Recall that the linear approximation for $f(x)$ at $x = a$ is $f(x) \approx f(a) + f'(a)(x - a)$. With $x = x^* + X_t$ and $a = x^*$, the linear approximation for $f(x^* + X_t)$ at x^* is

$$f(x^* + X_t) \approx f(x^*) + f'(x^*)X_t$$

Since, by our definition of $\{X_t\}$, $x_{t+1} = x^* + X_{t+1}$, we can therefore approximate Equation (5.25) by

$$x^* + X_{t+1} \approx f(x^*) + f'(x^*)X_t$$

So:

$$X_{t+1} \approx f'(x^*)X_t \qquad \text{$x^* = f(x^*)$ so first terms cancel.} \tag{5.26}$$

This approximation should remind you of the equation $X_{t+1} = RX_t$ (exponential growth), whose solution is $X_t = R^t X_0$. We can solve the recursion for X_t by identifying the constant $f'(x^*)$ as R. The solution of this recursion is $X_t = (f'(x^*))^t X_0$. The solution converges to 0, meaning that the perturbation gets smaller and smaller with time, if $(f'(x^*))^t \to 0$ as $t \to \infty$; that is, if $|f'(x^*)| < 1$. If $|f'(x^*)| < 1$, then $X_t \to 0$, or equivalently $x_t \to x^*$ as $t \to \infty$. •

Looking back at Figures 5.83 and 5.84, we can see, in addition, that the equilibrium is approached without oscillations if $f'(x^*) > 0$ and with oscillations if $f'(x^*) < 0$.

EXAMPLE 1 Use the stability criterion to characterize the stability of the equilibria of

$$x_{t+1} = \frac{1}{4} - \frac{5}{4}x_t^2, \quad t = 0, 1, 2, \ldots$$

Solution To find the equilibria, we need to solve

$$x^* = \frac{1}{4} - \frac{5}{4}x^{*2}$$

$$\frac{5}{4}x^{*2} + x^* - \frac{1}{4} = 0$$

$$5x^{*2} + 4x^* - 1 = 0$$

The left-hand side can be factored into $(5x^* - 1)(x^* + 1)$, and we find that

$$(5x^* - 1)(x^* + 1) = 0 \quad \text{if} \quad x^* = \frac{1}{5} \quad \text{or} \quad x^* = -1$$

To determine stability, we need to evaluate the derivative of $f(x) = \frac{1}{4} - \frac{5}{4}x^2$ at the equilibria. Now,

$$f'(x^*) = -\frac{5}{2}x^*$$

so if $x^* = \frac{1}{5}$, then $|f'(\frac{1}{5})| = |-\frac{1}{2}| = \frac{1}{2} < 1$ and if $x^* = -1$, then $|f'(-1)| = |\frac{5}{2}| = \frac{5}{2} > 1$. Thus, $x^* = \frac{1}{5}$ is locally stable and $x^* = -1$ is unstable.

We can say a bit more, namely, that since $f'(\frac{1}{5}) = -\frac{1}{2} < 0$, the equilibrium $x^* = 1/5$ is approached with oscillations. •

EXAMPLE 2 Use the stability criterion to characterize the stability of the equilibria of

$$x_{t+1} = \frac{x_t}{0.1 + x_t}, \quad t = 0, 1, 2, \ldots$$

Solution To find the equilibria, we need to solve

$$x^* = \frac{x^*}{0.1 + x^*}.$$

This immediately yields $x^* = 0$ as a solution. If $x^* \neq 0$, then after dividing by x^*, we have

$$1 = \frac{1}{0.1 + x^*} \Rightarrow 0.1 + x^* = 1 \Rightarrow x^* = 0.9.$$

With $f(x) = \frac{x}{0.1+x}$, we find that

$$f'(x^*) = \frac{(0.1 + x^*) - x^*}{(0.1 + x^*)^2} = \frac{0.1}{(0.1 + x^*)^2}$$

Since $f'(0) = \frac{1}{0.1} = 10 > 1$, we conclude that $x^* = 0$ is unstable. Because $f'(0.9) = 0.1$, $0 < f'(0.9) < 1$ so we conclude that $x^* = 0.9$ is stable and is approached without oscillations. •

5.7.3 Population Growth Models

In the remaining examples in this section, we will revisit the models for density-dependent population growth that we discussed in Section 2.3. There, we analyzed these models by simulations, and we could only conjecture under what conditions the solutions to these models converged to stable final populations. We can now determine stability analytically, using the stability criterion.

EXAMPLE 3 **Beverton–Holt Model** Denote by N_t the size of a population at time t, $t = 0, 1, 2, \ldots$. Find all equilibria and determine their stability for the Beverton–Holt model:

$$N_{t+1} = \frac{R_0 N_t}{1 + aN_t}$$

where we assume that the parameters R_0 and a satisfy $R_0 > 1$ and $a > 0$. Recall that $R_0 - 1$ is the reproductive rate when N is far below the carrying capacity of the habitat, so we expect $R_0 - 1 > 0$.

Solution To find the equilibria, we set

$$N^* = \frac{R_0 N^*}{1 + aN^*}$$

and solve for N. This gives immediately the trivial solution $N^* = 0$ and, after division by N^*:

$$1 = \frac{R_0}{1 + aN^*}.$$

Solving the latter expression for N^* yields the nontrivial solution $N^* = \frac{R_0 - 1}{a}$.

To determine the stability of the two equilibria, we need to differentiate

$$f(N) = \frac{\overbrace{R_0 N}^{P(N)}}{\underbrace{1 + aN}_{q(N)}} \qquad \text{Use quotient rule.}$$

$$f'(N) = \frac{R_0\,(1 + aN) - R_0 N a}{(1 + aN)^2}$$

$$= \frac{R_0}{(1 + aN)^2}$$

To determine the stability of the trivial equilibrium $N^* = 0$, we compute

$$f'(0) = R_0 > 1$$

(since we assumed that $R_0 > 1$). Thus, $N^* = 0$ is unstable. To determine the stability of the nontrivial equilibrium $N^* = \frac{R_0 - 1}{a}$ we compute:

$$f'\left(\frac{R_0 - 1}{a}\right) = \frac{R_0}{\left(1 + a \cdot \frac{R_0 - 1}{a}\right)^2} = \frac{1}{R_0}$$

Hence, $\left|f'\left(\frac{R_0-1}{a}\right)\right| < 1$ because $R_0 > 1$. Consequently, $N^* = \frac{R_0-1}{a}$ is locally stable provided $R_0 > 1$. Since $f'\left(\frac{R_0-1}{a}\right) > 0$, the equilibrium is approached without oscillations. •

EXAMPLE 4 **Logistic Growth** Denote by N_t the size of a population at time t, $t = 0, 1, 2, \ldots$. Find all equilibria and determine their stability for the discrete logistic growth equation

$$N_{t+1} = R_0 N_t - b N_t^2$$

where we assume that $R_0 > 1$ and $b > 0$. $R_0 - 1$ is the reproductive rate when the population is far below the carrying capacity of its habitat, so $R_0 - 1 > 0$. (Previously we set $N_{t+1} = 0$ if $N > b/R_0$, but that does not affect our analysis of the model's fixed points.)

Solution To find the equilibria, we set

$$N = R_0 N - b N^2.$$

This equation yields the trivial solution $N^* = 0$ and the nontrivial solution $N^* = \frac{R_0-1}{b}$.

To determine stability of the equilibria, we need to differentiate

$$f(N) = R_0 N - b N^2$$

$$\Rightarrow f'(N) = R_0 - 2bN.$$

Since $f'(0) = R_0 > 1$, the equilibrium point $N^* = 0$ is unstable. Now,

$$f'\left(\frac{R_0 - 1}{b}\right) = R_0 - 2b\left(\frac{R_0 - 1}{b}\right) = 2 - R_0.$$

The equilibrium point at $N^* = \frac{R_0-1}{b}$ is stable if $\left|f'\left(\frac{R_0-1}{b}\right)\right| < 1$; that is if $|2 - R_0| < 1$, or equivalently $1 < R_0 < 3$. We conclude that $N^* = \frac{R_0-1}{b}$ is locally stable if $1 < R_0 < 3$; proving what we found by numerical simulations in Section 2.3. Furthermore if $1 < R_0 < 2$, then the sequence converges to $N^* = \frac{R_0-1}{b}$ without oscillations, since $f'\left(\frac{R_0-1}{b}\right) > 0$; while if $2 < R_0 < 3$, then the sequence converges to $N^* = \frac{R_0-1}{b}$ *with* oscillations, since $f'\left(\frac{R_0-1}{b}\right) < 0$. •

Section 5.7 Problems

1. **(a)** Find all equilibria of

$$N_{t+1} = 1.4N_t, \quad t = 0, 1, 2, \ldots$$

(b) Use cobwebbing to determine the stability of the equilibria you found in (a).

2. **(a)** Find all equilibria of

$$N_{t+1} = 0.7N_t, \quad t = 0, 1, 2, \ldots$$

(b) Use cobwebbing to determine the stability of the equilibria you found in (a).

3. **(a)** Find all equilibria of

$$N_{t+1} = 2N_t, \quad t = 0, 1, 2, \ldots$$

(b) Use cobwebbing to determine the stability of the equilibria you found in (a).

4. Find the equilibria of

$$x_{t+1} = \frac{2}{3} - \frac{2}{3}x_t^2, \quad t = 0, 1, 2, \ldots$$

and use the stability criterion for an equilibrium point to determine whether they are stable or unstable.

5. Find the equilibria of

$$x_{t+1} = \frac{3}{5}x_t^2 - \frac{2}{5}, \quad t = 0, 1, 2, \ldots$$

and use the stability criterion for an equilibrium point to determine whether they are stable or unstable.

6. Find the equilibria of

$$x_{t+1} = \frac{1}{4}x_t^2 + x_t - \frac{1}{4}, \quad t = 0, 1, 2, \ldots$$

and use the stability criterion for an equilibrium point to determine whether they are stable or unstable.

7. Find the equilibria of

$$x_{t+1} = \frac{1}{6}(x_t^2 + x_t + 4), \quad t = 0, 1, 2, \ldots$$

and use the stability criterion for an equilibrium point to determine whether they are stable or unstable.

8. Find the equilibria of

$$x_{t+1} = \sqrt{x_t + 1} + 1, \quad t = 0, 1, 2, \ldots$$

and use the stability criterion for an equilibrium point to determine whether they are stable or unstable (you may assume that $x_t \geq -1$).

9. Find the equilibria of

$$x_{t+1} = 4x_t^2 + x_t - 1, \quad t = 0, 1, 2, \ldots$$

and use the stability criterion for an equilibrium point to determine whether they are stable or unstable.

10. Find the equilibrium point of

$$x_{t+1} = x_t + \frac{1}{2}e^{-x_t} - \frac{1}{2}, \quad t = 0, 1, 2, \ldots$$

and use the stability criterion for an equilibrium point to determine whether it is stable or unstable.

11. Find the equilibria of

$$x_{t+1} = \frac{2x_t}{1 + x_t}, \quad t = 0, 1, 2, \ldots$$

and use the stability criterion for an equilibrium point to determine whether they are stable or unstable.

12. Find the equilibria of

$$x_{t+1} = \frac{x_t}{0.3 + x_t}, \quad t = 0, 1, 2, \ldots$$

and use the stability criterion for an equilibrium point to determine whether they are stable or unstable.

13. **(a)** Use the stability criterion to characterize the stability of the equilibria of

$$x_{t+1} = \frac{5x_t^2}{4 + x_t^2}, \quad t = 0, 1, 2, \ldots$$

(b) Use cobwebbing to find the limit that x_t converges to as $t \to \infty$ if **(i)** $x_0 = 0.5$ and **(ii)** $x_0 = 2$.

14. **(a)** Use the stability criterion to characterize the stability of the equilibria of

$$x_{t+1} = \frac{10x_t^2}{9 + x_t^2}, \quad t = 0, 1, 2, \ldots$$

(b) Use cobwebbing to decide the limit x_t converges to as $t \to \infty$ if (i) $x_0 = 0.5$ and (ii) $x_0 = 3$.

15. The dynamics of a population of fish is modeled using the Beverton-Holt model:

$$N_{t+1} = \frac{2N_t}{1 + \frac{N_t}{100}}$$

(a) Calculate the first ten terms of the sequence when $N_0 = 10$.

(b) Calculate the first ten terms of the sequence when $N_0 = 150$.

(c) Find all equilibria of the system, and use the stability criterion to determine which of them (if any) are stable.

(d) Explain why your answers from (a) and (b) are consistent with what you have determined about the equilibria of the system.

16. The dynamics of a population of fish is modeled using the Beverton-Holt model:

$$N_{t+1} = \frac{3N_t}{1 + \frac{N_t}{30}}$$

(a) Calculate the first ten terms of the sequence when $N_0 = 10$.

(b) Calculate the first ten terms of the sequence when $N_0 = 120$.

(c) Find all equilibria of the system, and use the stability criterion to determine which of them (if any) are stable.

(d) Explain why your answers from (a) and (b) are consistent with what you have determined about the equilibria of the system.

In Problems 17–19, consider the following discrete logistic model for the change in the size of a population over time:

$$\boldsymbol{N_{t+1} = R_0 N_t - \frac{1}{100}N_t^2}$$

for $t = 0, 1, 2, \ldots$.

17. **(a)** Find all equilibria when $R_0 = 1.5$ and calculate which (if any) are stable.

(b) Calculate the first ten terms of the sequence when $N_0 = 10$ and describe what you see.

18. (a) Find all equilibria when $R_0 = 2.5$ and calculate which (if any) are stable.

(b) Calculate the first ten terms of the sequence when $N_0 = 10$ and describe what you see.

19. (a) Find all equilibria when $R_0 = 3.5$ and calculate which (if any) are stable.

(b) Calculate the first ten terms of the sequence when $N_0 = 10$ and describe what you see.

20. Density-Dependent Population Growth A generalization of the Beverton-Holt model for population growth was created by ***Hassell (1975)***. Under Hassell's model the population N_t at discrete times $t = 0, 1, 2, \ldots$ is modeled by a recurrence equation:

$$N_{t+1} = \frac{R_0 N_t}{(1 + aN_t)^c}$$

where R_0, a, and c are all positive constants.

(a) Explain why you would expect that $R_0 > 1$.

(b) Assume that $c = 2$, $R_0 = 9$, and $a = \frac{1}{10}$. Find all possible equilibria of the system.

(c) Use the stability criterion for equilibria to determine which, if any, of the equilibria of the recursion relation are stable.

21. A population is modeled using Hassell's equation (introduced in Problem 20):

$$N_{t+1} = \frac{R_0 N_t}{(1 + aN_t)^2}$$

where $R_0 > 1$ and $a > 0$ are both constants, that *take different values* for different populations and $t = 0, 1, 2, \ldots$.

(a) Show that the equilibria for this population are the trivial equilibrium point $N = 0$ and another non-trivial equilibrium point, which you will need to calculate.

(b) Use the stability criterion for equilibria to show that, provided $R_0 > 1$, the non-trivial equilibrium point of this system is stable.

22. Power Law Model for Population Growth A commonly used model for the density-dependent dynamics of a population is the recurrence equation:

$$N_{t+1} = RN_t^b$$

where $b, R > 0$ are constants that take different values depending on the species of organism that is being modeled and its habitat. When $b = 1$, the model predicts that the population will grow exponentially.

(a) Show that the population has a non-trivial equilibrium point (that is $N \neq 0$), that you should determine.

(b) Show that if $-1 < b < 1$ then the non-trivial equilibrium point is locally stable, no matter what the value of R is.

(c) What happens if $b > 1$? Let $b = 2$ and $R = 1$, so $N_{t+1} = N_t^2$. Find all of the equilibria of the recursion relation, and determine which (if any) are stable.

(d) Calculate the first ten terms of the recursion relation $N_{t+1} = N_t^2$ if (i) $N_0 = 0.5$ and (ii) $N_0 = 2$?

(e) If $N_{t+1} = N_t^2$, what are the possible behaviors of the population as $t \to \infty$?

23. Density-Dependent Population Growth The Ricker model was introduced by ***Ricker (1954)*** as an alternative to the discrete logistic equation to describe the density-dependent growth of a population. Under the Ricker model the population N_t sampled at discrete times $t = 0, 1, 2, \ldots$ is modeled by a recurrence equation

$$N_{t+1} = R_0 N_t \exp(-aN_t)$$

where R_0 and a are positive constants that will vary between different species and between different habitats.

(a) Explain why you would expect $R_0 > 1$ (*Hint*: consider the population growth when N_t is very small.)

(b) Show that the recursion relation has two equilibria, a trivial equilibrium (that is, $N = 0$) and another equilibrium, which you should find.

(c) Show that if $R_0 > 1$ then use the stability criterion for equilibria to show that the trivial equilibrium point is unstable.

(d) Use the stability criterion for equilibria to show that the non-trivial equilibrium point is stable if $0 < \ln R_0 < 2$.

(e) If $R_0 > 1$ then $\ln R_0 > 0$, so most populations will meet the first inequality condition. What happens if $\ln R_0 > 2$? Let's try some explicit values: $R_0 = 10$, $a = 1$, $N_0 = 1$. Calculate the first ten terms of the sequence, and describe in words how the sequence behaves.

24. Suppose that the size of a fish population at generation t is given by the Ricker model (introduced in Problem 23)

$$N_{t+1} = 1.5 N_t e^{-0.001 N_t}$$

for $t = 0, 1, 2, \ldots$.

(a) Assume that $N_0 = 100$. Find the size of the fish population at generation t for $t = 1, 2, \ldots, 20$.

(b) Assume that $N_0 = 800$. Find the size of the fish population at generation t for $t = 1, 2, \ldots, 20$.

(c) Determine all fixed points. On the basis of your computations in (a) and (b), make a guess as to what will happen to the population in the long run, starting from **(i)** $N_0 = 100$ and **(ii)** $N_0 = 800$.

(d) Use the cobwebbing method to illustrate your answer in (a).

(e) Explain why the population size converges to the nontrivial fixed point.

25. Suppose that the size of a fish population at generation t is given by the Ricker model (introduced in Problem 23)

$$N_{t+1} = 10 N_t e^{-0.01 N_t}$$

for $t = 0, 1, 2, \ldots$.

(a) Assume that $N_0 = 100$. Find the size of the fish population at generation t for $t = 1, 2, \ldots, 20$.

(b) Show that if $N_0 = 100 \ln 10$, then $N_t = 100 \ln 10$ for $t = 1, 2, 3, \ldots$; that is, show that $N = 100 \ln 10$ is a nontrivial fixed point, or equilibrium. How would you find N? Are there any other equilibria?

(c) On the basis of your computations in (a), make a prediction about the long-term behavior of the fish population when $N_0 = 100$. How does your answer compare with that in (b)?

(d) Use the cobwebbing method to illustrate your answer in (c).

26. Medications in the Body In Chapter 2, we modeled the concentration of ibuprofen in a patient's blood using recursion relations. If the patient takes an ibuprofen pill once every 6 hours, we showed that the concentration of ibuprofen in her blood one hour after each pill is taken (that is, after 1, 7, 13, 19 hours, and so on) is given by a recurrence equation:

$$C_{n+1} = (0.7575)^6 C_n + 40$$

where C_n is the concentration in the patient's blood one hour after the n-th pill was taken (we have to wait one hour so that there is time for the drug to be absorbed from the pill). Find the equilibrium point of this recurrence equation and show that it is locally stable.

27. If a patient takes ibuprofen every T hours, rather than every 6 hours then the concentration of ibuprofen in their blood one hour after each pill is taken (that is, after $1, 1+T, 1+2T$, hours, and so on) is given by a recurrence equation:

$$C_{n+1} = (0.7575)^T C_n + 40$$

(a) Find the equilibrium point of this recurence equation, and show that it is locally stable for any value of $T > 0$.

(b) Assume that $T = 1$ and $C_1 = 40$. Make a cobweb plot to illustrate the behavior of the sequence $C_1, C_2, C_3, \ldots$.

5.8 Numerical Methods: The Newton–Raphson Method

Frequently we need to solve equations of the form $f(x) = 0$ in which finding an exact solution is impossible. Then we must use numerical methods. In Section 3.5 we introduced the bisection method as one method for finding solutions. The bisection method requires only that we be able to evaluate the function $f(x)$, and that this function is continuous. If $f(x)$ is differentiable, and we are able to calculate $f'(x)$, then we may use the **Newton–Raphson** method, which will be introduced in this section. The Newton–Raphson method can usually find solutions much more quickly than the bisection method.

The Newton–Raphson method allows us to find solutions of equations of the form

$$f(x) = 0$$

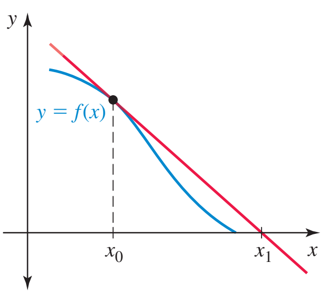


Figure 5.85 By approximating the graph of $y = f(x)$ by the linearized form of the function, we can improve on our initial guess of the value of the root.

The idea behind the method can be best explained graphically. (See Figure 5.85.)

We start by guessing a value of the x that is close to the root. Call this value x_0 (see Figure 5.85). We want to generate a *better* guess; that is a value for x that is closer to the root. If we know the derivative of the function, $f'(x)$, then we know that close to x_0 we can approximate $f(x)$ by its linearized form:

$$f(x) \approx f(x_0) + (x - x_0) f'(x_0)$$

The graph of $y = f(x_0) + (x - x_0) f'(x_0)$ is a straight line that is tangent to $y = f(x)$ at the point $x = x_0$. This line will intersect the x-axis at some point $x = x_1$, provided that $f'(x_0) \neq 0$. To find x_1, we set $x = x_1$ and $f(x_1) = 0$; that is: $f(x_0) + (x_1 - x_0) f'(x_0) = 0$ and solve for x_1:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \tag{5.27}$$

As we can see from Figure 5.85, the point x_1 is closer to the true root than the point x_0 is. We can then repeat the procedure that we followed above by approximating $f(x)$ by a tangent line at $x = x_1$, to get a new point x_2 that is even closer to the root. Replacing x_0 by x_1 and x_1 by x_2 in (5.27) we get:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

In general, if we have a value x_n that we believe is close to a root of the equation $f(x) = 0$, we can follow the same procedure to generate a better guess, x_{n+1}:

Newton–Raphson rule for calculating roots of the equation $f(x) = 0$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad \text{for } n = 0, 1, 2, \ldots \tag{5.28}$$

Starting with x_0 we use the Newton–Raphson rule to obtain x_1. From x_1 we apply the Newton–Raphson rule again to obtain x_2 and so on. The Newton–Raphson rule can therefore be used as a recurrence equation. Repeatedly applying the recurrence equation produces a sequence of numbers $x_1, x_2, \ldots$ that converges to the root as $n \to \infty$. In practice a few iterations of the Newton–Raphson rule will almost always give the root very accurately. However, the method will not always converge. In particular it requires that $f'(x)$ be defined and non-zero at the location of the root, and the success also depends on starting with a guess x_0 that is close enough to the real root. We will explore the problems that can arise if either condition is not met in the Examples.

EXAMPLE 1 Use the Newton–Raphson method to find a numerical approximation to a solution of the equation

$$x^2 - 3 = 0.$$

Solution We can solve this equation exactly, the roots are $x = \sqrt{3}$ and $-\sqrt{3}$. So in this case we can compare the exact root with the numeral approximation produced by the Newton–Raphson method. If we define $f(x) = x^2 - 3$, our task is to find the roots of $f(x) = 0$. First we need the derivative of $f(x)$:

$$f'(x) = 2x$$

Using (5.28), we find that

$$x_{n+1} = x_n - \frac{x_n^2 - 3}{2x_n}$$
$$= x_n - \frac{x_n}{2} + \frac{3}{2x_n} = \frac{x_n}{2} + \frac{3}{2x_n} \quad \text{for } n = 0, 1, 2, \ldots$$

To use this recurrence equation we need an initial guess for the value of the solution. We know that the root is near $x = 2$, so we will take $x_0 = 2$ for our initial guess. (See Figure 5.86 for the first step of the approximation.)

The following table shows the results of the procedure, as well as the distance of our numerically computed root, x_n, from the true value of the root.

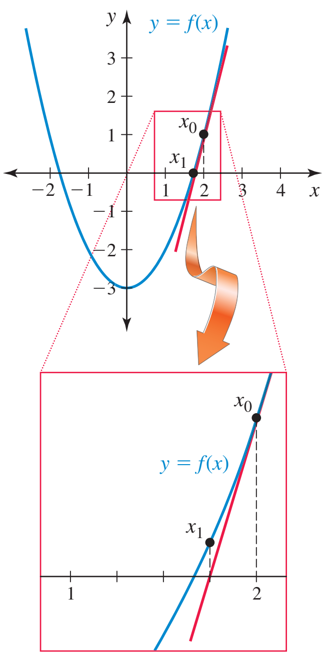


Figure 5.86 The graph of $f(x) = x^2 - 3$ and the first step in the Newton–Raphson method.

n	x_n	$x_{n+1} = \dfrac{x_n}{2} + \dfrac{3}{2x_n}$	$\lvert\sqrt{3} - x_{n+1}\rvert$
0	2	1.75	0.0179
1	1.75	1.7321429	9.2×10^{-5}
2	1.7321429	1.7320508	2.45×10^{-9}

With the starting value $x_0 = 2$, after three steps we obtain the approximation $x_3 = 1.73205081001$, which is within 8 decimal places of the true value of the root: $\sqrt{3} = 1.73205080757$ •

In general we do not have an exact solution to compare against; indeed the usefulness of the Newton–Raphson method comes from its ability to solve equations that cannot be solved by algebraic methods. Since we cannot directly measure the distance between the numerically calculated root and the true value of the root, as we did in the previous example, we then have to examine the values of the terms in the sequence x_n to determine when the sequence has converged to the solution, as the next example shows.

EXAMPLE 2 Solve the equation

$$e^x = 4x.$$

Solution First we need to turn this problem into an equation of the form $f(x) = 0$. To do that we rewrite the equation as $e^x - 4x = 0$, which is of the form needed for the Newton–Raphson method, if $f(x) = e^x - 4x$. We cannot solve this equation exactly; we need to use a numerical method. Plotting the graph of the function (see Figure 5.87), we see that the function has two roots, one between 0 and 1, and one between 2 and 3. We will calculate the smaller root (you will be asked to calculate the larger root for

Problem 3). Since we know that the root is in the interval (0, 1) we will choose $x_0 = 0.5$ for a starting guess. We also need the derivative:

$$f'(x) = e^x - 4$$

So the Newton–Raphson rule gives us the following recursion equation:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{e^{x_n} - 4x_n}{e^{x_n} - 4} \quad \text{with} \quad x_0 = 0.5$$

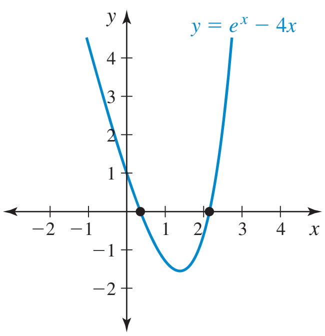


Figure 5.87 The function $f(x) = e^x - 4x$ has two roots, one between 0 and 1, and one between 2 and 3.

The first few steps of the Newton–Raphson method are summarized in the table below:

n	x_n	$f(x_n) = e^{x_n} - 4x_n$	$x_{n+1} = x_n - \frac{e^{x_n} - 4x_n}{e^{x_n} - 4}$
0	0.5	−0.351279	0.350601
1	0.350601	0.017517	0.357390
2	0.357390	0.000032	0.357403
3	0.357403	1.16×10^{-10}	0.357403

The first step in the approximation is shown in Figure 5.88.

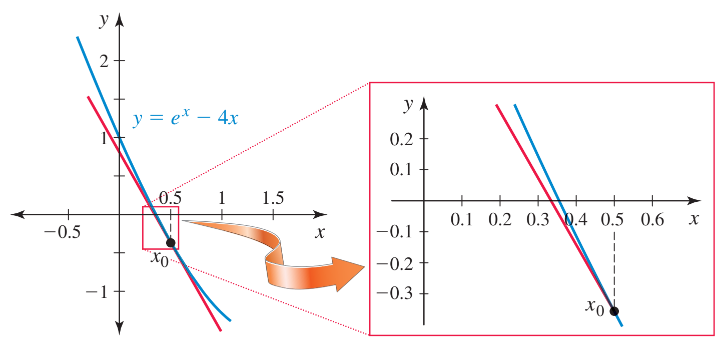


Figure 5.88 The first step in approximating the smaller root of $e^x - 4x = 0$ using the Newton–Raphson method.

In this case we do not have an exact solution to compare with. We have two methods to check that we have converged to the true root of the equation. First, we can check that successive terms of the sequence are close together. From our table we can see that x_3 and x_4 agree in their first 6 decimal places. This agreement suggests that our solution has been found to 6 decimal places by the third iteration of the method. Second, in the table we also track the value of $f(x_n)$. By the third iteration, $f(x_3)$, is only 1.16×10^{-10}, so x_3 is very close to being a solution of the equation. •

The Newton–Raphson method is a powerful tool for fitting a mathematical model to real data, as the next example shows.

EXAMPLE 3 The concentration of a particular drug in a patient's blood (measured in μg/ml) as a function of time (measured in hours) is believed to be described by a mathematical model:

$$c(t) = c_\infty - c_1 e^{-kt}$$

where c_∞, c_1, and k are all constants that we do not initially know. The patient's blood is sampled at time points $t = 0$, 1, and 2.5 hrs and the following concentrations are measured:

t	$c(t)$
0	1
1	2
2.5	2.3

Find the values for the coefficients c_∞, c_1, and k that would make the model fit all the data.

Solution Rather than having a single equation with a single variable to solve for, we are given three values of $c(t)$:

$$c(0) = 1, \quad c(1) = 2, \quad \text{and} \quad c(2.5) = 2.3,$$

which constitute 3 equations with 3 unknowns to solve for. Writing the equations out

$$c_\infty - c_1 = 1 \tag{5.29}$$

$$c_\infty - c_1 e^{-k} = 2 \tag{5.30}$$

$$c_\infty - c_1 e^{-2.5k} = 2.3 \tag{5.31}$$

To solve for c_∞, c_1, and k we manipulate the equations to reduce them from 3 equations and 3 unknowns, first to 2 equations and 2 unknowns, and from there to one equation in one unknown: c_∞ appears in the same way in Equations (5.29–5.31); by subtracting the equations we can eliminate it

$$c_1(1 - e^{-k}) = 1 \qquad \text{(5.30)–(5.29):} \tag{5.32}$$

$$c_1(1 - e^{-2.5k}) = 1.3 \qquad \text{(5.31)–(5.29):} \tag{5.33}$$

Now we reduce to one equation in one unknown; since c_1 multiplies the left-hand side of both (5.32) and (5.33) we can eliminate it by dividing the two equations:

$$\frac{1 - e^{-2.5k}}{1 - e^{-k}} = 1.3 \qquad \text{(5.33)/(5.32):} \tag{5.34}$$

This is a single equation with a single unknown, k, but it cannot be solved using algebraic manipulations. We need to solve it numerically. We will use the Newton–Raphson method. First we will rearrange the equation into a form $f(k) = 0$:

$$1 - e^{-2.5k} = 1.3(1 - e^{-k}) \quad \Longrightarrow \quad f(k) = 1.3e^{-k} - e^{-2.5k} - 0.3 = 0 \tag{5.35}$$

We could also rewrite the equation as: $f(k) = \frac{1-e^{-2.5k}}{1-e^{-k}} - 1.3 = 0$, but this function is harder to differentiate.

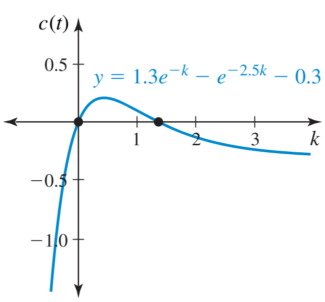


Figure 5.89 The equation $1.3e^{-k} - e^{-2.5k} - 0.3 = 0$ has two roots; one at $k = 0$ and one somewhere between $k = 1$ and $k = 2$.

Because we have multiplied both sides by $1 - e^{-k}$ we must be careful about the case where $1 - e^{-k} = 0$. When we plot the function $f(k)$ we see that there are two roots, one at $k = 0$ and one between $k = 1$ and $k = 2$ (Figure 5.89). The root at $k = 0$ is a spurious root that was introduced because when $k = 0$, $1 - e^{-k} = 0$ so in going from Equation (5.34) to Equation (5.35) we end up multiplying both sides by 0. You can check that $k = 0$ does not satisfy the original Equation (5.34). Accordingly we focus on finding the other root. Because we know that the root lies somewhere between $k = 1$ and $k = 2$, our initial guess for the value of the root is $k_0 = 1.5$. To apply the Newton–Raphson rule, we also need the derivative:

$$f'(k) = -1.3e^{-k} + 2.5e^{-2.5k}$$

So the Newton–Raphson method gives a recursion equation:

$$k_{n+1} = k_n - \frac{f(k_n)}{f'(k_n)} = k_n + \frac{1.3e^{-k_n} - e^{-2.5k_n} - 0.3}{1.3e^{-k_n} - 2.5e^{-2.5k_n}}$$

The first few steps of the Newton–Raphson method are summarized in the table below:

n	k_n	k_{n+1}	$f(k_n)$
0	1.5	1.355373	−0.0334485
1	1.355373	1.3611340	0.00144482
2	1.3611340	1.361142	0.0000021
3	1.361142	1.361142	4.2×10^{-12}

We see that after three iterations of the method the guess k_n has converged to 6 decimal places (another sign that we are very close to the root is that $|f(k_n)| < 10^{-11}$).

To complete our solution of the problem, we need to also obtain c_0 and c_∞. To do that we substitute for k in one of our earlier equations

$$c_1 = \frac{1}{1 - e^{-k}} = 1.345 \qquad \text{Rearrange (5.32)}$$

Then:

$$c_\infty = 1 + c_1 = 2.345 \qquad \text{Substitute for } c_1 \text{ in (5.29).}$$

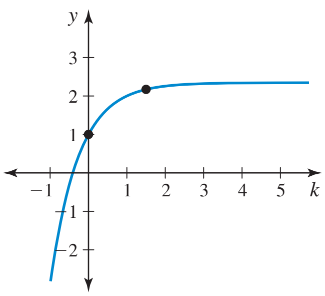


Figure 5.90 When the values for c_∞, c_1, and k are chosen correctly the model $c(t) = c_\infty - c_0 e^{-kt}$ fits experimentally measured concentration data.

So the mathematical model fits the data if $(c_\infty, c_1, k) = (2.345, 1.345, 1.361)$. We can also make a plot of the mathematical model using these coefficient values, to confirm that it goes through all three of the data points that are given to us (see Figure 5.90)

It is not our goal here to provide a complete description of the Newton–Raphson method (when it works, how quickly it converges, etc.). But we give one cautionary example to show a problem that you may encounter when applying the Newton–Raphson method.

EXAMPLE 4 Use the Newton–Raphson method to find the root of

$$x^{1/3} = 0$$

when the starting value is $x_0 = 1$.

Solution We can guess one root of the equation straightaway: $x = 0$. We set $f(x) = x^{1/3}$. Then $f'(x) = \frac{1}{3}x^{-2/3}$, and

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^{1/3}}{\frac{1}{3}x_n^{-2/3}} = x_n - 3x_n = -2x_n$$

The following table shows successive values of x_n:

n	x_n	$x_{n+1} = -2x_n$
0	1	−2
1	−2	4
2	4	−8
3	−8	16

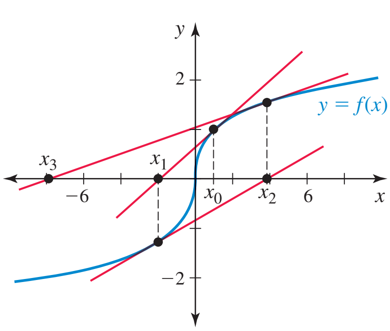


Figure 5.91 The Newton–Raphson method does not converge to the root $x = 0$ for the equation $x^{1/3} = 0$.

The successive values do not converge to the root $x = 0$, but grow successively larger and larger. The situation is graphically illustrated in Figure 5.91. The problem in this example is that $f'(x) = \frac{1}{3}x^{-2/3}$ is not defined if $x = 0$, i.e., the function f is not differentiable at the root $x = 0$.

Section 5.8 Problems

1. Use the Newton–Raphson method to find a numerical approximation to the solution of

$$x^2 - 7 = 0$$

that is correct to six decimal places.

2. The equation

$$x^2 - 5 = 0$$

has two solutions. Use the Newton–Raphson method to approximate the two solutions to four decimal places.

3. Use the Newton–Raphson method to find a numerical approximation to the solution of

$$e^x = 4x$$

in the interval $(2, 3)$ correct to six decimal places.

4. Use the Newton–Raphson method to find a numerical approximation to the solution of

$$e^x + x = 2$$

that is correct to six decimal places.

5. Use the Newton–Raphson method to find a numerical approximation to the solution of

$$x^2 + \ln x = 0, x > 0$$

that is correct to six decimal places.

6. Use the Newton–Raphson method to find a numerical approximation for all of the solutions of:

$$x^3 + x^2 + 1 = x$$

correct to six decimal places.

7. Use the Newton–Raphson method to find a numerical approximation for all of the solutions of:

$$x^4 + x^3 + 1 = x^2 + 2x$$

correct to six decimal places.

8. Use the Newton–Raphson method to solve the equation

$$\sin x = \frac{1}{2}x$$

in the interval $(0, \pi)$.

9. Use the Newton–Raphson method to solve the equation

$$\sin x = x^2$$

in the interval $(0, \pi)$.

10. In Example 4, we discussed the case of finding the root of $x^{1/3} = 0$.

(a) Given x_0, find a formula for $|x_n|$.

(b) Find

$$\lim_{n\to\infty} |x_n|$$

(c) Graph $f(x) = x^{1/3}$ and illustrate what happens when you apply the Newton–Raphson method.

11. Insect Population Explosion The size of an insect population (measured in millions of insects) is given by the function:

$$N(t) = N_0 t^2 e^{-mt}$$

with $N_0 = 10$, $m = 0.14$. Calculate using the Newton–Raphson method, and accurate to four decimal places,

(a) the time at which the size of the insect population first decreases to below 1 million insects.

(b) the time at which the size of the insect population first decreases to below 0.1 million insects.

12. Drug Absorption The blood concentration of a particular drug is given by a mathematical model: $c(t) = c_0 + c_\infty e^{-kt}$ Suppose that you measure, for a single patient, the following data:

time (t)	concentration ($c(t)$)
0	0
1	2.5
1.5	2.8

Use the Newton–Raphson method to calculate, correct to 3 s.f., the constants c_0, c_∞, and k that fit the model to your measured data.

13. Spread of an Epidemic You are trying to fit the following model for the spread of flu in the UCLA dormitories. At time t (measured in days) the number of sick students is:

$$N(t) = N_0 + N_1 e^{-rt}$$

where N_0, N_1, r are all constants that need to be fit to the data. You also have the following data, provided by campus health services:

Day	Number of sick students
0	10
1	200
3	300

(a) Use the data to estimate the constants N_0, N_1, and r in your epidemiological model.

(b) What is the maximum number of people who will get sick (i.e., what is the maximum value of $N(t)$) according to your model?

14. The growth of a particular population is described by a power law model, in which the rate of growth is given by a function:

$$r(t) = \frac{A}{(t+a)^m}$$

where A, m, and a are all unknown constants. Given the following data for the size of the population, calculate the value for these constants that would fit the model to the data:

t	$r(t)$
0	1.89
1	1.31
3	0.988

Hint: Eliminate A first. It may help to then take logarithms of the equations that you derive after eliminating A.

15. Tumor Growth The Gompertz function has been used to model the growth of tumors. It predicts that the size, $s(t)$, of the tumor is given as a function of time, t, by:

$$s(t) = a \exp\left(-be^{-ct}\right)$$

where $a, b, c > 0$ are all unknown, positive constants. Given the following data for the size of a tumor, calculate the value for these constants that would fit the model to the data:

t	$s(t)$
0	0.37
1	1.59
2	2.59

Hint: Eliminate a first. It may help to then take logarithms of the equations that you derive after eliminating a.

5.9 Modeling Biological Systems Using Differential Equations

One of the most important uses of calculus is to build mathematical models. These models find applications throughout biology, from predicting the growth of populations, to the spread of disease, to the passage of drugs through a patient's body, and we will discuss many examples of mathematical models in Chapter 8. In Section 2.3, we built models for the growth of a population, if the population is measured at a discrete set of times $t = 0, 1, 2, \ldots$. Discrete models for population growth usually take the form of a recurrence equation in which the population at time step $t+1$ (i.e., N_{t+1}), can be calculated by applying some function to the population at the previous time step, so that $N_{t+1} = f(N_t)$. The hard work of building the model is often finding the right function $f(N_t)$ for each system. In practice, it is often more useful to have a model that allows time to vary continuously, and that predicts the population size not just at discrete intervals, but at any possible time. If, instead of trying to predict a sequence, $\{N_t\}$, that gives the population size at discrete times, we are trying to calculate a function $N(t)$ that would enable the population size to be predicted for any time, t, then our model will usually take the form of a differential equation for the rate of change of the population size, $\frac{dN}{dt}$. In this section we will focus on building these differential equation models for population growth, and for the passage of a drug through a patient's body (many more examples will be introduced in Chapter 8). Solving differential equations, that is, calculating $N(t)$ if we have a formula for $\frac{dN}{dt}$, requires a completely different set of techniques than solving recurrence equations. In this section, we will concentrate on the math involved in building the models. For each model, we will then give you the solution. In Chapter 8 you will learn techniques to solve the differential equations for yourself.

5.9.1 Modeling Population Growth

Suppose that you are modeling the growth of a population of cells that is growing in a flask. At time t the population size is given by $N(t)$. Our goal in this subsection is to calculate the function, N. To do this, we make a similar argument to the one that we made when modeling population growth in Section 2.3. That is, rather than trying to get $N(t)$ directly, we ask what processes might lead $N(t)$ to change over time. Specifically, let's take a time t and a later time $t+h$. How would $N(t)$ change between those two time points? We can write down a word equation

$$N(t+h) = N(t) + \begin{array}{l}\text{Number of cells}\\ \text{born between}\\ t \text{ and } t+h\end{array} - \begin{array}{l}\text{Number of cells that}\\ \text{die between}\\ t \text{ and } t+h.\end{array} \tag{5.36}$$

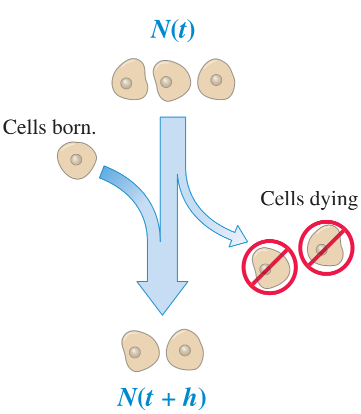


Figure 5.92 In the time interval $(t, t+h)$, the population size may increase because of births or decrease because of death.

See Fig. 5.92 for a graphical representation of this equation.

We use the word "birth" here to describe the division of one cell into two cells. We will moreover assume that h is a small interval of time (ultimately, to write a differential equation, we will need to let $h \to 0$). Typically the number of births (and deaths) will be proportional to h; that is, if we double the time interval between t and $t+h$, we would expect that approximately twice as many births (and deaths) should occur. Additionally, we expect the number of births (and deaths) to be proportional to the number of cells are present; if we take the same amount of time, and double the population size, then the number of births (or deaths) would also double. Put another way, if one cell out of 10 dies in a certain time interval, then we would expect 10 cells out of 100 to die in the same interval, or 100 cells out of 1000, and so on. Because of these proportionality rules, we expect that

$$\text{Number of births between } t \text{ and } t+h \approx bNh$$

for some constant b. b, which is often called the **birth rate**, tells us how many births occur (that is, how many new cells are created by division) per cell of the population and per unit time. For example, if $b = 0.2$ /hr, then a fraction 0.2 (or 2 in 10) of the cells will divide to create new cells in one hour. So if $h = 0.1$ hr, and $N = 100$, we expect the number of cells born in the interval between t and $t + h$ to be $0.2 \times 100 \times 0.1 = 2$.

Similarly:

$$\text{Number of deaths between } t \text{ and } t + h \ \approx mNh$$

for some constant m. m, which is often called the **mortality rate**, tells us how many deaths occur, per cell of the population at time t, and per unit time. For example, if $m = 0.01$ /hr then a fraction 0.01 (or 1 in 100) of the cells will die in one hour.

Why do we use $\approx$ instead of $=$ in the above equations? The assumption that the number will only hold if h is sufficiently small. If h is large, then the changes in $N(t)$ over the time interval between t and $t + h$ will start to affect the numbers of births or deaths. To make this idea more concrete, assume that $b = 0.2$/hr, and $N(t) = 1000$ and $m = 0$ (no deaths occur during the experiment). Then if $h = 1$ hr, the number of births between t and $t + h$ will be $bN(t)h = 200$, according to our formula. So if we neglect cell deaths, $N(t + h) = N(t) + 200 = 1200$. What if we waited another hour? Then in the interval between $t + h$ and $t + 2h$ the number of births will be $bN(t + h)h = 0.2 \times 1200 \times 1 = 240$. In total $200 + 240 = 440$ births are predicted between t and $t + 2h$. But we could also use the formula for the number of births directly, replacing h by $2h$. Using the formula directly, we predict $bN(t) \cdot 2h = 400$ births. We get different answers depending on whether we split the interval into two hour-long intervals, or treat it as a single two-hour interval. The formulas are valid only for very small time intervals, h, (in fact in the limit as $h \to 0$).

Putting the equations for the number of births and the number of deaths together, and rearranging Equation (5.36) we obtain:

$$N(t + h) - N(t) = bN(t)h - mN(t)h$$

As explained above, this equation is only valid if the limit as $h \to 0$ is taken. In that limit, both sides of the equation become smaller and smaller. However, we can rearrange terms to put all dependence on h on the left-hand side of the equation:

$$\frac{N(t + h) - N(t)}{h} = bN - mN$$

We can see that now the right-hand side doesn't depend on h. But h does appear on the left-hand side. As $h \to 0$ both sides of this equation must therefore converge to well-defined limits. The left-hand side becomes $\frac{dN}{dt}$, and since the right-hand side has no h-dependence its limit is just $bN - mN$. We have therefore derived the differential equation:

$$\frac{dN}{dt} = bN - mN. \tag{5.37}$$

This equation alone, even if we knew the values of b and m, would not be enough to completely determine the function $N(t)$. We could imagine a scenario where two flasks contain the same type of cell, with the same birth rate b and the same death rate m. In both cases $\frac{dN}{dt}$ would satisfy the same differential equation, but if one flask were started with 100 cells (that is, $N(0) = 100$) and the other flask with 1000 cells (that is $N(0) = 1000$) we would expect the population size at any time t to be given by different functions in the two cases. So to completely solve for the function $N(t)$ we need both the differential equation (Equation (5.37)) and the size of the population at one point in time, for example, the population size at $t = 0$, i.e., $N(0)$. In fact knowing $N(t_0)$ at any time $t = t_0$ will be sufficient to completely determine $N(t)$. We call this extra piece of data $N(t_0)$, the **initial condition** for the differential equation.

Given the differential equation (5.37) and an initial condition $N(0) = N_0$, you can use the techniques from Chapter 8 to solve for $N(t)$. (The solution to this differential equation is also discussed in Problem 56 of Section 5.1.) Rather than waiting until then, we will just give the solution. It is the exponential:

$$N(t) = N_0 e^{(b-m)t} \tag{5.38}$$

We can check that this function truly does satisfy both the differential equation and its initial condition. On the left-hand side of (5.37):

$$\frac{dN}{dt} = (b - m)N_0 e^{(b-m)t}$$

Recall that N_0, b, m are all constants.

while on the right-hand side: $(b-m)N = (b-m)N_0e^{(b-m)t}$. So substituting into both sides of the equation we get the same function, and the differential equation is therefore satisfied. Also $N(0) = N_0 \cdot e^0 = N_0$, so our initial condition is also satisfied.

5.9.2 Interpreting the Mathematical Model

As a sanity check, let's make sure that the population growth predicted by our model depends, in the way that we would expect, on all of the constants: N_0, b, and m. The size of the population remains proportional to N_0 at all times, t, so if two flasks are started at the same time and initially contain 10 and 100 cells respectively, then at all future times there will be 10 times more cells in the second flask. Birth and death rates are both constants, so, over time, the cell population grows (or decays) exponentially – just like the discrete model of density-independent growth in Chapter 2. Larger values of b, meaning higher birth rates, lead to faster exponential growth, and larger values of m, meaning higher death rates, lead to slower exponential growth. When the birth rate exceeds the mortality rate, i.e., $b > m$, the population grows exponentially. When the mortality rate is large, i.e., $b < m$, the population decays exponentially.

As described above, b and m are both rates; they have units of $(\text{time})^{-1}$. We have interpreted b as the per cell rate of births – that is the number of times each cell in the flask will divide in one unit of time. How do we interpret these rate in terms of quantities that we could measure? Suppose on average a cell takes time t_b to divide (ignore death for now). How is b related to t_b? Suppose we apply our model to a flask that contains just 1 cell initially, and neglect cell deaths. Then (5.37) becomes

$$\frac{dN}{dt} = b \cdot N \qquad \text{with} \quad N(0) = 1.$$

with solution $N(t) = e^{bt}$. After time t_b has elapsed the starting cell will have divided once, so that there are now $N(t_b) = 2$ cells in the colony. That is, $e^{bt_b} = 2$, which implies $t_b = \frac{\ln 2}{b}$ or $b = \frac{\ln 2}{t_b}$.

Under optimal laboratory conditions yeast cells divide approximately every 90 minutes. So $b = \frac{\ln 2}{1.5\text{ hours}} = 0.46\text{ hours}^{-1}$.

The constant m represents the per cell death rate. How would we relate this to the lifetime t_m of a cell? If we ignore births and apply our model to a flask that contains just one cell initially, then (5.37) becomes:

$$\frac{dN}{dt} = -m \cdot N \qquad \text{with} \qquad N(0) = 1.$$

with solution $N(t) = e^{-mt}$. Our first instinct is to identify t_m as the time at which the cell has died, so that $N(t_m) = 0$. But $N(t)$ approaches 0 only asymptotically as $t \to \infty$; the model never predicts that the cell dies out (we will discuss next how to interpret the values from the model when they are not integers, i.e., how to make sense of a prediction like $N(t) = 0.01$). To understand the relationship between m and t_m, we must realize that cells, like people, don't have fixed lifetimes. However, we can speak of the average lifetime of a cell. There are multiple ways to construct this average, but one of the most straightforward is to identify t_m as the time it takes half of all cells to die. Thus if there are only two cells in the flask initially, t_m is defined to be the time at which the number of cells will have decreased to 1. According to our model, if $N(0) = 2$ then $N(t) = 2e^{-mt}$. So if half of the cells die in time t_m, then $N(t_m) = 1$, which implies that $2e^{-mt_m} = 1$ or $m = \frac{\ln 2}{t_m}$. Because the population size is proportional to N_0 at all times, t, we would get the same answer if $N(0) = 100$ and we define t_m by $N(t_m) = 50$, or for any other starting number of cells.

Under typical laboratory conditions yeast cells die after 1 week, i.e., their lifetime, is $t_l = 168$ hrs, and so $m = 0.004\text{ hr}^{-1}$. We can put these ingredients together to predict the growth of a yeast cell population:

$$N(t) = N(0)e^{0.458t}. \tag{5.39}$$

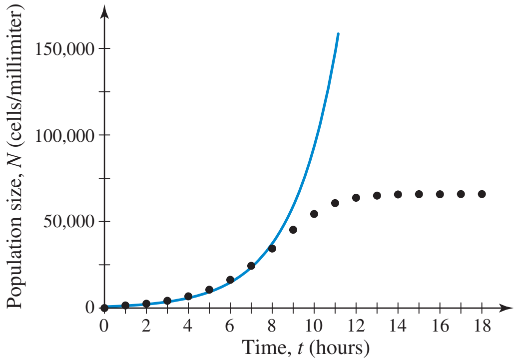


Figure 5.93 A mathematical model for the growth of a population of cells captures the initial growth of a real population of yeast cells grown in a flask. However, the model predicts that the population will grow indefinitely, while real yeast populations eventually plateau.

If $N(0)$ is chosen to agree with a real experiment, this equation agrees quite well with real measurements of the growth of a yeast population, at least if t is not too large. We compare our model to data collected by ***Carlson (1913)*** for yeast cells grown in a flask in Figure 5.93. The model and the real data agree very well during the first few hours of population growth. But the model predicts that the yeast population will grow indefinitely. In reality the flask, like all habitats, has a carrying capacity. As the population size approaches this carrying capacity, population growth slows, and eventually the population size plateaus.

By contrast, our model says that growth should continue indefinitely: after one week (t =168 hr) it predicts there are 2.6×10^{33} cells in the colony, while after one month ($t = 720$ hr) it predicts that $N = 10^{143}$, which is larger than the number of atoms in the observable universe. This comparison shows that why we need to be careful when interpreting models – the models may not capture all of the features of the real life systems that they are created to represent. In real populations, the cells eventually exhaust the nutrients available to them and then they cease to divide. We already encountered density-dependent reproductive rates in discrete models in Section 2.3. In Chapter 8 we will describe how density-dependent growth can be incorporated into differential equation models.

Another feature of the model that we need to unpack is the representation of the population size, $N(t)$, by a continuous function. Real populations are made up of a discrete number of individuals, that is, the size of a single population must be an integer like $0, 1, 2, \ldots$. However, our model predicts that at most times $N(t)$ will take a non-integer value. For example, if our yeast cell population starts with 100 cells, and is described by the data-fitted model (Equation 5.39), then $N(t) = 100e^{0.458t}$, so after one hour the model predicts that the population size is $N(1) = 158.09$, and after two hours it is $N(2) = 249.93$. What does it mean for the model to predict that there is 0.09 of a cell? Real population growth is a random process; cells divide at random times, and b is the average rate of division (and similarly, m is the average rate of deaths). The model describes only what the population does on average. If we were to start many populations of yeast cells in identical, replicate flasks and all with 100 cells at time $t = 0$, then at time $t = 1$ there would be a range of different population sizes over all of the flasks. Some of the flasks would have 158 cells, some would have 159, some would have 160, and so on. Our model predicts that if an average population size were taken over all of the replicate flasks, this average would be 158.09 cells; no flask needs to contain exactly 158.09 cells.

Although Equation (5.38) provides a formula for the growth of a population of cells, to apply it to real cell populations we need to know the parameters N_0 and $(b-m)$. Sometimes we can measure these parameters directly. Or they may be inferred by fitting the model to experimental data, as we show in the next two examples.

EXAMPLE 1 A yeast colony initially contains 1000 cells. After 2 hours elapse there are 3000 cells present. At what time will the colony size reach 10000 cells?

Solution If t is measured in hours, we are given data $N(0) = 1000$ and $N(2) = 3000$ so

$$N_0 e^0 = 1000 \quad \text{and} \quad N_0 e^{2(b-m)} = 3000.$$

From the first equation we read off $N_0 = 1000$ and from the second we then obtain

$$b - m = \frac{1}{2} \ln\left(\frac{3000}{N_0}\right) \qquad \text{Divide by } N_0 \text{ and take ln of both sides of the equation.}$$

$$= \frac{1}{2} \ln\left(\frac{3000}{1000}\right) = 0.549 \text{ hr}^{-1}$$

so the colony size reaches 10000 cells at time $t = t^*$ where

$$N(t^*) = 10000 \implies N_0 e^{t^*(b-m)} = 10000$$

or

$$t^* = \frac{1}{(b-m)} \ln\left(\frac{10000}{N_0}\right) = \frac{\ln 10}{0.549\ \mathrm{hr}^{-1}} = 4.2\ \mathrm{hrs}.$$

EXAMPLE 2 Bacteria in a flask grow exponentially in time. Two hours after the beginning of the experiment you find that there are 10^4 cells in the flask. Four hours later there are 10^5 cells in the flask. Infer how many cells were present at the beginning of the experiment.

Solution Just as in Example 1 we have two data points: if t is measured in hours $N(2) = 10^4$ and $N(6) = 10^5$. In this case, we want to calculate $N(0)$.

$$N_0 e^{2(b-m)} = 10^4 \quad \text{and} \quad N_0 e^{6(b-m)} = 10^5.$$

By dividing the two expressions we can eliminate N_0:

$$\frac{N_0 e^{6(b-m)}}{N_0 e^{2(b-m)}} = \frac{10^5}{10^4}$$

$$\Longrightarrow \quad e^{4(b-m)} = 10,$$

so $e^{(b-m)} = 10^{1/4}$. There is no need to solve for $b - m$ directly, since these coefficients always show up together as e^{b-m}. Then we can solve for N_0 using:

$$N_0 e^{2(b-m)} = 10^4 \quad \Longrightarrow \quad N_0 = 10^4 e^{-2(b-m)} = 10^4 \times (10^{-1/2}) = 3.1 \times 10^3 \text{ cells}.$$

so at time $t = 0$ there are $N(0) = N_0 = 3.1 \times 10^3$ cells in the flask.

5.9.3 Passage of Drugs Through the Human Body

A second major area for which mathematical models are tremendously important is the study of how drugs enter, are utilized by, and then are eliminated from the human body. Let $M(t)$ be the total amount of drug present in the blood at time t; $M(t)$ might represent a concentration of drug (for example in mg/liter of blood) or a total amount, measured in grams for example. We want to derive a differential equation model for how M changes with time. Just as for the population growth example, we start by considering the processes that would cause M to change in the small time interval between t and $t + h$. We can write down a word equation:

$M(t+h) = M(t)$	$+$	Amount of drug entering the blood (either from the gut, or directly, by an intravenous line) between t and $t+h$	$-$	Amount of drug leaving the blood because it is absorbed by cells, or filtered and passed out of the body between t and $t+h$.

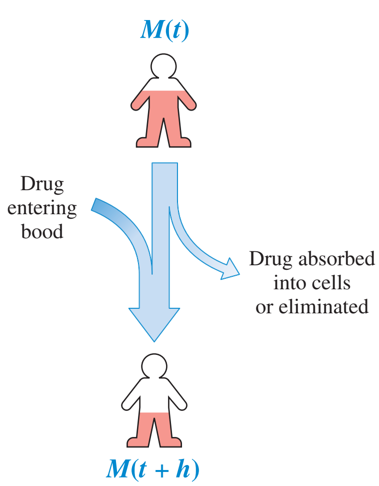


Figure 5.94 In the time interval $(t, t+h)$, the amount of drug present in the blood may increase if the drug enters the blood from the gut, or decrease if the drug is absorbed or eliminated from the blood.

This equation is represented graphically in Figure 5.94. The amount of drug that enters the blood in the time between t and $t+h$ will depend on how the drug is administered (whether it is taken as a pill, or by injection or by an intravenous line). We will discuss how to model some of these processes in the examples below. For now, let's focus on building a model for the amount that leaves the blood; this process, whether it is due to the drug being filtered out by the kidneys or by entering the cells, is called **elimination**. In general the amount of drug that is eliminated will be proportional to the length of the time interval h; that is, if we double the time interval between t and $t+h$, we would expect twice as much drug to be eliminated from the blood. The amount eliminated will depend on whether the drug has zeroth order or first order elimination kinetics. We met these concepts for discrete models in Chapter 2; but we will rewrite the definitions below in a way that is appropriate for differential equation models.

A drug has **zeroth order elimination kinetics** if the amount of drug that is eliminated between t and $t + h$ is independent of the amount of drug that is present. That is, if there is 10 mg present in the blood, or 100 mg present in the blood, the same amount is eliminated in the time interval h.

A drug has **first order elimination kinetics** if the amount of drug that is eliminated between t and $t+h$ is proportional to the amount of drug that is present. That is, a fixed fraction of the drug is eliminated in the time interval h. If there is 100 mg present in the blood, then 10 times more drug will be eliminated than if there were 10 mg present in the blood.

The particular type of kinetics that the drug has (zeroth order or first order) depends on the chemical composition of the drug, and how it is used by the body. This kind of chemical information will need to be given to you before you build the model. For a drug with zeroth order elimination kinetics, since the amount eliminated does not depend on M and is proportional to h:

$$\text{Amount eliminated between } t \text{ and } t+h \approx k_0 h$$

where k_0 is the rate of elimination (amount of drug eliminated in unit time).

For a drug with first order kinetics, since the amount eliminated must be proportional to both M and h:

$$\text{Amount eliminated between } t \text{ and } t+h \approx k_1 M h$$

where k_1, still called the rate of elimination, represents the fraction of drug eliminated in unit time.

The amount absorbed to the blood between time t and time $t+h$ will typically also be proportional to h, but its dependence on time may be complicated. Accordingly we write:

$$\text{Amount of drug entering the blood between } t \text{ and } t+h \approx A(t)h$$

where $A(t)$ is the rate of drug absorption; we will discuss different functions to represent drug absorption in the Examples below.

Just as in the population growth modeling, the amounts of drug that enter or are eliminated from the blood are given only approximately (hence the use of $\approx$ rather than $=$) by these expressions. In particular the amount of drug that is absorbed or eliminated is only proportional to h in the limit as $h \to 0$.

Putting together the expressions for drug absorption and elimination we obtain the following equations:

$$\frac{M(t+h)-M(t)}{h} = A(t) - k_0$$

for a drug with zeroth order elimination kinetics, and:

$$\frac{M(t+h)-M(t)}{h} = A(t) - k_1 M(t)$$

for a drug with first order elimination kinetics. Again we consider the limit as $h \to 0$ and we obtain a differential equation for $\frac{dM}{dt}$:

$$\frac{dM}{dt} = A(t) - k_0 \tag{5.40}$$

for a drug with zeroth order elimination kinetics, and

$$\frac{dM}{dt} = A(t) - k_1 M \tag{5.41}$$

for a drug with first order elimination kinetics. To solve these differential equations (that is, to find $M(t)$) we must also be given the value of M at some time t_0: i.e., the initial condition for the differential equation. In the following examples we will focus on different forms that the differential equation can take, depending on the function $A(t)$.

EXAMPLE 3 Acetaminophen (often sold under the brand name Tylenol) is a pain-killer and fever-reducing drug. A patient takes a pill, and sometime later, the contents of the pill have been completely absorbed into her blood. The concentration of drug in blood plasma is 10 μg/ml once the drug is completely absorbed. Acetaminophen has first order elimination kinetics, and the rate of elimination is $k_1 = 0.347$ hr^{-1} according to ***Prescott (1980)***. Determine the function $M(t)$ that represents the concentration of acetaminophen in the patient's blood. At what time does the concentration drop to one-tenth of its starting value?

Solution Let $t = 0$ be the time at which the drug has been totally absorbed into the patient's blood. There is no further absorption after this time, so $A(t) = 0$ in our differential equation. We are also told that the drug has first order elimination kinetics, so we model its elimination using Equation (5.41)

$$\frac{dM}{dt} = -k_1 M \quad \text{with initial condition} \quad \mathrm{M}(0) = 10\mu\mathrm{g/ml}.$$

where $k_1 = 0.347$ hr^{-1}. This is actually the same equation that we encountered when we modeled population growth in 5.9.1. Based off our solution to that equation in Equation (5.38) we can write down the solution now:

$$M(t) = M(0)e^{-k_1 t} = 10e^{-0.347t}$$

where t is measured in hours and $M(t)$ in μg/ml. We are asked to calculate when the concentration of drug drops to one-tenth of its starting value, that is, to find t^*, for which $M(t^*) = \frac{1}{10}M(0)$. Finding t^* requires that we solve:

$$M(0)e^{-k_1 t^*} = \frac{1}{10}M(0) \quad \Longrightarrow \quad t^* = -\frac{1}{k_1}\ln\left(\frac{1}{10}\right) = \frac{\ln 10}{0.347} = 6.6 \text{ hrs.}$$

As context for this result, typically the directions for acetaminophen recommend taking one dose no more than every four to six hours. •

EXAMPLE 4 Type I diabetes is a condition in which the body stops making insulin, a hormone that controls the level of glucose in the blood. Traditionally patients with this condition would inject themselves with insulin after each meal. During the past twenty years, more and more patients have been treated using an insulin pump. This is an implanted device that automatically infuses insulin into the patient's blood. We will build a simple model for the action of this device. First we will assume that the pump infuses insulin into the patient's blood at a constant rate (in most insulin pumps, the insulin output can be varied so that it is highest around meals). That is $A(t) = a$, a constant. According to ***Lauritzen et al. (1983)*** a typical value for a is $a = 0.84$ IU/hr. (IU or International Units are a unit of the amount of a drug that is used in pharmacology. For insulin 1 IU = 0.0347 mg.) Also according to *Lauritzen et al.*, insulin has first order elimination kinetics, with $k_1 = 6.8$ hr^{-1}. Using this information, and assuming that at time $t = 0$ there is no insulin in the patient's blood, find a function, $M(t)$, representing how the level of insulin in the patient's blood changes with time. What happens to the patient's insulin level as $t \to \infty$?

Solution First we input all of the information into our differential equation. Since the drug is absorbed at constant rate, and has first order elimination kinetics, we model the amount of insulin in the patient's blood using (5.41):

$$\frac{dM}{dt} = a - k_1 M \text{ with initial condition } M(0) = 0. \tag{5.42}$$

where $a = 0.84$ IU/hr and $k_1 = 6.8$ hr^{-1}. This equation can be solved using the methods from Chapter 8. We will state the solution here:

$$M(t) = \frac{a}{k_1}\left(1 - e^{-k_1 t}\right).$$

We can check that this function satisfies the differential equation by substituting it into the left- and right-hand sides of (5.42):

$$\frac{dM}{dt} = ae^{-k_1 t} \quad \text{and} \quad a - k_1 M = a - k_1 \cdot \frac{a}{k_1}\left(1 - e^{-k_1 t}\right) = ae^{-k_1 t},$$

so we obtain the same answer when we substitute M into the left-hand side as for the right-hand side. This function $M(t)$ also satisfies the initial condition: $M(0) = \frac{a}{k_1}(1 - e^0) = 0$.

We see that unlike Example 3, constant addition of insulin by the pump means that $M(t) \not\to 0$ as $t \to \infty$; instead the level of insulin grows over time, and converges to $\frac{a}{k_1}$. Notice that the dependence of this long time level upon the coefficients a and k_1 makes sense based on the roles that they play in the model. The long time level of insulin is proportional to a: a represents the rate that insulin is added to the blood by the pump. If insulin is added at a higher rate, then the level goes up. The level of insulin is inversely proportional to k_1, which represents the rate at which insulin is broken down or eliminated by the body. If insulin is broken down faster (k_1 is increased) then the level of insulin goes down. Using the values given above, the long time limit is:

$$\frac{a}{k_1} = \frac{0.84 \text{ IU/hr}}{6.8/\text{hr}} = 0.12 \text{ IU}$$

Section 5.9 Problems

5.9.1 and 5.9.2

1. A population of *E. coli* bacteria grows exponentially with time. You believe that the mean time between divisions is $t_b = 40$ min, and that cell death occurs on average after $t_m = 100$ hr. The population starts with 1000 cells.

(a) Use Equation (5.38) to predict how many cells are present after 3 hours.

(b) In fact you measure that there are 40000 cells present in the population after 3 hours. Can the discrepancy between the model and data be explained by your estimate for t_b being wrong? Can it be explained by the estimate for t_m being wrong?

(c) Assuming that t_m is correct, calculate a revised estimate for t_b to make the mathematical model fit your experimental data.

2. A population of bacteria grows exponentially and has mean time between divisions $t_b = 2$ hours. Assume that cell death can be ignored (that is, $m = 0$).

(a) Sketch on the same axes (of $N(t)$ against t) the size of the population over the interval $0 < t < 6$ hours for two populations: **(i)** One that starts with $N(0) = 1000$ cells and **(ii)** a population that starts with $N(0) = 3000$ cells.

(b) Redraw the same plots from (a) but on semilogarithmic axes.

(c) A mutant strain of the bacterium divides two times slower than the wild type (original) strain. On the same axes (of $N(t)$ against t) sketch the number of cells in the population as a function of time t, for two populations: **(i)** one population of wild type cells and **(ii)** one population of mutant cells. Both populations start with 1000 cells.

(d) Redraw the same plots from (c) but on semilogarithmic axes.

3. A population of cells initially contains 1000 cells. Two hours later the population contains 3000 cells.

(a) Estimate the division time t_b for this population (you can assume that mortality may be neglected; that is, $m = 0$).

(b) At what time would we expect the size of the population to reach 6000 cells?

(c) If we did not neglect cell death (that is, $m \neq 0$), would our estimate for the division time t_b increase or decrease from the value given in (a)?

(d) If we did not neglect cell death (that is, $m \neq 0$), would our estimate for the time taken by the population to reach 6000 cells increase or decrease from the value given in (b)?

4. A population of cells starts growing at time $t = 0$. After $t = 2$ hours the population contains 2000 cells. After $t = 3$ hours it contains 5000 cells.

(a) Assuming that the population grows according to Equation (5.38), how many cells were present when the population started to grow?

(b) How many cells would you predict the population to contain after $t = 5$ hours?

5. A population of cells is grown in the presence of an antibiotic; the antibiotic stresses the cells and alters their division time and their life time. For unstressed cells the division time is $t_b = 1$ hour. Initially you assume that the division time remains the same in the presence of the antibiotic, and that stressing the cells only decreases their lifetime.

(a) You start a population with 2000 cells. Two hours later there are still 2000 cells present. Calculate the lifetime of the cells.

(b) In fact you find in an independent measurement that the real lifetime of cells in the presence of antibiotic is $t_m = 4$ hours. How would you explain the data from part (a)?

6. Initially you measure that a colony of bacterial cells contains 1000 cells. 2 hours later you measure the colony again, and count 2000 cells.

(a) How many cells would you expect the colony to contain 4 hours after the start of the experiment?

(b) In fact, you realize that the hemocytometer that you used to count the cells for both measurements is only accurate to 10%, meaning that if you count 1000 cells, the real number of cells is somewhere between $1000 - 100 = 900$ cells and $1000 + 100 = 1100$ cells. What is the largest possible number of cells in the colony 4 hours after the start of the experiment? And what is the smallest possible number of cells at 4 hours?

7. Initially you measure that a colony of bacterial cells contains 2000 cells. 2 hours later you measure the colony again, and count 4000 cells.

(a) How many cells would you expect the colony to contain 3 hours after the start of the experiment?

(b) In fact, you realize that the hemocytometer that you used to count the cells for both measurements is only accurate to 10%, meaning that if you count 1000 cells, the real number of cells is somewhere between $1000 - 100 = 900$ cells and $1000 + 100 = 1100$ cells. What is the largest possible number of cells in the colony 3 hours after the start of the experiment? And what is the smallest possible number of cells at 3 hours?

8. **Effect of Antibiotic on Bacterial Population Growth** You are exploring the effect that an antibiotic has on the growth of a population of bacteria. The bacteria are grown initially without any antibiotic. Then antibiotic is added after four hours. The growth in the number of bacteria is shown in Figure 5.95.

(a) Two hours after the antibiotics are applied the bacteria start to divide normally again. Sketch how you would expect $\log N$ to vary with time over an interval $0 \le t \le 10$ hours.

(b) During the interval $4 \le t \le 6$ hours the size of the population doesn't change. Which of the following statements must be true?

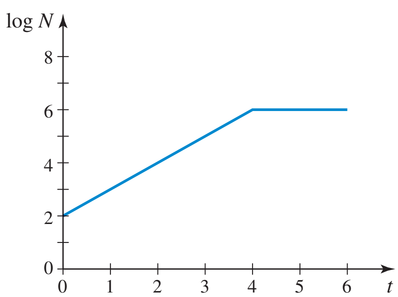


Figure 5.95 Bacterial population size for Problem 8.

(i) Cells are not dividing during this interval

(ii) Cells are not dying during this interval

(iii) Both (i) and (ii)

(iv) Neither (i) nor (ii)

5.9.3

9. **Concentration of Adderall in Blood** The drug Adderall (a proprietary combination of amphetamine salts) is used to treat ADHD. Adderall has first order kinetics for elimination, with elimination rate constant $k_1 = 0.08$ hr. We will assume that pills are quickly absorbed into the blood; that is, when the patient takes a pill their blood concentration of the drug immediately jumps.

(a) Assuming that the patient takes one pill at 8 am, the concentration in their blood after taking the pill is 33.8 ng/ml. Assuming that they take no other pills during the day, write down and then solve the differential equation that gives the concentration of drug in the blood, $M(t)$, over the course of a day. (*Hint*: You may find it helpful to define time t by the number of hours elapsed since 8 am.)

(b) What is the blood concentration just before the patient takes their next dose of the drug, at 8 am the next day?

(c) At what time during the day does the blood concentration fall to half of its initial value?

(d) In an alternative treatment regimen, the patient takes two half pills, one every 12 hours. They take the first pill at 8 am, with no drug in their system. Why would we expect their drug concentration to be 16.9 ng/ml immediately after taking the pill?

(e) What is the concentration in their blood at 8 pm?

(f) At 8 pm, the patient takes the other half pill. This increases the concentration of Adderall in their blood by 16.9 ng/ml (i.e., the concentration increases at 8 pm by 16.9 ng/ml). Derive a formula for the concentration in their blood as a function of time elapsed since 8 am. (You will need different expressions for the concentration for $0 < t < 12$ hrs and $12 < t < 24$ hrs.)

(g) Sketch the graph of $M(t)$ over this 24-hour period.

(h) What is the minimum blood concentration of Adderall over the entire 24-hour period?

10. In Problem 9 we neglected to consider the time delay between a pill being taken and the drug entering the patient's blood. In Chapter 8 we will introduce compartment models as models for drug absorption. We will show that a good model for a drug being absorbed from the gut is that the rate of drug absorption, $A(t)$, varies with time according to:

$$A(t) = Ce^{-kt} \quad , \quad t \ge 0$$

where $C > 0$ and $k > 0$ are coefficients that will depend on the type of drug, as well as varying between patients.

(a) Assume that the drug has first order elimination kinetics, with elimination rate k_1. Show that the amount of drug in the patient's blood will obey a differential equation:

$$\frac{dM}{dt} = Ce^{-kt} - k_1 M.$$

(b) Verify that a solution of this differential equation is:

$$M(t) = \frac{Ce^{-kt}}{k_1 - k} + ae^{-k_1 t},$$

where a is any coefficient, and we assume $k_1 \ne k$.

(c) To determine the coefficient a, we need to apply an initial condition. Assume that there was no drug present in the patient's blood when the pill first entered the gut (that is, $M(0) = 0$). Find the value of a.

(d) Let's assume some specific parameter values. Let $C = 2$, $k = 3$, and $k_1 = 1$. Show that $M(t)$ is initially increasing, and then starts to decrease. Find the maximum level of drug in the patient's blood.

(e) Show that $M(t) \to 0$ as $t \to \infty$.

(f) Using the information from (d) and (e), make a sketch of $M(t)$ as a function of t.

11. **Blood Alcohol Content** This question is about elimination of alcohol from the blood. The elimination of ethanol from the blood is generally believed to have **zeroth order** kinetics, meaning that the rate of elimination is a constant, independent of the concentration.

(a) Explain why it follows if no alcohol is being drunk, the blood alcohol concentration is described by the differential equation:

$$\frac{dc}{dt} = -k_0 \quad , \quad c(0) = c_0$$

where k_0 and c_0 are both constants, which vary from person to person and depend on how much alcohol the patient has consumed.

(b) Verify that a solution to this equation is $c(t) = c_0 - k_0 t$.

(c) At midnight after a boozy party in the math department Prof. R.'s blood alcohol concentration is $c = 1.500$ g/liter. Assuming $k_0 = 0.186$ g/(liter hr) (see Problem 12 for more discussion of this value), what is Prof. R.'s blood concentration at 2 am?

(d) When does Prof. R.'s blood alcohol concentration return to 0 g/liter?

(e) In the United States, the legal limit for driving is 0.8 g/liter. At what time can Prof. R. legally drive?

12. Blood Alcohol Content This question is about modeling blood alcohol levels (using the solutions of Problem 11). A driver is pulled over at 9 pm, suspected of driving drunk. The driver refuses to take a breathalyzer test for measuring blood concentration based on the alcohol content of the breath. They are taken to the police station, and their blood alcohol concentration is measured at 10 pm to be 0.65 g/liter.

(a) Were they driving over the safe limit? That is, was their blood alcohol concentration above 0.8 g/liter when they were pulled over? Assume that $k_0 = 0.186$ g/(liter hr).

(b) In reality values of k_0 vary from person to person. ***Al-Lanqawi et al. (1992)*** found that different individuals had different rates of elimination, k_0, ranging from 0.150 g/(liter hr) in some individuals to 0.238 g/(liter hr) in others. Based on the 10 pm measurement and using this range of values for k_0, what are the maximum and minimum possible values for the blood concentration of alcohol in the driver when they were pulled over?

(c) Explain how you could use another measurement, taken at 10:30 pm, to estimate k_0 and from there determine the real concentration in the driver's blood when they were pulled over.

13. Ibuprofen in Blood You are modeling the concentration of the drug ibuprofen (Advil) in a person's blood after they take one pill. We assume that after they take the pill the drug enters their blood effectively instantaneously. Ibuprofen has first order elimination kinetics.

(a) Explain why the concentration of drug in their blood satisfies a differential equation:

$$\frac{dc}{dt} = -k_1 c \quad \text{with} \quad c(0) = c_0$$

and explain what the constants k_1 and c_0 represent. (You do not need to solve the differential equation.)

(b) You measure the following data for the concentration of ibuprofen in a patient's blood

t (hrs)	$c(t)$ (mg/ liter)
0	40
1	30.3

Write down the solution to the differential equation from part (a). Then calculate the parameters c_0 and k_1 that fit the model to this data.

5.10 Antiderivatives

In Section 5.9 we showed that mathematical models for biological systems very often take the form of differential equations. Up until now, we have written down the solution to any differential equation that we encountered. In this section, we will discuss a particular type of differential equation and address two important general questions: First, given a differential equation, how can we find its solutions? Second, given a solution of a differential equation, how do we know if it is the only one?

We will consider differential equations of the form

$$\frac{dy}{dx} = f(x)$$

That is, the rate of change of y with respect to x depends *only* on x. Our goal is to find functions y that satisfy $y' = f(x)$. We will see that if we can find one such function, then there is a whole family of functions with this property. If we want to pick out one of these functions, we need to specify an **initial condition**—that is we also need to be given a single point (x_0, y_0) on the graph of the function. If $y(x)$ satisfies both the differential equation and the initial condition, then we say it is the **solution** of the **initial-value problem**

$$\frac{dy}{dx} = f(x) \quad \text{with } y = y_0 \text{ when } x = x_0. \tag{5.43}$$

Let's look at an example before we begin a systematic treatment of the solution of differential equations of the form (5.43). Consider a population whose size at time t is denoted by $N(t)$, and assume that the growth rate is given by

$$\frac{dN}{dt} = t \quad \text{for } t > 0 \tag{5.44}$$

and $N(0) = 20$. Where t is the independent variable and N is the dependent variable. The same methods work whatever these variables are called. Then

$$N(t) = \frac{1}{2}t^2 + 20 \quad \text{for } t \geq 0 \tag{5.45}$$

is a solution of the differential equation (5.44) that satisfies the initial condition $N(0) = 20$. This is easy to check: First, note that $N(0) = \frac{1}{2} \cdot 0^2 + 20 = 20$. Second, by differentiating $N(t)$, we find that

$$\frac{dN}{dt} = \frac{d}{dt}\left(\frac{1}{2}t^2 + 20\right) = t \quad \text{for } t > 0$$

That is, $N(t) = \frac{1}{2}t^2 + 20$ satisfies (5.44) with $N(0) = 20$.

This example shows that if we have a function that we think solves a differential equation, we can verify that the function indeed satisfies the differential equation by differentiating it. The method also suggests that, in order to find solutions, we need to reverse the process of differentiation. This leads us to what is called an antiderivative, which is defined as follows:

> **Definition** A function F is called an **antiderivative** of f on an interval I if $F'(x) = f(x)$ for all $x \in I$.

How can we find antiderivatives? Let

$$f(x) = 3x^2 \quad \text{for } x \in \mathbf{R}$$

To find the antiderivative of $f(x) = 3x^2$, we need to find a function whose derivative is $3x^2$. We can guess an answer, namely,

$$F(x) = x^3 \quad \text{for } x \in \mathbf{R}$$

which certainly satisfies $F'(x) = 3x^2$. But this is not the only answer. For example, take $F(x) = x^3 + 4$. Then $F'(x) = 3x^2$; hence, $x^3 + 4$ is also an antiderivative of $3x^2$. In fact, $F(x) = x^3 + C$, $x \in \mathbf{R}$, where C is any constant, is an antiderivative of $3x^2$. (We will soon show that all antiderivatives of $f(x)$ are of the form $F(x) = 3x^2 + C$.) The function $f(x)$ and some of its antiderivatives are shown in Figure 5.96. Since all of these antiderivatives are the same up to the value of the constant C, they can be obtained from each other by vertical shifts.

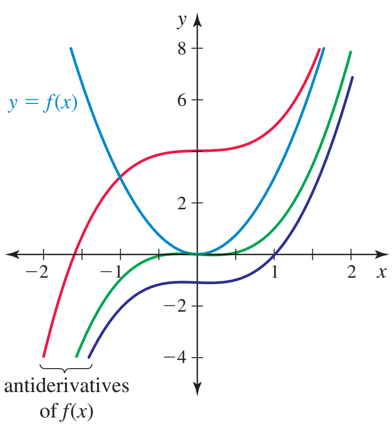


Figure 5.96 The function $f(x) = 3x^2$ and some of its antiderivatives. All antiderivatives are related to each other by vertical shifts.

Although we will learn rules that allow us to compute antiderivatives, this process is typically much more difficult than finding derivatives, and sometimes it takes ingenuity to come up with the correct answer; in addition, there are even cases where it is impossible to find an expression for an antiderivative.

We begin by recalling two corollaries of the mean-value theorem that will help us in finding antiderivatives. The first of these is Corollary 2 from Section 5.1:

> **Corollary 2** If f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), with $f'(x) = 0$ for all $x \in (a, b)$, then f is constant on $[a, b]$.

Note that Corollary 2 is the converse of the rule which states that $f'(x) = 0$ when $f(x) = c$, where c is a constant. It tells us that *all* antiderivatives of a function that is identically 0 are constant functions.

The next corollary, which you have not yet seen, tells us that functions with identical derivatives differ only by a constant.

> **Corollary 3** If $F(x)$ and $G(x)$ are antiderivatives of the continuous function $f(x)$ on an interval I, then there exists a constant C such that
>
> $$G(x) = F(x) + C \quad \text{for all } x \in I$$

Proof $F(x)$ and $G(x)$ are both antiderivatives of $f(x)$, meaning that $F'(x) = f(x)$ and $G'(x) = f(x)$. So: $F'(x) = G'(x)$, or $G'(x) - F'(x) = 0$. Since

$$0 = G'(x) - F'(x) = \frac{d}{dx}[G(x) - F(x)]$$

it follows from applying Corollary 2 to the function $G - F$, that $G(x) - F(x) = C$, where C is a constant. •

EXAMPLE 1 Find general antiderivatives for the given functions. Assume that all functions are defined for $x \in \mathbf{R}$.

(a) $f(x) = x^2$ **(b)** $f(x) = \cos x$ **(c)** $f(x) = e^x$.

Solution We find the antiderivatives by trying to guess a function that when differentiated will give $f(x)$.

(a) If $F(x) = x^3$, then $F'(x) = 3x^2$. That's almost what we need, but with the wrong coefficient in front. To fix the coefficient, we try instead: $F(x) = \frac{1}{3}x^3$. (By the rules of differentiation, if we multiply $F(x)$ by $1/3$, then $F'(x)$ is multiplied by $1/3$ also, canceling the unwanted 3.) Then $F'(x) = x^2$, which was what is needed. So $F(x) = \frac{1}{3}x^3$ is a particular antiderivative of x^2. Using Corollary 3, we find the general antiderivative simply by adding a constant; that is, the general antiderivative of $f(x) = x^2$ is the function $G(x) = \frac{1}{3}x^3 + C$, where C is a constant.

(b) If $F(x) = \sin x$, then $F'(x) = \cos x$. Hence, the general antiderivative of $f(x) = \cos x$ is the function $G(x) = \sin x + C$, where C is a constant.

(c) If $F(x) = e^x$, then $F'(x) = e^x$ so the general antiderivative of $f(x) = e^x$ is the function $G(x) = e^x + C$, where C is a constant. •

EXAMPLE 2 Find general antiderivatives for the given functions. (Assume the largest possible domain.)

(a) $f(x) = 3x^5$ **(b)** $f(x) = x^2 + 2x - 1$

(c) $f(x) = e^{2x}$ **(d)** $f(x) = \sec^2(3x)$

Solution Again we use some guesswork to find the solutions.

(a) $\frac{d}{dx}(x^6) = 6x^5$, which is a factor of 2 too large. To fix the coefficient we multiply by $\frac{1}{2}$, then: $\frac{d}{dx}\left(\frac{1}{2}x^6\right) = 3x^5$, so $F(x) = \frac{1}{2}x^6 + C$ is the general antiderivative of $f(x) = 3x^5$.

(b) Treat each term separately; i.e., find the antiderivatives of x^2, $2x$, and -1 separately. We already know that $\frac{d}{dx}(\frac{1}{3}x^3) = x^2$ (see Example 1(a)). Then we notice that $\frac{d}{dx}(x^2) = 2x$. Finally $\frac{d}{dx}(-x) = -1$. We can add all three of these terms together to show that $F(x) = \frac{1}{3}x^3 + x^2 - x$ is an antiderivative of $f(x) = x^2 + 2x - 1$ and $F(x) = \frac{1}{3}x^3 + x^2 - x + C$ is the general antiderivative. Note that there is only one overall constant C, there is no need to keep one constant for each term.

(c) $\frac{d}{dx}(e^{2x}) = 2e^{2x}$, which is a factor of 2 too large. To fix the coefficient, we multiply by $\frac{1}{2}$, then $\frac{d}{dx}\left(\frac{1}{2}e^{2x}\right) = e^{2x}$. So $F(x) = \frac{1}{2}e^{2x} + C$ is the general antiderivative of $f(x) = e^{2x}$.

(d) $\frac{d}{dx}(\tan(3x)) = 3\sec^2(3x)$ (apply the chain rule, with $u = 3x$) which is a factor of 3 too large. We fix the coefficient by multiplying by $\frac{1}{3}$. $F(x) = \frac{1}{3}\tan(3x) + C$ is therefore the general antiderivative of $f(x) = \sec^2(3x)$. •

Table 5-1 summarizes some of the rules for finding antiderivatives. We denote functions by $f(x)$ and $g(x)$ and their particular antiderivatives by $F(x)$ and $G(x)$, respectively. The general antiderivative is then obtained simply by adding a constant. The quantities a and k denote nonzero constants.

TABLE 5-1 Examples of antiderivatives

Function	***Particular Antiderivative***
$kf(x)$	$kF(x)$
$f(x)+g(x)$	$F(x)+G(x)$
$x^n,\ n \neq -1$	$\frac{1}{n+1}x^{n+1}$
$\frac{1}{x}$	$\ln\lvert x\rvert$
e^{ax}	$\frac{1}{a}e^{ax}$
$\sin(ax)$	$-\frac{1}{a}\cos(ax)$
$\cos(ax)$	$\frac{1}{a}\sin(ax)$
$\tan(ax)$	$-\frac{1}{a}\ln\lvert\cos(ax)\rvert$
$\sec^2(ax)$	$\frac{1}{a}\tan(ax)$

We can prove each of these statements by differentiating the expression on the right-hand side. For example, to show that $F(x) = -\frac{1}{a}\ln|\cos(ax)|$ is the antiderivative of $f(x) = \tan(ax)$, use the chain rule with $u = \cos(ax)$:

$$\frac{d}{dx}\left(-\frac{1}{a}\ln|\cos(ax)|\right) = \underbrace{-\frac{1}{a}\frac{1}{\cos(ax)}}_{F'(u)}\underbrace{(-a\sin(ax))}_{du/dx} = \tan(ax) \qquad F(u) = -\tfrac{1}{a}\ln|u|.$$

We can now return to our initial question, namely, how do we solve differential equations of the form (5.43)?

EXAMPLE 3 Find the general solution of

$$\frac{dy}{dx} = \frac{3}{x^2} - 2x^2, \quad x \neq 0. \qquad \text{Right-hand side is undefined if } x = 0.$$

Solution Finding the general solution of this differential equation means finding the antiderivative of the function $f(x) = \frac{3}{x^2} - 2x^2$. Using Table 5-1, we obtain

$$F(x) = \frac{3}{-1}x^{-1} - \frac{2}{3}x^3 = -\frac{3}{x} - \frac{2}{3}x^3$$

as a particular antiderivative. That is, the general solution is

$$y = -\frac{3}{x} - \frac{2}{3}x^3 + C, \quad x \neq 0$$

In Example 3, we found the general solution of the given differential equation. Often, we wish to select a particular solution; for instance, we may know that the solution has to pass through a specific point (x_0, y_0). Such a problem is called an *initial-value problem*, as explained at the beginning of this section. We return to the differential equation in (5.44).

EXAMPLE 4 Solve the differential equation that models the growth of a population with time:

$$\frac{dN}{dt} = t \quad \text{for } t \geq 0 \text{ with } N(0) = 20.$$

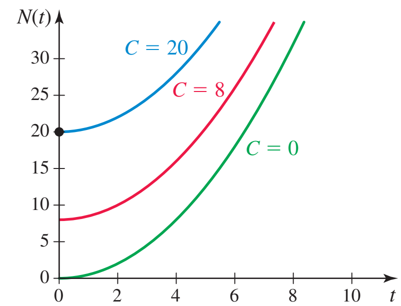


Figure 5.97 If $\frac{dN}{dt} = t$, then $N(t) = \frac{1}{2}t^2 + C$. Among these functions, the only one that also satisfies the initial condition $N(0) = 20$ is $N(t) = \frac{1}{2}t^2 + 20$.

Solution The general antiderivative of $f(t) = t$ is $F(t) = \frac{1}{2}t^2 + C$. Since $N(0) = 20$, we have

$$N(0) = \frac{1}{2}0^2 + C = 20, \qquad \text{or} \qquad C = 20$$

That is, the function

$$N(t) = \frac{1}{2}t^2 + 20, \quad t \geq 0$$

solves the initial-value problem, and because of Corollary 3, it is the only solution of this initial value problem. (See Figure 5.97.) •

EXAMPLE 5 Solve the initial-value problem

$$\frac{dy}{dx} = -2x^2 + 3 \quad \text{for } x \in \mathbf{R} \text{ and } y(3) = 10$$

Solution The general antiderivative of $f(x) = -2x^2 + 3$ is $F(x) = -\frac{2}{3}x^3 + 3x + C$. Since we require:

$$F(3) = -\frac{2}{3}3^3 + (3)(3) + C = -9 + C = 10$$

it follows that $C = 19$. That is,

$$y = -\frac{2}{3}x^3 + 3x + 19, \quad x \in \mathbf{R}$$

solves the initial-value problem, and it is the only solution. •

EXAMPLE 6 An object that falls freely in a vacuum, close to the surface of the earth, has a constant acceleration of $g = 9.81\text{m/s}^2$. If the object is dropped from rest, find its velocity and the distance it has fallen t seconds after it was released.

Solution If the distance traveled by the object is $s(t)$, then the velocity $v(t)$ is given by $v(t) = \frac{d}{dt}s(t)$ and its acceleration is given by $a(t) = \frac{d}{dt}v(t) = \frac{d^2}{dt^2}s(t)$ (see Section 4.7). We wish to solve the initial-value problem

$$\frac{d}{dt}v(t) = g \quad \text{with } v(0) = 0.$$

$v(0) = 0$ since object starts from rest

A general solution is

$$v(t) = gt + C$$

Since $v(0) = 0$, it follows that $C = 0$. Hence $v(t) = gt$ for $t \geq 0$. To find the distance traveled, note that $s(0) = 0$ and $\frac{ds}{dt} = v(t) = gt$. A general solution of this differential equation is

$$s(t) = \frac{1}{2}gt^2 + C$$

Since $s(0) = 0$, it follows that $C = 0$. Thus,

$$s(t) = \frac{1}{2}gt^2, \quad t \geq 0$$

•

EXAMPLE 7 In Section 5.9 we met the following model for the total amount of insulin contained in the blood of a patient whose type I diabetes is being treated with an implanted insulin pump:

$$\frac{dM}{dt} = a - k_1 M, \qquad M(0) = 0.$$

Solve this initial value problem.

Solution When we derived this initial value problem in Section 5.9, we were not able to solve it directly — we were given the solution and could only check that it satisfied both the differential equation and its initial condition. Initially this does not look like the sort of differential equation that antiderivatives can be used on. The differential equation gives $\frac{dM}{dt}$ as a function of M and not as a function of t. Although the differential equation is not written in the right form to be solved for $M(t)$, let us instead note that if M is monotonic, that is only increasing or only decreasing for $t \geq 0$, then we can write t as a function of M: $t = t(M)$. We will use the rules of implicit differentiation to rewrite our differential equation as an equation for $\frac{dt}{dM}$. First recall from Chapter 4 that if $t = t(M)$, then $\frac{dt}{dM} = \frac{1}{\frac{dM}{dt}}$, so we can rewrite our equation as:

$$\frac{1}{\frac{dt}{dM}} = a - k_1 M \quad \Longrightarrow \quad \frac{dt}{dM} = \frac{1}{a - k_1 M}$$

with our initial condition being that the solution must pass through $(t, M) = (0, 0)$. We must guess the antiderivative: remembering that the antiderivative of $f(x) = \frac{1}{x}$ is $F(x) = \ln|x|$, our first guess for the antiderivative is $\ln|a - k_1 M|$. Now:

$$\frac{d}{dM} \ln|a - k_1 M| = \frac{-k_1}{a - k_1 M} \qquad \text{Chain rule with } u = a - k_1 M, F(u) = \ln|u|$$

So we are correct except for an additional factor of $-k_1$. We fix the coefficient by dividing our antiderivative by $-k_1$, giving a general antiderivative

$$t(M) = -\frac{1}{k_1} \ln|a - k_1 M| + C$$

Since our solution is for t in terms of M, and we want M in terms of t we rearrange the equation:

$$\ln|a - k_1 M| = -k_1 t - k_1 C = -k_1 t + C_1$$

Because C is an arbitrary constant, we can define a new arbitrary constant $C_1 = -k_1 C$. Then, by exponentiating both sides, we obtain:

$$|a - k_1 M(t)| = e^{-k_1 t + C_1} = e^{-k_1 t} e^{C_1} = C_1 e^{-k_1 t}$$

Because C_1 is an arbitrary constant, we can rewrite e^{C_1} as C_2, which is still an arbitrary constant. To remove the absolute value signs rewrite the equation as:

$$a - k_1 M = \pm C_2 e^{-k_1 t} = C_3 e^{-k_1 t}$$

Again C_2 is an arbitrary constant, so we define $C_3 = \pm C_2$, which is still an arbitrary constant. On rearranging we obtain:

$$M(t) = \frac{a}{k_1} - \frac{C_1}{k_1} e^{-k_1 t}$$

To calculate the unknown constant C_3, apply the initial condition $M(0) = 0$, which gives: $0 = \frac{a}{k_1} - \frac{C_3}{k_1} e^{-k_1 0} = \frac{a}{k_1} - \frac{C_3}{k_1}$ or $C_3 = a$, so:

$$M(t) = \frac{a}{k_1}\left(1 - e^{-k_1 t}\right).$$

•

Section 5.10 Problems

In Problems 1–40, find the general antiderivative of the given function.

1. $f(x) = x^2 - 4x$

2. $f(x) = 5(1 - x^2)$

3. $f(x) = x^2 + 3x - 4$

4. $f(x) = 3(x^2 - x^4)$

5. $f(x) = (x - 1)(x + 1)$

6. $f(x) = x^3 + x^2 - 5x$

7. $f(x) = 4x^3 - 2x + 3$

8. $f(x) = x - 3x^2 + 3x^3 - x^4$

9. $f(x) = 1 - \dfrac{1}{x} + \dfrac{1}{x^2}$

10. $f(x) = x^2 - \dfrac{2}{x^2} + \dfrac{3}{x^3}$

11. $f(x) = 1 + \dfrac{1}{x^2}$

12. $f(x) = x^3 - \dfrac{1}{x^3}$

13. $f(x) = \dfrac{1}{1 + x}$

14. $f(x) = \dfrac{x}{1 + x}$

15. $f(x) = 5x^4 + \dfrac{5}{x^4}$

16. $f(x) = x^7 + \dfrac{1}{x^7}$

17. $f(x) = \dfrac{1}{1 + 2x}$

18. $f(x) = \dfrac{1}{1 + 3x}$

19. $f(x) = e^{-3x}$

20. $f(x) = e^{-x/2} - e^{-2x}$

21. $f(x) = 2e^{2x}$

22. $f(x) = -3e^{-4x}$

23. $f(x) = \dfrac{1}{e^{2x}}$

24. $f(x) = \dfrac{3}{e^{-x}}$

25. $f(x) = \sin(2x)$

26. $f(x) = \cos(3x)$

27. $f(x) = \sin\left(\dfrac{x}{3}\right) + \cos\left(\dfrac{x}{3}\right)$

28. $f(x) = \cos\left(\dfrac{x}{5}\right) - \sin\left(\dfrac{x}{5}\right)$

29. $f(x) = 2\sin\left(\dfrac{\pi}{2}x\right) - 3\cos\left(\dfrac{\pi}{2}x\right)$

30. $f(x) = -3\sin\left(\dfrac{\pi}{3}x\right) + 4\cos\left(-\dfrac{\pi}{4}x\right)$

31. $f(x) = \sec^2(2x)$

32. $f(x) = \sec^2(4x)$

33. $f(x) = \tan\left(\dfrac{x}{3}\right)$

34. $f(x) = \tan\left(\dfrac{x}{4}\right)$

35. $f(x) = \cos^2 x + 1$

36. $f(x) = \cos^2 x - \sin^2 x$

37. $f(x) = x^{-7} + 3x^5 + \sin(2x)$

38. $f(x) = 2e^{-3x} + \sec^2\left(\dfrac{x}{2}\right)$

39. $f(x) = \sec^2(3x - 1) + \dfrac{x^2 - 3}{x}$

40. $f(x) = 5e^{3x} - \sec^2(x - 3)$

In Problems 41–46, assume that a is a positive constant. Find the general antiderivative of the given function.

41. $f(x) = \dfrac{e^{(a+1)x}}{a}$

42. $f(x) = \sin^2(ax + 1)$

43. $f(x) = \dfrac{1}{ax + 3}$

44. $f(x) = \dfrac{a}{a + x}$

45. $f(x) = e^{ax}$

46. $f(x) = \dfrac{e^{-ax} + e^{ax}}{2a}$

In Problems 47–58, find the general solution of the differential equation.

47. $\dfrac{dy}{dx} = \dfrac{2}{x} - x,\ x > 0$

48. $\dfrac{dy}{dx} = \dfrac{2}{x^3} - x^3,\ x > 0$

49. $\dfrac{dy}{dx} = x(1 + x),\ x > 0$

50. $\dfrac{dy}{dx} = e^{x+1},\ x > 0$

51. $\dfrac{dy}{dt} = t(1 - t),\ t \geq 0$

52. $\dfrac{dy}{dt} = t^2(1 + t^2),\ t \geq 0$

53. $\dfrac{dy}{dt} = e^{-t/2},\ t \geq 0$

54. $\dfrac{dy}{dt} = 1 - e^{-2t},\ t \geq 0$

55. $\dfrac{dy}{ds} = \sin(\pi s),\ 0 \leq s \leq 1$

56. $\dfrac{dy}{ds} = \cos(2\pi s),\ 0 \leq s \leq 1$

57. $\dfrac{dy}{dx} = \frac{1}{1-x},\ x > 1$

58. $\dfrac{dy}{dx} = \frac{1}{x+1},\ x > -1$

In Problems 59–72, solve the initial-value problem.

59. $\dfrac{dy}{dx} = 3x^2$, for $x \geq 0$ with $y(0) = 1$

60. $\dfrac{dy}{dx} = \dfrac{x^2}{3}$, for $x \geq 0$ with $y(0) = 2$

61. $\dfrac{dy}{dx} = \sqrt{x}$, for $x \geq 0$ with $y(1) = 2$

62. $\dfrac{dy}{dx} = \dfrac{2}{\sqrt{x}}$, for $x \geq 1$ with $y(4) = 3$

63. $\dfrac{dN}{dt} = \dfrac{1}{t}$, for $t \geq 1$ with $N(1) = 10$

64. $\dfrac{dN}{dt} = \dfrac{t + 2}{t}$, for $t \geq 1$ with $N(1) = 2$

65. $\dfrac{dW}{dt} = e^t$, for $t \geq 0$ with $W(0) = 1$

66. $\dfrac{dW}{dt} = e^{-3t}$, for $t \geq 0$ with $W(0) = 2$

67. $\dfrac{dW}{dt} = \exp(t + 1)$, for $t \geq 0$ with $W(0) = 2/3$

68. $\dfrac{dW}{dt} = e^{-5t}$, for $t \geq 0$ with $W(0) = 1$

69. $\dfrac{dT}{dt} = \sin(\pi t)$, for $t \geq 0$ with $T(0) = 3$

70. $\dfrac{dT}{dt} = \cos(\pi t)$, for $t \geq 0$ with $T(0) = 3$

71. $\dfrac{dy}{dx} = \dfrac{e^{-x} + e^x}{2}$, for $x \geq 0$ with $y = 0$ when $x = 0$

72. $\dfrac{dN}{dt} = t^{-1/3}$, for $t > 0$ with $N(0) = 60$

Solve Problems 73–78 by rewriting the differential equation as an equation for $\frac{dx}{dy}$:

73. $\frac{dy}{dx} = \frac{1}{y}$, for $x \geq 1$ with $y(1) = 1$

74. $\frac{dy}{dx} = 1 - y$, for $x \geq 0$ with $y(0) = 0$

75. $\frac{dy}{dx} = e^y$, for $x \geq 0$ with $y(0) = 0$

76. $\frac{dy}{dx} = \frac{1}{1-y}$, for $x \geq 0$ with $y(0) = 0$

77. $\frac{dy}{dx} = \frac{y}{y^2+1}$, for $x \geq 0$ with $y(0) = 1$

78. $\frac{dy}{dx} = \frac{y}{y+1}$, for $x \geq 0$ with $y(0) = 1$

79. Suppose that the length of a certain organism at age t is given by $L(t)$, which satisfies the differential equation

$$\frac{dL}{dt} = e^{-0.1t}, \quad t \geq 0$$

Find $L(t)$ if the limiting length L_∞ is given by

$$L_\infty = \lim_{t\to\infty} L(t) = 25$$

How big is the organism at age $t = 0$?

80. Fish Growth Fish are indeterminate growers; that is, their length $L(t)$ increases with age t throughout their lifetime. If we plot the growth rate dL/dt versus age t on semilog paper, a straight line with negative slope results, meaning that:

$$\frac{dL}{dt} = Ae^{-kt}$$

where $A > 0$ and $k > 0$ are both coefficients that depend on the species of fish, and the habitat that it is growing in.

(a) Find the solution for this differential equation (your solution will include A and k as unknown constants, as well as one additional unknown constant C from the antiderivative).

(b) Find the values for the constants A, k, C, that would fit the solution to the following data $L(0) = 5$, $L(1) = 10$, and

$$\lim_{t\to\infty} L(t) = 30.$$

(c) Graph the solution $L(t)$ as a function of t.

81. Recall the model of population growth from Section 5.9:

$$\frac{dN}{dt} = rN \quad \text{with} \quad N(0) = N_0$$

We have rewritten Equation (5.37) by defining $r = b - m$.

By rewriting the differential equation in terms of $\frac{dt}{dN}$, solve this initial value problem.

82. Blood Alcohol Content Elimination of ethanol from the blood is known to have zeroth order kinetics. Provided no more ethanol enters the blood, the concentration of ethanol in a person's blood will therefore obey the following differential equation:

$$\frac{dM}{dt} = -k_0$$

where for a typical adult $k_0 = 0.186$ g/liter/hr (***al-Lanqawi et al. 1992***).

(a) Explain why $M(t)$ can only obey the above differential equation if $M > 0$ (once M drops to 0, it is usual to assume that $\frac{dM}{dt} = 0$).

(b) If a person's blood alcohol concentration is 1.6 g/liter at midnight, what will their blood alcohol concentration be at 2 am? You may assume that she drinks no more alcohol after midnight.

(c) At what time will their blood alcohol concentration drop to 0 g/liter?

83. Circadian Rhythms Some microbes regulate their growth according to a circadian clock. This clock means that their growth rate fluctuates predictably over the course one day. For example the filamentous fungus *Neurospora crassa* grows almost twice as fast at night-time than during the day. ***Gooch, Freeman and Lakin-Thomas (2004)*** measured the change of growth rate over time. If growth rate is measured in mm/day then their data can be fit by the following relationship:

$$\frac{dL}{dt} = 38.4 + 2.4\cos(4\pi t) - 12\sin(2\pi t)$$

where $L(t)$ is the total size of the fungus, measured in mm, and t is the time measured in hours. Calculate:

(a) the total extra length added to the fungus between $t = 0$ and $t = 1$ hours.

(b) the total extra length added to the fungus between $t = 0$ and $t = 12$ hours.

(c) the total extra length added to the fungus between $t = 0$ and $t = 24$ hours.

84. Suppose that a drug is eliminated so slowly from the blood that its elimination kinetics can be essentially ignored. Then according to Section 5.9 the total amount of drug in the blood is given by a differential equation:

$$\frac{dM}{dt} = A(t)$$

where $A(t)$ is the rate of absorption. We will show in Chapter 8 that if the drug is absorbed into the blood from a pill in the patient's gut, then $A(t)$ is given by a function

$$A(t) = Ce^{-kt}$$

where $C > 0$ and $k > 0$ are constants that depend on the type of the drug being administered. Assume that at $t = 0$ there is no drug present in the patient's blood (i.e., $M(0) = 0$). Solve this initial value problem, and, using the methods from Section 5.6, sketch the graph of $M(t)$ against t.

Chapter 5 Review

Key Terms

Discuss the following definitions and concepts:

1. Global or absolute extrema
2. Local or relative extrema: local minimum and local maximum
3. Extreme-Value Theorem
4. Fermat's theorem
5. Mean-Value Theorem
6. Rolle's theorem
7. Increasing and decreasing function
8. Monotonicity and the first derivative
9. Concavity: concave up and concave down
10. Concavity and the second derivative
11. Diminishing return
12. Candidates for local extrema
13. Monotonicity and local extrema
14. The second-derivative test for local extrema
15. Inflection points
16. Inflection points and the second derivative
17. Optimization
18. L'Hôpital's rule

Review Problems

1. Suppose that

$$f(x) = xe^{-x}, \quad x \geq 0.$$

(a) Show that $f(0) = 0$, $f(x) > 0$ for $x > 0$, and

$$\lim_{x\to\infty} f(x) = 0.$$

(b) Find local and absolute extrema.

(c) Find inflection points.

(d) Use the information from parts (a)–(c) to graph $f(x)$.

2. Suppose that

$$f(x) = x \ln x, \quad x > 0.$$

(a) What value should be assigned to $f(0)$ at $x = 0$ so that $f(x)$ is continuous at $x = 0$?

(b) Find the intervals on which $f(x) > 0$ and the intervals on which $f(x) < 0$.

(c) Find extrema and inflection points of this function.

(d) Use the information from parts (b) and (c) to graph $f(x)$.

3. Consider the hyperbolic functions, which are defined as follows:

$$\sinh x = \frac{e^x - e^{-x}}{2}, \quad x \in \mathbf{R},$$

$$\cosh x = \frac{e^x + e^{-x}}{2}, \quad x \in \mathbf{R},$$

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad x \in \mathbf{R}.$$

(a) Show that $f(x) = \tanh x$, $x \in \mathbf{R}$, is a strictly increasing function on $\mathbf{R}$. Evaluate

$$\lim_{x\to-\infty} \tanh x$$

and

$$\lim_{x\to\infty} \tanh x.$$

(b) Use your results in (a) to explain why $f(x) = \tanh x$, $x \in \mathbf{R}$, is invertible, and show that its inverse function $f^{-1}(x) = \tanh^{-1} x$ is given by

$$f^{-1}(x) = \frac{1}{2} \ln \frac{1+x}{1-x}.$$

What is the domain of $f^{-1}(x)$?

(c) Show that

$$\frac{d}{dx} f^{-1}(x) = \frac{1}{1-x^2}.$$

(d) Use your result in (c) and the facts that

$$\tanh x = \frac{\sinh x}{\cosh x}$$

and

$$\cosh^2 x - \sinh^2 x = 1,$$

to show that

$$\frac{d}{dx} \tanh x = \frac{1}{\cosh^2 x}.$$

4. Oxygen Binding in Blood Hill's equation is used as a model for how hemoglobin in blood takes up oxygen from the tissues. According to Hill's equation the fraction of hemoglobin molecules that are bound to oxygen varies with the partial pressure (that is concentration) of oxygen, P, according to the following formula:

$$f(P) = \frac{P^m}{k^m + P^m}, \quad P \geq 0$$

where $k > 0$ and $m > 0$ are constants that vary from one animal to another, and also can change depending on the environment that the animal is living in.

(a) Show that whatever the values of k and m are, f is an increasing function of P.

(b) Show that $f(P) \to 1$ as $P \to \infty$.

(c) Suppose that $m \leq 1$: show that $f(P)$ is a concave down function.

(d) Suppose instead that $m > 1$: show that the curve of $f(P)$ against P has an inflection point, and determine whether the function goes from concave down to concave up, or conversely, at this inflection point.

(e) Using the information from parts (a)–(d), make a sketch of the shape of the curve of $f(P)$ against P, when $m = 3$ (the usual value for hemoglobin). In particular, explain how changing the value of k would change the shape of the function.

5. Let

$$f(x) = \frac{x}{1+e^{-x}}, \quad x \in \mathbf{R}$$

(a) Show that $y = 0$ is a horizontal asymptote as $x \to -\infty$.

(b) Show that $y \to \infty$ as $x \to \infty$.

(c) Show that

$$f'(x) = \frac{1 + e^{-x}(1+x)}{(1+e^{-x})^2}$$

(d) Use your result in (c) to show that $f(x)$ has exactly one local extremum at $x = c$, where c satisfies the equation

$$1 + c + e^c = 0$$

[*Hint*: Show that $f'(x) = 0$ if and only if $1 + e^{-x}(1+x) = 0$. Let $g(x) = 1 + e^{-x}(1+x)$. Show that $g(x)$ is strictly increasing for $x < 0$, that $g(-2) < 0$, and $g(0) > 0$. So $g(x) = 0$ has exactly one solution on $(-2, 0)$.

(e) (This part requires a numerical root finding method, such as the one introduced in Section 5.8.) There is no exact solution for the equation $1 + c + e^c = 0$. Solve this equation numerically. [*Hint*: From (d), you know that $c \in (-2, 0)$.]

(f) Show that $f(x) < 0$ for $x < 0$. [This implies that, for $x < 0$, the graph of $f(x)$ is below the horizontal asymptote $y = 0$.]

(g) Show that $x - f(x) > 0$ for $x > 0$, and that $x - f(x) \to 0$ as $x \to \infty$ [That is, as $x \to \infty$, the curve of $y = f(x)$ approaches the line $y = x$.]

(h) Use your results in (a)–(g) and the fact that $f(0) = 0$ and $f'(0) = \frac{1}{2}$ to sketch the graph of $f(x)$.

6. Tumor Growth The *Gompertz growth curve* is used to model the growth of tumors. It is assumed that tumor growth starts slow, then speeds up and then slows again as the tumor approaches some maximum size. If the size of the tumor is $L(t)$, the tumor growth rate varies with tumor size. The tumor growth rate is given by a function of L:

$$r(L) = aL \ln\left(\frac{K}{L}\right), \quad L > 0$$

where a and K are both positive constants that are different for different tumor types.

(a) Show that $\lim_{L \to 0+} r(L) = 0$.

(b) Show that $r(L)$ is positive for small values of L, and negative for large values of L. At what value of L does $r(L)$ switch from being negative to positive?

(c) The size of a tumor that grows according to the Gompertz model must obey a differential equation:

$$\frac{dL}{dt} = r(L) = aL \ln\left(\frac{K}{L}\right), \quad t \geq 0$$

Verify that the function:

$$L(t) = K \exp[-ce^{-at}]$$

solves this differential equation. Here $c > 0$ is any positive constant.

(d) We need an additional piece of data to pin down the value of the coefficient c. This information comes from the initial condition (the tumor size at $t = 0$). Assume that the initial tumor size is $L(0) = L_0$. Show that $L(0) = Ke^{-c}$ and, hence,

$$c = \ln \frac{K}{L_0}$$

(e) Show that $L(t) = K$ is a horizontal asymptote and explain why **(i)** $L(t)$ is increasing if $L(t) < K$, and **(ii)** $L(t)$ is decreasing if $L(t) > K$. (*Hint*: Use your answer from part (b) to determine whether the function is increasing or strictly decreasing).

(f) Draw a sketch of the trajectories of the function $L(t)$ for initial conditions **(i)** $L_0 < K$ and **(ii)** $L_0 > K$.

7. Monod Growth Model The Monod growth curve (introduced by ***Monod (1949)***) is a model for how the rate of growth of a population of organisms depends on the amount of nutrients that are available to them. If the amount of nutrient is denoted by x, then according to the Monod growth curve, the reproductive rate of the organisms will be:

$$r(x) = \frac{cx}{k + x}, \quad x \geq 0$$

where $c > 0$ and $k > 0$ are constants that take different values depending on the species of organism and on the type of nutrient.

(a) Show that $r(x) \to c$ as $x \to \infty$. The constant c is called the *saturation value*.

(b) Show that $r(x)$ is strictly increasing. Explain why this implies that $r(x) < c$.

(c) Show that if $x = k$, then $r(x) = c/2$ (that is, half the saturation value). (For this reason, the constant k is called the *half-saturation constant*.)

(d) Show that $r(x)$ is concave down for $x \geq 0$.

(e) Draw a sketch of $r(x)$ and mark on your sketch the location of the saturation value and the half saturation value.

(f) How is the shape of the curve for $r(x)$ affected by changing the constants c and k? Draw graphs for two different values of c and two different values of k.

8. Density-Dependent Population Growth In Section 5.9, we introduced a model for population growth. Our model predicts that a population of organisms will grow according to the differential equation:

$$\frac{dN}{dt} = rN \quad \text{with } N(0) = N_0 \tag{5.46}$$

where r is the reproductive rate; the net number of organisms added (that is # births – # deaths) in unit time, per member of the population.

(a) By rewriting this equation as a differential equation for $\frac{dt}{dN}$, derive the solution to this initial value problem: $N(t) = N_0 e^{rt}$

(b) Equation (5.46) predicts that the population will grow exponentially indefinitely. It overlooks the fact that real habitats have carrying capacities, that is, a maximum number of organisms that they can support, that eventually limit population growth. To incorporate the carrying capacity in the model, we can modify Equation (5.46) so that reproductive rate r depends on population size. In fact if the carrying capacity is 1 (measured in some units) let's assume that the reproductive rate is directly proportional to how far the population is below its carrying capacity; that is, $r(N) \propto 1 - N$. The logistic equation is then written as:

$$\frac{dN}{dt} = rN(1 - N) \quad \text{with } N(0) = N_0 \tag{5.47}$$

for some constant $r > 0$. Verify that the solution to this equation is the function:

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

for $t \geq 0$ (you may assume that $N_0 \neq 0$). You do not need to derive the solution; only show that it solves the initial value problem (5.47).

(c) Show $N \to 1$ as $t \to \infty$.

(d) Show that $N(t)$ is an increasing function of t if $N_0 < 1$. If $N_0 > 1$ then show that $N(t)$ is a decreasing function of t. What happens if $N_0 = 1$?

(e) Using the rules for implicit differentiation, show that:

$$\frac{d^2N}{dt^2} = r\frac{dN}{dt}(1 - 2N) = r^2 N(1 - N)(1 - 2N)$$

(f) Show that if $N_0 < 1/2$, then $N(t)$, $t \geq 0$, has exactly one inflection point; namely $(t^*, N(t^*))$, with $t^* > 0$ and $N(t^*) = \frac{1}{2}$ (that is, when the population is at half the carrying capacity). What happens if $1/2 < N_0 < 1$? What if $N_0 > 1$? Where is the function $N(t)$, concave up, and where is it concave down?

(g) Use the information from parts (c)–(f) to make a sketch of the graph of $N(t)$ for three different initial conditions: $N_0 < 1/2$, $1/2 < N_0 < 1$, and $N_0 > 1$. Mark the inflection point clearly if it exists, and show the $t \to \infty$ asymptote of the function.

(h) Explain in words what role the parameter r plays in your model. Then draw two graphs, showing the population growth for two populations having the same initial condition, but different values of r.

9. Evolution of Cooperative Behaviors

Bio Info • How cooperative behaviors can evolve is an open problem in biology. A cooperative behavior is one in which each organism performs a costly task, but the reward from performing this task is shared somewhat indiscriminately among organisms. For example, emperor penguins huddle together to keep warm during the cold Antarctic winter. The penguins at the edge of the flock endure the worst wind, but their bodies shield the penguins inside the flock. If penguins take it in turns to be on the outside, the cost (shield other penguins with your body) and the reward (being warmed and shielded from the wind) are shared quite evenly between penguins. Although cooperation benefits the population as a whole, it is also vulnerable to an effect called **defection**. If a single organism decides not to cooperate then it will still share in the rewards from cooperation, but will not need to pay the cost. For example, a penguin could choose to not spend any time at the edge of the flock, so it is shielded from the wind by the other penguins but does not take its turn shielding the others.

We will analyze a simple example of a cooperative system (we will study cooperative dynamics again in Chapter 8). Two organisms must choose whether to cooperate or to defect (not cooperate). If either organism cooperates then both organisms get a reward b. If neither organism cooperates then there is no reward. If both organisms cooperate then each must pay a cost $c/2$, but if only one organism cooperates, that organism pays a cost c, and the defecting organism pays no cost. If organism A interacts with organism B then the net reward to organism A can be summarized using a pay-off matrix.

Net Benefit to A	**B cooperates**	**B defects**
A cooperates	$b - c/2.$	$b - c.$
A defects	b	$0.$

Let's assume that a fraction p of the time B decides to cooperate, and a fraction $1 - p$ of the time it defects. Then if A cooperates the average net reward that it receives is:

$$\text{reward for cooperating} = p(b - c/2) + (1 - p)(b - c)$$

$$\begin{matrix}\text{reward for}\\ \text{cooperating}\end{matrix} = \begin{matrix}\text{Prob. } B\\ \text{cooperates}\end{matrix} \times \begin{matrix}\text{Net reward}\\ \text{to } A \text{ from } B\\ \text{cooperating}\end{matrix} + \begin{matrix}\text{Probability}\\ B \text{ defects}\end{matrix} \times \begin{matrix}\text{Net reward}\\ \text{to } A \text{ from } B\\ \text{defecting}\end{matrix}$$

Whereas if A defects, the average net reward that it receives is:

$$\begin{matrix}\text{Probability that}\\ B \text{ cooperates}\end{matrix} \times \begin{matrix}\text{Average reward}\\ \text{for cooperating}\end{matrix} + \begin{matrix}\text{Probability that}\\ B \text{ defects}\end{matrix} \times \begin{matrix}\text{Reward}\\ \text{for defecting}\end{matrix}$$

$$\text{reward for defecting} = pb + (1 - p) \cdot 0 = pb$$

So if A cooperates a fraction of the time, q, and defects a fraction of the time, $1 - q$, its average net reward is:

$$\begin{aligned} r(q) &= q \times \text{reward for cooperating} + (1 - q) \times \text{reward for defecting} \\ &= q\,(p(b - c/2) + (1 - p)(b - c)) + (1 - q)pb, \; 0 \le q \le 1 \end{aligned}$$

(a) What should A do? One possible goal for A may be to choose the value of q (that is, the fraction of times that it cooperates) that maximizes $r(q)$.

(i) Show that if $p < \frac{b-c}{b-c/2}$ then the optimal value of q (which optimizes $r(q)$) is $q = 1$: that is, A should always cooperate.

(ii) Show that if $p > \frac{b-c}{b-c/2}$ then the optimal value of q (which optimizes $r(q)$) is $q = 0$: that is, A should always defect.

(b) Let's suppose now that A and B are from the same species and that they follow exactly the same strategy (that is, $p = q$). Show that the average reward to A is then given by:

$$r(p) = \left(b - \frac{c}{2}\right)(2 - p)p$$

If $b > \frac{c}{2}$, show that the optimal strategy for each organism to follow is $p = 1$ (always cooperate).

(c) What happens if $b < \frac{c}{2}$? Explain in words why $b > \frac{c}{2}$ and $b < \frac{c}{2}$ might lead to different optimal strategies.

10. Drug Concentration You are modeling how the amount $M(t)$ of a particular drug in the blood changes with time t. This drug has zeroth order elimination kinetics, meaning that provided $M(t) > 0$, the drug is eliminated at a constant rate, k_0. The drug is absorbed from a pill in the patient's gut. We will build a model for the absorption process in Chapter 8. It will be shown then that the rate of absorption has the following form:

$$A(t) = Ce^{-at}$$

where $C > 0$ and $a > 0$ are coefficients that depend on the type of drug being absorbed, its dose, and the composition of the pill.

(a) Explain why, given the above information, the amount of drug in the patient's blood will obey a differential equation:

$$\frac{dM}{dt} = Ce^{-at} - k_0$$

provided $M \ge 0$.

(b) Solve the differential equation under the assumption that at time $t = 0$ there is no drug present in the patient's blood, i.e., $M(0) = 0$.

(c) Show that $M(t)$ is increasing for $t < \frac{1}{a}\ln(C/k_0)$ and decreasing for $t > \frac{1}{a}\ln(C/k_0)$.

(d) Is $M(t)$ concave up or concave down?

(e) Assume that $C = 6$, $k_0 = 1$, and $a = 2$. Make a sketch of your solution $M(t)$ as a functional time.

(f) (This part of the question requires use of a numerical root finding method, such as the one given in Section 5.8.) Assuming the combination of parameters from part (c), calculate the time at which the blood concentration of the drug drops to 0.

11. Michaelis-Menten Elimination Kinetics For many drugs, the kinetics of drug elimination are somewhere between zeroth order and first order: when the blood concentration of the drug is low a fixed fraction of the drug is eliminated each hour (first order kinetics), but when drug concentrations are high, a fixed amount of the drug is eliminated each hour (zeroth order kinetics). One model for varying rate of drug elimination is that the rate of elimination of a drug (amount removed in unit time), of which an amount $M(t)$ is present in the blood, is given by the Michaelis-Menten function:

$$r(M) = \frac{k_0 M}{M + \frac{k_0}{k_1}}, \quad M \ge 0$$

where $k_0 > 0$ and $k_1 > 0$ are coefficients that depend on the type of drug, and can also vary from patient to patient.

(a) Show that for all values of k_0 and k_1, $r(M)$ is an increasing function of M.

(b) Show that as $M \to \infty$, $r(M) \to k_0$. (k_0 is the saturation rate for drug elimination).

(c) Show that for small M, $r(M) \approx k_1 M$.

(d) Explain why the Michaelis-Menten function can be thought of as producing drug elimination kinetics that are intermediate between zeroth and first order (in the sense described in the introduction to this Problem).

(e) Show that $r(M) < k_0$ for all $M \geq 0$.

(f) Show that $r(M)$ has no inflection points for $M \geq 0$: is the function concave up or concave down?

(g) Using the information from parts (a)–(e), draw a sketch of the function $r(M)$ against M. What happens to the shape of the curve if we vary the coefficient k_1? What happens if k_0 is varied? Draw sketches of the curve for two different values of k_1, and for two different values of k_0.

(h) Assuming that no extra drug is absorbed into the blood, explain why the amount of drug in the blood must obey a differential equation:

$$\frac{dM}{dt} = -r(M) = -\frac{k_0 M}{M + \frac{k_0}{k_1}}$$

(i) Assuming parameter values $k_0 = k_1 = 1$ and $M(0) = 2$, solve this differential equation. Your solution may take the form of a function that implicitly relates t and M (that is, you don't need to write M explicitly as a function of t).

(j) Calculate the time value at which $M(t)$ drops to 10% of its starting values (that is, find t for which $M(t) = 0.2$).

(k) (This part of the question requires a numerical equation solving method, such as the one introduced in Section 5.8.) Find the value of $M(t)$ when $t = 2$.

12. **Velocity and Distance** A ball is thrown directly upward, with initial velocity 1 m/s at $t = 0$. At time t after it was thrown, the ball has risen to a height $s(t)$. Neglecting air resistance, the weight of the ball will cause it to decelerate at a constant rate $-g$, that is:

$$\frac{d^2 s}{dt^2} = -g$$

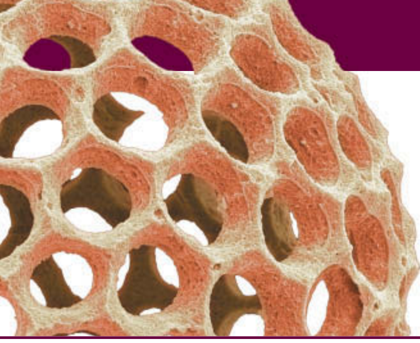
– sign because acceleration due to gravity is downward and height is measured upward

where $g = 9.81$ m/s^2 is the earth's gravitational constant.

(a) Find the time at which the ball reaches its maximum height.

(b) Find the maximum height.

(c) At what time does the ball return to $s = 0$?



CHAPTER

6

Integration

In this chapter, we introduce integration. Specifically, we will learn how to

- calculate areas;
- use integration to undo the effect of differentiation;
- use integration to calculate cumulative rates of change and average values.

The problem of finding the area of a region bounded by curves has been studied by mathematicians since the ancient Greeks. In this section we will introduce **integration**, which is a method for approximating the area bounded by curves using many tiny rectangles. The area under the curve made by a particular function $f(x)$ is then known as the **integral** of $f(x)$. This method was invented by Georg Bernhard Riemann in the nineteenth century. Calculating the area under a curve might seem like a strange topic to focus so much attention on. But we will show in Section 6.2 that integration is in some sense the opposite of differentiation—that if we know a function's derivative, we can use integration to calculate the function itself.

Once we have the techniques to calculate integrals (which will be described in this chapter and in Chapter 7), we will be able to then use integration to solve the differential equations introduced in Chapter 5 and many others, besides.

6.1 The Definite Integral

6.1.1 The Area Problem

We wish to find the surface area of the lake shown in Figure 6.1. To find the lake's area, we could overlay a grid on a map of the lake and count the number of squares that contain some part of the lake. The sum of the areas of these squares will then approximate the area of the lake. The finer the grid, the closer our approximation will be to the true area of the lake.

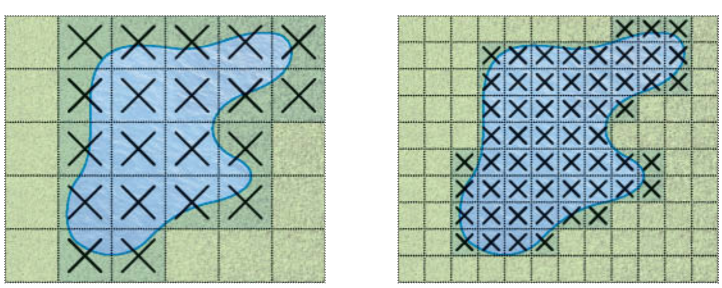


Figure 6.1 The outline of a lake with superimposed grid. Boxes with Xs contain some part of the lake. The finer the grid, the more accurately the area of the lake can be determined.

Dividing a region into smaller regions of known area is the basic principle we will employ in this section to find the area of a region bounded by a curve.

EXAMPLE 1

We will try to find the area of the region below the linear function $f(x) = x$ and above the x-axis between 0 and 1 by approximating it by rectangles. (See Figure 6.2.) Because the region whose area we want to calculate is triangular, we can calculate the area directly using the formula for the area of a triangle.

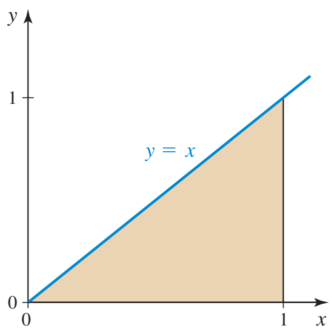


Figure 6.2 The region under the curve $y = x$ in Example 1.

$$\text{Area } = \frac{1}{2} \times 1 \times 1 = \frac{1}{2}. \qquad \text{Area} = \tfrac{1}{2} \times \text{base} \times \text{height}.$$

We will be able to compare this exact value to the value that we estimate by approximating the area using rectangles. To do this, we divide the interval $[0, 1]$ into n subintervals of equal length and approximate the area under $y = x$ by a sum of the areas of rectangles, whose widths are equal to the lengths of the subintervals and whose heights are the values of the function at the left endpoints of these subintervals. This technique is illustrated in Figure 6.3 with $n = 5$.

In the figure, each rectangle has width $1/5 = 0.2$. The height of the first rectangle is $f(0) = 0$, the height of the second rectangle is $f(0.2) = 0.2$, the height of the third rectangle is $f(0.4) = 0.4$, and so on. The area of a rectangle is the product of its width and height; adding up the areas of the approximating rectangles in Figure 6.3 yields

$$\begin{aligned}&(0.2)(0) + (0.2)(0.2) + (0.2)(0.4) + (0.2)(0.6) + (0.2)(0.8)\\ &= (0.2)\,[0 + (0.2) + (0.4) + (0.6) + (0.8)] = (0.2)(2.0) = 0.4.\end{aligned}$$

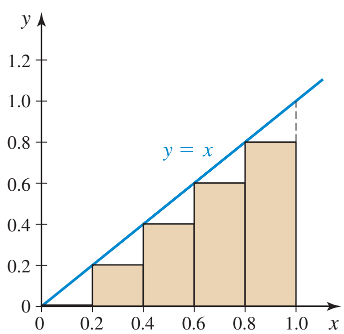


Figure 6.3 Approximation of the area under the curve $y = x$ from $x = 0$ to $x = 1$ by five rectangles.

Thus, an approximation of the area between 0 and 1 is 0.4 when we use five subintervals.

Our estimate for the area is not far from the actual area (0.5). Our estimate of the area is a little low because there are gaps between the curves $y = x$ and the rectangles that we use to approximate it. If we were to use more rectangles, then the gaps would be smaller, and we expect that our estimate for the area would become more and more accurate (see Figure 6.4).

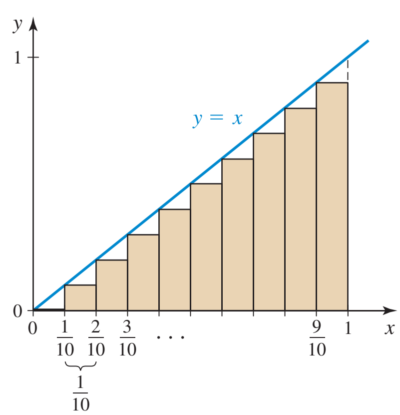


Figure 6.4 Approximating the same area as in Figure 6.3 using ten rectangles.

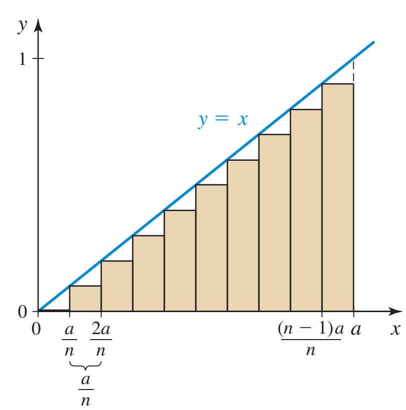


Figure 6.5 Approximation of the area under the curve $y = x$ from $x = 0$ to $x = a$ by n rectangles.

To see what happens as the number of rectangles becomes larger and larger we derive a general expression when the area under the curve $y = x$ between $x = 0$ and $x = a$ is approximated by n rectangles (illustrated in Figure 6.5). Since the interval $[0, a]$ has length a and the number of subintervals is n, each subinterval has length a/n. The left endpoints of successive subintervals are therefore $0, a/n, 2a/n, 3a/n, \ldots, (n-1)a/n$. The heights of the successive rectangles are then $f(0) = 0$, $f(a/n) = (a/n)$, $f(2a/n) = (2a/n), \ldots, f((n-1)a/n) = ((n-1)a/n)$. We denote the sum of the areas of the n rectangles by S_n, where S stands for "sum" and the subscript n denotes the number of subintervals. We find that

$$\begin{aligned} S_n &= \frac{a}{n} f(0) + \frac{a}{n} f\left(\frac{a}{n}\right) + \frac{a}{n} f\left(\frac{2a}{n}\right) + \cdots + \frac{a}{n} f\left(\frac{(n-1)a}{n}\right) \\ &= \frac{a}{n} \cdot 0 + \frac{a}{n} \cdot \frac{a}{n} + \frac{a}{n} \cdot \frac{2a}{n} + \cdots + \frac{a}{n} \cdot \frac{(n-1)a}{n} \\ &= \frac{a^2}{n^2}[1 + 2 + \cdots + (n-1)] \end{aligned}$$

To calculate S_n in a form that can be easily evaluated, we need an expression for the sum $T_n = 1+2+3+\cdots+(n-1)$ that appears in S_n. We can get this expression using a trick that is often attributed to the mathematician Carl Friedrich Gauss.

First (a little unintuitively), consider the sum $T_n + T_n$:

$$T_n + T_n = (1 + 2 + 3 + \cdots + (n-1)) + (1 + 2 + 3 + \cdots + (n-1))$$

Now we will pair off the terms in the two sums (i.e., take one number from the first sum and one number from the second sum). Instead of pairing the first term of the first sum with the first term of the second, we pair the first term of the first sum with the last term of the second sum, then the second term with the second from last term, the third term with the third from last term, and so on. So:

$$T_n + T_n = \underbrace{(1 + n - 1)}_{\text{first and last terms}} + \underbrace{(2 + n - 2)}_{\text{second and second from last terms}} + \underbrace{(3 + n - 3)}_{\text{third and third from last terms}} + \cdots + \underbrace{(k + n - k)}_{k\text{th and }k\text{th from last terms}} + \cdots + \underbrace{(n - 1 + 1)}_{\text{last and first terms}}$$

The kth pair in this sum consists of the kth term from the first sum and the kth from the last term from the second sum, these terms are respectively k and $n-k$. These two terms sum to $k + (n-k) = n$. Thus, each of the pairs sums to exactly n. There are $n-1$ pairs, so

$$\begin{aligned} T_n + T_n &= (n-1)n \qquad \text{Number of pairs} \times \text{sum of each pair.} \\ \Rightarrow T_n &= \frac{1}{2}(n-1)n \end{aligned}$$

Using the formula that we have derived for T_n, we can now calculate the total area of the rectangles, S_n:

$$S_n = \frac{a^2}{n^2} T_n = \frac{a^2}{n^2} \frac{(n-1)n}{2} = \frac{a^2}{2}\left(1 - \frac{1}{n}\right)$$

The finer the subdivision of $[0, a]$ (i.e., the larger n), the more accurate is the approximation, as we illustrated in Figures 6.3 and 6.4. Choosing finer and finer subdivisions means that we let n go to infinity. We find that

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{a^2}{2}\left(1 - \frac{1}{n}\right) = \frac{a^2}{2}(1) = \frac{a^2}{2} \qquad \lim_{n\to\infty}\left(1 - \tfrac{1}{n}\right) = 1$$

That is, the area under $y = x$ from 0 to a is equal to $a^2/2$. So when $a = 1$, the area is $\frac{1}{2}$, just as we obtain from the triangle area formula. •

6.1.2 The General Theory of Riemann Integrals

We will now develop a more systematic solution to the area problem. Although our approach will be similar to that in the previous subsection, we will look at more general situations. We will now allow the function whose graph makes up the boundary of the region of interest to take on negative values as well as positive ones.

Suppose we want to calculate the area between the curve $y = f(x)$ and the x-axis, between points $x = a$ and $x = b$ (see Figure 6.6). We can approximate this area by a series of n rectangles, each of which has width $\frac{b-a}{n}$.

The bases of these rectangles are a set of subintervals: $[x_0, x_1]$, $[x_1, x_2]$, $[x_2, x_3], \ldots [x_{n-1}, x_n]$, where $x_0 = a$, $x_n = b$, and $x_k - x_{k-1} = \frac{b-a}{n}$ (that is, each subinterval has the same width $w = \frac{b-a}{n}$). Our approximation to the area under the curve is then:

$$S_n = \underbrace{f(x_0)w}_{\substack{\text{height of} \\ \text{first rectangle} \\ \times \text{ width}}} + \underbrace{f(x_1)w}_{\substack{\text{height of} \\ \text{second rectangle} \\ \times \text{ width}}} + f(x_2)w + \cdots + f(x_{n-1})w$$

$$= w\,(f(x_0) + f(x_1) + f(x_2) + \cdots + f(x_{n-1})) \tag{6.1}$$

Notice that $f(x)$ can be either positive or negative. This is an important feature of the way the area under the curve is defined—regions of the function that are under the x-axis contribute negative area, while regions above the x-axis contribute positive area. So we count rectangles in which $f(x) < 0$ as having negative area and rectangles in which $f(x) > 0$ as having positive area (see Figure 6.7).

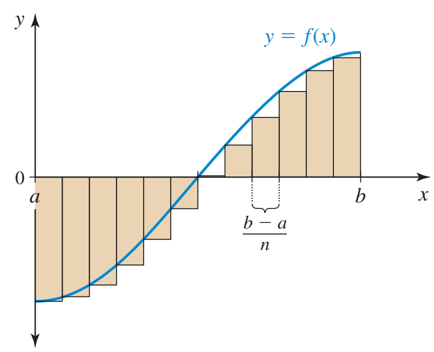


Figure 6.6 Approximating the area between the curve $y = f(x)$ and the x-axis.

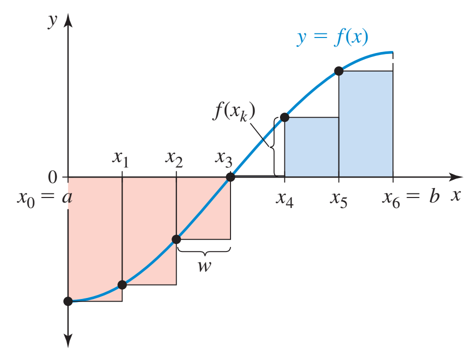


Figure 6.7 Approximating the area between $f(x)$ and the x-axis by a Riemann sum. Red rectangles have negative area, and blue rectangles have positive area.

Equation (6.1) is called a **Riemann sum** for f on $[a, b]$. Different numbers of rectangles will produce different approximations to the area between the curve and x-axis.

EXAMPLE 2 Use (a) two, (b) five, (c) ten subintervals to find the Riemann sum for $f(x) = x^2$ on $[0, 1]$.

Solution We partition $[0, 1]$ into equal subintervals. In case (a) there are two subintervals, each of length 0.5: $[0, 0.5]$ and $[0.5, 1]$. In case (b) the five subintervals are $[0, 0.2]$, $[0.2, 0.4]$, $[0.4, 0.6]$, $[0.6, 0.8]$, $[0.8, 1]$. In case (c) the ten subintervals are $[0, 0.1]$, $[0.1, 0.2]$, $[0.2, 0.3], \ldots$ etc. The Riemann sums in the three cases are:

(a) $S_2 = 0.5 \times (f(0) + f(0.5))$

$= 0.5 \times (0 + (0.5)^2) = 0.125$ Apply Eqn (6.1)

(b) $S_5 = 0.2 \times (f(0) + f(0.2) + f(0.4) + f(0.6) + f(0.8))$

$= 0.2 \times (0 + (0.2)^2 + (0.4)^2 + (0.6)^2 + (0.8)^2)$

$= 0.24$

(c)
$$\begin{aligned} S_{10} &= 0.1 \times (f(0) + f(0.1) + f(0.2) + \cdots + f(0.9)) \\ &= 0.1 \times (0 + (0.1)^2 + (0.2)^2 + \cdots + (0.9)^2) \\ &= 0.285 \end{aligned}$$

To obtain a better approximation, we need to choose finer and finer divisions of $[a, b]$ so that the rectangles fill out the region between the curve and the x-axis more and more accurately. A finer division means that both the number of rectangles becomes larger and their width becomes smaller. (Compare for example Figures 6.7 and 6.8.)

In the table below we calculate the area between $y = x^2$ and the x-axis from $x = 0$ to $x = 1$ using different numbers of rectangles (the cases $n = 2, 5, 10$ were discussed in Example 2). To calculate these areas we used a computer—and we will discuss how to do this in Section 7.5.

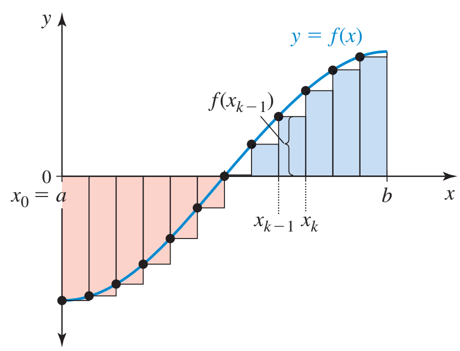


Figure 6.8 Approximating the area under the curve with a larger number of rectangles.

Number of rectangles, n	S_n
2	0.125
5	0.240
10	0.285
20	0.309
50	0.323
100	0.328
1000	0.333

The Riemann sums $\{S_n : n = 1, 2, 3, \ldots, \}$ form a sequence, and in this case the sequence appears to converge as $n \to \infty$. If the sequence S_n, converges, then we call the limit of the sequence the **definite integral** of f from a to b. In Example 1 we proved that the sequence S_n converged for the function $y = x$, but the argument for $y = x^2$ relies on calculating the sequence for some large values of n. We did not prove that the sequence converges. In Problems 7 and 8, you will prove that the limit exists using an argument similar to that made in Example 1.

Definite Integral Let the points $x_0 = a < x_1 < x_2 < x_3 < \cdots < x_n = b$ divide the interval $[a, b]$ into n even subintervals, of width $w = (b - a)/n$. The **definite integral** of f from a to b is

$$\begin{aligned} \int_a^b f(x)\,dx &= \lim_{n\to\infty} (f(x_0) + f(x_1) + f(x_2) + \cdots + f(x_{n-1}))w \\ &= \lim_{n\to\infty} \sum_{k=0}^{n-1} f(x_k)\,w \qquad \text{Using } \textstyle\sum\text{-notation from Chapter 2} \end{aligned}$$

if the limit exists, in which case f is said to be **(Riemann) integrable** on the interval $[a, b]$.

The symbol $\int$ is an elongated S (as in "sum") and was introduced by Leibniz. It is called the **integral sign**. In the notation $\int_a^b f(x)\,dx$ [read "the integral from a to b of $f(x)\,dx$"], $f(x)$ is called the **integrand**, the number a is the **lower limit of integration**, and b is the **upper limit of integration**. Although the symbol dx by itself has no meaning, it should remind you that, as we take the limit, the widths of the subintervals become ever smaller. The x in dx indicates that x is the independent variable and that we integrate with respect to x.

We have proved that the limit exists for $f(x) = x$ and argued (but not proved) that it exists for $f(x) = x^2$; that is, both of these functions are integrable. An important result tells us that if f is continuous on $[a, b]$, the definite integral of f on $[a, b]$ exists.

Theorem Integrability of continuous functions All continuous functions are Riemann integrable; that is, if $f(x)$ is continuous on $[a, b]$, then

$$\int_a^b f(x)\,dx$$

exists.

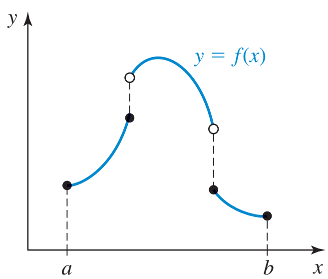


Figure 6.9 The function $y = f(x)$ is piecewise continuous and bounded on $[a, b]$.

The class of functions that are Riemann integrable is quite a bit larger than the set of continuous functions; for instance, functions that are both bounded (so that there exists an $M < \infty$ such that $|f(x)| < M$ for all x over which we wish to integrate) and piecewise continuous (continuous except for a finite number of discontinuities) are integrable. (See Figure 6.9.) We will be concerned primarily with continuous functions in this text; knowing that continuous functions are Riemann integrable will therefore suffice for the most part.

Note that $\int_a^b f(x)\,dx$ is a number that does not depend on x. We could have written $\int_a^b f(u)\,du$ (or any other letter in place of x) and meant the same thing.

EXAMPLE 3 Express the definite integral

$$\int_3^7 (x^2 - 1)\,dx$$

as a limit of Riemann sums.

Solution We have

$$\int_3^7 (x^2 - 1)\,dx = \lim_{n\to\infty} ((x_0^2 - 1) + (x_1^2 - 1) + \cdots + (x_{n-1}^2 - 1))\,w$$

where $w = 4/n$ and $x_0 = 3, x_1, x_2, \ldots, x_n = 7$ all divide the interval $[3, 7]$ up into n even subintervals, all of width w. •

EXAMPLE 4 Express the limit

$$\lim_{n\to\infty} (\sqrt{x_0 - 1} + \sqrt{x_1 - 1} + \cdots + \sqrt{x_{n-1} - 1})\,w$$

as a definite integral, where $x_0 = 2$, $x_n = 4$, and $x_0, x_1, x_2, \ldots, x_n$ all divide $[2, 4]$ into n subintervals of width $w = 2/n$.

Solution

$$\lim_{n\to\infty} (\sqrt{x_0 - 1} + \sqrt{x_1 - 1} + \cdots + \sqrt{x_{n-1} - 1})\,w = \int_2^4 \sqrt{x - 1}\,dx$$

Note that we cannot (yet!) calculate the numerical value of this definite integral. •

EXAMPLE 5 Evaluate

$$\int_0^2 x\,dx.$$

Solution We evaluated the Riemann sum and its limit for $y = x$ from 0 to a in Example 1 and found that

$$\lim_{n\to\infty} S_n = \int_0^a x\,dx = \frac{a^2}{2}.$$

With $a = 2$, we therefore have

$$\int_0^2 x\,dx = 2.$$

Geometric Interpretation of Definite Integrals. In Example 1, we computed the area of the region below $y = x$ and above the x-axis between 0 and a by approximating the region with n rectangles of equal width and then taking the limit as $n \to \infty$. More generally, we can now define the **area** of a region A above the x-axis, as shown in Figure 6.10, as the limiting value (if it exists) of the Riemann sum of approximating rectangles. (Note that an area is always a positive number.) This definition allows us to interpret the definite integral of a nonnegative function as an area.

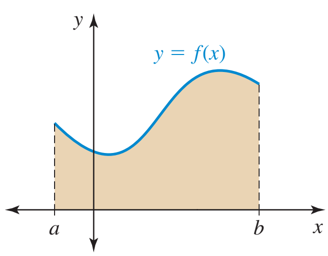


Figure 6.10 The area of a region under the curve of a positive function is given by the definite integral $\int_a^b f(x)\,dx$.

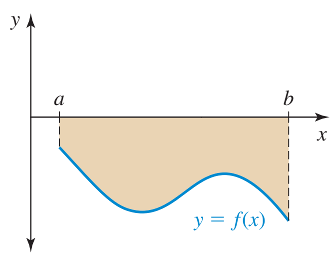


Figure 6.11 The area of a region above the curve of a negative function is given by $-\int_a^b f(x)\,dx$.

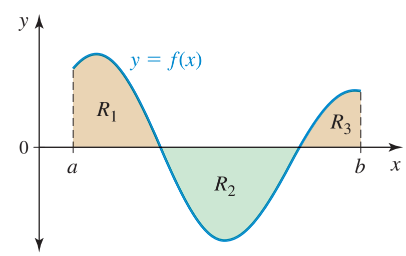


Figure 6.12 $\int_a^b f(x)\,dx = A_+ - A_-$, where $A_+ = R_1 + R_3$, and $A_-\ R_2$.

If $f(x) \le 0$ on $[a, b]$, then the definite integral $\int_a^b f(x)\,dx$ is less than or equal to 0 and its value is the negative of the area of the region above the graph of f and below the x-axis between a and b. (See Figure 6.11.) We refer to the integral as a "signed area." (A signed area may be either positive or negative.)

In general, a definite integral can thus be interpreted as a difference of areas, as illustrated in Figure 6.12. If A_+ denotes the total area of the region above the x-axis and below the graph of f (where $f \ge 0$) and A_- denotes the total area of the region below the x-axis and above the graph of f (where $f \le 0$), then

$$\int_a^b f(x)\,dx = A_+ - A_-$$

1. If f is integrable on $[a, b]$ and $f(x) \ge 0$ on $[a, b]$, then

$$\int_a^b f(x)\,dx = \left[\begin{array}{l}\text{the area of the region between the}\\ \text{graph of } f \text{ and the } x\text{-axis from } a \text{ to } b\end{array}\right]$$

2. If f is integrable on $[a, b]$, then

$$\int_a^b f(x)\,dx = [\text{area above } x\text{-axis}] - [\text{area below } x\text{-axis}]$$

EXAMPLE 6 Find the value of $\int_{-2}^{3}(2x+1)\,dx$ by interpreting it as the signed area of an appropriately chosen region.

Solution We graph $y = 2x + 1$ between -2 and 3. (See Figure 6.13.) The line intersects the x-axis at $x = -1/2$. The area of the region to the left of $-1/2$ between the graph of $y = 2x+1$ and the x-axis is denoted by A_-; the area of the region to the right of $-1/2$ between the graph of $y = 2x+1$ and the x-axis is denoted by A_+. Both areas can be calculated using the formula for the area of a triangle.

Then:

$$A_- = \frac{1}{2}\cdot\frac{3}{2}\cdot 3 = \frac{9}{4} \qquad \tfrac{1}{2}\times\text{base}\times\text{height}$$

$$A_+ = \frac{1}{2}\cdot\frac{7}{2}\cdot 7 = \frac{49}{4} \qquad \tfrac{1}{2}\times\text{base}\times\text{height}$$

Therefore,

$$\int_{-2}^{3}(2x+1)\,dx = A_+ - A_- = \frac{49}{4} - \frac{9}{4} = \frac{40}{4} = 10. \qquad \bullet$$

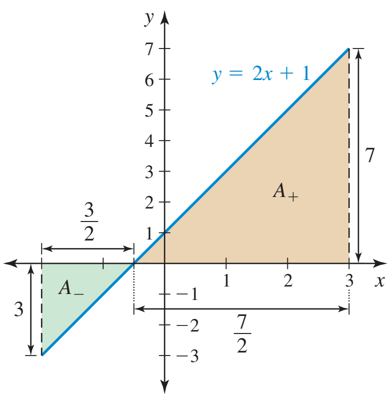


Figure 6.13 The area of R_1 is A_-; the area of R_2 is A_+.

EXAMPLE 7 Find the value of $\int_0^{2\pi}\sin x\,dx$ by interpreting it as the signed area of an appropriately chosen region.

Solution We graph $y = \sin x$ from 0 to 2π. (See Figure 6.14.) The function $f(x) = \sin x$ is symmetric about $x = \pi$. It follows from this symmetry that the area of the region below the graph of f and above the x-axis between 0 and π (denoted by A_+) is the same as the area of the region above the graph of f and below the x-axis between π and 2π (denoted by A_-). Therefore, $A_+ = A_-$ and

$$\int_0^{2\pi}\sin x\,dx = A_+ - A_- = 0. \qquad \bullet$$

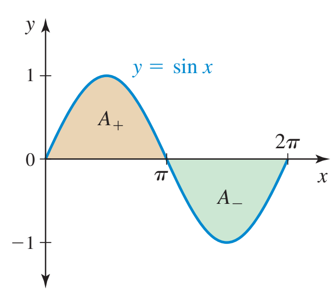


Figure 6.14 The graph of $f(x) = \sin x$, $0 \le x \le 2\pi$, in Example 7.

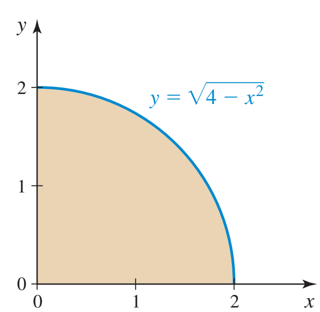


Figure 6.15 The graph of $f(x) = \sqrt{4-x^2}$ is a quarter-circle. The area of the shaded region is equal to $\int_0^2\sqrt{4-x^2}\,dx$.

EXAMPLE 8 Find the value of $\int_0^2\sqrt{4-x^2}\,dx$ by interpreting it as the signed area of an appropriately chosen region.

Solution The graph of $y = \sqrt{4-x^2}$, $0 \le x \le 2$, is the quarter-circle with center at $(0, 0)$ and radius 2 in the first quadrant. (See Figure 6.15.) Since the area of a circle with radius 2 is $\pi(2)^2 = 4\pi$, the area of a quarter-circle is $4\pi/4 = \pi$. Hence,

$$\int_0^2\sqrt{4-x^2}\,dx = \pi. \qquad \bullet$$

In this section we have simplified the definition of the definite integral as a limit of Riemann sums to consider only one type of Riemann sum, in which all rectangles have the same width, and the height of the kth rectangle is $f(x_{k-1})$ (i.e., the value of the function at the left end of the interval $[x_{k-1}, x_k]$ that is covered by the kth rectangle). A more careful definition of the definite integral allows rectangles to have different widths, provided the width of the largest rectangle goes to 0 as $n \to \infty$, and the height of the kth rectangle can be equal to $f(c_k)$ where c_k is *any* number in the interval $[x_{k-1}, x_k]$. However, Riemann sums are very seldom used to actually calculate definite integrals; their most important application is to derive the Fundamental Theorem of Calculus (which we will do in Section 6.2), and our simplified treatment is enough to be able to do that.

6.1.3 Properties of the Riemann Integral

In this subsection, we collect important properties that will help us to evaluate definite integrals.

Properties of the Definite Integral Assume that f is integrable over $[a, b]$. Then

1. $\int_a^a f(x)\,dx = 0$ and

2. $\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$

The first property says that the signed area between a and a is equal to 0; given that the width of the interval is equal to 0, we expect the area to be 0. The second property gives an orientation to the integral; for instance, if $f(x)$ is nonnegative on $[a, b]$, then $\int_a^b f(x)\,dx$ is nonnegative and can be interpreted as the area of the region between the graph of $f(x)$ and the x-axis from a to b. If we reverse the direction of the integration—that is, compute $\int_b^a f(x)\,dx$—we want the integral to be negative.

The next three properties follow from the definition of the definite integral as the limit of a sum of areas of approximating rectangles.

Properties of the Definite Integral Assume that f and g are integrable over $[a, b]$.

3. If k is a constant, then

$$\int_a^b k f(x)\,dx = k\int_a^b f(x)\,dx$$

4. $\int_a^b [f(x) + g(x)]\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$

5. If f is integrable over an interval containing the three numbers a, b, and c, then

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

We prove Properties (4) and (5) to illustrate how the definition of the definite integral can be used to prove its properties. We leave the Proof of (3) as an Exercise (see Problem 48).

Proof of (4) Divide the interval $[a, b]$ into n even subintervals using points $x_0 = a < x_1 < x_2 < \cdots < x_n = b$, with each subinterval having width $w = \frac{b-a}{n}$. We then use the definition of the definite integral. If $h(x) = f(x) + g(x)$, then

$$\begin{aligned}\int_a^b (f(x)+g(x))\,dx &= \int_a^b h(x)dx \\ &= \lim_{n\to\infty} w\,(h(x_0)+h(x_1)+\cdots+h(x_{n-1})) \quad \text{Provided this limit exists.} \\ &= \lim_{n\to\infty} w\,[(f(x_0)+g(x_0))+(f(x_1)+g(x_1))+\cdots+(f(x_{n-1})+g(x_{n-1}))] \\ &= \lim_{n\to\infty} [w(f(x_0)+f(x_1)+\cdots+f(x_{n-1}))+w(g(x_0)+g(x_1)+\cdots+g(x_{n-1}))] \quad \text{Separate terms involving } f \text{ and terms involving } g.\end{aligned}$$

Now if the separate limits $\lim_{n\to\infty} w(f(x_0) + f(x_1) + \cdots + f(x_{n-1}))$ and $\lim_{n\to\infty} w(g(x_0) + g(x_1) + \cdots + g(x_{n-1}))$ both exist, i.e., f and g are both integrable, then we can separate them out, so:

$$\begin{aligned}\int_a^b (f(x)+g(x))\,dx &= \lim_{n\to\infty} w\,(f(x_0)+f(x_1)+\cdots+f(x_{n-1})) + \lim_{n\to\infty} w\,(g(x_0)+g(x_1)+\cdots+g(x_{n-1})) \\ &= \int_a^b f(x)dx + \int_a^b g(x)\,dx. \quad \text{Use the definitions of } \textstyle\int_a^b f(x)dx \text{ and } \int_a^b g(x)dx \text{ as Riemann sums.}\end{aligned}$$

•

Proof of (5). (5) is an addition property. We will prove two special cases, namely $a < c < b$ and $a < b < c$ (illustrated in Figures 6.16 and 6.17). The definite integral $\int_a^b f(x)\,dx$ can be interpreted as the area between the graph of $f(x)$ and the x-axis from a to b if $a < c < b$. We see from Figure 6.16 that this area is composed of two areas: the area between the graph of $f(x)$ and the x-axis from a to c and the area between the graph of $f(x)$ and the x-axis from c to b. We can express this relationship mathematically as

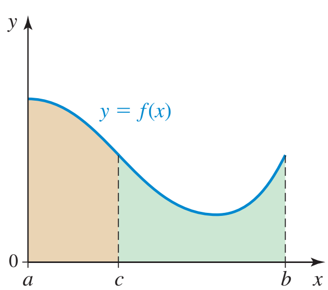


Figure 6.16 Property (5) when $a < c < b$.

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

which is Property (5).

If $a < b < c$ (Figure 6.17),

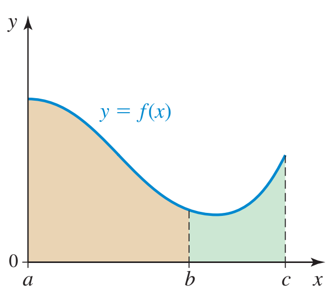


Figure 6.17 Property (5) when $a < b < c$.

$$\underbrace{\int_a^c f(x)\,dx}_{\text{Area from } a \text{ to } c} = \underbrace{\int_a^b f(x)\,dx}_{\text{Area from } a \text{ to } b} + \underbrace{\int_b^c f(x)\,dx}_{\text{Area from } b \text{ to } c}$$

So

$$\int_a^b f(x)dx = \int_a^c f(x)dx - \int_b^c f(x)dx$$

But because of Property (2),

$$\int_b^c f(x)\,dx = -\int_c^b f(x)\,dx$$

Therefore,

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

as stated in Property (5).

In fact Property (5) holds no matter how a, b, c are ordered on the number line (not just if $a < c < b$, as in Figure 6.16, or if $a < b < c$, as in Figure 6.17). You will consider some of the other cases in Problem 49. The next example shows how to use this property.

EXAMPLE 9 Given that $\int_0^a x\,dx = a^2/2$, evaluate

$$\int_1^4 2x\,dx$$

Solution

$$\int_1^4 2x\,dx = 2\int_1^4 x\,dx \qquad \text{Use Property (3)}$$

To evaluate $\int_1^4 x\,dx$, we use property (5) and write

$$\int_1^4 x\,dx = \int_1^0 x\,dx + \int_0^4 x\,dx$$

Since

$$\int_1^0 x\,dx = -\int_0^1 x\,dx$$

it follows that

$$\int_1^4 x\,dx = -\int_0^1 x\,dx + \int_0^4 x\,dx$$

and each of the integrals on the right hand side can be evaluated with the use of $\int_0^a x\,dx = a^2/2$.

$$\begin{aligned}\int_1^4 2x\,dx &= 2\left[\int_0^4 x\,dx - \int_0^1 x\,dx\right]\\ &= 2\left(\frac{4^2}{2} - \frac{1^2}{2}\right)\\ &= 2\left(8 - \frac{1}{2}\right) = 15.\end{aligned}$$

We can check the result from Example 9 using a geometrical interpretation of the integral. The area given by the integral is a trapezoid (see Problem 15). So its area is given by $\frac{1}{2}$ × base × (height at left end + height at right end) or:

$$\begin{aligned}\int_1^4 2x\,dx &= \frac{1}{2} \times \underbrace{3}_{\text{base}} \times \Big(\underbrace{2}_{\text{height at left}} + \underbrace{8}_{\text{height at right}}\Big)\\ &= \frac{1}{2} \times 3 \times 10 = 15.\end{aligned}$$

6.1.4 Order Properties of the Riemann Integral

The next three properties are called order properties. They allow us either to compare definite integrals or say something about how big or small a particular definite integral can be. We first state the properties and then explain what they mean geometrically.

Properties of Definite Integrals Assume that f and g are continuous on the interval $[a, b]$, and $a < b$.

6. If $f(x) \geq 0$ on $[a, b]$, then $\int_a^b f(x)\,dx \geq 0$.

7. If $f(x) \leq g(x)$ on $[a, b]$, then $\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx$.

8. If $m \leq f(x) \leq M$ on $[a, b]$, then

$$m(b-a) \leq \int_a^b f(x)\,dx \leq M(b-a)$$

with equality in the first inequality if and only if $f(x) = m$, and in the second inequality if and only if $f(x) = M$.

Property (6), illustrated in Figure 6.18, says that if f is nonnegative over the interval $[a, b]$, then the **definite integral over that interval** is also nonnegative. We can understand this statement from its geometric interpretation: If $f(x) \geq 0$ for $x \in [a, b]$, then $\int_a^b f(x)\,dx$ is the area between the curve and the x-axis between a and b. But an area must be a nonnegative number.

Property (7) is explained in Figure 6.19 when both f and g are positive functions on $[a, b]$. In this case both definite integrals can be interpreted as areas. We see that the function f has a smaller area than g has. Property (7) holds without the assumption that both f and g are positive, and we can draw an analogous figure for the general case as well. The definite integral then needs to be interpreted as a signed area.

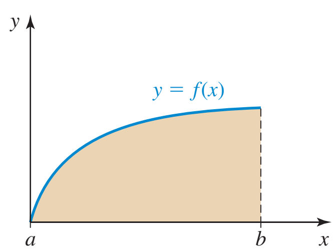


Figure 6.18 An illustration of Property (6).

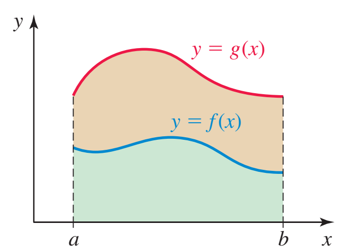


Figure 6.19 An illustration of Property (7).

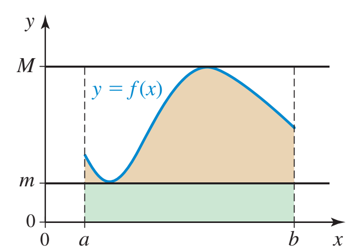


Figure 6.20 An illustration of Property (8).

Property (8) is explained in Figure 6.20 for $f(x) \geq 0$ in $[a, b]$. We see that the rectangle with height m is contained in the area between the graph of f and the x-axis, which in turn is contained in the rectangle with height M. Since $m(b-a)$ is the area of the small rectangle, $\int_a^b f(x)\,dx$ is the area between the graph of f and the x-axis for nonnegative f, and $M(b-a)$ is the area of the big rectangle, the inequalities in (8) follow. Note that the statement does not require that f be nonnegative; you can draw an analogous figure when f is negative on parts or all of $[a, b]$.

The next example illustrates how the order properties (6)–(8) are used.

EXAMPLE 10 Show that

$$0 \leq \int_0^\pi \sin x\,dx \leq \pi$$

Solution Note that $0 \leq \sin x \leq 1$ for $x \in [0, \pi]$. Using Property (6), we find that

$$\int_0^\pi \sin x\,dx \geq 0$$

Using Property (8), we obtain

$$\int_0^{\pi} \sin x\, dx \le (1)(\pi) = \pi.$$

(See Figure 6.21.)

The next example will help us to deepen our understanding of signed areas; it also uses the order properties.

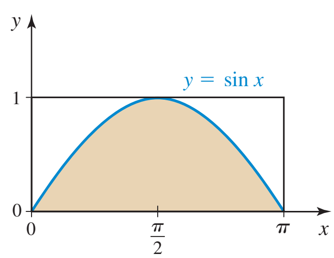


Figure 6.21 An illustration of the integral in Example 10. The shaded area is nonnegative and less than the area of the rectangle with base length π and height 1.

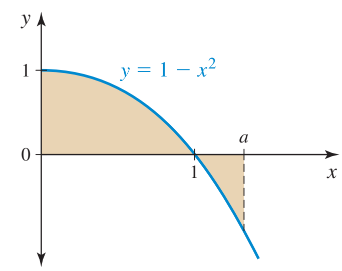


Figure 6.22 An illustration of the integral in Example 11.

EXAMPLE 11 Find the value of $a \ge 0$ that maximizes

$$\int_0^a (1 - x^2)\, dx$$

Solution We graph the integrand $f(x) = 1 - x^2$ for $x \ge 0$ in Figure 6.22. Using the interpretation of the definite integral as the signed area, we see from the graph of $f(x)$ that $a = 1$ maximizes the integral, since $f(x)$ is positive for $x < 1$ and negative for $x > 1$. So if $a > 1$, then:

$$\int_0^a f(x)dx = \int_0^1 f(x)dx + \int_1^a f(x)dx \qquad \text{Use Property (5)}$$

but since $f(x) < 0$ for $a > 1$, $\int_1^a f(x)dx < 0$ by Property (8) (since $f(x) \ne 0$, we may use the strict inequality). So if $a > 1$:

$$\int_0^a f(x)dx = \int_0^1 f(x)dx + \underbrace{\int_1^a f(x)dx}_{<0} < \int_0^1 f(x)dx.$$

Whereas if $0 \le a < 1$, then also by Property (5):

$$\int_0^1 f(x)dx = \int_0^a f(x)dx + \int_a^1 f(x)dx$$

But $f(x) > 0$ for $x \in (a, 1)$, so if $a < 1$:

$$\int_0^a f(x)dx = \int_0^1 f(x)dx - \underbrace{\int_a^1 f(x)dx}_{>0 \text{ by Property (8)}}$$

$$< \int_0^1 f(x)dx.$$

So if $a > 1$ or $a < 1$, then $\int_0^a f(x)dx < \int_0^1 f(x)dx$, which shows that the integral takes its largest value when $a = 1$.

Section 6.1 Problems

6.1.1

1. Approximate the area under the parabola $y = x^2$ from 0 to 1, using four equal subintervals.

2. Approximate the area under the parabola $y = x^2$ from 0 to 1, using five equal subintervals.

3. Approximate the area under the curve $y = x^3$ from 0 to 1, using six equal subintervals.

4. Approximate the area under the parabola $y = 1 - x^2$ from 0 to 1, using five equal subintervals.

5. Approximate the area under the curve $y = x^3 - x$ from 0 to 1 using five equal subintervals.

6. Approximate the area under the curve $y = x^2 - x$ from 0 to 1 using six equal subintervals.

In Problems 7 and 8, you will use Riemann sums to prove that $\int_0^a x^2 dx = \frac{1}{3}a^3$.

7. (a) Let the points $x_0 = 0 < x_1 < x_2 < \cdots < x_n = a$ divide the interval into n rectangles of equal width w. What is the width of each rectangle w?

(b) Show that the total area of the rectangles used to approximate the area under $y = x^2$ between $x = 0$ and $x = a$ is:

$$S_n = \frac{a^3}{n^3}\left(1^2 + 2^2 + 3^2 + \cdots + (n-1)^2\right)$$

(c) Assume (the proof is worked out in Problem 8) that $1^2 + 2^2 + 3^2 \cdots + k^2 = \frac{k(k+1)(2k+1)}{6}$. Use this formula to show that your expression for S_n can be rewritten as:

$$S_n = \frac{a^3}{6}\left(1 - \frac{1}{n}\right)\left(2 - \frac{1}{n}\right)$$

(d) Evaluate the limit: $\lim_{n\to\infty} S_n$, and derive the formula for $\int_0^a x^2 dx$.

8. To prove the sum formula that we used in Problem 7, namely:

$$1^2 + 2^2 + 3^2 + \cdots + k^2 = \frac{k(k+1)(2k+1)}{6}$$

we will consider two sums:

$$T_k = 2^3 + 3^3 + 4^3 + \cdots + k^3 + (k+1)^3$$

and

$$U_k = 1^3 + 2^3 + 3^3 + \cdots + k^3.$$

Notice that these sums have exactly the same number of terms.

(a) Explain, by comparing which terms show up in both sums, why:

$$T_k - U_k = (k+1)^3 - 1^3.$$

(b) Form the difference between the two sums by subtracting the first term of U_k from the first term of T_k, the second term of U_k from the second term of T_k, and so on. Then:

$$T_k - U_k = (2^3 - 1^3) + (3^3 - 2^3) + (4^3 - 3^3) + \cdots + ((k+1)^3 - k^3).$$

In this way of writing the sum, the mth term is:

$$(m+1)^3 - m^3$$

By expanding $(m+1)^3$, explain why the mth term can be written as $3m^2 + 3m + 1$. So:

$$T_k - U_k = \left(3\cdot(1)^2 + 3(1) + 1\right) + \left(3\cdot(2)^2 + 3(2) + 1\right) + (3\cdot(3)^2 + 3(3) + 1) + \cdots + (3\cdot k^2 + 3k + 1).$$

(c) Collect together the first terms in all the sums, then the second terms in all the sums, and finally the third terms in all the sums:

$$T_k - U_k = 3\left(\overbrace{1^2 + 2^2 + 3^2 + \cdots + k^2}^{\text{A}}\right) + 3\left(\underbrace{1 + 2 + 3 + \cdots + k}_{\text{B}}\right) + (\underbrace{1 + 1 + \cdots + 1}_{\text{C}}).$$

We have labeled the three terms A, B, and C.
A is the sum that we want to calculate.
Explain why $\text{B} = \frac{1}{2}k(k+1)$ and $\text{C} = k$.

(d) By combining the results of (a) and (c) show that:

$$3\left(1^2 + 2^2 + 3^2 + \cdots + k^2\right) + \frac{3}{2}k(k+1) + k = (k+1)^3 - 1. \quad (6.2)$$

Then show, by rearranging and simplifying (6.2), that:

$$1^2 + 2^2 + 3^2 + \cdots + k^2 = \frac{k(k+1)(2k+1)}{6}$$

6.1.2

9. Approximate

$$\int_{-1}^{1} (1 - x^2)\, dx$$

using five equal subintervals.

10. Approximate

$$\int_{-1}^{1} (1 + x^2)\, dx$$

using five equal subintervals.

11. Approximate

$$\int_{-1}^{1} (2 + x^2)\, dx$$

using five equal subintervals.

12. Approximate

$$\int_{-2}^{2} (2 + x^2)\, dx$$

using six equal subintervals.

13. Approximate

$$\int_{-1}^{2} e^{-x}\, dx$$

using three equal subintervals.

14. Approximate

$$\int_0^{\pi} \sin x\, dx$$

using four equal subintervals.

15. **(a)** Assume that $a, b > 0$. Evaluate $\int_a^b x\,dx$, using the fact that the region bounded by $y = x$ and the x-axis between a and b is a trapezoid. (See Figure 6.23.)

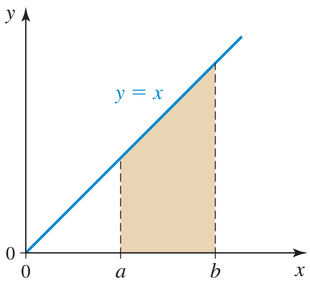


Figure 6.23 The region for Problem 15.

(b) Approximate $\int_1^2 x\,dx$ using (i) three and (ii) six equal subintervals; compare your approximations to the exact answer that you calculated in part (a).

16. Assume that $a < 0 < b$. Use a geometric argument to show that

$$\int_a^b x\,dx = \frac{b^2 - a^2}{2}$$

Express the limits in Problems 17–23 as definite integrals. In each case the n points $x_0 < x_1 < x_2 < \cdots < x_n$ evenly divide the interval $[a, b]$ into n subintervals, each of width $w = (b - a)/n$. You do not need to evaluate the definite integrals.

17.

$$\lim_{n\to\infty} w\left(\frac{x_0^2}{2} + \frac{x_1^2}{2} + \frac{x_2^2}{2} + \cdots + \frac{x_{n-1}^2}{2}\right)$$

where $a = 1, b = 2$.

18.

$$\lim_{n\to\infty} w\left(\sqrt{x_0+1} + \sqrt{x_1+1} + \sqrt{x_2+1} + \cdots + \sqrt{x_{n-1}+1}\right)$$

where $a = 0, b = 3$.

19.

$$\lim_{n\to\infty} w\left(\left(2x_0 - \frac{1}{2}\right) + \left(2x_1 - \frac{1}{2}\right) + \cdots + \left(2x_{n-1} - \frac{1}{2}\right)\right)$$

where $a = -1, b = 1$.

20.

$$\lim_{n\to\infty} w\left(1 + \frac{1}{x_0}\right) + \left(1 + \frac{1}{x_1}\right) + \cdots + \left(1 + \frac{1}{x_{n-1}}\right)$$

where $a = 1, b = 5$.

21.

$$\lim_{n\to\infty} w\,(2^{x_0} + 2^{x_1} + 2^{x_2} + \cdots + 2^{x_{n-1}})$$

where $a = 0$ and $b = 1$.

22.

$$\lim_{n\to\infty} w\,(\cos(x_0) + \cos(x_1) + \cdots + \cos(x_{n-1}))$$

23.

$$\lim_{n\to\infty} w\left(\frac{1}{e^{-x_0}+1} + \frac{1}{e^{-x_1}+1} + \cdots + \frac{1}{e^{-x_{n-1}}+1}\right)$$

where $a = 0$ and $b = 2$.

In Problems 24–29, express the definite integrals as limits of Riemann sums.

24. $\displaystyle\int_{-2}^{-1} \frac{x^2}{1+x^2}\,dx$ **25.** $\displaystyle\int_1^3 (x+1)^{1/3}\,dx$

26. $\displaystyle\int_1^3 e^{-2x}\,dx$ **27.** $\displaystyle\int_1^e \ln x\,dx$

28. $\displaystyle\int_0^{\pi} \cos\frac{2x}{\pi}\,dx$ **29.** $\displaystyle\int_0^5 x^3\,dx$

In Problems 30–36, use a graph to interpret the definite integral in terms of areas. Do not compute the integrals.

30. $\displaystyle\int_0^5 e^{-x}\,dx$ **31.** $\displaystyle\int_{-1}^2 (x^2-1)\,dx$

32. $\displaystyle\int_{-2}^2 \frac{1}{2}x^3\,dx$ **33.** $\displaystyle\int_0^3 (2x+1)\,dx$

34. $\displaystyle\int_{-\pi}^{\pi} \cos x\,dx$ **35.** $\displaystyle\int_{-3}^2 \left(1 - \frac{1}{2}x\right)dx$

36. $\displaystyle\int_{1/2}^4 \ln x\,dx$

In Problems 37–47, use an area formula from geometry to find the value of each integral by interpreting it as the (signed) area under the graph of an appropriately chosen function.

37. $\displaystyle\int_{-2}^3 |x|\,dx$ **38.** $\displaystyle\int_{-3}^3 \sqrt{9-x^2}\,dx$

39. $\displaystyle\int_2^5 \left(\frac{1}{2}x - 4\right)dx$ **40.** $\displaystyle\int_{1/2}^1 \sqrt{1-x^2}\,dx$

41. $\displaystyle\int_{-2}^2 (\sqrt{4-x^2} - 2)\,dx$ **42.** $\displaystyle\int_0^1 \sqrt{2-x^2}\,dx$

43. $\displaystyle\int_{-3}^0 (4 - \sqrt{9-x^2})\,dx$ **44.** $\displaystyle\int_{-2}^1 \sqrt{4-x^2}\,dx$

45. $\displaystyle\int_{-1}^2 (2-|x|)\,dx$ **46.** $\displaystyle\int_0^2 |x-1|\,dx$ **47.** $\displaystyle\int_0^2 (2+x)\,dx$

6.1.3

48. Use the definition of the Riemann integral in terms of Riemann sums to prove property (3) of definite integrals. That is, if $f(x)$ is continuous on $[a, b]$ and k is any constant, then:

$$\int_a^b k f(x)\,dx = k\int_a^b f(x)\,dx$$

49. Use a diagram to explain why, if $f(x)$ is continuous on an interval that contains all of the points a, b, c, then

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

That is, derive property (5) of definite integrals. You should consider the cases

(a) $b < a < c$ **(b)** $c < b < a$.

50. Given that $\displaystyle\int_0^a x^2\,dx = \frac{1}{3}a^3$ evaluate the following:

(a) $\displaystyle\int_0^1 \frac{1}{2}x^2\,dx$ **(b)** $\displaystyle\int_0^{-1} 3x^2\,dx$ **(c)** $\displaystyle\int_{-1}^2 \frac{1}{3}x^2\,dx$

(d) $\displaystyle\int_1^1 3x^2\,dx$ **(e)** $\displaystyle\int_{-2}^3 (x+1)^2\,dx$ **(f)** $\displaystyle\int_2^4 (x-2)^2\,dx$

51. Find $\int_2^2 \cos(3x^2)\,dx$.

52. Find $\int_{-3}^{-3} e^{-x^2/2}\,dx$.

53. Find $\int_{-1}^{1} 3x\,dx$.

54. Find $\int_{-1}^{1} 3x^5\,dx$.

55. Find $\int_0^2 (x-1)^3\,dx$.

56. Explain geometrically why

$$\int_1^2 x^2\,dx = \int_0^2 x^2\,dx - \int_0^1 x^2\,dx \tag{6.3}$$

and show that (6.3) can be written as

$$\int_1^2 x^2\,dx = \int_1^0 x^2\,dx + \int_0^2 x^2\,dx \tag{6.4}$$

Relate (6.4) to addition property (5).

57. Given that $\int_0^a x^3\,dx = \frac{1}{4}a^4$, evaluate the following integrals:

(a) $\int_0^2 x^3\,dx$ **(b)** $\int_0^1 2x^3\,dx$ **(c)** $\int_{-1}^{1} 2x^3\,dx$

(d) $\int_{-1}^{1} (x+1)^3\,dx$ **(e)** $\int_1^2 2(x+2)^3\,dx$.

58. Given that $\int_0^a x^4\,dx = \frac{1}{5}a^5$ evaluate the following integrals

(a) $\int_0^2 x^4\,dx$ **(b)** $\int_0^1 \frac{x^4}{2}\,dx$

(c) $\int_{-1}^{1} \frac{x^4}{2}dx$ **(d)** $\int_{-2}^{0} (x+2)^4 dx$

(e) $\int_{-3}^{0} (x+1)^4\,dx$ **(f)** $\int_0^2 2(x-2)^4\,dx$.

6.1.4

In Problems 59–63, verify each inequality without evaluating the integrals.

59. $\int_0^1 x\,dx \geq \int_0^1 x^2\,dx$

60. $\int_2^4 x\,dx \leq \int_2^4 x^2\,dx$

61. $0 \leq \int_0^9 \sqrt{x}\,dx \leq 27$

62. $\sqrt{3} \leq \int_0^1 \sqrt{4-x^2}\,dx \leq 2$.

63. $\frac{\pi}{3} \leq \int_{\pi/6}^{5\pi/6} \sin x\,dx \leq \frac{2\pi}{3}$

64. Find the value of $a \geq 0$ that maximizes $\int_0^a (4-x^2)\,dx$.

65. Find the value of $a \in [0, 2\pi]$ that maximizes $\int_0^a \cos x\,dx$.

66. Find $a \in (0, 2\pi]$ such that $\int_0^a \sin x\,dx = 0$

67. Find $a > 1$ such that $\int_1^a (x-3)^3\,dx = 0$

68. Find $a > 0$ such that $\int_0^a (1-x)\,dx = 0$

Problems 69 and 70 show some real-life applications for integrals.

69. Total Rainfall A rain gauge is set up to measure the amount of rainfall occurring in 1 hr on the UCLA campus (the readout from the rain gauge is in mm/hr). Assume that the following data is collected in a 6 hour window.

Time, t	**Rainfall rate, $r(t)$ in mm/hr**
0	1
1	2
2	3
3	1
4	1
5	0
6	0

The total amount of rainfall between $t = 0$ and $t = 6$ is given by

$$\int_0^6 r(t)\,dt$$

(see Section 6.3 for an explanation of this formula).

(a) Use *six* even subintervals to approximate the total rainfall between $t = 0$ and $t = 6$ as a sum of areas of rectangles and evaluate this sum using the data in the table.

(b) Repeat part (a) but now use *three* even subintervals.

70. Total Mortality You are measuring the ability of an antibiotic to kill harmful bacteria. You measure the rate at which the antibiotic kills bacteria (i.e., number of bacteria killed in one hour); this is called the *mortality rate*. You measure the following data for the number of bacteria killed in a 12 hour time period starting at $t = 0$, and ending at $t = 12$.

Time, t	**Mortality rate, per hour $m(t)$**
0	20
1	300
2	350
3	400
4	500
5	450
6	410
7	350
8	320
9	300
10	200
11	100
12	110

The total number of deaths between times $t = a$ and $t = b$ is given by

$$\int_a^b m(t)\,dt$$

(see Section 6.3 for an explanation of this formula).

(a) Use six even subintervals to approximate the total number of deaths between $t = 0$ and $t = 6$ and evaluate this sum using the data in the table.

(b) Use six even subintervals to approximate the total number of deaths between $t = 0$ and $t = 12$ and evaluate this sum using the data in the table.

(c) Use four even subintervals to approximate the total number of deaths between $t = 4$ and $t = 12$ and evaluate this sum using the data in the table.

6.2 The Fundamental Theorem of Calculus

In Section 6.1, we used the definition of definite integrals to compute $\int_0^a x\,dx$. This required the summation of a large number of terms, which was facilitated by the explicit summation formula for $1+2+3+\cdots+(n-1)$.

In Problems 7 and 8 of Section 6.1 you used a similar process to evaluate the integral $\int_0^a x^2\,dx$, this time making use of a formula for the sum $1^2+2^2+3^2+\cdots+(n-1)^2$. Based on these problems you may feel that finding general formulas that can be used to calculate definite integrals is dauntingly difficult. Indeed for centuries the problem of finding areas under curves proceeded very slowly with different mathematicians adding one-by-one formulas for different specific functions $f(x)$. However, Newton and Leibniz, who discovered the calculus of derivatives and rates of change, quickly realized that the problem of calculating areas under curves is the inverse problem to finding a rate of change. This result is known as the **Fundamental Theorem of Calculus**.

The fundamental theorem of calculus has two parts: The first part links antiderivatives and integrals, and the second part provides a method for computing definite integrals. Although we will first approach the Fundamental Theorem as a powerful tool to calculating integrals, its real significance will not be seen until Chapter 8, when we will use integration to undo differentiation and to solve differential equations that model population growth, cooperation among organisms, ecosystem dynamics, chemical reactions, and more.

6.2.1 The Fundamental Theorem of Calculus (Part I)

Let $f(x)$ be a continuous function on $[a, b]$, and let

$$F(x) = \int_a^x f(t)\,dt$$

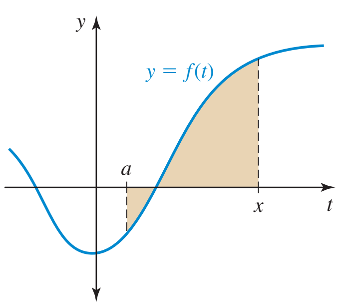


Figure 6.24 The shaded signed area is $F(x)$.

Geometrically, $F(x) = \int_a^x f(t)\,dt$ represents the signed area between the graph of $f(t)$ and the horizontal axis between a and x. (See Figure 6.24.) Note that the independent variable x appears as the upper limit of integration. We can now ask how the signed area $F(x)$ changes as x varies. To answer this question, we compute $\frac{dF}{dx}$; that is,

$$\frac{d}{dx}F(x) = \lim_{h\to 0}\frac{F(x+h)-F(x)}{h} \qquad \text{Use definition of a derivative.}$$

$$= \lim_{h\to 0}\frac{1}{h}\left[\int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,du\right] \tag{6.6}$$

$$= \lim_{h\to 0}\frac{1}{h}\int_x^{x+h} f(t)\,dt \qquad \text{Use Property (5) of subsection 6.1.3.}$$

To evaluate

$$\lim_{h\to 0}\frac{1}{h}\int_x^{x+h} f(t)\,dt$$

we will resort to the geometric interpretation of definite integrals. Our argument is illustrated in Figure 6.25: Note that

$$\int_x^{x+h} f(t)\,dt$$

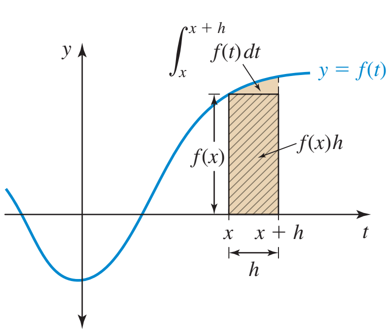


Figure 6.25 Approximating $\int_x^{x+h} f(t)dt$ by a rectangle.

is the signed area of the region bounded by the graph of $f(t)$ and the horizontal axis between x and $x+h$. If h is small, then this area is closely approximated by the area of a rectangle with height $f(x)$ and width h. The signed area of this rectangle is $f(x)h$. Hence,

$$\int_x^{x+h} f(t)\,dt \approx f(x)h$$

If we divide both sides by h and let h tend to 0, then

$$\lim_{h\to 0}\frac{1}{h}\int_x^{x+h} f(t)\,du = f(x) \tag{6.7}$$

Combining (6.6) and (6.7), we arrive at the remarkable result

$$\frac{d}{dx}F(x) = f(x)$$

In addition, we see that $F(x)$ is continuous since it is differentiable.

The preceding argument relies on geometric intuition. To show how we can make the argument mathematically rigorous, we give the complete proof in the (optional) Subsection 6.2.2. The result is summarized in the following theorem:

Theorem The Fundamental Theorem of Calculus (FTC) (Part I) If f is continuous on $[a, b]$, then the function F defined by

$$F(x) = \int_a^x f(u)\,du, \quad a \le x \le b$$

is continuous on $[a, b]$ and differentiable on (a, b), with

$$\frac{d}{dx}F(x) = f(x)$$

Simply stated, the FTC (part I) says that if we first integrate $f(x)$ and then differentiate the result, we get $f(x)$ again. In this sense, it shows that integration and differentiation are inverse operations.

We begin with an example that can be immediately solved by using the FTC.

EXAMPLE 1 Compute $\frac{d}{dx}\int_0^x (\sin t - e^{-t})\,dt$ for $x > 0$.

Solution First, note that $f(x) = \sin x - e^{-x}$ is continuous for $x \ge 0$. If we set $F(x) = \int_0^x (\sin t - e^{-t})\,dt$ and apply the FTC then

$$\frac{d}{dx}F(x) = \frac{d}{dx}\int_0^x (\sin t - e^{-t})\,dt = f(x) = \sin x - e^{-x}.$$

•

EXAMPLE 2 Compute $\frac{d}{dx}\int_3^x \frac{1}{1+t^2}\,dt$ for $x > 3$.

Solution First, note that $f(x) = \frac{1}{1+x^2}$ is continuous for $x \ge 3$. If we set $F(x) = \int_3^x \frac{1}{1+t^2}\,dt$ and apply the FTC then

$$\frac{d}{dx}F(x) = \frac{d}{dx}\int_3^x \frac{1}{1+t^2}\,dt = \frac{1}{1+x^2}.$$

•

6.2.2 Leibniz's Rule and a Rigorous Proof of the Fundamental Theorem of Calculus

Leibniz's Rule. Combining the chain rule and the FTC (part I), we can differentiate integrals with respect to x when the upper and/or lower limits of integration are functions of x.

In the first example, the upper limit of integration is a function of x.

EXAMPLE 3 Compute

$$\frac{d}{dx}\int_0^{x^2} (t^3 - 2)\,dt, \quad x > 0$$

Solution Note that $f(t) = t^3 - 2$ is continuous for all $t \in \mathbf{R}$. We set $F(u) = \int_0^u (t^3 - 2)\,dt$, $u > 0$. Here $F(u)$ evaluates to the integral of interest to us if

$$u = x^2.$$

That is we wish to compute $\frac{d}{dx}F(x^2)$. To do so, we need to apply the chain rule. We set $u(x) = x^2$. Then

$$\frac{d}{dx}F(x^2) = \frac{dF(u)}{du}\frac{du}{dx}$$

To evaluate $\frac{d}{du}F(u) = \frac{d}{du}\int_0^u (t^3 - 2)\,dt$, we use the FTC:

$$\frac{d}{du}\int_0^u (t^3 - 2)\,dt = u^3 - 2$$

Since $\frac{du}{dx} = \frac{d}{dx}(x^2) = 2x$, it follows that

$$\begin{aligned}\frac{d}{dx}F(x^2) &= (u^3 - 2)\cdot 2x = [(x^2)^3 - 2]\cdot 2x \qquad \text{Substitute } u = x^2\\ &= 2x(x^6 - 2).\end{aligned}$$

We can make a similar argument to differentiate integrals in which the lower limit of integration depends on x.

EXAMPLE 4 Compute

$$\frac{d}{dx}\int_{x^2}^1 t^2\,dt$$

Solution Note that $f(t) = t^2$ is continuous for all $t \in \mathbf{R}$. We use the fact that

$$\int_{x^2}^1 t^2\,dt = -\int_1^{x^2} t^2\,dt$$

The upper limit now depends on x, but we introduced a minus sign. Hence

$$\begin{aligned}\frac{d}{dx}\int_{x^2}^1 t^2\,dt &= -\frac{d}{dx}\int_1^{x^2} t^2\,dt\\ &= -(x^2)^2\cdot 2x = -2x^5\end{aligned}$$

where, as in Example 3, we used the chain rule in the last step, by letting $F(u) = \int_1^u t^2 dt$ and $u = x^2$ with $F'(u) = u^2$ and $\frac{du}{dx} = 2x$.

The preceding example makes an important point: We need to be careful about whether the upper or the lower limit of integration depends on x. In the next example, we show what we must do when both limits of integration depend on x.

EXAMPLE 5 For $x \in \mathbf{R}$, compute

$$\frac{d}{dx}\int_x^{x^2} e^t\,dt$$

Solution Note that $f(t) = e^t$ is continuous for all $t \in \mathbf{R}$. The given integral is therefore defined for all $x \in \mathbf{R}$, and we can split it into two integrals at any $a \in \mathbf{R}$. We choose $a = 0$, which yields

$$\int_x^{x^2} e^t\,dt = \int_x^0 e^t\,dt + \int_0^{x^2} e^t\,dt = -\int_0^x e^t\,dt + \int_0^{x^2} e^t\,dt$$

The right-hand side is now written in a form that we know how to differentiate, and we find that

$$\begin{aligned}\frac{d}{dx}\int_x^{x^2} e^t\,dt &= -\frac{d}{dx}\int_0^x e^t\,dt + \frac{d}{dx}\int_0^{x^2} e^t\,dt\\ &= -[e^x] + \left[e^{x^2}\frac{d}{dx}x^2\right]\\ &= -e^x + 2xe^{x^2}\end{aligned}$$

The preceding example illustrates the most general case that we can encounter, namely, when both limits of integration are functions of x. We summarize this case in the following box, in a property known as Leibniz's rule:

> **Leibniz's Rule** If $g(x)$ and $h(x)$ are differentiable functions and $f(t)$ is continuous for t between $g(x)$ and $h(x)$, then
>
> $$\frac{d}{dx}\int_{g(x)}^{h(x)} f(t)\,dt = f[h(x)]h'(x) - f[g(x)]g'(x)$$

We can check that Examples 3–5 can be solved with the preceding formula; for instance, in Example 4, we have $f(t) = t^2$, $g(x) = x^2$, and $h(x) = 1$. Then $g'(x) = 2x$ and $h'(x) = 0$. We therefore find that

$$f[h(x)]h'(x) - f[g(x)]g'(x) = 0 - (x^2)^2 \cdot 2x = -2x^5.$$

which is the answer we obtained in Example 4.

Proof of the Fundamental Theorem of Calculus (Part I) (Optional) In subsection 6.2.1 we showed that if

$$F(x) = \int_a^x f(t)\,dt$$

then

$$\frac{d}{dx}F(x) = \lim_{h\to 0}\frac{1}{h}\int_x^{x+h} f(t)\,dt. \tag{6.8}$$

In subsection 6.2.1 we then approximated the right hand side of (6.8) by a rectangle to prove, non-rigorously, the Fundamental Theorem of Calculus. A rigorous proof relies on material from subsection 6.1.4, so you will need to study that subsection before this one.

We now give a mathematically rigorous argument which will show that

$$\frac{d}{dx}F(x) = f(x)$$

We begin with the observation that, according to the extreme-value theorem, the continuous function $f(t)$ defined on the closed interval $[x, x+h]$ attains an absolute minimum and an absolute maximum on $[x, x+h]$. That is, there exist m and M such that m is the minimum of f on $[x, x+h]$ and M is the maximum of f on $[x, x+h]$, which implies that

$$m \le f(t) \le M \quad \text{for all } t \in [x, x+h] \tag{6.9}$$

So

$$mh \le \int_x^{x+h} f(u)\,du \le Mh \qquad \text{Apply Property (8) from Subsection 6.1.4}$$

Dividing by h, we obtain

$$m \le \frac{1}{h}\int_x^{x+h} f(t)\,dt \le M \tag{6.10}$$

We set

$$I = \frac{1}{h}\int_x^{x+h} f(t)\,dt$$

Then (6.10) becomes $m \le I \le M$; that is, I is a number between m and M. We compare this inequality with (6.9), which says that $f(t)$ also lies between m, the minimum of f on $[x, x+h]$, and M, the maximum of f on $[x, x+h]$, for all $u \in [x, x+h]$. The intermediate-value theorem applied to $f(t)$ tells us that any value between m and M is attained by $f(u)$ for some number on the interval $[x, x+h]$. Hence, since I lies between m and M, there must exist a number $c_h \in [x, x+h]$ such that $f(c_h) = I$; that is, there exists $c_h \in [x, x+h]$ for which

$$f(c_h) = \frac{1}{h}\int_x^{x+h} f(t)\,dt \tag{6.11}$$

Because $x \le c_h \le x+h$, it follows that

$$\lim_{h\to 0} c_h = x$$

Since f is continuous,

$$\lim_{h\to 0} f(c_h) = f\left(\lim_{h\to 0} c_h\right) = f(x) \tag{6.12}$$

Combining (6.8), (6.11), and (6.12) yields the result that we are looking for:

$$\frac{d}{dx}F(x) = f(x)$$

•

6.2.3 Antiderivatives and Indefinite Integrals

The first part of the fundamental theorem of calculus tells us that if

$$F(x) = \int_a^x f(t)\,dt$$

then $F'(x) = f(x)$ [provided that $f(x)$ is continuous over the range of integration]. Equivalently, this statement says that $F(x)$ is an *antiderivative* of $f(x)$. (We introduced antiderivatives in Section 5.10.) Now, if we let

$$F(x) = \int_a^x f(t)\,dt \qquad \text{and} \qquad G(x) = \int_b^x f(t)\,dt$$

where a and b are two numbers, then both integrals have the same derivative, namely, $F'(x) = G'(x) = f(x)$ [again, provided that $f(x)$ is continuous over the range of integration]. That is, both $F(x)$ and $G(x)$ are antiderivatives of $f(x)$. We saw in Section 5.10 that antiderivatives of a given function differ only by a constant. We can identify the constant, namely

$$F(x) - G(x) = \int_a^x f(t)\,dt - \int_b^x f(t)dt = \int_a^x f(t)\,dt + \int_x^b f(t)\,dt$$

Use Property (2) from Subsection 6.1.3

$$= \int_a^b f(t)\,dt$$

Use Property (5) from Subsection 6.1.3

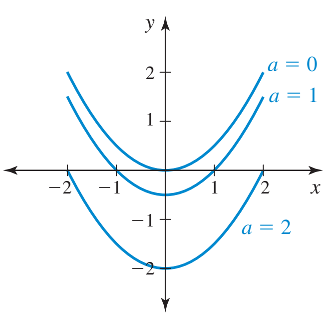


Figure 6.26 The function $F(x) = \int_a^x t\,dt$ produces the same curve up to vertical shifts for any value of a.

That is, $F(x) = G(x) + C$, where $C = \int_a^b f(t)\,dt$ is the integral of $f(t)$ between the start point of the integration that defines $F(x)$ and the start point of the $G(x)$ integration. However, the precise value of the constant is not as important as the observation that all integrals of the form $F(x) = \int_a^x f(t)\,dt$ give the same function up to some arbitrary constant. As an example of this, we show the functions $F(x) = \int_a^x t\,dt$ for some different values of a in Figure 6.26.

The general antiderivative of a function $f(x)$ is $F(x)+C$, where $F'(x) = f(x)$ and C is an arbitrary constant. It follows that $C + \int_a^x f(t)\,dt$ is the general antiderivative of $f(x)$. We will use the notation $\int f(x)\,dx$ to denote both the general antiderivative of $f(x)$ and the function $C + \int_a^x f(t)\,dt$; that is:

Definition of the Indefinite Integral

$$\int f(x)\,dx = C + \int_a^x f(u)\,du \tag{6.13}$$

Thus, the first part of the FTC says that indefinite integrals and antiderivatives are the same.

When we write $\int_a^x f(t)\,dt$, we use a letter other than x in the integrand because x already appears as the upper limit of integration. However, in the symbolic notation $\int f(x)\,dx$ we write x in the indefinite integral. This notation is to be interpreted as in (6.13); it is a convenient shorthand for $C + \int_a^x f(t)\,dt$. The choice of value of a for the lower limit of integration on the right side of (6.13) is not important, because different indefinite integrals of the same function $f(x)$ differ only by an additive constant that can be absorbed into the constant C.

Examples 6–8 show how to compute indefinite integrals. You may also find it helpful to review the Examples in Section 5.10 (which covered how to find antiderivatives) now we know the two problems are equivalent.

EXAMPLE 6 Compute $\int x^4\,dx$.

Solution We need to find a function $F(x)$ such that $F'(x) = x^4$. The solution is

$$\int x^4\,dx = \frac{1}{5}x^5 + C$$

where C is a constant. We check that, indeed,

$$\frac{d}{dx}\left(\frac{1}{5}x^5 + C\right) = x^4$$

•

EXAMPLE 7 Compute $\int (e^x + \sin x)\,dx$.

Solution We need to find an antiderivative of $f(x) = e^x + \sin x$. Since

$$\frac{d}{dx}(e^x - \cos x) = e^x - (-\sin x) = e^x + \sin x$$

it follows that

$$\int (e^x + \sin x)\,dx = e^x - \cos x + C$$

•

When we compute the indefinite integral $\int f(x)\,dx$, we want to know the general antiderivative of $f(x)$; this is why we added the constant C in the previous two examples.

EXAMPLE 8 Show that

$$\int \frac{1}{x}\,dx = \ln|x| + C \quad \text{for } x \neq 0$$

Solution Since the absolute value of x appears on the right-hand side, we split our discussion into two parts, according to whether $x \geq 0$ or $x \leq 0$. Recall that

$$|x| = \begin{cases} x & \text{for } x \geq 0 \\ -x & \text{for } x < 0 \end{cases}$$

Since $\ln x$ is not defined at $x = 0$, we consider the two cases $x > 0$ and $x < 0$.

(i) $x > 0$: $\ln|x| = \ln x$ when $x > 0$, and $\frac{d}{dx}\ln x = \frac{1}{x}$
Hence, $\int \frac{1}{x}\,dx = \ln x + C \quad \text{for } x > 0$

(ii) $x < 0$: $\ln|x| = \ln(-x)$ when $x < 0$, and $\frac{d}{dx}\ln(-x) = \frac{1}{-x}(-1) = \frac{1}{x}$
Hence, $\int \frac{1}{x}\,dx = \ln(-x) + C \quad \text{for } x < 0$

Combining (i) and (ii), we obtain

$$\int \frac{1}{x}\,dx = \ln|x| + C \quad \text{for } x \neq 0$$

•

We have seen that in order to evaluate indefinite integrals, we must find antiderivatives. Table 6-1 gives a list of indefinite integrals. (The table is a slightly expanded form of the table of antiderivatives from Section 5.10.) As Examples 10 and 11 show we may need to manipulate our integrand to put it in the form of one of the functions on this list.

TABLE 6-1 A Collection of Indefinite Integrals

$\int x^n\,dx = \frac{x^{n+1}}{n+1} + C \quad (n \neq -1)$	$\int \frac{1}{x}\,dx = \ln\lvert x\rvert + C$
$\int e^x\,dx = e^x + C$	$\int a^x\,dx = \frac{a^x}{\ln a} + C$
$\int \cos x\,dx = \sin x + C$	$\int \sin x\,dx = -\cos x + C$
$\int \sec^2 x\,dx = \tan x + C$	$\int \csc^2 x\,dx = -\cot x + C$
$\int \sec x \tan x\,dx = \sec x + C$	$\int \csc x \cot x\,dx = -\csc x + C$
$\int \tan x\,dx = \ln\lvert \sec x\rvert + C$	$\int \cot x\,dx = -\ln\lvert \csc x\rvert + C$
$\int \frac{1}{1+x^2}\,dx = \tan^{-1} x + C$	$\int \frac{1}{\sqrt{1-x^2}}\,dx = \sin^{-1} x + C$

EXAMPLE 9 Evaluate $\int x^3\,dx$.

Solution From Table 6-1:

$$\int x^3\,dx = \frac{1}{4}x^4 + C \qquad \int x^n\,dx \text{ with } n = 3$$

EXAMPLE 10 Evaluate $\int \frac{1}{\sin^2 x - 1}\,dx$.

Solution We first work on the integrand. Using the fact that $\sin^2 x + \cos^2 x = 1$, we find that

$$\frac{1}{\sin^2 x - 1} = -\frac{1}{\cos^2 x} = -\sec^2 x$$

Hence,

$$\int \frac{1}{\sin^2 x - 1}\,dx = -\int \sec^2 x\,dx = -\tan x + C \qquad \int \sec^2 x\,dx = \tan x + C_1 \text{ define } C = -C_1$$

EXAMPLE 11 Evaluate $\int \frac{x^2}{x^2+1}\,dx$.

Solution We first do long division on the integrand:

$$\frac{x^2}{x^2+1} = \frac{x^2+1-1}{x^2+1} = \frac{x^2+1}{x^2+1} - \frac{1}{x^2+1} = 1 - \frac{1}{x^2+1}$$

Then, from Table 6-1,

$$\int \frac{x^2}{x^2+1}\,dx = \int \left(1 - \frac{1}{x^2+1}\right) dx = x - \tan^{-1} x + C$$

•

6.2.4 The Fundamental Theorem of Calculus (Part II)

The first part of the FTC allows us to compute integrals of the form $\int_a^x f(t)\,dt$ only up to an additive constant; for instance,

$$F(x) = \int_1^x t^2\,du = \frac{1}{3}x^3 + C$$

To evaluate the definite integral $F(2) = \int_1^2 t^2\,dt$, which represents the area under the graph of $f(x) = x^2$ between $x = 1$ and $x = 2$, we would need to know the value of the constant. This value is provided by the second part of the FTC, which allows us to evaluate definite integrals.

Suppose we want to calculate $\int_a^b f(t)\,dt$. Let $F(x)$ be any antiderivative of $f(x)$, i.e.,

$$F(x) = \int_a^x f(t)\,dt + C \tag{6.14}$$

for some constant C. Then $F(b) = \int_a^b f(t)\,dt + C$, so $F(b)$ is almost the definite integral that we want, but it also contains the constant C. To evaluate the definite integral from $F(x)$ we would need to know the constant C. How can we find C? The solution is to evaluate both sides of (6.14) when $x = a$:

$$\begin{aligned} F(a) &= \underbrace{\int_a^a \quad f(t)\,dt \quad}_{=0} + C \\ &= 0 + C = C \end{aligned}$$

so:

$$F(x) - F(a) = \int_a^x f(t)\,dt \qquad \text{Pull } C = F(a) \text{ to the left-hand side of (6.14)}$$

Thus, if we set $x = b$, we obtain:

$$F(b) - F(a) = \int_a^b f(t)\,dt.$$

This formula allows us to evaluate definite integrals and is the content of the second part of the FTC.

Theorem The Fundamental Theorem of Calculus (Part II) Assume that f is continuous on $[a, b]$; then

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

where $F(x)$ is any antiderivative of $f(x)$; i.e., $F'(x) = f(x)$.

So how do we use this part of the FTC? To compute the definite integral $\int_a^b f(x)\,dx$ when f is continuous on $[a, b]$, we first need to find an antiderivative $F(x)$ of $f(x)$ (*any* antiderivative will do) and then compute $F(b) - F(a)$. This number is then equal to $\int_a^b f(x)\,dx$. Table 6-1 will help us find the required antiderivative. (Note that an indefinite integral is a function, whereas a definite integral is simply a number.)

Using the FTC (Part II) to Evaluate Definite Integrals.

EXAMPLE 12 Evaluate $\int_{-1}^2 (x^2 - 3x)\,dx$.

Solution Note that $f(x) = x^2 - 3x$ is continuous on $[-1, 2]$. We need to find an antiderivative of $f(x) = x^2 - 3x$; for instance, $F(x) = \frac{1}{3}x^3 - \frac{3}{2}x^2$ is an antiderivative of $f(x)$ since

$F'(x) = f(x)$. We then must evaluate $F(2) - F(-1)$:

$$F(2) = \frac{1}{3} \cdot 2^3 - \frac{3}{2} \cdot 2^2 = \frac{8}{3} - 6 = -\frac{10}{3}$$

$$F(-1) = \frac{1}{3}(-1)^3 - \frac{3}{2}(-1)^2 = -\frac{1}{3} - \frac{3}{2} = -\frac{11}{6}$$

We find that $F(2) - F(-1) = -\frac{10}{3} - (-\frac{11}{6}) = -\frac{9}{6} = -\frac{3}{2}$. Therefore,

$$\int_{-1}^{2} (x^2 - 3x)\, dx = F(2) - F(-1) = -\frac{3}{2}$$ •

In Example 12, we chose the simplest antiderivative—that is, the one in which the constant C is equal to 0. We could have chosen any $C \neq 0$, and the answer would have been the same. Let's see why. The general antiderivative of $f(x) = x^2 - 3x$ is $G(x) = \frac{1}{3}x^3 - \frac{3}{2}x^2 + C$. We can write this as $G(x) = F(x) + C$, where $F(x)$ is the antiderivative we used in our solution to Example 12. Then, using $G(x)$ to evaluate the integral, we find that

$$\begin{aligned}\int_{-1}^{2} (x^2 - 3x)\, dx &= G(2) - G(-1) \\ &= [F(2) + C] - [F(-1) + C] = F(2) - F(-1)\end{aligned}$$

which is the same answer as before, since the constant C cancels out. We thus see that we can use the simplest antiderivative (the one in which $C = 0$), and we will do so from now on.

EXAMPLE 13 Evaluate $\int_0^{\pi} \sin x\, dx$.

Solution Note that $\sin x$ is continuous on $[0, \pi]$. Since $F(x) = -\cos x$ is an antiderivative of $\sin x$, we have

$$\int_0^{\pi} \sin x\, dx = F(\pi) - F(0) = -\cos \pi - (-\cos 0) = -(-1) + 1 = 2.$$ •

EXAMPLE 14 Evaluate $\int_{-5}^{-1} \frac{1}{x}\, dx$

Solution Note that $\frac{1}{x}$ is continuous on $[-5, -1]$. Now $\ln |x|$ is an antiderivative of $\frac{1}{x}$. We use this antiderivative to evaluate the integral and obtain

$$\int_{-5}^{-1} \frac{1}{x}\, dx = \ln|-1| - \ln|-5| = -\ln 5$$

since $\ln|-1| = \ln 1 = 0$ and $|-5| = 5$. •

We now introduce some more notation. The notation $F(x)\big]_a^b$ indicates that we evaluate the function $F(x)$ at b and a, respectively, and compute the difference $F(b) - F(a)$. Then $F(x)$ is an antiderivative of $f(x)$, we can write

$$\int_a^b f(x)\, dx = F(x)\big]_a^b = F(b) - F(a)$$

For instance,

$$\int_{-5}^{-1} \frac{1}{x}\, dx = \ln|x|\big]_{-5}^{-1} = \ln|-1| - \ln|-5|$$

In this book we will consistently use the notation $F(x)\big]_a^b$ to denote the difference $F(b) - F(a)$. In other references you may see the alternate notations

$$[F(x)]_{x=a}^{b}, \quad [F(x)]_a^b, \quad F(x)|_{x=a}^{b}, \quad \text{or} \quad F(x)|_a^b.$$

It is not necessary for you to use any of these other notations, but you should recognize that they all stand for the same quantity, i.e., $F(b) - F(a)$.

EXAMPLE 15 Evaluate

$$\int_0^3 2xe^{x^2}\,dx$$

Solution Observe that $2xe^{x^2}$ is continuous on $[0, 3]$ and that $F(x) = e^{x^2}$ is an antiderivative of $f(x) = 2xe^{x^2}$, since, applying the chain rule, we find that

$$F'(x) = \underbrace{e^{x^2}}_{\frac{dF}{du}} \underbrace{\left(\frac{d}{dx}x^2\right)}_{\frac{du}{dx}} = e^{x^2} \cdot 2x \qquad F(x) = e^u \text{ and } u = x^2$$

Therefore,

$$\int_0^3 2xe^{x^2}\,dx = e^{x^2}\Big]_0^3 = e^9 - e^0 = e^9 - 1.$$

Don't worry if you did not spot that the antiderivative of $2xe^{x^2}$ is e^{x^2}; in Chapter 7 we will learn a method (integration by substitution) for finding antiderivatives of this kind.

EXAMPLE 16 Evaluate

$$\int_1^4 \frac{2x^2 + \sqrt{x}}{\sqrt{x}}\,dx.$$

Solution The integrand is continuous on $[1, 4]$. We first simplify the integrand:

$$f(x) = \frac{2x^2 + \sqrt{x}}{\sqrt{x}} = 2x^{3/2} + 1$$

An antiderivative of $f(x)$ is, therefore,

$$F(x) = 2 \cdot \frac{2}{5}x^{5/2} + x = \frac{4}{5}x^{5/2} + x$$

which can be checked by differentiating $F(x)$. We can now evaluate the integral:

$$\begin{aligned}\int_1^4 \frac{2x^2 + \sqrt{x}}{\sqrt{x}}\,dx &= \frac{4}{5}x^{5/2} + x\Big]_1^4 \\ &= \left(\frac{4}{5} \cdot 4^{5/2} + 4\right) - \left(\frac{4}{5} \cdot 1^{5/2} + 1\right) = \left(\frac{4}{5} \cdot 32 + 4\right) - \left(\frac{4}{5} + 1\right) \\ &= \frac{148}{5} - \frac{9}{5} = \frac{139}{5}.\end{aligned}$$

Finding an Integrand.

EXAMPLE 17 Suppose that

$$\int_0^x f(t)\,dt = \cos(2x) + C$$

where C is a constant. Find $f(x)$.

Solution Use the FTC, part I, which implies that

$$\frac{d}{dx}\int_0^x f(t)\,dt = f(x).$$

Hence,

$$f(x) = \frac{d}{dx}[\cos(2x) + C] = -2\sin(2x).$$

EXAMPLE 18 Suppose that $\int_1^x f(t)\,dt = xe^x + C$ where C is a constant. Find $f(x)$ and the constant C.

Solution Unlike Example 17, here we need to find both the integrand function and the constant C. To find $f(x)$ use the FTC, part I:

$$\frac{d}{dx}\int_1^x f(t)\,dt = f(x)$$

so:

$$\frac{d}{dx}(xe^x + C) = f(x)$$
$$(x+1)e^x = f(x)$$

To calculate the constant C that matches this particular lower bound on the integral, note that if $x = 1$, the integral vanishes by Property (1) of Subsection 6.1.3. But:

$$\underbrace{\int_1^1 f(t)\,dt}_{0} = 1 \cdot e^1 + C$$
$$\Rightarrow C = -e.$$

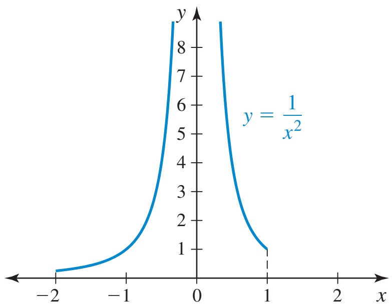


Figure 6.27 The graph of $y = \frac{1}{x^2}$ between $x = -2$ and $x = 1$. The function is discontinuous at $x = 0$.

Discontinuous Integrand. You might wonder why we always check that the integrand is continuous on the interval between the lower and the upper limit of integration. The next example shows what can go wrong when the integrand is discontinuous.

EXAMPLE 19 Evaluate

$$\int_{-2}^{1} \frac{1}{x^2}\,dx$$

Solution An antiderivative of $f(x) = 1/x^2$ is $F(x) = -\frac{1}{x}$. We find that $F(1) = -1$ and $F(-2) = \frac{1}{2}$. When we compute $F(1) - F(-2)$, we get $-\frac{3}{2}$. This is obviously not equal to $\int_{-2}^{1} \frac{1}{x^2}\,dx$, since $f(x) = \frac{1}{x^2}$ is positive on $[-2, 1]$ (see Figure 6.27) and, therefore, the integral of $f(x)$ between -2 and 1 should not be negative. The function $f(x)$ is discontinuous at $x = 0$ (where it has a vertical asymptote). Hence, the second part of the FTC therefore cannot be applied. We will learn how to deal with such discontinuities in Section 7.4. In any case, before you evaluate an integral, always check whether the integrand is continuous between the limits of integration. •

Section 6.2 Problems

6.2.1

In Problems 1–14, find $\frac{dy}{dx}$.

1. $y = \int_0^x 2t^2 dt$

2. $y = \int_0^x (4 - \frac{t^4}{2})dt$

3. $y = \int_0^x (4t - 3)dt$

4. $y = \int_0^x (3 + t^4)dt$

5. $y = \int_0^x \sqrt{1 + 2t}\,dt,\ x > \frac{-1}{2}$

6. $y = \int_0^x \sqrt{1 + t^2}\,dt$

7. $y = \int_0^x \sqrt{\sin 2t}\,dt, 0 < x < \frac{\pi}{2}$

8. $y = \int_0^x \cos(t + 1)dt$

9. $y = \int_3^x te^{4t}\,dt$

10. $y = \int_1^x te^{-t^2}\,dt$

11. $y = \int_0^x \frac{1}{t+1}\,dt,\ x > -1$

12. $y = \int_{-1}^x \frac{2}{t^2 + t}\,dt$

13. $y = \int_{\pi/2}^x \sin(t^2 + 1)\,dt$

14. $y = \int_{\pi/4}^x \cos^2(t - 3)\,dt$

6.2.2

In Problems 15–38, use Leibniz's rule to find $\frac{dy}{dx}$.

15. $y = \int_0^{3x} (1 + t^2)\,dt$

16. $y = \int_0^{2x-1} (t^3 - 1)\,dt$

17. $y = \int_0^{1-4x} (2t^2 + 1)\,dt$

18. $y = \int_0^{3x+2} t(1 + t)\,dt$

19. $y = \int_4^{x^2+1} \sqrt{t}\,dt$

20. $y = \int_2^{x^2-2} \sqrt{3 + u}\,du$

21. $y = \int_0^{3x} (1 + e^t)\,dt$

22. $y = \int_0^{2x^2-1} (e^{-2t} + e^{2t})\,dt$

23. $y = \int_1^{3x^2+x} (1 + te^t)\, dt$

24. $y = \int_2^{\ln x} e^{-t}\, dt, x > 0$

25. $y = \int_x^3 (1 + t)\, dt$

26. $y = \int_x^5 (1 + e^t)\, dt$

27. $y = \int_{2x}^3 (1 + \cos t)\, dt$

28. $y = \int_{2x^2}^0 \frac{t}{t+1}\, dt$

29. $y = \int_x^5 \frac{1}{u^2}\, du, x > 0$

30. $y = \int_{x^2}^3 \frac{1}{1+t}\, dt$

31. $y = \int_{x^2}^1 \sec t\, dt, -\sqrt{\frac{\pi}{2}} < x < \sqrt{\frac{\pi}{2}}$

32. $y = \int_{2+x^2}^2 \cot t\, dt, -\sqrt{\pi - 2} < x < \sqrt{\pi - 2}$

33. $y = \int_x^{2x} (1 + t^2)\, dt$

34. $y = \int_{-x}^x u\, du$

35. $y = \int_{x^2}^{x^3} \ln(t - 3)\, dt, x > \sqrt{3}$

36. $y = \int_{x^3}^{x^4} \ln(1 + t^2)\, dt$

37. $y = \int_{2-x^2}^{x+x^3} (t^2 - 1) dt$

38. $y = \int_{1+x^2}^{x^3-2x} (t + 1) dt$

6.2.3

In Problems 39–96, compute the indefinite integrals.

39. $\int (1 + 3x^2)\, dx$

40. $\int (x^3 - 4)\, dx$

41. $\int \left(\frac{1}{3}x^2 - \frac{1}{2}x\right) dx$

42. $\int (4x^3 + 5x^2)\, dx$

43. $\int \left(\frac{1}{2}x^2 + 3x - \frac{1}{3}\right) dx$

44. $\int \left(\frac{1}{2}x^5 + 2x^3 - 1\right) dx$

45. $\int \frac{2x^2 + x}{\sqrt{x}}\, dx$

46. $\int \frac{x^3 + 3x^2}{2\sqrt{x}}\, dx$

47. $\int x^2 \sqrt{x}\, dx$

48. $\int (1 + x^3)\sqrt{x}\, dx$

49. $\int (x^{7/2} + x^{2/7})\, dx$

50. $\int (x^{3/5} + x^{5/3})\, dx$

51. $\int \left(\sqrt{x} + \frac{1}{\sqrt{x}}\right) dx$

52. $\int \left(3x^{1/3} + \frac{1}{3x^{1/3}}\right) dx$

53. $\int (x - 1)(x + 1)\, dx$

54. $\int (x - 1)^2\, dx$

55. $\int x(x + 1)\, dx$

56. $\int (x + 1)x^2\, dx$

57. $\int e^{2x}\, dx$

58. $\int 2e^{3x}\, dx$

59. $\int 3e^{-x}\, dx$

60. $\int 2e^{-x/3}\, dx$

61. $\int e^x(e^x + 1)\, dx$

62. $\int e^x(1 - e^{-x})\, dx$

63. $\int \sin(2x)\, dx$

64. $\int \sin \frac{x}{3}\, dx$

65. $\int \cos(3x)\, dx$

66. $\int \cos(2 + x)\, dx$

67. $\int \sin(2x - 1)\, dx$

68. $\int \cos(2x + 1)\, dx$

69. $\int \frac{\sin x}{1 - \sin^2 x}\, dx$

70. $\int \frac{\cos x}{1 - \cos^2 x}\, dx$

71. $\int \cos x \sin x\, dx$

72. $\int (\cos^2 x - \sin^2 x)\, dx$

73. $\int (\cos x + \cos^2 x)\, dx$

74. $\int (\sin x - \sin^2 x)\, dx$

75. $\int \frac{4}{1 + x^2}\, dx$

76. $\int \left(\frac{x^2}{1 + x^2}\right) dx$

77. $\int \frac{1}{\sqrt{1 - x^2}}\, dx$

78. $\int \frac{5}{\sqrt{1 - x^2}}\, dx$

79. $\int \frac{1}{x + 2}\, dx$

80. $\int \frac{1}{x - 3}\, dx$

81. $\int \frac{2x - 1}{3x}\, dx$

82. $\int \frac{2x + 5}{x}\, dx$

83. $\int \frac{1}{2x + 1}\, dx$

84. $\int \frac{1}{3x - 3}\, dx$

85. $\int \frac{1}{x^2 + 4}\, dx$

86. $\int \frac{1}{x^2}\, dx$

87. $\int \frac{2x^2}{x^2 + 1}\, dx$

88. $\int \frac{2x^2}{4 + x^2}\, dx$

89. $\int 3^x\, dx$

90. $\int 2^x\, dx$

91. $\int 4^{-x}\, dx$

92. $\int 3^{-2x}\, dx$

93. $\int (x^2 + 2^x)\, dx$

94. $\int (x^{-3} + 3^{-x})\, dx$

95. $\int (\sqrt{x} + \sqrt{e^x})\, dx$

96. $\int \left(\frac{1}{\sqrt{x}} + \frac{1}{\sqrt{e^x}}\right) dx$

6.2.4

In Problems 97–122, evaluate the definite integrals.

97. $\int_1^4 (3 + 2x)\, dx$

98. $\int_0^3 (2x^2 - 1)\, dx$

99. $\int_0^1 (x^3 - x^{1/3})\, dx$

100. $\int_1^2 x^{5/2}\, dx$

101. $\int_1^8 x^{-2/3}\, dx$

102. $\int_1^9 \frac{1 + \sqrt{x}}{\sqrt{x^3}}\, dx$

103. $\int_0^2 t(t + 3)\, dt$

104. $\int_0^2 (2 + 3t)^2\, dt$

105. $\int_0^{\pi/4} \sin(2x)\, dx$

106. $\int_{-\pi/3}^{\pi/3} 2 \cos\left(\frac{x}{2}\right) dx$

107. $\int_0^{\pi/4} \sin\left(x - \frac{\pi}{4}\right) dx$

108. $\int_0^1 \sin(\pi(x + 1))\, dx$

109. $\int_0^1 \frac{1}{1 + x^2}\, dx$

110. $\int_{-1}^1 \frac{4}{1 + x^2}\, dx$

111. $\int_0^{1/2} \frac{1}{\sqrt{1 - x^2}}\, dx$

112. $\int_{-1/2}^{1/2} \frac{2}{\sqrt{1 - x^2}}\, dx$

113. $\int_0^{\pi/2} \cos 2x\, dx$

114. $\int_0^{\pi} \sin 2x\, dx$

115. $\int_{-1}^0 e^{3x}\, dx$

116. $\int_0^2 2te^{t^2}\, dt$

117. $\int_0^{\pi/2} x^2 \exp(x^3)\, dx$

118. $\int_0^{\pi/4} \sec^2 x\, dx$

119. $\int_1^e \frac{1}{x}\,dx$

120. $\int_0^1 \frac{1}{z+1}\,dz$

121. $\int_0^1 \frac{1}{1+2u}\,du$

122. $\int_1^2 \frac{2}{3t+1}\,dt$

123. Suppose that

$$\int_0^x f(t)\,dt = 2x^2$$

Find $f(x)$.

124. Suppose that

$$\int_0^x f(t)\,dt = e^x - 1$$

Find $f(x)$.

125. Suppose $\int_\pi^x f(t)\,dt = \sin x + C$ for some constant C. Find the function $f(x)$ and the constant C.

126. Suppose $\int_0^x f(t)\,dt = \frac{1}{x^2+1} + C$ for some constant C. Find the function $f(x)$ and the constant C.

6.3 Applications of Integration

In this section, we will discuss a number of applications of integrals. The first application is to use an integral to calculate the cumulative (or net) change in a function; the second is to use integrals to calculate the average of a function (e.g., the average rate of rainfall over a time interval). The third and fourth applications are optional; we will use integrals to calculate areas between curves, and the volumes, lengths, and surface areas of objects that can be described using functions. The last two applications are good practice for thinking of integrals as sums of many small increments; they are useful for applications of calculus in physics and engineering, but you are less likely to encounter these kinds of problems in life sciences, which is why we make these topics optional. The most important application of integration comes from the Fundamental Theorem of Calculus: Integration can be used to "undo" differentiation and to solve differential equations. That application will be covered at length in Chapter 8.

6.3.1 Cumulative Change

Consider a population whose size at time t, $t \geq 0$, is $N(t)$ and that grows at a rate $r(t)$. Referring back to interpretations of the derivative in Section 4.2, we can say that:

$$\frac{dN}{dt} = r(t) \tag{6.15}$$

because dN/dt is the rate of growth of the population represented using derivatives. Using the Fundamental Theorem of Calculus we can see that, since $N(t)$ is an antiderivative of the function $r(t)$,

$$N(t) = \int_0^t r(s)\,ds + C \tag{6.16}$$

for some constant C. (We choose 0 as the lower limit of integration for convenience—the lower limit is arbitrary, and any value could be used.) If we know the function $r(t)$, then (6.16) can be used to calculate $N(t)$, but this only gives us $N(t)$ up to an unknown constant C. To calculate C, we need to know the value of $N(t)$ at a specific time; that is, we need an **initial condition**. The initial condition, say $N(0) = N_0$ for some known value of N_0, and the differential (6.15) together make an initial value problem

$$\frac{dN}{dt} = r(t) \text{ and } N(0) = N_0$$

whose solution is:

$$N(t) = \int_0^t r(s)\,ds + N_0. \qquad \text{Substitute } t = 0 \text{ into (6.16) to show } C = N_0 \tag{6.17}$$

EXAMPLE 1 A population grows at rate $r(t) = \frac{1}{2}t^2$ and at time $t = 0$ contains 200 individuals. Find the size of the population (number of individuals) as a function of time.

Solution Let $N(t)$ represent the size of the population at time t. Then we are given that

$$\frac{dN}{dt} = \frac{1}{2}t^2 \quad \text{and} \quad N(0) = 200.$$

We can solve this initial value problem using (6.17):

$$N(t) = \int_0^t \frac{1}{2}s^2\,ds + 200$$
$$= \frac{1}{6}t^3 + 200.$$

Although it is possible to memorize (6.17), it is better to derive the equation when you need it, and the following argument gives a direct way to do that. If:

$$\frac{dN}{dt} = r(t)$$

then, if we integrate both sides of the equation from 0 to t, we obtain:

$$\int_0^t \frac{dN}{ds}\,ds = \int_0^t r(s)\,ds \qquad \text{Replace } t \text{ by a different variable in the integrands.}$$
$$N(t) - N(0) = \int_0^t r(s)\,ds$$

For the last line we use the Fundamental Theorem of Calculus, noting that $N(t)$ is the antiderivative of dN/dt.

This way of writing the equation tells us that the integral $\int_0^t r(s)\,ds$ is equal to the total number of individuals added to the population between times 0 and t. This count is also referred to as the **net change** or **cumulative change** in the population over this interval. That is, more generally, if a quantity y changes at a rate dy/dt, then:

$$\begin{matrix}\text{cumulative change in} \\ y \text{ over interval } [a, b]\end{matrix} = y(b) - y(a) = \int_a^b \frac{dy}{dt}\,dt.$$

EXAMPLE 2 A vehicle travels along a straight road at a velocity that changes with time, t. Suppose that

$$v(t) = t - t^2.$$

Find the total distance traveled by the vehicle in time t.

Solution Let $s(t)$ represent the distance traveled by the vehicle at time t. Then the velocity of the vehicle represents the rate at which its distance along the road changes, i.e., when written using derivatives:

$$\frac{ds}{dt} = v(t)$$

so:

$$\int_0^t \frac{ds}{du}\,du = \int_0^t v(u)\,du \qquad \text{Integrate both sides from 0 to } t$$

i.e., $s(t) - s(0) = \int_0^t v(u)du$, and the left-hand side of this equation gives the cumulative distance that the vehicle travels from time 0 to time t. Hence:

$$\text{distance traveled } = s(t) - s(0) = \int_0^t (u - u^2)\,du$$
$$= \left.\frac{1}{2}u^2 - \frac{1}{3}u^3\right]_0^t$$
$$= \frac{1}{2}t^2 - \frac{1}{3}t^3$$

EXAMPLE 3 A rain gauge can be used to measure the rate of rainfall (e.g., in millimeters per hour) at a weather station. At a tropical weather station, rain falls at approximately the same time each day. We model this by approximating the rate of rainfall by a periodic function of time t:

$$r(t) = 3(1 - \cos(2\pi t))$$

where t is measured in fractions of a day since midnight (so for example, $t = 0.5$ is noon, $t = 0.75$ is 6 pm, $t = 1$ is midnight again). Calculate the total (cumulative) rainfall between times 0 and t.

Solution If R is the total rainfall between time 0 and time t, then, since r is the rate of rainfall:

$$\frac{dR}{dt} = r(t) \tag{6.18}$$

Note also that $R(0) = 0$ (the rainfall between time 0 and time $t = 0$ is zero). So integrating both sides of (6.18), we obtain:

$$\int_0^t \frac{dR}{ds}\,ds = \int_0^t r(s)\,ds$$

$$R(t) - \underbrace{R(0)}_{0} = \int_0^t r(s)\,ds$$

thus

$$R(t) = \int_0^t 3(1 - \cos(2\pi s))\,ds$$

$$= 3s - \frac{3}{2\pi}\sin(2\pi s)\Big]_0^t \qquad \frac{d}{ds}\left(\frac{1}{2\pi}\cdot\sin(2\pi s)\right) = \cos(2\pi s)$$

$$R(t) = 3t - \frac{3}{2\pi}\sin 2\pi t.$$

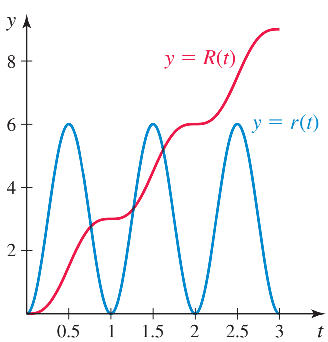


Figure 6.28 The rate of rainfall $r(t)$, and total cumulative rainfall $R(t)$ from Example 3.

We show a graph of $R(t)$ in Figure 6.28. Note that although $r(t)$ is periodic with period t, $R(t)$ is monotonic increasing and so not periodic. This makes sense because the cumulative amount of rain that has fallen must always increase over time.

6.3.2 Average Values

The concentration of soil nitrogen in g/m^3 was measured every meter along a transect (i.e., a straight line path) in moist tundra and yielded the following data:

Distance along Transect **(m)**	***Concentration*** **(g/m^3)**
0	589.3
1	602.7
2	618.5
3	667.2
4	641.2
5	658.3
6	672.8
7	661.2
8	652.3
9	669.8

If we denote the concentration at distance x by $c(x)$, then the average concentration, denoted by $\bar{c}$ (read "c bar"), is the arithmetic average

$$\bar{c} = \frac{1}{10}(c(0) + c(1) + c(2) + \cdots + c(9)) = 643.3\ \text{g/m}^3$$

More generally, to find the average concentration between two points a and b along a transect, we measure the concentration at equal distances. To formulate this notation mathematically, we evenly divide $[a, b]$ using $n+1$ points: $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ with each pair of points separated by a distance $w = \frac{b-a}{n}$ (that is, $x_k - x_{k-1} = w$ for $k = 1, 2, \ldots, n$). We then measure the concentration at each point. If the concentration at location x_k is denoted by $c(x_k)$, then the average concentration $\overline{c}$ is

$$\overline{c} = \frac{1}{n+1}\left(c(x_0) + c(x_1) + c(x_2) + \cdots + c(x_n)\right) \tag{6.19}$$

This is similar to our equation for the integral, defined in Section 6.1, using Riemann sums, and we will manipulate (6.19) into a form that makes the similarity even stronger:

$$\begin{aligned}\overline{c} &= \frac{1}{n+1}\left(c(x_0) + c(x_1) + \cdots + c(x_{n-1})\right) + \frac{c(x_n)}{n+1} && \text{Separate off last term}\\ &= \frac{w\left(c(x_0) + c(x_1) + \cdots + c(x_{n-1})\right)}{(n+1)w} + \frac{c(b)}{n+1} && \text{Multiply numerator and denominator by } w\end{aligned}$$

As $n \to \infty$ the last term, $c(b)/(n+1)$, converges to 0. In the first term, the numerator is the approximation of the area under the curve $c(x)$ between $x = a$ and $x = b$, i.e., it converges to $\int_a^b c(x)\,dx$ as $n \to \infty$. Meanwhile, in the denominator, $(n+1)w = \frac{(n+1)(b-a)}{n} = (b-a)\left(1 + \frac{1}{n}\right)$, which converges to $(b-a)$ as $n \to \infty$. Hence, as the number, n, of sampled points on the transect increases, $\overline{c}$ converges to:

$$\overline{c} = \frac{1}{b-a}\int_a^b c(x)\,dx.$$

That is, the average concentration can be expressed as an integral over $c(x)$ between a and b, divided by the length of the interval $[a, b]$:

Average Value of a Function Assume that $f(x)$ is a continuous function on $[a, b]$. The average value of f on the interval $[a, b]$ is

$$\overline{f} = \frac{1}{b-a}\int_a^b f(x)\,dx \tag{6.20}$$

EXAMPLE 4 Find the average value of $f(x) = 4 - x^2$ on the interval $[-2, 2]$.

Solution We use (6.20). Note that $f(x) = 4 - x^2$ is continuous on $[-2, 2]$. Then the average value of f on the interval $[-2, 2]$ is

$$\begin{aligned}f_{\text{avg}} &= \frac{1}{2-(-2)}\int_{-2}^{2}(4 - x^2)\,dx = \frac{1}{4}\left[4x - \frac{1}{3}x^3\right]_{-2}^{2}\\ &= \frac{1}{4}\left[8 - \frac{8}{3} + 8 - \frac{8}{3}\right] = \frac{1}{4}\cdot\frac{32}{3} = \frac{8}{3}\end{aligned}$$

•

EXAMPLE 5 **Average Rainfall.** Rainfall in Los Angeles varies seasonally, with most rain occurring at the beginning and end of the year. We might model this seasonal variation using the following formula for the monthly precipitation (in inches/month).

$$p(t) = 1.6 + 1.6\cos(2\pi t)$$

where t is the fraction of the year elapsed since January 1, so $t = 0$ is January 1, $t = 0.5$ is exactly halfway through the year (which turns out to be July 2), $t = 0.75$ is three-quarters of the way through the year (which turns out to be October 1). What is the average monthly rainfall in one year?

Solution We use (6.20):

$$\begin{aligned}\overline{P} &= \int_0^1 (1.6 + 1.6\cos(2\pi t))\,dt \\ &= 1.6\int_0^1 1\cdot dt + 1.6\int_0^1 \cos(2\pi t)\,dt \\ &= 1.6 + 1.6\left(\frac{1}{2\pi}\sin 2\pi t\Big]_0^1\right) \qquad b-a=1 \\ &= 1.6. \qquad \sin 0 = \sin 2\pi = 0\end{aligned}$$

EXAMPLE 6 **Surfactant in the Lungs.** The small airways of the lungs are coated by a thin film of water. This water helps keep the airways clean and protects the cells in the lungs from dust and from drying out. However, the surface tension of the water can close small airways. So the body makes surfactants, biological soaps, that lower the surface tension of the airway water. Human fetuses start to make these surfactants while still in the womb, but babies born before the 32nd week may not have enough surfactant for their airways to open. Doctors must administer the surfactant directly. As one model for how surfactant enters the lungs of a prematurely born baby, imagine surfactant spreading on a thin film of water covering a surface that represents the airway wall. The film thickness is h, and the velocity of the water will increase with height y, from $y = 0$ (the wall) to $y = h$ (the top of the film). The surfactant lies on the top of the film (at $y = h$), so one model for the flow of water in the film is that the water at height y travels at a velocity

$$u(y) = U_0 y/h$$

where U_0 is the velocity of the **Marangoni** flow that the surfactant creates. And because water "sticks" to the airway wall, the speed of flow at the wall is zero. Show that the surfactant travels at twice the average speed of water in the airway.

Solution We are told that the surfactant travels along the top of the film, i.e., has velocity $u(h) = U_0 h/h = U_0$. The average velocity of water in the film is

$$\overline{U} = \frac{1}{h}\int_0^h U(y)\,dy = \frac{1}{h}\left(\frac{U_0 y^2}{2h}\Big]_0^h\right) = \frac{1}{2}U_0$$

i.e., the surfactant travels twice as fast as the water. This is useful therapeutically, because it means that surfactant can be added to a prematurely born baby's lungs without adding much water at the same time.

6.3.3 The Mean Value Theorem

If $\overline{f}$ is the average value of a function $f(x)$ over an interval $[a, b]$, does $f(x)$ need to attain the value $\overline{f}$ at any point? Under some conditions the answer is yes.

> **Theorem The Mean-Value Theorem for Definite Integrals** Assume that $f(x)$ is a continuous function on $[a, b]$. Then there exists a number $c \in [a, b]$ such that
>
> $$f(c)(b-a) = \int_a^b f(x)\,dx$$

That is, when we compute the average value of a function that is continuous on $[a, b]$, we find that there exists a number c such that $f(c) = \overline{f}$. We can understand this concept graphically when we look at the graph of a function f and at $\overline{f}$. For simplicity, let's assume that $f(x) \geq 0$. Since (1) $\int_a^b f(x)\,dx$ is then equal to the area between the graph of $f(x)$ and the x-axis, (2) $\overline{f} \cdot (b - a)$ is equal to the area of the rectangle with height $\overline{f}$ and width $b - a$, and (3) the two areas are equal, the horizontal line $y = \overline{f}$ must intersect the graph of $f(x)$ at some point on the interval $[a, b]$. (See Figure 6.29.) The x-coordinate of this point of intersection is then the value c in the MVT for definite integrals. (Note that there could be more than one such number.) A similar argument can be made when we do not assume that $f(x)$ is positive. In this case, "area" is replaced by "signed area." The proof of this theorem is short, and we supply it for completeness at the end of this subsection.

EXAMPLE 7 Find the average value of $f(x) = x^3$ on the interval $[0, 1]$, and determine $x \in [0, 1]$ such that $f(x)$ equals the average value.

Solution The function $f(x) = x^3$ is continuous on $[0, 1]$. Then

$$\overline{f} = \int_0^1 x^3\,dx = \left[\frac{1}{4}x^4\right]_0^1 = \frac{1}{4}(1)^4 = \frac{1}{4} \qquad b - a = 1$$

$f(x) = \frac{1}{4}$ when $x^3 = \frac{1}{4}$; that is, $f(x)$ takes on its average value at $x = \left(\frac{1}{4}\right)^{1/3} \simeq 0.630$. (See Figure 6.30).

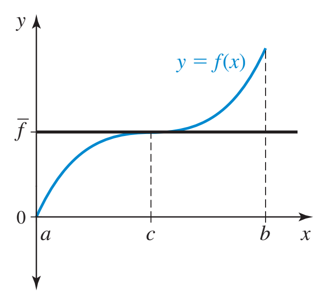


Figure 6.29 An illustration of the average value of a function: $\int_a^b f(x)\,dx = \overline{f}(b - a)$.

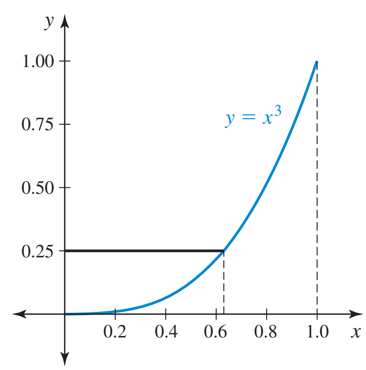


Figure 6.30 The graph of $y = x^3$, $0 \leq x \leq 1$. The average value is $\frac{1}{4}$, attained when $x \simeq 0.630$.

EXAMPLE 8 **Flow in a River** The speed of water in a channel varies considerably with depth. Because of friction, the velocity reaches zero at the bottom and along the sides of the channel; the velocity is greatest near the surface of the water. The average velocity of a stream is of interest in characterizing rivers. One way to obtain this average value would be to measure a stream's velocity at various depths along a vertical transect and then average the values obtained. In practice, however, a much simpler method is employed: The speed is measured at 60% of the depth from the surface, because the speed at that depth is very close to the average speed. Explain why it is possible that the measurement at just one depth would yield the average stream velocity.

Solution Assuming that the velocity profile of the stream along a vertical transect is a continuous function of depth, the MVT for definite integrals guarantees that there exists a depth h at which the velocity is equal to the average stream velocity.

The MVT only gives the existence of such a depth; it does not tell us where the velocity is equal to the average velocity. It is surprising and fortunate that the depth where the velocity reaches its average is quite universal; that is, it does not depend much on the specifics of the river. (The 60%-depth rule is derived in Problems 17–20 of Chapter 6 Review Problems.)

Proof of the Mean-Value Theorem for Definite Integrals Since $f(x)$ is continuous on $[a, b]$, we can apply the Extreme-Value Theorem to conclude that f attains an absolute maximum and an absolute minimum in $[a, b]$. If we denote the absolute maximum by M and the absolute minimum by m, then

$$m \le f(x) \le M \quad \text{for all } x \in [a, b]$$

and f takes on both m and M for some values in $[a, b]$. We therefore find that

$$m(b-a) \le \int_a^b f(x)\,dx \le M(b-a),$$

or

$$m \le \frac{1}{b-a}\int_a^b f(x)\,dx \le M.$$

We set $I = \frac{1}{b-a}\int_a^b f(x)\,dx$; then $m \le I \le M$.

Using the facts that $f(x)$ takes on all values between m and M in the interval $[a, b]$ (this follows from the intermediate-value theorem) and that I is a number between m and M, it follows (also from the intermediate-value theorem) that there must be a number $c \in [a, b]$ such that $f(c) = I$; that is,

$$f(c) = \frac{1}{b-a}\int_a^b f(x)\,dx.$$

6.3.4 Areas

We first introduced integrals as the area under curves. Specifically, if f is a nonnegative, continuous function on $[a, b]$, then

$$A = \int_a^b f(x)\,dx$$

represents the area of the region bounded by the graph of $f(x)$ between a and b, the vertical lines $x = a$ and $x = b$, and the x-axis between a and b. In all of the examples we have presented thus far, one of the boundaries of the region whose area we wanted to know has been the x-axis. We will now discuss how to find the geometric area between two arbitrary curves. We emphasize that we want to compute geometric areas; that is, the areas we compute in this subsection will always be positive.

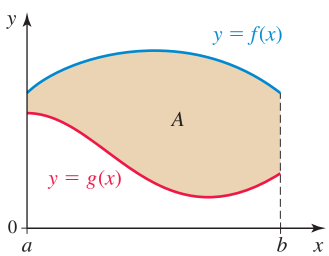


Figure 6.31 Computing the area between the two curves.

Suppose that $f(x)$ and $g(x)$ are continuous functions on $[a, b]$. We wish to find the area between the graphs of f and g. We assume for the moment that both f and g are nonnegative on $[a, b]$ and that $f(x) \ge g(x)$ on $[a, b]$. (See Figure 6.31.) From the figure, we see that

$$\begin{aligned} A &= \begin{bmatrix} \text{area between} \\ f \text{ and } x\text{-axis} \end{bmatrix} - \begin{bmatrix} \text{area between} \\ g \text{ and } x\text{-axis} \end{bmatrix} \\ &= \int_a^b f(x)\,dx - \int_a^b g(x)\,dx \end{aligned}$$

Using Property (4) of Subsection 6.1.3, we can write this equation as

$$A = \int_a^b [f(x) - g(x)]\,dx$$

We obtained this formula under the assumption that both f and g are nonnegative on $[a, b]$. But that assumption is not necessary. Assume that $f(x) \ge g(x)$ for all $x \in [a, b]$, but do not assume that $f(x)$ and $g(x)$ are always non-negative (see Figure 6.32).

Define $h(x) = f(x) - g(x)$. $h(x)$ gives the distance between the curves at the point x. Thus the height of the function, $y = h(x)$, is the same as the thickness of the gap between the curves. So the area under $y = h(x)$ must be equal to the area between the curves, i.e.,

$$A = \int_a^b h(x)\,dx = \int_a^b (f(x) - g(x))\,dx$$

just as before. Geometrically, drawing $y = h(x)$ is like drawing the thickness of the gap between the two curves with the lower curve, $y = g(x)$, straightened out along the x-axes.

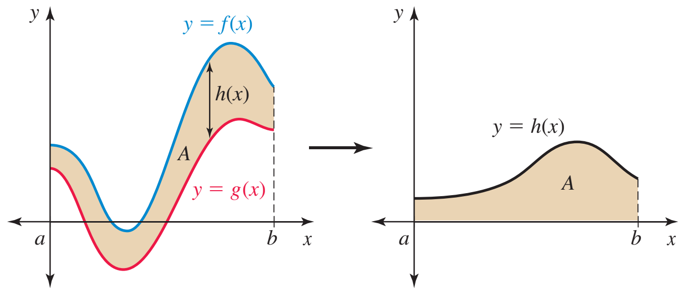


Figure 6.32 The area under $y = h(x) = f(x) - g(x)$ is equal to the area between the two curves $y = f(x)$ and $y = g(x)$.

The following box expresses this result:

> If f and g are continuous on $[a, b]$, with $f(x) \geq g(x)$ for all $x \in [a, b]$, then the area of the region between the curves $y = f(x)$ and $y = g(x)$ from a to b is equal to
> $$\text{Area} = \int_a^b [f(x) - g(x)]\,dx$$

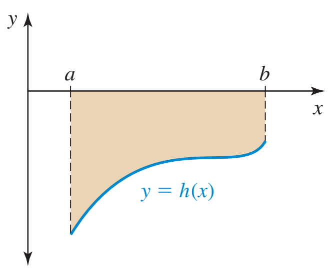


Figure 6.33 The definite integral $\int_a^b h(x)\,dx$ represents a signed area; it is negative here.

Before looking at a number of examples, we point out once more that this area formula always yields a nonnegative number since it computes the geometric area. To contrast this kind of area with the concept of a signed area, let us consider a function $h(x)$ on $[a, b]$, with $h(x) \leq 0$ for all $x \in [a, b]$. (See Figure 6.33.) The definite integral $\int_a^b h(x)\,dx$ represents a signed area and is negative in this case; more precisely, $\int_a^b h(x)\,dx$ is the negative of the geometric area of the region between the x-axis and the graph of $h(x)$ from $x = a$ to $x = b$. That this relationship is consistent with our definition of area can be seen as follows: The region of interest is bounded by the two curves $y = 0$ and $y = h(x)$. Since $h(x) \leq 0$ for $x \in [a, b]$, the area formula yields

$$\text{Area} = \int_a^b [0 - h(x)]\,dx = -\int_a^b h(x)\,dx$$

which is a positive number.

When you compute the area between two curves, you should always graph the bounding curves. This will show you how to set up the appropriate integral(s).

EXAMPLE 9 Find the area between the curves $y = \sec^2 x$ and $y = \cos x$ from $x = 0$ to $x = \pi/4$.

Solution We first graph the bounding curves, as shown in Figure 6.34. We see that both $\sec^2 x$ and $\cos x$ are continuous on $[0, \pi/4]$ and that $\sec^2 x \geq \cos x$ for $x \in [0, \frac{\pi}{4}]$. Therefore,

$$\begin{aligned}
\text{Area} &= \int_0^{\pi/4} [\sec^2 x - \cos x]\,dx \\
&= \tan x - \sin x\Big]_0^{\pi/4} \\
&= \left(\tan\frac{\pi}{4} - \sin\frac{\pi}{4}\right) - (\tan 0 - \sin 0) \\
&= \left(1 - \frac{1}{2}\sqrt{2}\right) - (0 - 0) = 1 - \frac{1}{2}\sqrt{2}.
\end{aligned}$$

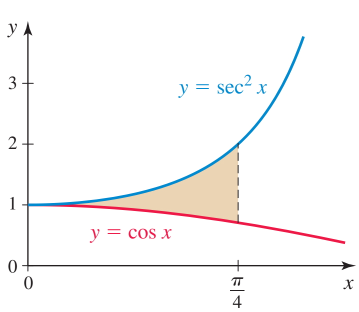


Figure 6.34 The region for Example 9.

EXAMPLE 10 Find the area of the region enclosed by $y = (x-1)^2 - 1$ and $y = -x + 2$.

Solution The bounding curves are graphed in Figure 6.35. To find the points where the two curves intersect, we solve

$$\begin{aligned}(x-1)^2 - 1 &= -x + 2\\ x^2 - 2x + 1 - 1 &= -x + 2\\ x^2 - x - 2 &= 0 \qquad \text{Bring all terms to one side}\\ (x+1)(x-2) &= 0 \qquad \text{Factorize}\end{aligned}$$

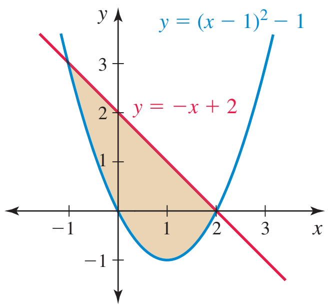


Figure 6.35 The region for Example 10.

Therefore,

$$x = -1 \qquad \text{and} \qquad x = 2$$

are the x-coordinates of the points of intersection. Note that both $y = (x-1)^2 - 1$ and $y = -x + 2$ are continuous on $[-1, 2]$. Since $-x + 2 \geq (x-1)^2 - 1$ for $x \in [-1, 2]$, the area of the enclosed region is

$$\begin{aligned}\text{Area} &= \int_{-1}^{2} [(-x+2) - ((x-1)^2 - 1)]\,dx\\ &= \int_{-1}^{2} [-x + 2 - x^2 + 2x - 1 + 1]\,dx\\ &= \int_{-1}^{2} [-x^2 + x + 2]\,dx = -\frac{1}{3}x^3 + \frac{1}{2}x^2 + 2x\Big]_{-1}^{2}\\ &= \left(-\frac{1}{3}(2)^3 + \frac{1}{2}(2)^2 + (2)(2)\right) - \left(-\frac{1}{3}(-1)^3 + \frac{1}{2}(-1)^2 + (2)(-1)\right)\\ &= -\frac{8}{3} + 2 + 4 - \frac{1}{3} - \frac{1}{2} + 2 = \frac{9}{2}.\end{aligned}$$

•

EXAMPLE 11 Find the area of the region bounded by $y = \sqrt{x}$, $y = x - 2$, and the x-axis.

Solution We first graph the bounding curves in Figure 6.36. We see that $y = \sqrt{x}$ and $y = x - 2$ intersect. To find the point of intersection, we solve

$$\begin{aligned}x - 2 &= \sqrt{x}\\ (x-2)^2 &= (\sqrt{x})^2 \qquad \text{Square both sides}\\ x^2 - 4x + 4 &= x\\ x^2 - 5x + 4 &= 0\end{aligned}$$

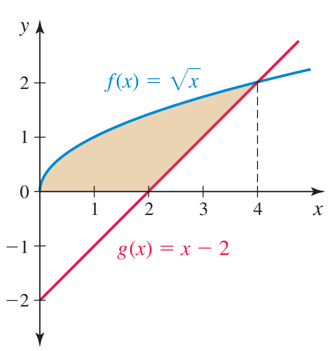


Figure 6.36 Evaluating the area between $f(x) = \sqrt{x}$ and $g(x) = x - 2$.

or

$$(x-4)(x-1) = 0 \qquad \text{Factorize}$$

which yields the solutions $x = 4$ and $x = 1$. Since squaring an equation can introduce extraneous solutions, we need to check whether the solutions satisfy $x - 2 = \sqrt{x}$. When $x = 4$, we find that $4 - 2 = \sqrt{4}$, which is a true statement; when $x = 1$, we find that $1 - 2 = \sqrt{1}$, which is a false statement. Hence, $x = 4$ is the only solution.

The graph of $y = x - 2$ intersects the x-axis at $x = 2$. To compute the area, we need to split the integral into two parts, because the lower bounding curve is composed of two parts: the x-axis from $x = 0$ to $x = 2$ and the line $y = x - 2$ from $x = 2$ to $x = 4$. We see from the graph that all bounding curves are continuous on their respective intervals. We get

$$\begin{aligned}\text{Area} &= \overbrace{\int_0^2 \sqrt{x}\,dx}^{\text{first interval}} + \overbrace{\int_2^4 [\sqrt{x} - (x-2)]\,dx}^{\text{second interval}}\\ &= \left[\frac{2}{3}x^{3/2}\right]_0^2 + \left[\frac{2}{3}x^{3/2} - \frac{1}{2}x^2 + 2x\right]_2^4\\ &= \frac{2}{3}\cdot 2^{3/2} + \frac{2}{3}\cdot 4^{3/2} - \frac{1}{2}\cdot 16 + 8 - \frac{2}{3}\cdot 2^{3/2} + 2 - 4 = \frac{10}{3}.\end{aligned}$$

•

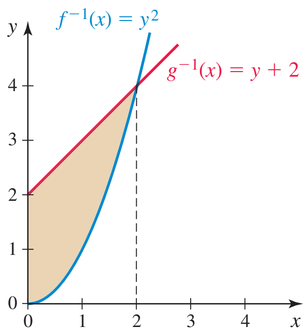


Figure 6.37 The area in Example 11 can be expressed as the area between $f^{-1}(y)$ and $g^{-1}(y)$.

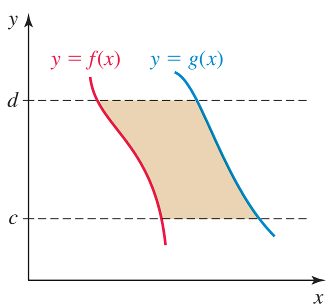


Figure 6.38 The area between $f(x)$ and $g(x)$.

In the preceding example, we needed to split the integral into two parts because the lower boundary of the area was composed of two different curves: one for $0 < x < 2$ and one for $2 < x < 4$. In these cases, it is sometimes more convenient to integrate over y, rather than over x. We show this approach in Figure 6.37, in which we flipped one axis so that x is the dependent variable. If we denote the curves in Figure 6.36 as $f(x) = \sqrt{x}$ and $g(x) = x - 2$, then when the area is converted into a y-integral, the area of interest is between $f^{-1}(y)$ and $g^{-1}(y)$.

Since the same two curves define the entire region, there is no need to split the integral. Using $g^{-1}(y) = y + 2$ and $f^{-1}(y) = y^2$, we obtain, for the total area,

$$\text{Area} = \int_0^2 (g^{-1}(y) - f^{-1}(y))dy \qquad \text{g^{-1} is the upper curve, f^{-1} is the lower curve}$$

$$= \int_0^2 (y + 2 - y^2)\,dy = \frac{1}{2}y^2 + 2y - \frac{1}{3}y^3\Big]_0^2$$

$$= 2 + 4 - \frac{8}{3} = \frac{10}{3},$$

which is the same as the result in Example 11.

More generally, if a region is bounded on the left by $y = f(x)$, on the right by $y = g(x)$ below by $y = c$, and above by $y = d$ (see Figure 6.38). Its area is given by the following formula:

$$\boxed{\text{Area} = \int_c^d [g^{-1}(y) - f^{-1}(y)]\,dy}$$

6.3.5 The Volume of a Solid

From geometry, we know various formulas for computing the volumes of regular solids, like circular cylinders. To compute the volume of a less regularly shaped solid, we will use an approach that is similar to that for computing areas using integrals. These calculations provide good practice in constructing integrals by approximating a body using simple elements, and then letting the number of elements get larger and larger. The results themselves are very useful in engineering and physics, but we are aware of few examples of their use in life sciences problems. For that reason we regard this subject to be optional.

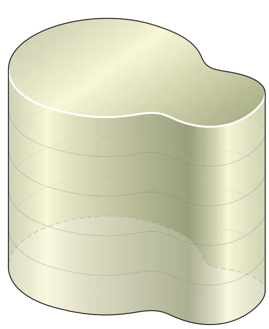


Figure 6.39 A right cylinder with an irregularly shaped base.

We begin with the volume of a generalized cylinder; the volume is the base area times the height. The base can be any arbitrarily shaped region. (See Figure 6.39, for example.)

If we denote the base area by A and the height of the cylinder by h, then the volume of the generalized cylinder is

$$V = Ah$$

As an example, consider the circular cylinder whose base is a disk. If the disk has radius r and the cylinder has height h, then the volume of the circular cylinder is $\pi r^2 h$. We will use cylinders to approximate volumes of more complicated solids rather like the way that we used rectangles to approximate the area under a curve in Section 6.1.

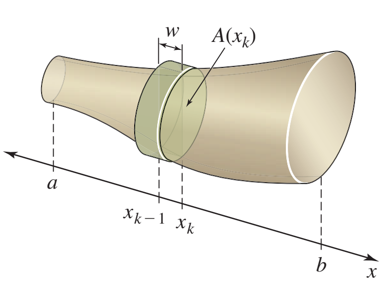


Figure 6.40 The volume of an irregularly shaped solid, found by the disk method.

Suppose that we wish to compute the volume of the solid shown in Figure 6.40. We can slice the solid into small slabs by cutting it perpendicular to the x-axis at points $x_0 = a < x_1 < x_2 < \cdots < x_n = b$ that evenly divide the interval $[a, b]$ into n subintervals and then slice the solid into planes. The intersection of such a plane and the solid is called a *cross section*. We denote the area of the cross section at x_k by $A(x_k)$. By cutting the solid along these planes, we obtain slices, just as we do when we cut bread. We will approximate the volume of a slice between x_{k-1} and x_k by the volume of a cylinder with base area equal to that of the slice at x_k and height $w = x_k - x_{k-1}$.

The volume of the slice between x_{k-1} and x_k is then approximately

$$A(x_{k-1})\,w$$

Adding the volumes of all the slices gives us an approximation for the total volume of the solid:

$$V \approx A(x_0)w + A(x_1)w + \cdots + A(x_{n-1})w = \sum_{k=0}^{n-1} A(x_k)\,w$$

By using more and more subintervals we can improve the approximation. As $n \to \infty$, the sum above becomes an integral over x:

Definition The **volume of a solid** of integrable cross-sectional area $A(x)$ between a and b is

$$\int_a^b A(x)\,dx$$

EXAMPLE 12 Find the volume of the sphere of radius r centered at the origin.

Solution The cross section at x is perpendicular to the x-axis. (See Figure 6.41.) It is a disk of radius $y = \sqrt{r^2 - x^2}$ and whose area is

$$A(x) = \pi y^2 = \pi(r^2 - x^2) \qquad \text{Area of a circle of radius } \sqrt{r^2 - x^2}$$

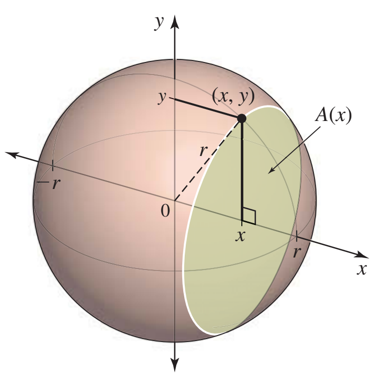


Figure 6.41 The volume of a sphere, calculated by integrating areas $A(x)$.

Since the solid is between $-r$ and r, it follows that

$$\text{Volume} = \int_{-r}^{r} \pi(r^2 - x^2)\,dx$$

The integrand is continuous on $[-r, r]$. Evaluating the integral yields

$$= \pi\left[r^2 x - \frac{1}{3}x^3\right]_{-r}^{r}$$

$$= \pi\left[\left(r^3 - \frac{1}{3}r^3\right) - \left(-r^3 + \frac{1}{3}r^3\right)\right] = \pi\left(\frac{2}{3}r^3 + \frac{2}{3}r^3\right) = \frac{4}{3}\pi r^3$$

This result agrees with the formula that we know from geometry. •

The cross sections of the sphere in the last example were all disks. We can think of a sphere as a **solid of revolution**—that is, a solid obtained by revolving a curve about the x-axis (or the y-axis). In this case, we rotate the curve $y = \sqrt{r^2 - x^2}$, $-r \le x \le r$, about the x-axis, which creates circular cross sections.

We can use other curves $y = f(x)$, rotate them about the x-axis, and obtain solids in the same way. We illustrate in Figure 6.42, in which we rotate the graph of $y = f(x)$, $a \le x \le b$, about the x-axis. A cross section through x perpendicular to the x-axis is then a disk with radius $f(x)$; hence, its cross-sectional area is $A(x) = \pi[f(x)]^2$. If we use the formula $\int_a^b A(x)\,dx$ to compute the volume of the solid, we find that the volume of the solid of revolution is

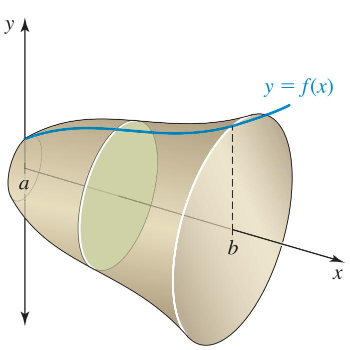


Figure 6.42 The solid of rotation when rotating $f(x)$ about the x-axis.

$$V = \int_a^b \pi[f(x)]^2\,dx \qquad (6.21)$$

Computing volumes by using (6.21) is called the **disk method**.

EXAMPLE 13 Compute the volume of the solid obtained by rotating $y = x^2$, $0 \le x \le 2$, about the x-axis.

Solution We illustrate the solid in Figure 6.43. When we rotate the graph of $y = x^2$ about the x-axis, we find that the cross section at x is a disk with radius $y = f(x) = x^2$. The cross-sectional area at x, $A(x)$, is then $\pi(x^2)^2 = \pi x^4$, which is integrable on $[0, 2]$. Thus, the volume is

$$V = \int_0^2 \pi[f(x)]^2\,dx = \int_0^2 \pi x^4\,dx$$
$$= \frac{\pi}{5}x^5\Big]_0^2 = \frac{32}{5}\pi.$$

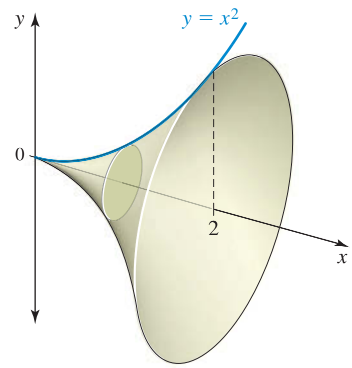


Figure 6.43 The solid of rotation for Example 13.

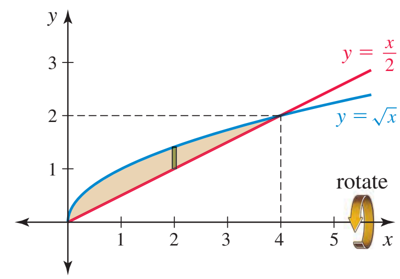


Figure 6.44 The plane region for Example 14.

EXAMPLE 14 Rotate the area bounded by the curves $y = \sqrt{x}$ and $y = x/2$ about the x-axis, and compute the volume of the solid of rotation.

Solution The curves $y = \sqrt{x}$ and $y = x/2$ are graphed in Figure 6.44, together with a vertical bar to indicate the cross section. We see from the graph that the curves intersect at $x = 0$ and $x = 4$. To find the points of intersection algebraically, we need to equate $\sqrt{x}$ and $x/2$ and solve for x:

$$\sqrt{x} = \frac{x}{2}$$

This equation immediately yields the solution $x = 0$. If $x > 0$, we can divide by $\sqrt{x}$ and find

$$1 = \frac{\sqrt{x}}{2}, \qquad \text{or} \qquad 2 = \sqrt{x}$$

Squaring yields $x = 4$. Thus, the curves intersect at $x = 0$ and $x = 4$. We can compute the volume of this solid of rotation by first rotating $y = \sqrt{x}$, $0 \le x \le 4$, about the x-axis and computing the volume of this solid, and then subtracting the volume of the solid obtained by rotating $y = x/2$, $0 \le x \le 4$. When we do so, we get

$$V = \int_0^4 \pi(\sqrt{x})^2\,dx - \int_0^4 \pi\left(\frac{1}{2}x\right)^2\,dx$$

Both integrands are continuous on $[0, 4]$ and we find that

$$V = \pi \cdot \frac{1}{2} \cdot x^2\Big]_0^4 - \pi \cdot \frac{1}{12} \cdot x^3\Big]_0^4$$
$$= 8\pi - \frac{16}{3}\pi = \frac{8}{3}\pi$$

Looking at Figure 6.45, we see that the cross-sectional area is that of a washer. In this case, the disk method is also referred to as the **washer method**.

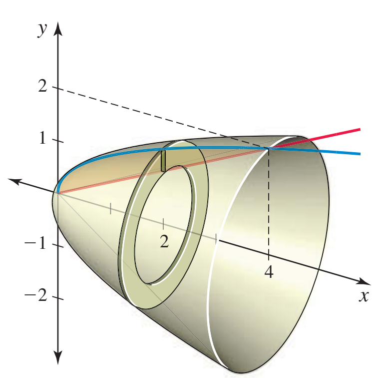


Figure 6.45 The solid of rotation for Example 14 can be made up of washer-like element.

In the next example, we rotate a curve about the y-axis.

EXAMPLE 15 Rotate the region bounded by $y = 2$, $x = 0$, $y = 0$, and $y = \ln x$ about the y-axis, and compute the volume of the resulting solid.

Solution When we rotate about the y-axis, the cross sections are perpendicular to the y-axis. At $y = \ln x$, the radius of the cross-sectional disk is x. (See Figure 6.46.) The solid is shown in Figure 6.47. Since we "sum" the slices along the y-axis, we must integrate with respect to y; because $y = \ln x$, we get $x = e^y$. Therefore, the cross-sectional area at y is $A(y) = \pi(e^y)^2$, which is integrable on $[0, 2]$, and the volume is

$$V = \int_0^2 \pi(e^y)^2\,dy = \int_0^2 \pi e^{2y}\,dy = \pi \frac{1}{2} e^{2y}\Big]_0^2 = \frac{\pi}{2}(e^4 - 1).$$

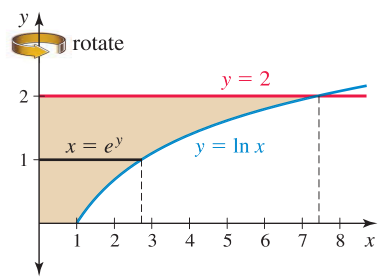


Figure 6.46 The plane region for Example 15.

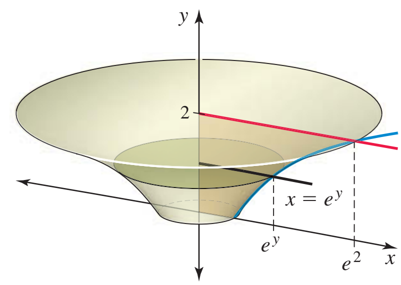


Figure 6.47 The solid of rotation for Example 15.

6.3.6 Rectification of Curves

In this subsection, we will study how to compute the lengths of curves in the plane. This is another example in which integrals appear as we add up a large number of small increments.

How would we **rectify** (i.e., determine the length of a curve) with a ruler? We could approximate the curve by short line segments, measure the length of each segment, and add up the measurements, as illustrated in Figure 6.48. By choosing smaller line segments, our approximation would improve. This is precisely the method that we will employ to find an exact formula.

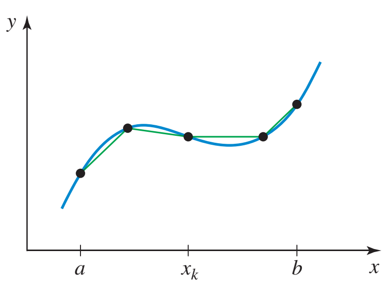


Figure 6.48 The length of a curve in the plane.

To find the exact formula, assume that the curve whose length we want to find is given by a function $y = f(x)$, $a \le x \le b$, which has a continuous first derivative on (a, b). We evenly divide the interval $[a, b]$ into subintervals by defining points $a = x_0 < x_1 < x_2 < \cdots < x_n = b$, and approximate the curve by a polygon that consists of the straight-line segments connecting neighboring points on the curve, as shown in Figure 6.48. A typical line segment connecting the points x_{k-1} and x_k is shown in Figure 6.49. Using the Pythagorean theorem, we can find its length. We set $w = x_k - x_{k-1}$ and $\Delta y_k = y_k - y_{k-1}$. Then the length of the line segment is given by

$$\sqrt{w^2 + (\Delta y_k)^2}$$

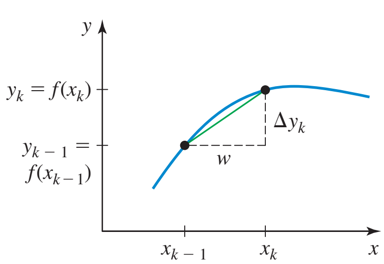


Figure 6.49 A typical line segment.

Summing each of these lengths, we obtain an estimate for the total length of the curve:

$$L = \sum_{k=1}^{n} \sqrt{w^2 + (\Delta y_k)^2} \tag{6.22}$$

With finer and finer subintervals, the length of the polygon will become a better and better approximation of the length of the corresponding curve. However, before we can take this limit, we need to work on the sum.

The difference Δy_k is equal to $f(x_k) - f(x_{k-1})$. If we linearize the function $f(x)$ we can approximate this difference $f(x_k) \approx f(x_{k-1}) + (x_k - x_{k-1})f'(x_{k-1})$.

Since $x_k - x_{k-1} = w$, it follows that

$$\Delta y_k \approx f'(x_k)(x_k - x_{k-1}) \approx f'(x_{k-1})\,w \tag{6.23}$$

And this approximation gets better and better as $w \to 0$. Replacing Δy_k in (6.22) by (6.23), we find that the length of the polygon is given by

$$\begin{aligned} L &\approx \sum_{k=1}^{n} \sqrt{w^2 + [f'(x_{k-1})\, w]^2} \\ &= \sum_{k=1}^{n} \sqrt{1 + [f'(x_{k-1})]^2}\, w \end{aligned} \tag{6.24}$$

This form allows us to take the limit as $n \to 0$ and obtain

$$\begin{aligned} L &= \lim_{n\to\infty} \sum_{k=1}^{n} \sqrt{1 + [f'(x_{k-1})]^2}\, w \\ &= \int_a^b \sqrt{1 + [f'(x)]^2}\, dx \qquad \text{Using the definition of a Riemann integral} \end{aligned}$$

Thus,

> If $f(x)$ is differentiable on (a, b) and $f'(x)$ is continuous on $[a, b]$, then the length of the curve $y = f(x)$ from a to b is given by
>
> $$L = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx$$

We can make this derivation rigorous by observing that the Mean Value Theorem guarantees that there is a point $c_k \in [x_{k-1}, x_k]$ for which $f'(c_k) = \frac{f(x_k)-f(x_{k-1})}{x_k - x_{k-1}}$. So if x_{k-1} is replaced by c_k in (6.24), then the formula for L becomes exact.

The first curve that was rectified was $y = f(x) = x^{3/2}$ whose length was calculated in 1657, by William Neile. We will choose this function for our first example of rectification.

EXAMPLE 16 Determine the length of the curve given by the graph of $y = f(x) = x^{3/2}$ between $a = 5/9$ and $b = 21/9$.

Solution To determine the length of the curve, we need to find $f'(x)$ first. We have

$$f'(x) = \frac{3}{2}x^{1/2}$$

Then

$$L = \int_{5/9}^{21/9} \sqrt{1 + \left[\frac{3}{2}x^{1/2}\right]^2}\, dx = \int_{5/9}^{21/9} \sqrt{1 + \frac{9}{4}x}\, dx$$

An antiderivative of $\sqrt{1 + \frac{9}{4}x}$ is $\frac{4}{9} \cdot \frac{2}{3} \left(1 + \frac{9}{4}x\right)^{3/2}$, which can be checked by differentiating the latter function with respect to x. Thus, the length is

$$\begin{aligned} L &= \left[\frac{4}{9} \cdot \frac{2}{3}\left(1 + \frac{9}{4}x\right)^{3/2}\right]_{5/9}^{21/9} = \frac{8}{27}\left[\left(1 + \frac{9}{4} \cdot \frac{21}{9}\right)^{3/2} - \left(1 + \frac{9}{4} \cdot \frac{5}{9}\right)^{3/2}\right] \\ &= \frac{8}{27}\left[\left(\frac{5}{2}\right)^3 - \left(\frac{3}{2}\right)^3\right] = \frac{8}{27}\left(\frac{125}{8} - \frac{27}{8}\right) = \frac{98}{27}. \end{aligned}$$

The length of the curve is therefore 98/27. •

Before we present another example, we discuss the formula in more detail. Using

$$\frac{dy}{dx} = f'(x)$$

we can write

$$L = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

for the length. Treating dy and dx as if they were numbers, we can rewrite this equation as

$$L = \int_a^b \sqrt{(dx)^2 + (dy)^2}$$

We call the expression $\sqrt{(dx)^2 + (dy)^2}$ the **arc length differential** and denote it by ds. We can think of ds as a typical infinitesimal line segment. The Pythagorean theorem in this infinitesimal form then becomes $(ds)^2 = (dx)^2 + (dy)^2$. "Adding up" these line segments (i.e., computing $\int_a^b ds$) then yields the length of the curve.

EXAMPLE 17 Determine the length of

$$f(x) = \frac{1}{4}x^2 - \frac{1}{2}\ln x \quad \text{from } x = 1 \text{ to } x = e$$

Solution Differentiating $f(x)$, we find that

$$f'(x) = \frac{x}{2} - \frac{1}{2x}$$

The length of the curve is then given by

$$\begin{aligned} L &= \int_1^e \sqrt{1 + \left(\frac{x}{2} - \frac{1}{2x}\right)^2}\, dx = \int_1^e \sqrt{1 + \left(\frac{x^2}{4} - \frac{1}{2} + \frac{1}{4x^2}\right)}\, dx \\ &= \int_1^e \sqrt{\frac{x^2}{4} + \frac{1}{2} + \frac{1}{4x^2}}\, dx \end{aligned}$$

We notice that the expression under the square root is a perfect square, namely,

$$\frac{x^2}{4} + \frac{1}{2} + \frac{1}{4x^2} = \left(\frac{x}{2} + \frac{1}{2x}\right)^2$$

Hence, the integral for the length simplifies to

$$\begin{aligned} L &= \int_1^e \sqrt{\left(\frac{x}{2} + \frac{1}{2x}\right)^2}\, dx \\ &= \int_1^e \left(\frac{x}{2} + \frac{1}{2x}\right) dx = \frac{1}{4}x^2 + \frac{1}{2}\ln|x|\Big]_1^e \\ &= \frac{1}{4}e^2 + \frac{1}{2} - \frac{1}{4} = \frac{1}{4}(e^2 + 1)\,. \end{aligned}$$

Because of the somewhat complicated form of the integrand in the computation of the length, we quickly run into problems when we actually try to compute the integral. In practice, the integrand rarely simplifies enough for easy computation, as it did in Examples 16 and 17.

Even seemingly simple looking functions, such as $y = 1/x$, quickly turn into complicated integrals when we compute the length of the curve.

EXAMPLE 18 Set up, but do not evaluate, the length of the curve of the hyperbola $f(x) = \frac{1}{x}$ between $a = 1$ and $b = 2$.

Solution To determine the length of the curve, we need to find $f'(x)$ first.

$$f'(x) = -\frac{1}{x^2}$$

Then the length of the curve is given by the integral

$$L = \int_1^2 \sqrt{1 + \left(-\frac{1}{x^2}\right)^2}\, dx = \int_1^2 \sqrt{1 + \frac{1}{x^4}}\, dx$$

The antiderivative of the integrand in Example 18 is quite complicated, and we will not be able to find it with the techniques available in this text. In Section 7.5, we will learn numerical methods for evaluating integrals, and some of these methods can be used to evaluate the integral in Example 18. Using these numerical methods we find that the length L is approximately 1.13.

Section 6.3 Problems

6.3.1

1. Consider a population whose size at time t is $N(t)$ and whose growth obeys the initial-value problem

$$\frac{dN}{dt} = e^{-t}$$

with $N(0) = 100$.

(a) Find $N(t)$ by solving the initial-value problem.

(b) Compute the cumulative change in population size between $t = 0$ and $t = 5$.

(c) Express the cumulative change in population size between time 0 and time t as an integral. Give a geometric interpretation of this quantity.

2. Suppose that a change in biomass $B(t)$ at time t during the interval $[0, 12]$ follows the equation

$$\frac{dB}{dt}(t) = \cos\left(\frac{\pi}{6}t\right)$$

for $0 \leq t \leq 12$.

(a) Graph $\frac{dB}{dt}$ as a function of t.

(b) Suppose that $B(0) = B_0$. Express the cumulative change in biomass during the interval $[0, t]$ as an integral. Give a geometric interpretation. What is the value of the biomass at the end of the interval $[0, 12]$ compared with the value at time 0? How are these two quantities related to the cumulative change in the biomass during the interval $[0, 12]$?

3. A particle moves along the x-axis with velocity

$$v(t) = -(t-2)^2 + 1$$

for $0 \leq t \leq 5$. Assume that the particle is at the origin at time 0.

(a) Graph $v(t)$ as a function of t.

(b) Use the graph of $v(t)$ to determine when the particle moves to the left and when it moves to the right.

(c) Find the location $s(t)$ of the particle at time t for $0 \leq t \leq 5$. Give a geometric interpretation of $s(t)$ in terms of the graph of $v(t)$.

(d) Graph $s(t)$ and find the leftmost and rightmost positions of the particle.

4. Recall that the acceleration $a(t)$ of a particle moving along a straight line is the instantaneous rate of change of the velocity $v(t)$; that is,

$$a(t) = \frac{d}{dt}v(t)$$

Assume that $a(t) = 32$ ft/s^2. Express the cumulative change in velocity during the interval $[0, t]$ as a definite integral, and compute the integral.

5. If $\frac{dl}{dt}$ represents the growth rate of an organism at time t (measured in months), explain what

$$\int_2^7 \frac{dl}{dt}\, dt$$

represents.

6. If $\frac{dw}{dx}$ represents the rate of change of the weight of an organism of age x, explain what

$$\int_3^5 \frac{dw}{dx}\, dx$$

means.

7. If $\frac{dB}{dt}$ represents the rate of change of biomass of a plant at time t, explain what

$$\int_1^6 \frac{dB}{dt}\, dt$$

means.

8. Let $N(t)$ denote the size of a population at time t, and assume that

$$\frac{dN}{dt} = f(t)$$

Express the cumulative change of the population size in the interval $[0, 3]$ as an integral.

9. Rainfall Suppose that rain falls at a rate $r(t)$, measured in mm/hr, which depends on time t, measured in hours.

(a) Interpret in words the quantity $\int_0^{24} r(t)\, dt$.

(b) If $r(t) = \frac{1.5}{1+t}$, calculate $\int_0^{24} r(t)\,dt$.

10. Fish Growth The rate of growth of a fish is sometimes modeled by the equation

$$dL/dt = L_0 e^{-kt}$$

where L is the length of the fish, and k and L_0 are positive constants.

(a) Interpret in words the quantity $\int_0^3 \frac{dL}{dt}\,dt$.

(b) Calculate the integral from part (a); your answer will include the constants k and L_0.

6.3.2

11. Let $f(x) = x^2 - 2$. Compute the average value of $f(x)$ over the interval $[0, 2]$.

12. Let $g(t) = \sin(\pi t)$. Compute the average value of $g(t)$ over the interval $[-1, 1]$.

13. Suppose that the temperature T (measured in degrees Fahrenheit) in a growing chamber varies over a 24-hour period according to

$$T(t) = 68 + \sin\left(\frac{\pi}{12}t\right)$$

for $0 \le t \le 24$.

(a) Graph the temperature T as a function of time t.

(b) Find the average temperature and explain your answer graphically.

14. Suppose that the concentration (measured in gm^{-3}) of nitrogen in the soil along a transect in moist tundra yields data points that follow a straight line with equation

$$y = 673.8 - 34.7x$$

for $0 \le x \le 10$, where x is the distance to the beginning of the transect. What is the average concentration of nitrogen in the soil along this transect?

15. Let $f(x) = \tan x$. Give a geometric argument to explain why the average value of $f(x)$ over $[-1, 1]$ is equal to 0.

16. Temperature The average daily temperature (measured in Fahrenheit) in New York city can be approximated by the following function of the time of year t. (t measures the fraction of the year that has elapsed since January 1.)

$$T(t) = 57.5 - 22.5\cos(2\pi t).$$

(a) Sketch the function $T(t)$ against t.

(b) What is the average daily temperature high, averaged over the course of one year?

(c) Explain how you could get your answer in part (b) without doing any integrations.

(d) What is the average summer temperature? You may assume that summer corresponds to the interval $0.47 \le t \le 0.73$. You will need to use a calculator to evaluate your answer.

17. Temperature The typical daily temperature high, measured in $^\circ C$ (degrees Celsius), in Los Angeles varies over the course of a year according to the formula $T(t) = 21.7 + 3.1\cos(2\pi(t - 0.75))$ (where t measures the fraction of the year that has elapsed since January 1).

(a) Sketch the graph of the function $T(t)$.

(b) What is the average daily temperature high, averaged over the course of one year?

(c) Explain how you could get your answer in part (b) without doing any integrations.

(d) What is the average winter temperature? You may assume that winter corresponds to the interval $0 \le t \le 0.22$ and $0.98 \le t \le 1$. You will need to use a calculator to evaluate your answer.

6.3.3

18. Suppose that you drive from St. Paul to Duluth and you average 50 mph. Explain why there must be a time during your trip at which your speed is exactly 50 mph.

19. Let $f(x) = 2x$, $0 \le x \le 2$. Use a geometric argument to find the average value of f over the interval $[0, 2]$, and find x such that $f(x)$ is equal to this average value.

20. A particle moves along the x-axis with velocity

$$v(t) = -(t-3)^2 + 5$$

for $0 \le t \le 6$.

(a) Graph $v(t)$ as a function of t for $0 \le t \le 6$.

(b) Find the average velocity of this particle during the interval $[0, 6]$.

(c) Find a time $t^* \in [0, 6]$ such that the velocity at time t^* is equal to the average velocity during the interval $[0, 6]$. Is it clear that such a point exists? Is there more than one such point in this case? Use your graph in (a) to explain how you would find t^* graphically.

21. Rainfall We can approximate the monthly rainfall in Los Angeles using the formula $p(t) = 1.6 + 1.6\cos 2\pi t$, where t is the fraction of the year elapsed since January 1 and $p(t)$ is the rainfall measured in inches/month.

(a) Explain (without doing any calculations) that there will be a value of t for which the monthly rainfall is exactly equal to the annual average rainfall.

(b) Find this value of t (there may be more than one).

(c) Assuming there are 12 months in a year and each month has the same duration, show that the total rainfall in one year is:

$$P_{\text{total}} = 12\int_0^1 p(t)\,dt$$

Explain in particular why the factor 12 is needed and what the units of P_{total} are.

(d) Calculate P_{total}.

22. Temperature The daily temperature high in Minneapolis–St. Paul can be modeled using a formula:

$$T(t) = 11.5 + 7.5\cos(2\pi(t - 0.6))$$

where T is measured in degrees Celsius and t measures the fraction of the year that has elapsed since January 1.

(a) Find the average annual daily temperature high.

(b) Find the time of year (value of t) at which this average daily temperature high is actually observed.

6.3.4

Find the areas of the regions bounded by the lines and curves in Problems 23–34.

23. $y = x^2 - 1$, $y = x + 1$

24. $y = 2x^2 - 1$, $y = 1$

25. $y = e^{x/2}$, $y = -x$, $x = 0$, $x = 2$

26. $y = \cos x, y = 0, x = 0, x = \dfrac{\pi}{2}$

27. $y = x^2 + 1, y = 2x, x = 0$

28. $y = x^2, y = 3x - 2$

29. $y = x^2, y = \dfrac{1}{x}, x = 1, x = 2$

30. $y = 1, y = \cos x$ from $x = 0$ to $x = \dfrac{\pi}{2}$

31. $y = \sin x, y = 1$ from $x = 0$ to $x = \dfrac{\pi}{4}$

32. $y = x^2, y = (x-2)^2, y = 0$ from $x = 0$ to $x = 1$

33. $y = x^2, y = x^3$ from $x = 0$ to $x = 2$

34. $y = e^{-x}, y = x + 1$ from $x = -1$ to $x = 1$

In Problems 35–38, find the areas of the regions bounded by the lines and curves by expressing x as a function of y and integrating with respect to y.

35. $y = x^2, y = (x-2)^2, y = 0$ from $x = 0$ to $x = 2$

36. $y = x, y = x^2$

37. $y = x, y = 0, y = 1 - x$, from $x = 0$ to $x = 1$

38. $y = \sqrt{x},\ y = 0,\ y = \sqrt{2-x}$ from $x = 0$ to $x = 2$

6.3.5

39. Find the volume of a right circular cone with base radius r and height h.

40. Find the volume of a pyramid with square base of side length a and height h.

In Problems 41–46, find the volumes of the solids obtained by rotating the region bounded by the given curves about the x-axis. In each case, sketch the region and a typical disk element.

41. $y = 4 - x^2, y = 0, x = 0$ (in the first quadrant)

42. $y = \sqrt{x}, y = 0, x = 1$

43. $y = x, 0 \le x \le 1$

44. $y = e^x, y = 0, x = 0, x = \ln 2$

45. $y = x^2, -1 \le x \le 1$

46. $y = \sqrt{1 - x^2}, 0 \le x \le 1, y = 0$

In Problems 47–58, find the volumes of the solids obtained by rotating the region bounded by the given curves about the x-axis. In each case, sketch the region together with a typical disk element.

47. $y = x^2, y = x, 0 \le x \le 1$

48. $y = 2 - x^3, y = 2 + x^3, 0 \le x \le 1$

49. $y = e^x, y = e^{-x}, 0 \le x \le 2$

50. $y = \sqrt{1 - x^2}, y = 1, -1 \le x \le 1$

51. $y = x, y = 1, 0 \le x \le 1$

52. $y = 1 - x^2, y = 1, -1 \le x \le 1$

53. $y = \sqrt{x}, y = 2, x = 0$

54. $y = x^2, y = 4, x = 0$ (in the first quadrant)

55. $y = |x|, y = 1, -1 \le x \le 1$

56. $y = \sqrt{x}, y = x, 0 \le x \le 1$

57. $y = x^3, y = x^2, 0 \le x \le 1$

58. $y = |x|, y = 0, -1 \le x \le 1$

6.3.6

59. Find the length of the straight line

$$y = 2x$$

from $x = 0$ to $x = 2$ by each of the following methods:

(a) planar geometry

(b) the integral formula for the lengths of curves, derived in Subsection 6.3.6

60. Find the length of the straight line

$$y = mx$$

from $x = 0$ to $x = a$, where m and a are positive constants, by each of the following methods:

(a) Pythagoras' theorem

(b) the integral formula for the lengths of curves, derived in Subsection 6.3.6

61. Find the length of the curve

$$y^2 = x^3$$

from $x = 1$ to $x = 4$.

62. Find the length of the curve

$$3y^2 = 4x^3$$

from $x = 0$ to $x = 1$.

63. Find the length of the curve

$$y = \frac{x^4}{4} + \frac{1}{8x^2}$$

from $x = 1$ to $x = 3$.

64. Find the length of the curve

$$y = \frac{x^3}{6} + \frac{1}{2x}$$

from $x = 2$ to $x = 4$.

In Problems 65–68, set up, but do not evaluate, the integrals for the lengths of the following curves:

65. $y = x^2, -1 \le x \le 1$

66. $y = x^2 + 1, -1 \le x \le 1$

67. $y = e^{-x}, 0 \le x \le 1$

68. $y = \frac{1}{x}, 1 \le x \le 2$

69. Find the length of the quarter-circle

$$y = \sqrt{1 - x^2}$$

for $0 \le x \le 1$, by each of the following methods:

(a) a formula from geometry

(b) the integral formula from Subsection 6.3.6

70. A cable that hangs between two poles at $x = -M$ and $x = M$ takes the shape of a catenary, with equation

$$y = \frac{1}{2a}(e^{ax} + e^{-ax})$$

where a is a positive constant. Compute the length of the cable when $a = 1$ and $M = \ln 3$.

71. Show that if

$$f(x) = \frac{e^x + e^{-x}}{2}$$

then the length of the curve $f(x)$ between $x = 0$ and $x = a$ for any $a > 0$ is given by $f'(a)$.

Chapter 6 Review

Key Terms

Discuss the following definitions and concepts:

1. Area
2. Signed areas
3. Approximating areas using rectangles
4. Riemann sum
5. Definite integral
6. Riemann integrable
7. Geometric interpretation of definite integrals
8. The constant-value and constant-multiple rules for integrals
9. The definite integral over a union of intervals
10. Comparison rules for definite integrals
11. The fundamental theorem of calculus, part I
12. Leibniz's rule
13. Antiderivatives
14. The fundamental theorem of calculus, part II
15. Evaluating definite integrals by using the FTC, part II
16. Cumulative change and definite integrals
17. Calculating mean of a function
18. The mean-value theorem for definite integrals
19. Computing the area between curves by using definite integrals
20. The volume of a solid and definite integrals
21. Volume of revalution
22. Length of a curve
23. Arc length differential

Review Problems

In Problems 1–6 calculate the indefinite integral of the function $f(x)$.

1. $f(x) = x^{3/2}$

2. $f(x) = e^{-2x}$

3. $f(x) = \frac{1}{x+1}$

4. $f(x) = \sin 2x$

5. $f(x) = \frac{1}{x^2+1}$

6. $f(x) = \frac{x+1}{x}$

In Problems 7–12 calculate the definite integrals.

7. $\int_0^1 (x^2+1)\,dx$

8. $\int_0^{\pi/2} \sin x\,dx$

9. $\int_0^{-1} x\,dx$

10. $\int_0^1 \frac{1}{x^2+1}\,dx$

11. $\int_0^{\pi} \sin 2x\,dx$

12. $\int_0^1 \frac{x}{x+1}\,dx$

13. Insect Flight A highly simplified model for how a small insect flies is to model the force exerted by a moving wing by the following formula:

$$F = C_D \rho U^2 \sin\alpha$$

where U is the velocity of the wing, α is the angle that the wing is held at by the insect, ρ is the density of the air around the insect, and C_D is a coefficient that depends on the shape of the wing. When hovering, the insect will beat its wing back and forth, so $U = U_0 \sin\omega t$ where U_0 and ω are positive constants.

(a) Interpret U_0 and ω in terms of the maximum speed of the wing and the period of its motion.

(b) By making a sketch of F as a function of t, explain why the average lift force is $\frac{1}{2}C_D \rho U_0^2 \sin\alpha$.

(c) For the insect to remain aloft, the average lift force that its wings generate must exceed the weight, W, of the insect. Calculate the minimum speed, U_0, that allows this (your answer will depend on α, ρ, C_D, and W).

14. Bacterial Biofilms In nature bacteria form biofilms—mats of cells stuck together by proteins. Biofilms stick to substrates, helping to prevent bacteria from being swept away. By sticking together the bacteria create a barrier that protects them from antibiotics. But the stickiness of biofilms makes it hard for bacteria to spread and move around within them. One strategy for bacterial spreading was investigated by ***Angelini et al. (2009)***, who found that biological soap produced by the bacteria set the biofilm into motion by altering its surface tension—an effect similar to adding dish soap to a basin of water. In this problem we will consider a simplified spreading scenario. Bacteria are initially concentrated at one end of a thin film of biofilm, contained in a Petri dish (see Figure 6.50).

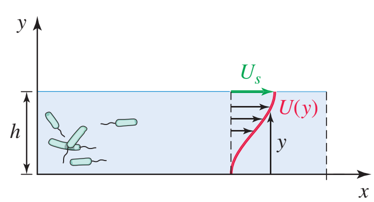


Figure 6.50 A thin biofilm of bacteria spreads when the bacteria create chemicals that alter the surface tension of the biofilm.

The biological soap created by the bacteria creates a surface flow, with speed U_s. Since the bottom of the biofilm remains stuck to the surface that it spreads on, the velocity of the film varies with height, y.

(a) Initially we model the speed of the spreading biofilm by:

$$U(y) = U_s y/h$$

Show that the average speed of the biofilm is $\frac{1}{2}U_s$.

(b) The biofilm is contained in a Petri dish, and the walls of the Petri dish prevent the entire biofilm from moving in the direction of the surface flow, by creating pressure within the biofilm. This pressure creates a flow $-Gy(2h-y)$, that opposes the surface tension flow, i.e.,

$$U(y) = U_s y/h - Gy(2h-y)$$

The constant G is set to ensure that the mean flow in the biofilm $\bar{u}$ is equal to 0. Find the constant G.

(c) When G takes the value that you calculated in part (b), for what range of values of y is $u(y) > 0$?

(d) Sketch the velocity in the biofilm when G takes the value that you calculated in (b).

15. Discharge of a River In studying the total flow of water in an open channel, such as a river in its bed, the total amount of water passing through a cross section per second—the discharge, Q—is of interest. The following formula is used to compute the discharge:

$$Q = \int_0^B \bar{v}(b)h(b)\,db \tag{6.25}$$

In this formula, b is the distance from one bank of the river to the point where the depth $h(b)$ of the river and the average velocity $\bar{v}(b)$ of the vertical velocity profile of the river at b were measured. The total width of the river is B. (See Figure 6.51.)

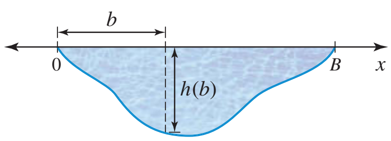


Figure 6.51 The river for Problem 15.

To evaluate the integral in (6.25), we would need to know $\bar{v}(b)$ and $h(b)$ at every location b across the river. In practice, the cross section is divided into a finite number of subintervals and measurements of $\bar{v}$ and h are taken at, say, the end of each subinterval. The following table contains an example of such measurements:

Distance from left bank	*h*	$\bar{v}$
0	0	0
2	0.28	0.172
4	0.76	0.213
6	1.34	0.230
8	1.57	0.256
10	1.42	0.241
12	1.21	0.206
14	0.83	0.187
16	0.42	0.116
18	0	0

The location 0 corresponds to the left bank, and the location $B = 18$ to the right bank, of the river. The units of the location and of h are meters, and of $\bar{v}$, meters per second. Approximate the integral in (6.25) by a Riemann sum, using the locations in the table, and find the approximate discharge, using the data from the table.

16. Pulse Chase Experiments A pulse chase experiment can be used to study how cells process a particular chemical. In the experiment the cells are fed a radioactive or chemically labeled version of the chemical for a certain length of time, then they are fed the non-radioactive version of the same chemical. Scientists then study how quickly the radioactive form of the chemical is used up by the cells.

Suppose that the radioactive form of the chemical is delivered to the cells at a rate $r(t)$.

(a) Interpret in words the quantity $\int_0^t r(s)\,ds$.

(b) Initially you model $r(t)$ using the following formula:

$$r(t) = a\,t\,e^{-kt}$$

where k and a are positive constants. (i) Sketch the function $r(t)$ against time.

(ii) Show that $R(t) = -a\left(\frac{t}{k} + \frac{1}{k^2}\right)e^{-kt}$ is an antiderivative of $r(t)$, and use this to evaluate $\int_0^t r(s)\,ds$.

(iii) Assuming $a = 1$, $k = 1$, sketch $\int_0^t r(s)\,ds$ as a function of t.

(c) You measure $r(t)$ directly for a specific pulse chase experiment and you find:

t (min)	**$r(t)$ (mg/min)**
0	0
1	0.21
2	0.35
3	0.42
4	0.37
5	0.35
6	0.33
7	0.30
8	0.26

By approximating the integral $\int_0^t r(s)\,ds$ by a Riemann sum, use the data in the table to calculate the value of the integral for: (i) $t = 3$, (ii) $t = 5$, (iii) $t = 9$.

Problems 17–20 discuss stream speed profiles and provide a justification for the two measurement methods described next.

(Adapted from ***Herschy, 1995***) The speed of water in a channel varies considerably with depth. Due to friction, the speed reaches zero at the bottom and along the sides of the channel. The speed is greatest near the surface of the stream. To find the average speed for the vertical speed profile, two methods are frequently employed in practice:

1. The 0.6 depth method: The speed is measured at 0.6 of the depth from the surface, and this value is taken as the average speed.
2. The 0.2 and 0.8 depth method: The speed is measured at 0.2 and 0.8 of the depth from the surface, and the average of the two readings is taken as the average speed.

The theoretical speed distribution of water flowing in an open channel is given approximately by

$$v(d) = \left(\frac{D-d}{a}\right)^{1/c} \tag{6.26}$$

where $v(d)$ is the speed at depth d below the water surface, c is a constant varying from 5 for rocky stream beds to 7 for smooth stream beds, D is the total depth of the channel, and a is a positive constant.

17. (a) Sketch the graph of $v(d)$ as a function of d for $D = 3$ m and $a = 1$ m for (i) $c = 5$ and (ii) $c = 7$.

(b) Show that the speed is equal to 0 at the bottom ($d = D$) and is maximal at the surface ($d = 0$).

18. **(a)** Show by integration that the average speed $\overline{v}$ in the vertical profile is given by

$$\overline{v} = \frac{c}{c+1}\left(\frac{D}{a}\right)^{1/c} \tag{6.27}$$

(b) What fraction of the maximum speed $v(0)$ is the average speed $\overline{v}$?

(c) If you knew that the maximum speed occurred at the surface of the river [as predicted in the approximate formula for $v(d)$], how could you find $\overline{v}$? That is, how is $\overline{v}$ related to $v(0)$?

In practice, however, the speed at the free surface is very difficult to measure, because there may be material (leaves, for example) floating at the free surface—the presence of this material means that (6.26) may not apply near the free surface. We therefore need alternate methods for estimating $\overline{v}$. We will evaluate two such methods in Problems 19 and 20.

19. Explain why the depth d_1, at which $v = \overline{v}$, the average speed, is given by the equation

$$\overline{v} = \left(\frac{D-d_1}{a}\right)^{1/c} \tag{6.28}$$

We can find d_1 by equating (6.27) and (6.28). Show that

$$\frac{d_1}{D} = 1 - \left(\frac{c}{c+1}\right)^c$$

and that d_1/D is approximately 0.6 for values of c between 5 and 7, thus resulting in the rule

$$\overline{v} \approx v_{0.6}$$

where $v_{0.6}$ is the speed at depth $0.6D$. (*Hint*: Graph $1-(c/(c+1))^c$ as a function of c for $c \in [5, 7]$, and investigate the range of this function.)

20. We denote by $v_{0.2}$ the speed at depth $0.2D$. We will now find the depth d_2 such that

$$\overline{v} = \tfrac{1}{2}(v_{0.2} + v_{d_2})$$

where $\overline{v}$ is the average speed across the depth.

(a) Show that d_2 satisfies

$$\frac{1}{2}\left[\left(\frac{D-0.2D}{a}\right)^{1/c} + \left(\frac{D-d_2}{a}\right)^{1/c}\right] = \frac{c}{c+1}\left(\frac{D}{a}\right)^{1/c}$$

[*Hint*: Use (6.27).]

(b) Show that

$$\frac{d_2}{D} = 1 - \left[\frac{2c}{c+1} - (0.8)^{1/c}\right]^c$$

and confirm that d_2/D is approximately 0.8 for values of c between 5 and 7, thus resulting in the rule

$$\overline{v} \approx \tfrac{1}{2}(v_{0.2} + v_{0.8}).$$

CHAPTER

7

Integration Techniques and Computational Methods

The primary focus of this chapter is on integration techniques. Specifically, we will learn how to

- integrate by using the substitution rule and integration by parts;
- integrate rational functions;
- integrate when either the integrand is discontinuous or the limits of integration are infinite;
- integrate numerically;
- integrate using tables; and
- approximate functions by polynomials.

So far we have only been able to integrate a function if we can guess or look up its anti-derivative. In this chapter we will learn more methods for integrating functions. Because of the connection between integration and differentiation, the first two techniques are derived directly from techniques for calculating derivatives. The first technique, called the substitution rule, is the chain rule applied backward; the second, called integration by parts, is the product rule applied backward. The chapter's first two sections are devoted to these integration techniques. An additional technique called the method of partial fractions is introduced in the third section. The fourth section deals with improper integrals, which are integrals for which the integrand goes to infinity somewhere over the interval of integration or for which the interval of integration is unbounded.

Even with this enlarged toolkit of integration techniques there are many functions that we are not able to integrate. Two methods are then available: integrating the function numerically, covered in the fifth section, analogous to finding the solution of equations numerically, or using tables of integrals, which is postponed to the seventh section. In the sixth section we cover a problem that seems, initially, unrelated to any of the above topics: how to approximate a function using polynomials. In Section 4.11 we learned how to use linearization—that is, the approximation of a function by a linear function—to estimate the value of a function. Section 7.7 shows how we can estimate the function more accurately by approximating it using a quadratic, cubic, or even higher order polynomial.

7.1 The Substitution Rule

7.1.1 Indefinite Integrals

Let's start with a motivating example. Suppose we want to calculate the following indefinite integral:

$$\int x \cos\left(3x^2 + 1\right)\, dx.$$

This integral is too complicated for us to be able to write down the antiderivative as we did in the examples we studied in Chapter 6. Let's start by identifying what makes the integral difficult to calculate: If instead we needed to evaluate $\int \cos u\, du$, we could immediately find the integral $\sin u + C$, that is, the integral becomes easier if we can replace the complicated function $3x^2 + 1$ by a single variable u, and integrate with respect to that variable. There is more to the integration than this, for example, what happens to the factor of x; but we will cover that too.

Keeping in mind that we will derive our rules for integration from rules for differentiation, we recall a similar problem that we encounter when we try to differentiate a function like:

$$f(x) = \sin(3x^2 + 1).$$

We need to use the chain rule to calculate $f'(x)$. We set $u = 3x^2 + 1$.

Then $f(u) = \sin u$, which we can differentiate more easily: $f'(u) = \frac{df}{du} = \cos u$. To differentiate $f(x)$ we must also differentiate the inner function $u = 3x^2 + 1$, for which:

$$\frac{du}{dx} = 6x$$

We obtain

$$\frac{df}{dx} = \frac{df}{du}\frac{du}{dx} = (\cos u)(6x) = \cos(3x^2 + 1) \cdot 6x \tag{7.1}$$

Let's integrate each of the terms in (7.1) with respect to x, which is the same as displaying the differentiation in reverse order

$$\begin{aligned}\int \cos(3x^2 + 1) \cdot 6x\, dx &= \int \cos \underbrace{u}_{u=(3x^2+1)} \cdot \underbrace{6x\, dx}_{\frac{du}{dx}\, dx} \\ &= \int \cos u\, du = \sin u + C \\ &= \sin(3x^2 + 1) + C.\end{aligned}$$

In the first step, we substituted u for $3x^2 + 1$ and used $du = 6x\, dx$. This substitution simplified the integrand. At the end, we substitute back $3x^2 + 1$ for u to get the final answer in terms of x.

To see the general principle behind this technique, if we write $u = g(x)$ [and hence $du = g'(x)\, dx$] and our integral can be arranged into the form

$$\int f[g(x)]g'(x)\, dx,$$

then we can integrate the function using the chain rule. If we use $F(x)$ to denote an antiderivative of $f(x)$ [i.e., $F'(x) = f(x)$], then, using the chain rule to differentiate $F[g(x)]$, we find that

$$\frac{d}{dx}F[g(x)] = F'[g(x)]g'(x) = f[g(x)]g'(x)$$

which shows that $F[g(x)]$ is an antiderivative of $f[g(x)]g'(x)$. We can therefore write

$$\int f[g(x)]g'(x)\, dx = F[g(x)] + C \tag{7.2}$$

If we set $u = g(x)$, then we can write the right-hand side of (7.2) as $F(u) + C$. But since $F(u)$ is an antiderivative of $f(u)$, it also has the representation

$$F(u) + C = \int f(u)\, du \tag{7.3}$$

Putting (7.2) and (7.3) together we obtain:

Substitution Rule for Indefinite Integrals If $u = g(x)$, then

$$\int f[g(x)]g'(x)\, dx = \int f(u)\, du \tag{7.4}$$

Equation (7.4) is helpful if you can recognize that your integral is of the form $f[g(x)]g'(x)$. In practice, that seldom happens. More often we use the sequence of steps that we followed to derive (7.4); defining a new variable, u, and turning our integration with respect to x into an integration with respect to u.

EXAMPLE 1 **Using Substitution** Evaluate $\int x\cos(3x^2+1)\,dx$.

Solution This is the integral that we used to motivate using the derivation of Equation (7.4). How would we analyze the integral if we did not know in advance what form its antiderivative would take?

The hardest part of calculating the antiderivative is $\cos(3x^2+1)$. Seeking to simplify this expression, we let $u = 3x^2+1$ so that our integral becomes $\int x\cos u\,dx$. We can integrate $\cos u$ with respect to u. But our integral currently is with respect to x. To convert it to an integral with respect to u we make use of the derivative:

$$\frac{du}{dx} = 6x$$

So $x = \frac{1}{6}\frac{du}{dx}$ and we can rewrite our integral as:

$$\begin{aligned}\frac{1}{6}\int \cos u \cdot \frac{du}{dx}dx &= \frac{1}{6}\int \cos u\,du \qquad \text{Use (7.4)}\\ &= \frac{1}{6}\sin u + C\\ &= \frac{1}{6}\sin(3x^2+1)+C.\end{aligned}$$

To derive this result we made use of the Leibniz notation, namely, if $u = g(x)$, then $g'(x) = du/dx$.

The Leibniz notation provides a convenient shorthand for converting from integrals over x to integrals over u:

$$\frac{du}{dx}\,dx = du \tag{7.5}$$

If we treat $\frac{du}{dx}$ as a real ratio, rather than as a derivative, then (7.5) can be derived by cancelling the factor dx on the left-hand side.

Integration by substitution therefore requires that we perform three steps:

1. Define a new variable $u = g(x)$.
2. Convert from an integral over x to an integral over u using $du = \frac{du}{dx}\,dx = g'(x)\,dx$.
3. Rewrite the integral so that x does not appear anywhere.

In practice the hardest part of integration by substitution is the first step: identifying a suitable change of variable $u = g(x)$. The following examples will show you some of the strategies used to find the right change of variables.

EXAMPLE 2 **Using Substitution** Evaluate $\int (2x+1)e^{x^2+x}\,dx$.

Solution This is a difficult integral because of the exponentiated function. So we set $u = x^2+x$, hoping to turn our integrand into a function like e^u.

To convert into an integral over u, we use.

$$\frac{du}{dx} = 2x+1 \quad \text{or} \quad du = (2x+1)\,dx$$

Hence,

$$\int \underbrace{e^{x^2+x}}_{e^u} \underbrace{(2x+1)\,dx}_{du} = \int e^u\,du = e^u + C = e^{x^2+x} + C.$$

In the last step, we substituted $x^2 + x$ back for u, since we want the final result in terms of x. •

EXAMPLE 3 **Using Substitution** Evaluate

$$\int \frac{1}{x \ln x}\,dx.$$

Solution The difficult part of this integral is the term $\frac{1}{\ln x}$.

So we set $u = \ln x$, hoping to convert the integrand into a function like $\frac{1}{u}$. To convert into an integral over u we use:

$$\frac{du}{dx} = \frac{1}{x} \quad \text{or} \quad du = \frac{1}{x}\,dx$$

Then

$$\int \underbrace{\frac{1}{\ln x}}_{\frac{1}{u}} \underbrace{\frac{1}{x}\,dx}_{du} = \int \frac{1}{u}\,du = \ln|u| + C = \ln|\ln x| + C$$ •

Examples 2 and 3 illustrate types of integrals that are frequently encountered, and we display those integrals as follows for ease of reference:

$$\int g'(x)e^{g(x)}\,dx = e^{g(x)} + C \qquad \text{In Example 2, } g(x) = x^2 + x$$

$$\int \frac{g'(x)}{g(x)}\,dx = \ln|g(x)| + C \qquad \text{In Example 3, } g(x) = \ln x$$

EXAMPLE 4 **Multiplicative Constant** Evaluate $\int 4x\sqrt{x^2+1}\,dx$.

Solution We set $u = x^2 + 1$, hoping to convert the integrand into a function like $\sqrt{u}$.

To convert into an integral over u, use

$$\frac{du}{dx} = 2x \qquad \text{or} \qquad du = 2x\,dx$$

So:

$$\int 4x\sqrt{x^2+1}\,dx = \int \underbrace{\sqrt{x^2+1}}_{\sqrt{u}} \cdot \underbrace{4x\,dx}_{2du} = \int 2\sqrt{u}\,du$$

$$= 2 \cdot \frac{2}{3}u^{3/2} + C = \frac{4}{3}(x^2+1)^{3/2} + C$$ •

Note that when we converted the integrand we did not get exactly $\int \sqrt{u}\,du$, but even with the additional factor of 2 we could still evaluate the integral.

EXAMPLE 5 **Rewriting the Integrand** Evaluate $\int \tan x\,dx$.

Solution It is not immediately apparent why a change of variables can help us to evaluate this integral, since in the way that it is written there is only one part to the integrand. A useful trick here is to recall that $\tan x$ is defined by $\frac{\sin x}{\cos x}$. To simplify the integrand we let $u = \cos x$ to try to turn the integrand into a function like $\frac{1}{u}$. Then to convert into an integral over u we use

$$\frac{du}{dx} = -\sin x \quad \text{or} \quad -du = \sin x\,dx$$

Then

$$\int \tan x\,dx = \int \frac{\sin x}{\cos x}\,dx = \int \underbrace{\frac{1}{\cos x}}_{\frac{1}{u}}\,\underbrace{\sin x\,dx}_{-du}$$

$$= -\int \frac{1}{u}\,du = -\ln|u| + C = -\ln|\cos x| + C$$

In Table 6-1 of Section 6.2, we listed $\int \tan x\,dx = \ln|\sec x| + C$. This is the same result that we have just obtained, since $-\ln|\cos x| = \ln|\cos x|^{-1} = \ln|\sec x|$. •

In Problem 59, you will evaluate $\int \cot x\,dx$ using the same trick that we used for Example 5.

Integrals of tan *x* and cot *x*.

$$\int \tan x\,dx = -\ln|\cos x| + C$$

$$\int \cot x\,dx = \ln|\sin x| + C$$

In fact both of these results may be regarded as special cases of the result $\int \frac{g'(x)}{g(x)}dx = \ln|g(x)| + C$. To integrate $\tan x$, take $g(x) = \cos x$, while to integrate $\cos x$, take $g(x) = \sin x$.

It is not always obvious that substitution will be successful; often we need to have some courage to convert the entire integral from x to u.

EXAMPLE 6 **Substitution and Square Roots** Evaluate

$$\int x\sqrt{2x-1}\,dx$$

Solution Set $u = 2x - 1$, hoping to turn the integrand into a function like $\sqrt{u}$. Then to integrate over u, we use

$$\frac{du}{dx} = 2 \quad \text{or} \quad dx = \frac{du}{2}$$

Since $u = 2x - 1$, we have $x = \frac{1}{2}(u+1)$. Making all the substitutions, we find that

$$\int x\sqrt{2x-1}\,dx = \int \frac{1}{2}(u+1)\sqrt{u}\,\frac{du}{2}$$

Initially it looks like we have not simplified the integral at all, but we notice that we may multiply out the integrand to obtain:

$$\int x\sqrt{2x-1}\,dx = \frac{1}{4}\int (u^{3/2} + u^{1/2})\,du = \frac{1}{4}\left(\frac{2}{5}u^{5/2} + \frac{2}{3}u^{3/2}\right) + C$$

$$= \frac{1}{10}(2x-1)^{5/2} + \frac{1}{6}(2x-1)^{3/2} + C$$ •

Functions in the integrand are not always explicitly given, as in the next example.

EXAMPLE 7 Assume that $g(x)$ is a differentiable function whose derivative $g'(x)$ is continuous. Evaluate

$$\int g'(x)\cos[g(x)]\,dx.$$

Solution We set $u = g(x)$ and transform to an integral over u using:

$$\frac{du}{dx} = g'(x) \quad \text{or} \quad du = g'(x)\,dx.$$

The integral then transforms into the form.

$$\int g'(x)\cos[g(x)]\,dx = \int \cos u\,du = \sin u + C = \sin[g(x)] + C$$

7.1.2 Definite Integrals

Part II of the FTC says that to evaluate a definite integral, we must find an antiderivative of the integrand and then evaluate the antiderivative at the limits of integration. When we use the substitution $u = g(x)$ to find an antiderivative of an integrand, the antiderivative will be given in terms of u at first. To complete the calculation, we can proceed in either of two ways: (1) We can leave the antiderivative in terms of u and change the limits of integration according to $u = g(x)$, or (2) we can write the antiderivative as a function of x by substituting $u = g(x)$ and then evaluate the antiderivative at the limits of integration in terms of x. We illustrate these two ways by evaluating

$$\int_0^4 2x\sqrt{x^2+1}\,dx.$$

Recall that when we compute definite integrals, we need to check whether the integrand is continuous over the interval of integration. This is still the case here.

First Way. We change the limits of integration along with the substitution. That is, we set $u = x^2 + 1$ and transform to an integral over u by the substitutions:

$$\frac{du}{dx} = 2x \quad \text{or} \quad du = 2x\,dx,$$

$$\text{if } x = 0, \quad \text{then } u = 1$$

$$\text{if } x = 4, \quad \text{then } u = 17.$$

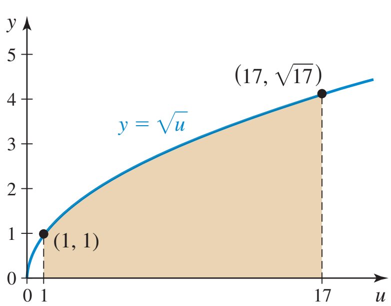


Figure 7.1 The definite integral $\int_0^4 2x\sqrt{x^2+1}\,dx$ becomes $\int_1^{17}\sqrt{u}\,du$ after the substitution $u = x^2 + 1$. The region corresponding to $\int_1^{17}\sqrt{u}\,du$ has an area of $\frac{2}{3}[17^{3/2} - 1]$.

Hence,

$$\int_0^4 2x\sqrt{x^2+1}\,dx = \int_1^{17}\sqrt{u}\,du = \frac{2}{3}u^{3/2}\Big]_1^{17} = \frac{2}{3}[17^{3/2} - 1]$$

Transform limits of integration

After substitution, the integrand is $\sqrt{u}$, and the limits of integration are $u = 1$ and $u = 17$. The region corresponding to the definite integral after substitution is shown in Figure 7.1.

The first way is the more common one, and is summarized as follows:

> **Substitution Rule for Definite Integrals** If $u = g(x)$, then
>
> $$\int_a^b f[g(x)]g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du$$

In practice this means that we add a fourth step to the three steps for evaluating an integral by substitution:

> **Evaluating a Definite Integral by Substitution.**
>
> **1–3.** Perform the three steps from Section 7.1.1.
>
> **4.** Convert the end points of the integration from x values to u values.

Second Way. We can also find the antiderivative of $f(x) = 2x\sqrt{x^2+1}$ first and then use part II of the FTC. To find an antiderivative of $f(x)$, we choose the substitution exactly as we did when evaluating the integral the first way, i.e.: $u = x^2 + 1$

$$\frac{du}{dx} = 2x \quad \text{or} \quad du = 2x\,dx$$

Then

$$\int 2x\sqrt{x^2+1}\,dx = \int \sqrt{u}\,du = \frac{2}{3}u^{3/2} + C = \frac{2}{3}(x^2+1)^{3/2} + C.$$

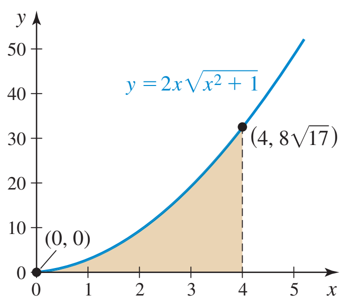


Figure 7.2 The region corresponding to $\int_0^4 2x\sqrt{x^2+1}\,dx$ before substitution has an area of $\frac{2}{3}[(17)^{3/2} - 1]$.

So $F(x) = \frac{2}{3}(x^2+1)^{3/2}$ is an antiderivative of $2x\sqrt{x^2+1}$. Using part II of the FTC, we can now compute the definite integral:

$$\int_0^4 2x\sqrt{x^2+1}\,dx = F(4) - F(0)$$

$$= \frac{2}{3}(17)^{3/2} - \frac{2}{3}(1)^{3/2} = \frac{2}{3}[17^{3/2} - 1].$$

The region corresponding to the definite integral before substitution is shown in Figure 7.2; note how different it looks from the area shown in Figure 7.1.

We will mostly use the first way when evaluating definite integrals, because, in our opinion, it is easier. But you can calculate any definite integral by substitution using either way of including the endpoints.

EXAMPLE 8 **Definite Integral** Compute

$$\int_1^2 \frac{3x^2+1}{x^3+x}\,dx$$

Solution The region corresponding to the definite integral is shown in Figure 7.3. The integrand is continuous on $[1, 2]$. We set $u = x^3 + x$ hoping to transform the integrand into a function like $\frac{1}{u}$. Then

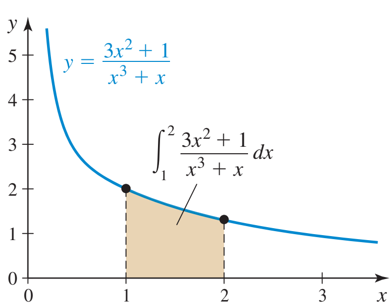


Figure 7.3 The region corresponding to the definite integral in Example 8.

$$\frac{du}{dx} = 3x^2 + 1 \quad \text{or} \quad du = (3x^2+1)\,dx$$

and the limits of integration transform to:

$$\text{if } x = 1, \quad \text{then } u = 2$$
$$\text{if } x = 2, \quad \text{then } u = 10.$$

Therefore,

$$\int_1^2 \frac{3x^2+1}{x^3+x}\,dx = \int_1^2 \underbrace{\frac{1}{x^3+x}}_{\frac{1}{u}} \cdot \underbrace{(3x^2+1)\,dx}_{du} = \int_2^{10} \frac{1}{u}\,du$$

$$= \ln|u|\Big]_2^{10} = \ln 10 - \ln 2 = \ln\frac{10}{2} = \ln 5.$$

EXAMPLE 9 **Substitution Function Is Decreasing** Compute $\int_{1/2}^1 \frac{1}{x^2}e^{1/x}\,dx$.

Solution The region corresponding to the definite integral is shown in Figure 7.4. The integrand is continuous on $[1/2, 1]$. We make the change of variable $u = \frac{1}{x}$, with the goal of transforming the integrand into a function like e^u. Then

$$\frac{du}{dx} = -\frac{1}{x^2} \text{ or } -du = \frac{1}{x^2}\,dx$$

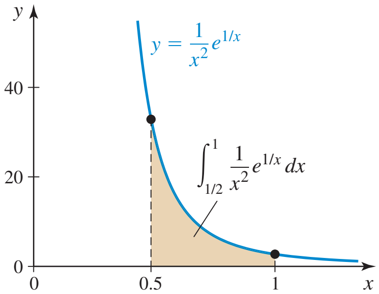


Figure 7.4 The region corresponding to the definite integral in Example 9.

and the limits of integration transform to:

$$\text{if } x = \frac{1}{2}, \quad \text{then } u = 2$$

$$\text{if } x = 1, \quad \text{then } u = 1.$$

Therefore,

$$\int_{1/2}^{1} \frac{1}{x^2}e^{1/x}\,dx = \int_{1/2}^{1} \underbrace{e^{1/x}}_{e^u} \cdot \underbrace{\frac{1}{x^2}dx}_{-du} = -\int_{2}^{1} e^u\,du = \int_{1}^{2} e^u\,du$$

$$= e^u\big]_1^2 = e^2 - e$$

Note that because $u = \frac{1}{x}$ is a decreasing function, the lower limit is greater than the upper limit of integration after the substitution. Fortunately because $\frac{du}{dx} < 0$, we also pick up a minus sign when we transform from dx to du. When we reversed the order of integration in the second step, we removed the minus sign.

However, you need not reverse the order of integration. Instead, you can compute directly:

$$-\int_{2}^{1} e^u\,du = -\,e^u\big]_2^1 = -(e^1 - e^2) = e^2 - e$$

EXAMPLE 10 **Rational Function** Compute $\displaystyle\int_4^9 \frac{2}{x-3}\,dx$.

Solution The region corresponding to the definite integral is shown in Figure 7.5. The integrand is continuous on the interval $[4, 9]$. We set $u = x - 3$ aiming to reduce our integrand to a function like $\frac{1}{u}$. To transform the integral we use $\frac{du}{dx} = 1$ or $du = dx$ and change the limits of integration:

$$\text{if } x = 4, \quad \text{then } u = 1$$

$$\text{if } x = 9, \quad \text{then } u = 6$$

Therefore,

$$\int_4^9 \underbrace{\frac{2}{x-3}}_{\frac{1}{u}} \underbrace{dx}_{du} = \int_1^6 \frac{2}{u}\,du = 2\ln|u|\big]_1^6 = 2(\ln 6 - \ln 1) = 2\ln 6$$

EXAMPLE 11 **Trigonometric Substitution** Compute

$$\int_0^{\pi/6} \cos x\, e^{\sin x}\,dx$$

Solution The region corresponding to the definite integral is shown in Figure 7.6. The integrand is continuous on the interval $[0, \pi/6]$. We make the substitution $u = \sin x$, hoping to turn our integrand into something like e^u

$$u = \sin x \quad \text{with} \quad \frac{du}{dx} = \cos x \quad \text{or} \quad du = \cos x\,dx$$

Now we change the limits of integration:

$$\text{if } x = 0, \quad \text{then } u = \sin 0 = 0$$

$$\text{if } x = \frac{\pi}{6}, \quad \text{then } u = \sin\frac{\pi}{6} = \frac{1}{2}$$

Therefore,

$$\int_0^{\pi/6} \cos x\, e^{\sin x}\,dx = \int_0^{1/2} e^u\,du = e^u\big]_0^{1/2} = e^{1/2} - 1$$

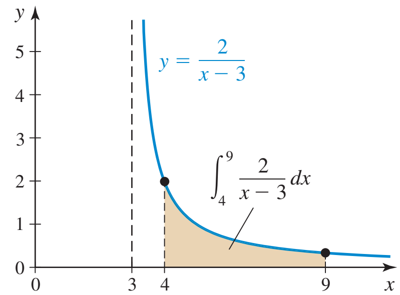


Figure 7.5 The region corresponding to the definite integral in Example 10.

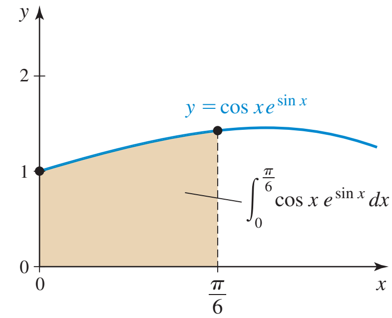


Figure 7.6 The region corresponding to the definite integral in Example 11.

EXAMPLE 12 Compute $\int_0^1 \frac{dx}{\sqrt{9-x^2}}$ using the substitution $x = 3\sin u$.

Solution Notice that in this case our change of variable is given in the form of a function $x = f(u)$, rather than $u = g(x)$. We could invert this function, but we will not do so—many of the details work out more readily when the function is in its current form. Now, it is not at all obvious how this particular substitution will simplify our integral, but we will try it nonetheless.

To transform the integral we differentiate x with respect to u.

$$\frac{dx}{du} = 3\cos u \Rightarrow dx = 3\cos u\,du.$$

And to transform the limits of the integration we must invert the function:

$$x = 3\sin u \Rightarrow u = \arcsin\left(\frac{x}{3}\right)$$

so:

$$\begin{aligned} &\text{if } x = 0, \quad \text{then} \quad u = \arcsin(0) = 0 \\ &\text{if } x = 1, \quad \text{then} \quad u = \arcsin(1/3). \end{aligned}$$

To transform the integral we substitute for x in the integrand:

$$\int_0^1 \underbrace{\frac{1}{\sqrt{9-x^2}}}_{\frac{1}{\sqrt{9-9\sin^2 u}}} \cdot \underbrace{dx}_{3\cos u\,du} = \int_0^{\arcsin(1/3)} \frac{1}{3\cos u} \cdot 3\cos u\,du \qquad \sqrt{9-9\sin^2 u} = 3\sqrt{1-\sin^2 u} = 3\cos u$$

$$= \int_0^{\arcsin(1/3)} 1\,du = u\Big]_0^{\arcsin(1/3)} = \arcsin(1/3)$$

We can easily spend a great deal of time on integration techniques. There are many substitutions like the one that was used in Example 12 that students of calculus used to need to memorize. But, now there are excellent software programs (such as *Mathematica*™ and MATLAB®) that can integrate symbolically. These programs do not render integration techniques useless; in fact, they use them. Understanding the basic techniques conceptually and being able to apply them in simple situations makes such software packages less of a "black box." Nevertheless, their availability has made it less important to acquire a large number of tricks.

So far, we have learned only one technique: substitution. Unless you can immediately recognize an antiderivative, substitution is the only method you can try at this point.

As we proceed, you will learn other techniques. You will then need to learn to recognize which technique to use. If you don't see right away what to do, just try something. Don't always expect the first attempt to succeed. With practice, you will see much more quickly whether or not your approach will succeed. If your attempt does not seem to work, try to determine the reason. That way, failed attempts can be quite useful for gaining experience in integration.

Section 7.1 Problems

7.1.1

In Problems 1–16, evaluate each indefinite integral by making the given substitution.

1. $\int 2x\sqrt{x^2+1}\,dx$, with $u = x^2+1$

2. $\int 4x^3\sqrt{x^4+1}\,dx$, with $u = x^4+1$

3. $\int 3x(1+x^2)^{1/4}\,dx$, with $u = 1+x^2$

4. $\int 4x^3(4-x^4)^{1/3}\,dx$, with $u = 4-x^4$

5. $\int 5\sin(2x)\,dx$, with $u = 2x$

6. $\int 5\sin(1-2x)\,dx$, with $u = 1-2x$

7. $\int 7x\sin(4x^2)\,dx$, with $u = 4x^2$

8. $\int x\cos(x^2-1)\,dx$, with $u = x^2-1$

9. $\int e^{2x+3}\,dx$, with $u = 2x+3$

10. $\int 3e^{1-x}\,dx$, with $u = 1-x$

11. $\int xe^{x^2/2}\,dx$, with $u = x^2/2$

12. $\int xe^{1-2x^2}\,dx$, with $u = 1-2x^2$

13. $\int \frac{x+2}{x^2+4x}\,dx$, with $u = x^2+4x$

14. $\int \frac{3x}{1-x^2}\,dx$, with $u = 1-x^2$

15. $\int \frac{3x}{x+2}\,dx$, with $u = x+2$

16. $\int \frac{x+1}{5-x}\,dx$, with $u = 5-x$

In Problems 17–36, use substitution to evaluate each indefinite integral.

17. $\int \sqrt{x+2}\,dx$

18. $\int (4+x)^{1/7}\,dx$

19. $\int (4x-1)\sqrt{2x^2-x+2}\,dx$

20. $\int (x^2-2x)(x^3-3x^2+2)^{2/3}\,dx$

21. $\int \frac{2x-2}{1+4x-2x^2}\,dx$

22. $\int \frac{x^2-1}{x^3-3x+1}\,dx$

23. $\int \frac{3x}{1+2x^2}\,dx$

24. $\int \frac{x^3-1}{x^4-4x}\,dx$

25. $\int 3x\,e^{x^2}\,dx$

26. $\int \cos x\,e^{-\sin x}\,dx$

27. $\int \frac{1}{x}((\ln x)^2+1)dx$

28. $\int \sec^2 x\,e^{\tan x}\,dx$

29. $\int \sin x\cos x\,dx$

30. $\int \cos(2x-1)\,dx$

31. $\int x\sqrt{1+x^2}dx$

32. $\int \sin^2 x\cos x\,dx$

33. $\int \frac{(\ln x)^2}{x}\,dx$

34. $\int \frac{dx}{(x+3)\ln(x+3)}$

35. $\int x^3\sqrt{1+x^2}\,dx$

36. $\int \frac{\sqrt{1+\ln x}}{x}\,dx$

In Problems 37–42, a, b, and c are constants and g(x) is a continuous function whose derivative g′(x) is also continuous. Use substitution to evaluate each indefinite integral.

37. $\int \frac{2ax+b}{ax^2+bx+c}\,dx$

38. $\int \frac{1}{ax+b}\,dx$

39. $\int g'(x)[g(x)]^n\,dx$

40. $\int g'(x)\cos[g(x)]\,dx$

41. $\int \frac{g'(x)}{[g(x)]^2+1}\,dx$

42. $\int g'(x)e^{-g(x)}\,dx$

7.1.2

In Problems 43–58, use substitution to evaluate each definite integral.

43. $\int_0^3 x\sqrt{x^2+1}\,dx$

44. $\int_0^2 x^5\sqrt{x^3+2}\,dx$

45. $\int_1^3 \frac{4x+6}{(x^2+3x)^3}\,dx$

46. $\int_0^2 \frac{2x}{(4x^2+2)^{1/3}}\,dx$

47. $\int_1^5 x\,e^{-x^2}\,dx$

48. $\int_{\ln 4}^{\ln 7} \frac{e^x}{(e^x+1)^2}\,dx$

49. $\int_0^{\pi/3} \sin x\cos x\,dx$

50. $\int_{-\pi/6}^{\pi/6} \sin^2 x\cos x\,dx$

51. $\int_0^{\pi/4} \tan x\sec^2 x\,dx$

52. $\int_0^{\pi/4} \frac{\sin x}{\cos^2 x}\,dx$

53. $\int_0^2 \frac{x}{x+2}\,dx$

54. $\int_4^9 \frac{x}{x-3}\,dx$

55. $\int_e^{e^2} \frac{dx}{x(\ln x)^2}$

56. $\int_1^2 \frac{x\,dx}{(x^2+1)\ln(x^2+1)}$

57. $\int_0^1 x^2\sqrt{x^3+1}\,dx$

58. $\int_0^2 x\sqrt{4-x^2}\,dx$

59. Use the fact that

$$\cot x = \frac{\cos x}{\sin x}$$

to evaluate

$$\int \cot x\,dx$$

7.2 Integration by Parts and Practicing Integration

7.2.1 Integration by Parts

As mentioned in the introduction to this chapter, integration by parts is the integration equivalent of the product rule in differentiation. Let $u = u(x)$ and $v = v(x)$ be two differentiable functions. Then, differentiating with respect to x yields

$$(uv)' = u'v + uv'$$

or, after rearranging,

$$uv' = (uv)' - u'v \tag{7.6}$$

Integrating both sides with respect to x, we find that

$$\int uv'\,dx = \int (uv)'\,dx - \int u'v\,dx.$$

Since uv is an antiderivative of $(uv)'$, it follows that

$$\int (uv)'\,dx = uv + C.$$

Therefore,

$$\int uv'\,dx = uv - \int u'v\,dx.$$

(Note that the constant C can be absorbed into the indefinite integral on the right-hand side.) Because $u' = du/dx$ and $v' = dv/dx$, we can write the preceding equation equivalently as:

$$\int u\,dv = uv - \int v\,du.$$

We summarize this result as follows:

Integration by Parts If $u(x)$ and $v(x)$ are differentiable functions, then

$$\int u(x)v'(x)\,dx = u(x)v(x) - \int u'(x)v(x)\,dx$$

or, in short form,

$$\int u\,dv = uv - \int v\,du$$

You are probably wondering how this technique will help, given that we traded one integral for another one. But in many cases the integral on the right-hand side can be simpler than the integral on the left-hand side. Here is a first example.

EXAMPLE 1 **Integration by Parts** Evaluate $\int x \sin x\,dx$.

Solution The integrand $x \sin x$ is a product of two functions, one of which will be designated as u, the other as v'. Since integration by parts will result in another integral of the form $\int u'v\,dx$, we must choose u and v' so that $u'v$ is of a simpler form than uv'. This suggests the following choices:

$$u = x \qquad \text{and} \qquad v' = \sin x$$

Because $v = -\cos x$ and $u' = 1$, the integral $\int u'v\,dx$ is of the form $-\int \cos x\,dx$, which is indeed simpler. We obtain

$$\int x \sin x\,dx = \underbrace{(x)}_{u}\underbrace{(-\cos x)}_{v} - \int \underbrace{(1)}_{u'}\underbrace{(-\cos x)}_{v}\,dx$$

$$= -x\cos x + \int \cos x\,dx$$

$$= -x\cos x + \sin x + C$$

If we had chosen $v' = x$ and $u = \sin x$, then we would have had $v = \frac{1}{2}x^2$ and $u' = \cos x$. The integral $\int u'v\,dx$ would have been of the form $\int \frac{1}{2}x^2 \cos x\,dx$, which is even more complicated than $\int x \sin x\,dx$.

If we use the short form $\int u\,dv = uv - \int v\,du$, we would write

$$u = x \qquad \text{and} \qquad dv = \sin x\,dx$$

Then

$$du = dx \qquad \text{and} \qquad v = -\cos x$$

and

$$\int \underbrace{x}_{u}\,\underbrace{\sin x\,dx}_{dv} = \underbrace{x}_{u}\,\underbrace{(-\cos x)}_{v} - \int \underbrace{(-\cos x)}_{v}\,\underbrace{dx}_{du}$$

$$= -x\cos x + \sin x + C$$

EXAMPLE 2 **Integration by Parts** Evaluate $\int x \ln x\,dx$.

Solution Since we do not know an antiderivative of $\ln x$, we try

$$u = \ln x \qquad \text{and} \qquad v' = x$$

Then

$$u' = \frac{1}{x} \qquad \text{and} \qquad v = \frac{1}{2}x^2$$

and

$$\int x \ln x\,dx = \underbrace{\ln x}_{u} \cdot \underbrace{\frac{1}{2}x^2}_{v} - \int \underbrace{\frac{1}{x}}_{u'} \cdot \underbrace{\frac{1}{2}x^2}_{v}\,dx$$

$$= \frac{1}{2}x^2 \ln x - \int \frac{1}{2}x\,dx = \frac{1}{2}x^2 \ln x - \frac{1}{4}x^2 + C$$

Before we present a few more useful "tricks," we show how to evaluate definite integrals with this method. Recall Equation (7.6): $uv' = (uv)' - u'v$. If we integrate both sides of this integration over the x-interval $a \le x \le b$, we obtain:

$$\int_a^b uv'\,dx = uv\Big]_a^b - \int_a^b u'v\,dx$$

This result is very similar to the result for indefinite integrals, but instead of having a function uv on the right-hand side (the original antiderivative) we must now calculate the difference in uv over the interval $[a, b]$.

EXAMPLE 3 **Definite Integral** Compute $\int_0^1 xe^{-x}\,dx$.

Solution The region representing the definite integral is shown in Figure 7.7. The integrand is continuous on $[0, 1]$. We set

$$u = x \qquad \text{and} \qquad \frac{dv}{dx} = e^{-x}$$

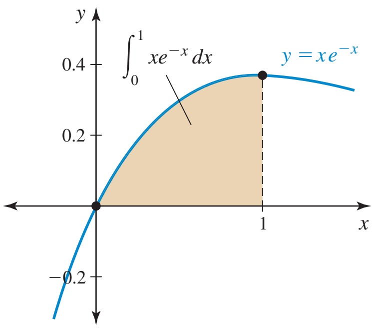


Figure 7.7 The region corresponding to the definite integral in Example 3.

Then

$$\frac{du}{dx} = 1 \qquad \text{and} \qquad v = -e^{-x}$$

Therefore,

$$\int_0^1 xe^{-x}\,dx = \underbrace{-xe^{-x}\Big]_0^1}_{uv]_a^b} - \int_0^1 \underbrace{1\cdot(-e^{-x})}_{u'v}\,dx$$

$$= -1\,e^{-1} - (-0\,e^{-0}) + \int_0^1 e^{-x}\,dx$$

$$= -e^{-1} + \left[-e^{-x}\right]_0^1 = -e^{-1} + (-e^{-1} - (-e^{-0}))$$

$$= -e^{-1} - e^{-1} + 1 = 1 - 2e^{-1}$$

In the next two examples, we demonstrate a trick that enables us to use integration by parts to evaluate integrals that do not appear to have two factors u and v. We multiply the integrand by 1 and treat the original integrand as u, and the factor 1 as v'.

EXAMPLE 4 **Multiplying by 1** Evaluate $\int \ln x\,dx$.

Solution The integrand $\ln x$ is not a product of two functions, but we can write it as $(1)(\ln x)$ and set

$$u = \ln x \qquad \text{and} \qquad v' = 1$$

Then

$$u' = \frac{1}{x} \qquad \text{and} \qquad v = x$$

We find that

$$\int \ln x\,dx = \int (\ln x)(1)\,dx = \ln x\cdot x - \int \frac{1}{x}\cdot x\cdot dx$$

$$= x\ln x - \int 1\,dx = x\ln x - x + C$$

Our choices for u and v' might surprise you, as we said that our goal was to make the integral look simpler, which often means that we try to reduce the power of functions of the form x^n. In this case, however, integrating 1 and differentiating $\ln x$ yielded a simpler integral. In fact, if we had chosen $u' = \ln x$ and $v = 1$, we would not have been able to carry out the integration by parts, since we would have needed the antiderivative of $\ln x$ to compute uv and $\int uv'\,dx$.

If you prefer the short-form notation, you don't need to multiply by 1, because

$$u = \ln x \qquad \text{and} \qquad dv = dx$$

together with

$$du = \frac{1}{x}dx \qquad \text{and} \qquad v = x$$

produces

$$\int \ln x\,dx = x\ln x - \int x\frac{1}{x}\,dx = x\ln x - \int dx$$

$$= x\ln x - x + C$$

EXAMPLE 5 **Multiplying by 1** Evaluate $\int \tan^{-1} x\,dx$.

Solution We write $\tan^{-1} x = (1)(\tan^{-1} x)$ and set

$$u = \tan^{-1} x \qquad \text{and} \qquad v' = 1$$

Then

$$u' = \frac{1}{x^2+1} \qquad \text{and} \qquad v = x$$

$x = \tan u \Rightarrow \frac{dx}{du} = \sec^2 u = \tan^2 u + 1 = x^2 + 1$

We find that

$$\int \tan^{-1} x\, dx = \tan^{-1} x \cdot x - \int \frac{x}{x^2+1}\, dx$$

We need to use substitution to evaluate the integral on the right-hand side. With

$$w = x^2 + 1 \qquad \text{and} \qquad \frac{dw}{dx} = 2x$$

we obtain

$$\int \frac{x}{x^2+1}\, dx = \frac{1}{2}\int \frac{dw}{w} = \frac{1}{2}\ln|w| + C_1 = \frac{1}{2}\ln(x^2+1) + C_1$$

where C_1 is the constant of integration. Hence,

$$\int \tan^{-1} x\, dx = x\tan^{-1} x - \frac{1}{2}\ln(x^2+1) - C_1.$$

We write the final answer as

$$\int \tan^{-1} x\, dx = x\tan^{-1} x - \frac{1}{2}\ln(x^2+1) + C$$

where C is the constant of integration with $C = -C_1$. It is not necessary to replace the constant of integration. We do it because we believe it looks cleaner. •

EXAMPLE 6 **Using Integration by Parts Repeatedly** Compute $\int_0^1 x^2 e^x\, dx$.

Solution When you integrate a definite integral by parts, it is often easier to integrate the indefinite integral first, i.e., find the antiderivative of the function, and then use part II of the FTC to evaluate the definite integral. To evaluate $\int x^2 e^x\, dx$, we set

$$u = x^2 \qquad \text{and} \qquad v' = e^x.$$

Then

$$u' = 2x \qquad \text{and} \qquad v = e^x.$$

Therefore,

$$\int x^2 e^x\, dx = x^2 e^x - \int 2x\, e^x\, dx. \tag{7.7}$$

To evaluate the integral $\int xe^x\, dx$, we must use integration by parts a second time. We set

$$u = x \qquad \text{and} \qquad v' = e^x.$$

Then

$$u' = 1 \qquad \text{and} \qquad v = e^x.$$

Therefore,

$$\int xe^x\, dx = xe^x - \int e^x\, dx = x\, e^x - e^x + C. \tag{7.8}$$

Combining (7.7) and (7.8), we find that

$$\begin{aligned}\int x^2 e^x\, dx &= x^2 e^x - 2(xe^x - e^x + C)\\ &= x^2 e^x - 2xe^x + 2e^x - 2C.\end{aligned}$$

After evaluating the indefinite integral, we can compute the definite integral. Note that the integrand is continuous on $[0, 1]$. We set $F(x) = x^2 e^x - 2xe^x + 2e^x$. Then

$$\begin{aligned}\int_0^1 x^2 e^x\, dx &= F(1) - F(0)\\ &= (e - 2e + 2e) - (0 - 0 + 2) = e - 2\end{aligned}$$

•

EXAMPLE 7 **Using Integration by Parts Repeatedly** Evaluate

$$\int e^x \cos x \, dx$$

Solution You can check that it does not matter which of the functions you call u and which v'. We set

$$u = \cos x \qquad \text{and} \qquad v' = e^x.$$

Then

$$u' = -\sin x \qquad \text{and} \qquad v = e^x.$$

Therefore,

$$\int e^x \cos x \, dx = e^x \cos x + \int e^x \sin x \, dx \tag{7.9}$$

At this point it does not look like we have simplified the integral at all by integrating by parts. But the trick here is to now integrate by parts a second time. This time, the choice of u and v' matters. We need to set

$$u = \sin x \qquad \text{and} \qquad v' = e^x.$$

Then

$$u' = \cos x \qquad \text{and} \qquad v = e^x.$$

Therefore,

$$\int e^x \sin x \, dx = e^x \sin x - \int e^x \cos x \, dx \tag{7.10}$$

Combining (7.9) and (7.10) yields

$$\int e^x \cos x \, dx = e^x \cos x + e^x \sin x - \int e^x \cos x \, dx$$

We see that the integral $\int e^x \cos x \, dx$ appears on both sides. Rearranging the equation, we obtain

$$2\int e^x \cos x \, dx = e^x \cos x + e^x \sin x + C_1 \qquad \text{Introduce a constant of integration, } C_1$$

or

$$\int e^x \cos x \, dx = \frac{1}{2} e^x (\cos x + \sin x) + C \qquad C = C_1/2$$

We said that the choices for u and v' in the second integration by parts matter. If we had designated $u = e^x$ and $v' = \sin x$, then $u' = e^x$ and $v = -\cos x$, yielding

$$\int e^x \sin x \, dx = -e^x \cos x + \int e^x \cos x \, dx$$

Combining this equation with (7.9), we then obtain

$$\int e^x \cos x \, dx = e^x \cos x - e^x \cos x + \int e^x \cos x \, dx = 0 + \int e^x \cos x \, dx$$

which is a correct, but useless, statement.

We conclude this subsection with a piece of practical advice: In integrals of the form $\int P(x) \sin(ax) \, dx$, $\int P(x) \cos(ax) \, dx$, and $\int P(x) e^{ax} \, dx$, where $P(x)$ is a polynomial and a is a constant, the polynomial $P(x)$ should be considered as u and the expressions $\sin(ax)$, $\cos(ax)$, and e^{ax} as v'. If an integral contains one of the functions $\ln x$, $\tan^{-1} x$, or $\sin^{-1} x$, then that function is usually treated as u. After practicing the problems at the end of the section, you can confirm this advice.

7.2.2 Practicing Integration

Thus far in this chapter, we have learned the two main integration techniques: substitution and integration by parts. One of the major difficulties in integration is deciding which rule to use. To evaluate some integrals you will need to make use of both techniques (e.g., Example 9, below). The best way to learn which techniques to use is to practice on as many different integrals as possible. If one approach does not work, try another.

EXAMPLE 8 Evaluate

$$\textbf{(a)} \quad \int xe^{2x}dx \qquad \textbf{(b)} \quad \int xe^{x^2}dx.$$

Solution These integrals look very similar, but they need to be evaluated using different techniques.

(a) This integral is similar to Examples 3 and 6. Even if we did not have these examples to work from, we can see that the integrand is made up of two factors, both of which may be integrated (or differentiated) in a straightforward way. So we use integration by parts to evaluate the integral, setting:

$$u = x \qquad v' = e^{2x}$$

Then

$$u' = 1 \qquad v = \frac{1}{2}e^{2x}$$

So:

$$\begin{aligned}\int xe^{2x}dx &= x \cdot \frac{1}{2}e^{2x} - \int 1 \cdot \frac{1}{2}e^{2x}\,dx \\ &= \frac{x}{2}e^{2x} - \frac{1}{4}e^{2x} + C.\end{aligned}$$

(b) This integral looks similar to (a), and it is tempting to try to evaluate it using integration by parts. However, we cannot set $v' = e^{x^2}$ in the same way as, in part (a), we set $v' = e^{2x}$, because we cannot integrate e^{x^2} to obtain v. If on the other hand we let:

$$u = e^{x^2} \qquad v' = x$$

then

$$u' = 2x \cdot e^{x^2} \qquad v = \frac{1}{2}x^2$$

and we obtain:

$$\begin{aligned}\int xe^{x^2}dx &= e^{x^2} \cdot \frac{1}{2}x^2 - \int 2xe^{x^2} \cdot \frac{1}{2}x^2dx \\ &= \frac{1}{2}x^2e^{x^2} - \int x^3e^{x^2}dx\end{aligned}$$

so our integration by parts has replaced a difficult integral by an even worse one. How then do we proceed? This integral is difficult because of the term e^{x^2}. If this were e^w integrated with respect to w, then we would have no problems evaluating the integral. (Notice that we now use w, rather than u, for our change of variable to avoid confusion with the factor $u(x)$ that appears when we integrate by parts.) So we let $w = x^2$, and to transform the integral we note that

$$\frac{dw}{dx} = 2x \quad \text{so} \quad dw = 2x\,dx.$$

And

$$\begin{aligned}\int xe^{x^2}dx &= \int \underbrace{e^{x^2}}_{e^w} \cdot \underbrace{x\,dx}_{\frac{1}{2}dw} = \frac{1}{2}\int e^w dw \\ &= \frac{1}{2}e^w + C = \frac{1}{2}e^{x^2} + C.\end{aligned}$$

•

As Example 8 shows, often the key to successful integration is persistence—if one technique doesn't succeed, then try the other, or try the technique again. Frequently, we must perform algebraic manipulations of the integrand before we can integrate.

EXAMPLE 9 Find $\displaystyle\int_0^{\sqrt{3}} \frac{1}{9+x^2}\,dx$.

Solution The region corresponding to the definite integral is shown in Figure 7.8. The integrand is continuous on $[0, \sqrt{3}]$. The integrand should remind you of the function $\frac{1}{1+u^2}$, whose antiderivative is $\tan^{-1} u$. However, we have a 9 in the denominator. To get a 1 there, we factor 9 in the denominator to obtain

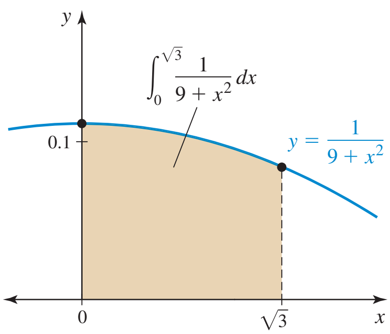


Figure 7.8 The region corresponding to the definite integral in Example 9.

$$\frac{1}{9+x^2} = \frac{1}{9(1+\frac{x^2}{9})} = \frac{1}{9(1+(\frac{x}{3})^2)}$$

The last expression now suggests that we should try the substitution

$$u = \frac{x}{3} \qquad \text{with} \qquad dx = 3\,du.$$

Since we wish to evaluate a definite integral, we must change the limits of integration as well. We find that $x = 0$ corresponds to $u = 0$ and $x = \sqrt{3}$ corresponds to $u = \frac{1}{3}\sqrt{3}$. We end up with

$$\begin{aligned}
\int_0^{\sqrt{3}} \frac{1}{9+x^2}\,dx &= \frac{1}{9}\int_0^{\sqrt{3}} \frac{1}{1+(\frac{x}{3})^2}\,dx = \frac{1}{9}\int_0^{\frac{1}{3}\sqrt{3}} \frac{3}{1+u^2}\,du \\
&= \frac{1}{3}\int_0^{\frac{1}{3}\sqrt{3}} \frac{1}{1+u^2}\,du = \frac{1}{3}\tan^{-1} u \Big]_0^{\frac{1}{3}\sqrt{3}} \\
&= \frac{1}{3}\left[\tan^{-1}\left(\frac{1}{3}\sqrt{3}\right) - \tan^{-1} 0\right] = \frac{1}{3}\left(\frac{\pi}{6} - 0\right) = \frac{\pi}{18}
\end{aligned}$$

EXAMPLE 10 Find $\int x^3 \sin(x^2+2)\,dx$.

Solution Although the integrand here has two factors (x^3 and $\sin(x^2+2)$), integration by parts will not work because, if we set $v' = \sin(x^2+2)$, then we cannot calculate v, but if we set $v' = x^3$, then $v = \frac{1}{4}x^4$, and the power of x in our integrand is increased, rather than decreased, by integration by parts. We therefore turn to substitution to simplify the integral. We notice that if, instead of integrating $\sin(x^2+2)$ with respect to x, we were to integrate $\sin w$ with respect to w, the integral would be greatly simplified. So we try for our substitution $w = x^2 + 2$. Then, to transform the integral, we use:

$$\frac{dw}{dx} = 2x \quad \text{so} \quad dw = 2x\,dx.$$

Hence under the change of variables our integral becomes:

$$\int x^3 \sin(x^2+2)\,dx = \int \underbrace{x^2}_{w-2}\, \underbrace{\sin(x^2+2)}_{\sin w} \cdot \underbrace{x\,dx}_{\frac{1}{2}dw} = \frac{1}{2}\int (w-2)\sin w\,dw$$

We are not yet at the point where we can write down the antiderivative. But the integral is now in a form where we may use integration by parts. Specifically we let:

$$\begin{aligned}
u = w - 2, &\quad \tfrac{dv}{dw} = \sin w \\
\frac{du}{dw} = 1, &\quad v = -\cos w
\end{aligned}$$

Then

$$\int (w-2)\sin w\,dw = -(w-2)\cos w + \int \cos w\,dw$$
$$= -(w-2)\cos w + \sin w + C.$$

So

$$\int x^3 \sin(x^2+2)\,dx = -\frac{1}{2}(w-2)\cos w + \frac{1}{2}\sin w + C$$
$$= -\frac{1}{2}x^2\cos(x^2+2) + \frac{1}{2}\sin(x^2+2) + C$$

Absorb the $\frac{1}{2}$ into C, since it is arbitrary

Section 7.2 Problems

7.2.1

In Problems 1–30, use integration by parts to evaluate each integral.

1. $\int x\cos x\,dx$

2. $\int 2x\sin x\,dx$

3. $\int 2x\cos 3x\,dx$

4. $\int 3x\cos(1-x)\,dx$

5. $\int 2x\sin\left(\frac{x}{2}\right)\,dx$

6. $\int x\sin(1-2x)\,dx$

7. $\int xe^x\,dx$

8. $\int 3xe^{-2x}\,dx$

9. $\int x^2e^x\,dx$

10. $\int x^2e^{-2x}\,dx$

11. $\int x\ln x\,dx$

12. $\int x^2\ln x\,dx$

13. $\int x\ln(3x)\,dx$

14. $\int x^2\ln x^2\,dx$

15. $\int x\sec^2 x\,dx$

16. $\int x\csc^2 x\,dx$

17. $\int_0^{\pi/3} x\sin x\,dx$

18. $\int_0^{\pi/4} x\cos 2x\,dx$

19. $\int_1^2 \ln x\,dx$

20. $\int_1^2 \ln(x+1)\,dx$

21. $\int_1^4 \ln\sqrt{x}\,dx$

22. $\int_1^4 \sqrt{x}\ln x\,dx$

23. $\int_0^1 xe^{-x}\,dx$

24. $\int_0^3 x^2e^{-x}\,dx$

25. $\int_0^{\pi/2} e^x\sin x\,dx$

26. $\int_0^{\pi/3} e^x\cos x\,dx$

27. $\int e^{-3x}\cos 2x\,dx$

28. $\int e^{-2x}\sin\left(\frac{x}{2}\right)\,dx$

29. $\int \sin(\ln x)\,dx$

30. $\int \cos(\ln x)\,dx$

31. Evaluating the integral $\int \cos^2 x\,dx$ requires two steps. First, write

$$\cos^2 x = (\cos x)(\cos x)$$

and integrate by parts to show that

$$\int \cos^2 x\,dx = \sin x\cos x + \int \sin^2 x\,dx$$

Then, use $\sin^2 x + \cos^2 x = 1$ to replace $\sin^2 x$ in the integral on the right-hand side, and complete the integration of $\int \cos^2 x\,dx$.

32. Evaluating the integral $\int \sin^2 x\,dx$ requires two steps. First, write

$$\sin^2 x = (\sin x)(\sin x)$$

and integrate by parts to show that

$$\int \sin^2 x\,dx = -\sin x\cos x + \int \cos^2 x\,dx$$

Then, use $\sin^2 x + \cos^2 x = 1$ to replace $\cos^2 x$ in the integral on the right-hand side, and complete the integration of $\int \sin^2 x\,dx$.

33. Evaluating the integral $\int \arcsin x\,dx$ requires two steps.

(a) Write

$$\arcsin x = 1\cdot\arcsin x$$

and integrate by parts once to show that

$$\int \arcsin x\,dx = x\arcsin x - \int \frac{x}{\sqrt{1-x^2}}\,dx$$

(b) Use substitution to compute

$$\int \frac{x}{\sqrt{1-x^2}}\,dx \tag{7.11}$$

and combine your result in (a) with (7.11) to complete the computation of $\int \arcsin x\,dx$.

34. Evaluating the integral $\int \arccos x\,dx$ requires two steps.

(a) Write

$$\arccos x = 1\cdot\arccos x$$

and integrate by parts once to show that

$$\int \arccos x\,dx = x\arccos x + \int \frac{x}{\sqrt{1-x^2}}\,dx$$

(b) Use substitution to compute

$$\int \frac{x}{\sqrt{1-x^2}}\,dx \tag{7.12}$$

and combine your result in (a) with (7.12) to complete the computation of $\int \arccos x\,dx$.

35. (a) Use integration by parts to show that, for $x > 0$,

$$\int \frac{1}{x}\ln x\,dx = (\ln x)^2 - \int \frac{1}{x}\ln x\,dx$$

(b) Use your result in (a) to evaluate

$$\int \frac{1}{x}\ln x\,dx$$

36. (a) Use integration by parts to show that

$$\int x^n e^x\,dx = x^n e^x - n\int x^{n-1}e^x\,dx$$

Such formulas are called **reduction formulas**, since they reduce the exponent of x by 1 each time they are applied.

(b) Apply the reduction formula in (a) repeatedly to compute $\int x^3 e^x\,dx$.

37. (a) Use integration by parts to verify the validity of the reduction formula

$$\int x^n e^{ax}\,dx = \frac{1}{a}x^n e^{ax} - \frac{n}{a}\int x^{n-1}e^{ax}\,dx$$

where a is a constant not equal to 0.

(b) Apply the reduction formula in (a) to compute $\int x^2 e^{-3x}\,dx$.

38. (a) Use integration by parts to verify the validity of the reduction formula

$$\int (\ln x)^n\,dx = x(\ln x)^n - n\int (\ln x)^{n-1}\,dx$$

(b) Apply the reduction formula in (a) repeatedly to compute

$$\int (\ln x)^3\,dx.$$

In Problems 39–48, first make an appropriate substitution and then use integration by parts to evaluate each integral.

39. $\int e^{\sqrt{x}}\,dx$ **40.** $\int e^{\sqrt{x+1}}\,dx$

41. $\int x^3 e^{-x^2/2}\,dx$ **42.** $\int x^5 e^{x^2}\,dx$

43. $\int x^3\sin(x^2)\,dx$ **44.** $\int \sin x\cos x\, e^{\sin x}\,dx$

45. $\int_0^{\pi/4}\cos\sqrt{x}\,dx$ **46.** $\int_0^{\pi^2}\sin\sqrt{x}\,dx$

47. $\int_1^4 \ln(\sqrt{x}+1)\,dx$ **48.** $\int_0^1 x^3\ln(x^2+1)\,dx$

7.2.2

In Problems 49–60, use either substitution or integration by parts to evaluate each integral.

49. $\int xe^{-2x}\,dx$ **50.** $\int xe^{-2x^2}\,dx$

51. $\int x(x+1)^{1/3}\,dx$ **52.** $\int \frac{x}{(1-x)^{1/4}}\,dx$

53. $\int 2x\sin(x^2)\,dx$ **54.** $\int 2x^2\sin x\,dx$

55. $\int \frac{1}{16+x^2}\,dx$ **56.** $\int \frac{x}{x^2+5}\,dx$

57. $\int \left(\frac{\tan^2 x+1}{\tan x+1}\right)dx$ **58.** $\int (\sin x+1)^2\cos x\,dx$

59. $\int x(\sin x+\cos x)\,dx$ **60.** $\int \frac{\cos 2x}{1+\sin 2x}\,dx$

61. The integral $\int \ln x\,dx$ can be evaluated in two ways.

(a) Write $\ln x = 1\cdot\ln x$ and use integration by parts to evaluate the integral.

(b) Use the substitution $u = \ln x$ and integration by parts to evaluate the integral.

62. Use an appropriate substitution followed by integration by parts to evaluate

$$\int x^3\ln(2x^2+1).$$

In Problems 63–68, evaluate each definite integral.

63. $\int_1^4 e^{\sqrt{x}}\,dx$ **64.** $\int_1^2 x\ln(x^2)\,dx$

65. $\int_{-1}^0 \frac{2}{1+x^2}\,dx$ **66.** $\int_1^2 x^2\ln x\,dx$

67. $\int_0^{\pi/2} e^x\sin x\,dx$ **68.** $\int_{-\pi/4}^{\pi/4}(1+\tan^2 x)\,dx$

Tumor Growth

The Gompertz equation is used to model the growth of a tumor. We will study it in Chapter 8. In this model the number of cells $N(t)$ in a tumor grows over time at a rate that depends on N, that is, tumors of different sizes grow at different rates, producing a differential equation:

$$\frac{dN}{dt} = aN\ln(b/N)$$

where a and b are positive constants that depend on the type of tumor, whether the tumor is being treated, and on the kind of treatment. In Chapter 8 we will see that the solution to this equation is given by evaluating the integral

$$t = \int \frac{dN}{aN\ln(b/N)} \tag{7.13}$$

69. Assume $a = b = 1$; then evaluate the integral $t = \int \frac{dN}{N\ln(1/N)}$. Your answer will contain an unknown constant of integration.

70. Evaluate the integral (7.13), keeping a and b as unknown constants

Fish Growth

von Bertalanffy's equation is used to model the growth of fish. The length of the fish, $L(t)$, grows at a rate that depends on its current length $L(t)$ (that is, big and small fish grow at different rates).

The growth of a fish is modeled using the differential equation:

$$\frac{dL}{dt} = k(L_\infty - L)$$

where k and L_∞ are both positive constants. To solve the equation, we will learn in Chapter 8 that it is necessary to calculate the following integral:

$$t = \int \frac{dL}{k(L_\infty - L)} \tag{7.14}$$

71. Assume $k = L_\infty = 1$; then evaluate the integral.

$$t = \int \frac{dL}{1-L}$$

Your answer will contain an unknown constant of integration.

72. Evaluate the integral (7.14), keeping k and L_∞ as unknown constants.

7.3 Rational Functions and Partial Fractions

A rational function f is the quotient of two polynomials. That is,

$$f(x) = \frac{P(x)}{Q(x)} \tag{7.15}$$

where $P(x)$ and $Q(x)$ are polynomials. To integrate rational functions, we use an algebraic technique, called the method of **partial fractions**, to write $f(x)$ as a sum of simpler rational functions. Such a sum is called a **partial-fraction decomposition**. These simpler rational functions, which can be integrated with the methods we have already learned, are either polynomials or of the form

$$\frac{A}{(ax+b)^n} \quad \text{or} \quad \frac{Bx+C}{(ax^2+bx+c)^n} \tag{7.16}$$

where A, B, C, a, b, and c are constants and n is a positive integer.

Why do we care about this kind of integral especially? As we will learn in Chapter 8, we often need to calculate integrals of rational functions to solve models of population growth, habitat change, and the flow of medication through the body. Although the theory of partial fractions is quite extensive, and includes many possible cases, we will focus only on the integrals that tend to arise from these mathematical models, which tend to take the form of the first type of partial fraction from (7.16). We will leave the second type of partial fraction to an optional subsection.

7.3.1 Proper Rational Functions

If the degree of $P(x)$ in (7.15) is greater than or equal to the degree of $Q(x)$, then the first step in the partial-fraction decomposition is to use long division to write $f(x)$ as a sum of a polynomial and a rational function, where the rational function is such that the degree of the polynomial in the numerator is less than the degree of the polynomial in the denominator. (Such rational functions are called **proper**.) We illustrate this step in the next two examples.

EXAMPLE 1 **Long Division Before Integration** Find

$$\int \frac{x}{x+2}\, dx.$$

Solution The degree of the numerator is equal to the degree of the denominator; using long division or writing the integrand in the form

$$\frac{x}{x+2} = \frac{x+2-2}{x+2} = 1 - \frac{2}{x+2} \qquad \text{Add and subtract 2 in numerator}$$

results in a polynomial of degree 0 and a proper rational function. We can integrate the integrand in this new form:

$$\int \frac{x}{x+2}\, dx = \int \left(1 - \frac{2}{x+2}\right) dx = x - 2\ln|x+2| + C$$

•

EXAMPLE 2 **Long Division Before Integration** Find

$$\int \left(\frac{3x^3+8x^2-x+7}{x+3}\right) dx$$

Solution Since the degree of the numerator is higher than the degree of the denominator, we use long division to simplify the integrand.

$$\begin{array}{r|l}
 & 3x^2 - x + 2 \\
\hline
x+3 & 3x^3 + 8x^2 - x + 7 \\
 & 3x^3 + 9x^2 \\
\hline
 & -x^2 - x \\
 & -x^2 - 3x \\
\hline
 & +2x + 7 \\
 & 2x + 6 \\
\hline
 & 1
\end{array}$$

That is:

$$\frac{3x^3 + 8x^2 - x + 7}{x+3} = 3x^2 - x + 2 + \frac{1}{x+3}$$

The integral that we wish to evaluate is therefore:

$$\begin{aligned}
\int \frac{3x^3 + 8x^2 - x + 7}{x+3}\, dx &= \int (3x^2 - x + 2)\, dx + \int \frac{dx}{x+3} \\
&= x^3 - \frac{1}{2}x^2 + 2x + \ln|x+3| + C
\end{aligned}$$

Assume now that the rational function is proper. Then, unless the integrand is already of one of the types in (7.16), we need to decompose it further.

7.3.2 Partial-Fraction Decomposition

Every polynomial can be **factorized**, written as a product of linear and quadratic factors.

We will first assume that $Q(x)$ can be factorized as a product of distinct factors, that is:

$$Q(x) = a(x - b_1)(x - b_2)(x - b_3)\ldots(x - b_n)$$

where $a, b_1, b_2, \ldots, b_n$ are all constants. This factorization is equivalent to a list of the roots (zeros) of the polynomial. The roots are $b_1, b_2, b_3, \ldots, b_n$, and the factors are distinct in the sense that no pair of factors is associated with the same root; so $b_1, b_2, \ldots, b_n$ are all different numbers. In this case we can expand the rational function $P(x)/Q(x)$ as a sum of partial fractions:

> If
>
> $$Q(x) = a(x - b_1)(x - b_2)\ldots(x - b_n)$$
>
> with the coefficients $b_1, b_2, \ldots, b_n$ all distinct, then any proper rational function can be written as
>
> $$\frac{P(x)}{Q(x)} = \frac{A_1}{x - b_1} + \frac{A_2}{x - b_2} + \cdots + \frac{A_n}{x - b_n} \tag{7.17}$$
>
> for some choice of constants $A_1, A_2, \ldots, A_n$.

Once you have put the rational function in the form of (7.17) the techniques that you have already learned will allow you to integrate it. For example, the integral of $\frac{A_1}{x-b_1}$ is:

$$\int \frac{A_1}{x - b_1}\, dx = A_1 \ln|x - b_1| + C$$

The work of the partial fraction expansion is to find all of the coefficients $A_1, A_2, \ldots A_n$. To do this, start by multiplying both sides of (7.17) by $Q(x)$.

$$
\begin{aligned}
P(x) \equiv A_1 & \underbrace{(x-b_2)(x-b_3)\ldots(x-b_n)}_{(x-b_1)\text{ missing}} \\
&+ A_2 \underbrace{(x-b_1)(x-b_3)\ldots(x-b_n)}_{(x-b_2)\text{ missing}} \\
&+ A_3 \underbrace{(x-b_1)(x-b_2)(x-b_4)\ldots(x-b_n)}_{(x-b_3)\text{ missing}} \\
&+ \cdots + A_n \underbrace{(x-b_1)(x-b_2)\ldots(x-b_{n-1})}_{(x-b_n)\text{ missing}}
\end{aligned}
\tag{7.18}
$$

$\equiv$ means "is identical to"

On the left-hand side we have just $P(x)$; on the right-hand side there is a sum of n terms. Each term is made up of all but one of the n factors of $Q(x)$. The expressions on both sides are identical. That is, they are different ways of writing the exact same polynomial. You should not think of this as an equation to be solved for x; if we substitute any value of x into the left- and right-hand sides of the equation, we should get the same answer on both sides. To show this we use the symbol $\equiv$; $\equiv$ means "is identical to" (as opposed to $=$ which means "is equal to").

There are two methods for calculating the values of the constants $A_1, A_2, \ldots, A_n$ on the right-hand side of the equation. Neither method is better than the other; as a matter of fact Professor Neuhauser likes the first method best, while Professor Roper prefers the second method. We will explain both methods. You should try both out and decide for yourself which one you prefer—we will use both methods in this chapter.

First Method. (Equate powers of x). Since the left-hand side and right-hand side of (7.18) are identical polynomials, each power of x must be the same; that is, the constant term must be the same, the term proportional to x, the term proportional to x^2, and so on. By comparing the coefficients for the term in $x^{n-1}, x^{n-2}, x^{n-3}, \ldots, x, 1$, we generate n separate equations, involving the n unknown coefficients $A_1, A_2, \ldots, A_n$. Solving these equations gives each of the coefficients.

Second Method. (Substitute in values of x) The two sides of equation (7.18) must evaluate to the same quantity, no matter what value of x is substituted in. We notice that each term on the right-hand side has $(n-1)$ zeros, and that each zero of $Q(x)$ is a zero of all but one of the terms on the right-hand side. For example, $x = b_1$ is not a zero of the first term, but it is a zero of the second, third, and all other terms. So if we substitute in $x = b_1$ into the right-hand side of (7.18), then all terms except the first will vanish. The equation will therefore yield the value of A_1 (since A_2, A_3, and so on, all vanish due to the substitution). To calculate A_2 we substitute $x = b_2$ into both sides of the equation (making every term except for the second vanish), while to calculate A_3 we substitute $x = b_3$ into the equation. By substituting $b_1, b_2, \ldots, b_n$ into (7.17) we can calculate $A_1, A_2, \ldots, A_n$ one after the other.

We have explained both methods quite abstractly. To get a sense of how to use them in practice, and to decide which method you prefer, it is helpful to work through some examples.

EXAMPLE 3 Find $\displaystyle\int \frac{1}{x(x-1)}\,dx$.

Solution The integrand is a proper rational function whose denominator is a product of two distinct linear functions. We claim that the integrand can be written in the form

$$\frac{1}{x(x-1)} = \frac{A}{x} + \frac{B}{x-1} \tag{7.19}$$

where A and B are constants that we need to determine. To find A and B, we write the right-hand side of (7.19) with a common denominator. That is,

$$\frac{A}{x} + \frac{B}{x-1} = \frac{A(x-1)+Bx}{x(x-1)}$$

Since this must be equal to $\frac{1}{x(x-1)}$, we conclude that

$$1 = A(x-1) + Bx. \qquad \text{Both sides are identical for all values of } x. \tag{7.20}$$

To calculate A and B we will use the first method (equating powers of x). To use this method we first collect together like powers of x on the right-hand side of (7.20):

$$1 = (A+B)x - A.$$

Therefore,

$$0 = A + B \quad x \text{ term} \qquad \text{and} \qquad 1 = -A \quad \text{Constant term}$$

This pair of equations yields $A = -1$ and $B = -A = 1$. So:

$$\frac{1}{x(x-1)} = -\frac{1}{x} + \frac{1}{x-1}.$$

The integrand can now be written as a sum of two rational functions, whose antiderivatives we learned in Chapter 6:

$$\int \frac{1}{x(x-1)}\,dx = \int \left(\frac{1}{x-1} - \frac{1}{x}\right) dx = \int \frac{1}{x-1}\,dx - \int \frac{1}{x}\,dx$$
$$= \ln|x-1| - \ln|x| + C = \ln\left|\frac{x-1}{x}\right| + C$$

•

EXAMPLE 4 Integrate $\displaystyle\int_1^2 \frac{x-1}{x^2+2x}\,dx$.

Solution Just as in Example 3, our integrand is a proper rational function. Although this is a definite integral, the first step to evaluating it is still to find the antiderivative, for which the partial fraction expansion is needed. We start by factorizing the denominator; since $x^2 + 2x = x(x+2)$, our partial fraction expansion is of the form:

$$\frac{x-1}{x^2+2x} = \frac{x-1}{x(x+2)} = \frac{A}{x} + \frac{B}{x+2} \tag{7.21}$$

for some pair of constants A and B, that we need to determine.

$$\frac{x-1}{x(x+2)} = \frac{A(x+2)+Bx}{x(x+2)} \qquad \text{Write right-hand side of (7.21) with a common denominator.}$$

so

$$x - 1 = A(x+2) + Bx. \tag{7.22}$$

This time, to find A and B we will follow method 2 by substituting for x in (7.22):

$$\begin{aligned} x = 0: \qquad & 0 - 1 = A(0+2) \qquad Bx = 0 \\ & -1 = 2A \Rightarrow A = -1/2 \\ x = -2: \qquad & -2 - 1 = -2B \Rightarrow B = 3/2 \qquad A(x+2) = 0 \end{aligned}$$

Thus:

$$\frac{x-1}{x^2+2x} = -\frac{1}{2x} + \frac{3}{2(x+2)}.$$

So:

$$\int_1^2 \frac{x-1}{x^2+2x}dx = \int_1^2 \left(-\frac{1}{2x} + \frac{3}{2(x+2)}\right) dx$$

$$= -\frac{1}{2}\ln|x| + \frac{3}{2}\ln|x+2|\Big]_1^2 \qquad \text{Using antiderivatives}$$

$$= \left(-\frac{1}{2}\ln 2 + \frac{3}{2}\ln 4\right) - \left(-\frac{1}{2}\ln 1 + \frac{3}{2}\ln 3\right)$$

$$= -\frac{1}{2}\ln 2 + 3\ln 2 - \frac{3}{2}\ln 3 \qquad \tfrac{3}{2}\ln 4 = 3\ln\sqrt{4} = 3\ln 2$$

$$= \frac{5}{2}\ln 2 - \frac{3}{2}\ln 3.$$

As the next example shows, we should be careful to check that our integrand truly is a proper rational function before attempting to find its expansion in partial fractions.

EXAMPLE 5 Evaluate $\int \frac{x^2}{x^2+3x+2}\,dx$.

Solution First we notice that the integrand is not a proper rational function because both numerator and denominator polynomials have the same degree (2). So we must use long division to reduce the rational function to its proper form:

$$\begin{array}{r r}
 & 1 \\
x^2+3x+2 \,\big) & \overline{x^2+0x+0} \\
 & \underline{x^2+3x+2} \\
 & -3x-2
\end{array} \qquad \text{Write } x^2 = x^2+0x+0$$

So

$$\frac{x^2}{x^2+3x+2} = 1 - \frac{3x+2}{x^2+3x+2} = 1 - \frac{3x+2}{(x+2)(x+1)} \qquad \text{Factorize denominator}$$

To find the partial fraction expression of the second term, we need to find constants A and B such that:

$$\frac{3x+2}{(x+2)(x+1)} = \frac{A}{x+2} + \frac{B}{x+1}$$

$$3x+2 = A(x+1) + B(x+2) \qquad \text{Put over a common denominator and compare numerators.}$$

We will again use method 2 (substituting for x) to find A and B:

$$x = -2 \quad : \quad 3(-2)+2 = -A \quad \Rightarrow \quad A = 4$$

$$x = -1 \quad : \quad 3(-1)+2 = B \quad \Rightarrow \quad B = -1$$

So:

$$\frac{3x+2}{(x+2)(x+1)} = \frac{4}{x+2} - \frac{1}{x+1}$$

So our integral can be expanded as:

$$\int \frac{x^2}{x^2+3x+2}dx = \int \left(1 - \left(\frac{4}{x+2} - \frac{1}{x+1}\right)\right) dx$$

$$= \int \left(1 - \frac{4}{x+2} + \frac{1}{x+1}\right) dx$$

$$= x - 4\ln|x+2| + \ln|x+1| + C.$$

7.3.3 Repeated Linear Factors

In our derivation of the partial fraction expansion in the previous subsection we assumed that the factors that make up the numerator polynomial $Q(x)$ were all distinct—that is, we assumed that $Q(x)$ has no repeated roots. If $Q(x)$ does have repeated roots, then you will not be able to apply Equation (7.17). For example, if $b_1 = b_2$, then the first two terms on the right-hand side of (7.17) will have the same form (constant times $\frac{1}{x-b_1}$): You will not be able to use either method from the previous subsection to calculate the constants A_1 and A_2 independently. This is a sign that the partial fraction expansion must be modified.

Partial Fraction Expansion with Repeated Linear Factors If $\frac{P(x)}{Q(x)}$ is a proper rational function and $Q(x)$ has a repeated root $x = b$ with multiplicity n, then our partial fraction expansion must include the following terms:

$$\frac{B_1}{x-b}, \quad \frac{B_2}{(x-b)^2}, \quad \ldots, \quad \frac{B_n}{(x-b)^n}$$

We must add these terms for each repeated root that $Q(x)$ has. For example, the rational function $\frac{x^3+1}{(x+2)^2x^3(x-1)}$ has partial fraction expansion:

$$\frac{x^3+1}{(x+2)^2x^3(x-1)} = \underbrace{\frac{A_1}{x+2} + \frac{A_2}{(x+2)^2}}_{\text{repeated factor } (x+2)^2} + \underbrace{\frac{B_1}{x} + \frac{B_2}{x^2} + \frac{B_3}{x^3}}_{\text{repeated factor } x^3} + \frac{C}{(x-1)}$$

To find the coefficients A_1, A_2, and so on, we will always use method one (compare like powers of x) from the previous subsection. Method two (substitution for x) can still be used, but in a slightly altered form, and you will explore that in Problems 47 and 48. However, comparing powers of x always works without any alterations, and for this reason we favor it for partial fractions in which $Q(x)$ cannot be factorized into distinct linear factors.

EXAMPLE 6 Evaluate $\displaystyle\int \frac{x}{(x+1)^2}\,dx$.

Solution The integrand is a proper rational function whose denominator is a linear factor that is repeated. We therefore write the integrand in the form

$$\frac{x}{(x+1)^2} = \frac{A}{x+1} + \frac{B}{(x+1)^2}. \qquad \text{A and B are constants to be determined}$$

Writing the right-hand side with a common denominator yields

$$\frac{x}{(x+1)^2} = \frac{A}{x+1} + \frac{B}{(x+1)^2} = \frac{A(x+1)+B}{(x+1)^2} = \frac{Ax+(A+B)}{(x+1)^2} \tag{7.23}$$

So, comparing numerators between the first and last terms in (7.23), we derive that:

$$x = Ax + (A+B)$$

To calculate A and B compare powers of x:

$$1 = A \qquad \text{x terms}$$

$$0 = A + B \qquad \text{constant terms}$$

So $A = 1$ and $B = -1$.

Therefore, $\frac{x}{(x+1)^2} = \frac{1}{x+1} - \frac{1}{(x+1)^2}$ and each term in the partial fraction expansion can be integrated:

$$\int \frac{x}{(x+1)^2}\,dx = \int \left(\frac{1}{x+1} - \frac{1}{(x+1)^2}\right) dx$$

$$= \ln|x+1| + \frac{1}{x+1} + C$$

EXAMPLE 7 Evaluate $\int \frac{dx}{x^2(x+1)}$.

Solution The integrand is a proper rational function whose denominator is a product of three linear functions: x, x (again), and $x+1$. The factor x is repeated, whereas $x+1$ only occurs once. We write the integrand in the form

$$\frac{1}{x^2(x+1)} = \frac{A}{x} + \frac{B}{x^2} + \frac{C}{x+1} \qquad \text{A, B, C are constants to be determined}$$

As in the previous example, we find A, B, and C by writing the right-hand side with a common denominator. This yields

$$\begin{aligned}\frac{1}{x^2(x+1)} = \frac{A}{x} + \frac{B}{x^2} + \frac{C}{x+1} &= \frac{Ax(x+1) + B(x+1) + Cx^2}{x^2(x+1)} \\ &= \frac{Ax^2 + Ax + Bx + B + Cx^2}{x^2(x+1)} && \text{Expand all terms} \\ &= \frac{(A+C)x^2 + (A+B)x + B}{x^2(x+1)} && \text{Collect powers of } x\end{aligned}$$

So, comparing numerators on the left- and right-hand sides, we obtain

$$1 = (A+C)x^2 + (A+B)x + B$$

Again, compare powers of x to find all of the unknown constants:

$$\begin{aligned} 0 &= A + C && x^2 \text{ terms} \\ 0 &= A + B && x \text{ terms} \\ 1 &= B && \text{constant terms}\end{aligned}$$

So $A = -1$, $B = 1$, and $C = 1$, giving a partial fraction expansion:

$$\frac{1}{x^2(x+1)} = -\frac{1}{x} + \frac{1}{x^2} + \frac{1}{x+1}$$

in which each term may be integrated:

$$\begin{aligned}\int \frac{1}{x^2(x+1)}\,dx &= \int \left(-\frac{1}{x} + \frac{1}{x^2} + \frac{1}{x+1}\right) dx \\ &= -\ln|x| - \frac{1}{x} + \ln|x+1| + C\end{aligned}$$

•

7.3.4 Irreducible Quadratic Factors

What if the denominator of our rational functions were x^2+1? This polynomial has no real roots. But we can use integration by substitution to evaluate the integral of any rational function with this denominator.

EXAMPLE 8 Evaluate: $\int \frac{x+2}{x^2+1}\,dx$.

Solution We will attempt to simplify the integral by substituting $u = x^2+1$, and $du = 2x\,dx$. We separate the x and 2 terms in the numerator, since the term in x will give us $x\,dx$, which is needed to change from an integral over x to an integral over u:

$$\begin{aligned}
\int \frac{x+2}{x^2+1} \cdot dx &= \int \frac{xdx}{x^2+1} + \int \frac{2dx}{x^2+1} \\
&= \int \underbrace{\frac{1}{x^2+1}}_{\frac{1}{u}} \cdot \underbrace{x\,dx}_{\frac{1}{2}du} + 2\int \frac{1}{x^2+1}\,dx \\
&= \int \frac{1}{2u}\,du + 2\int \frac{1}{x^2+1}\,dx \\
&= \frac{1}{2}\ln|u| + 2\tan^{-1}x + C \\
&= \frac{1}{2}\ln(x^2+1) + 2\tan^{-1}x + C \qquad \tfrac{d}{dx}\tan^{-1}x = \tfrac{1}{x^2+1}
\end{aligned}$$

•

The method of partial fractions can be used to evaluate integrals when the quotient function has an **irreducible quadratic factor** (that is, a factor that is a quadratic polynomial with no real roots) by adding the following rules to the rules from Sections 7.3.2 and 7.3.3.

> **Partial Fraction with Irreducible Quadratic Factor** If $P(x)/Q(x)$ is a proper rational function, and $Q(x)$ has the irreducible quadratic $(ax^2 + bx + c)$ as a (non-repeated) factor, then the partial fraction expansion for $\frac{P(x)}{Q(x)}$ must include a term of the form $\frac{Ax+B}{ax^2+bx+c}$ for some constants A and B.

To find the constants A and B, follow the first method given in section 7.3.2 (matching powers of x). If $Q(x)$ has multiple irreducible factors, then you will need to include one term of this form for each factor. An exception to this is if the same irreducible factor appears multiple times. We will discuss what to do in that case in the next subsection. To show that the quadratic factor $ax^2 + bx + c$ is irreducible, remember (see Section 1.2.6) that we must calculate its discriminant: $b^2 - 4ac$. If the discriminant is negative, then the quadratic is irreducible. Once we have calculated the partial fraction expansion, we may use substitution to integrate terms like $\frac{Ax+B}{ax^2+bx+c}$, just as we did in Example 8.

EXAMPLE 9 **Rational Function with Irreducible Quadratic Factor** Evaluate $\displaystyle\int \frac{x^2}{x^3-1}dx$.

Solution The integrand is a proper rational function. We need to factor the denominator. Since the denominator is a cubic polynomial, there is no straightforward way to find its factors. Instead, we guess possible roots: substituting in a few small values of x, we find that $x = 1$ is a root of the denominator. So $(x - 1)$ must be a factor. In fact:

$$x^3 - 1 = (x-1)(x^2+x+1)$$

And since $b^2 - 4ac = 1 - 4(1)(1) = -3 < 0$, the second factor is an irreducible polynomial. Hence our partial fraction expansion must take the form:

$$\frac{x^2}{x^3-1} = \frac{A}{x-1} + \frac{Bx+C}{x^2+x+1}$$

for some constants A, B, and C that we must determine.

$$\begin{aligned}
x^2 &= A(x^2+x+1) + (Bx+C)(x-1) && \text{Put right-hand side over common denominator. Compare numerators.} \\
&= (A+B)x^2 + (A-B+C)x + (A-C) && \text{Collect like powers of } x.
\end{aligned}$$

So comparing powers of x we obtain three equations for three unknowns A, B, and C:

$$0 = A - C \quad \text{constants} \tag{R$_1$}$$

$$0 = A - B + C \quad \text{terms in } x \tag{R$_2$}$$

$$1 = A + B \quad \text{terms in } x^2 \tag{R$_3$}$$

Unlike our previous examples, there is no equation that involves just one of the variables; instead we must manipulate the equations to eliminate one of the variables. We discuss in more depth how to do this in Chapter 9. If you haven't solved simultaneous equations before, we recommend that you read Section 9.1 before returning to the current subsection. We eliminate C from the equations by adding together (R_1) and (R_2):

$$0 + 0 = (A - C) + (A - B + C)$$

$$0 = 2A - B \tag{R$_4$}$$

Now, along with (R_3) we have two equations in two unknowns. Adding (R_4) to (R_3) eliminates B, giving:

$$1 = 3A \quad \text{or} \quad A = 1/3.$$

Then from (R_4) we have: $0 = 2A - B$ or $B = 2A = 2/3$.

Finally from (R_1): $0 = A - C$ or $C = A = 1/3$.

Hence:

$$\int \frac{x^2}{x^3 - 1}\,dx = \int \left(\frac{1}{3(x-1)} + \frac{2x+1}{3(x^2+x+1)} \right) dx.$$

Unlike partial fractions involving only linear factors, we cannot simply write down the integral once we have obtained the partial fraction expansion. To integrate the last term, use substitution: We want an integral like $\frac{1}{u}$, so we start by letting $u = x^2 + x + 1$. Then $du = (2x+1)\,dx$. So:

$$\int \frac{x^2}{x^3-1}\,dx = \frac{1}{3}\ln|x-1| + \frac{1}{3}\int \underbrace{\frac{1}{x^2+x+1}}_{\frac{1}{u}} \cdot \underbrace{(2x+1)\,dx}_{du}$$

$$= \frac{1}{3}\ln|x-1| + \frac{1}{3}\ln|u| + C$$

$$= \frac{1}{3}\ln|x-1| + \frac{1}{3}\ln|x^2+x+1| + C \quad \text{Separate the two integrals}$$

•

There is one more tool that we must add to our toolset to evaluate integrals with irreducible quadratic factors. To get this result we need to differentiate $f(x) = \frac{1}{a}\tan^{-1}\left(\frac{x-b}{a}\right)$ for some pair of constants a and b (a must be non-zero). To calculate $f'(x)$, use the chain rule with $u = \frac{x-b}{a}$, $f(u) = \frac{1}{a}\tan^{-1} u$. Then:

$$\frac{df}{dx} = \frac{df}{du} \cdot \frac{du}{dx} \quad \text{Chain rule}$$

$$= \underbrace{\frac{1}{a} \cdot \frac{1}{u^2+1^2}}_{\frac{df}{du}} \cdot \underbrace{\frac{1}{a}}_{\frac{du}{dx}} = \frac{1}{a^2} \cdot \frac{1}{\left(\frac{x-b}{a}\right)^2 + 1}$$

$$= \frac{1}{(x-b)^2 + a^2}$$

So by the fundamental theorem of calculus:

Generalized Arctan Integral

$$\int \frac{dx}{(x-b)^2 + a^2} = \frac{1}{a}\tan^{-1}\left(\frac{x-b}{a}\right) + C$$

EXAMPLE 10 Evaluate $\displaystyle\int \frac{x^2+3x+4}{(x+1)(x^2+2x+5)}$.

Just as for linear factors, we must include one term in our partial fraction expansion for each of the irreducible quadratic factors of the quotient function.

Solution The integrand is a proper partial fraction since the degree of the numerator is 2, and the degree of the denominator is 3. The factor x^2+2x+5 is irreducible because its discriminant is $b^2-4ac = (2)^2-4(1)(5) = -16 < 0$. Hence our partial fraction expansion will be of the form:

$$\frac{x^2+3x+4}{(x+1)(x^2+2x+5)} = \frac{A}{x+1} + \frac{Bx+C}{x^2+2x+5}$$

where A, B, and C are constants that we will need to determine.

$$\frac{x^2+3x+4}{(x+1)(x^2+2x+5)} = \frac{A(x^2+2x+5)+(Bx+C)(x+1)}{(x+1)(x^2+2x+5)}$$ Put right-hand side over common denominator

So, comparing numerators on both sides, we find

$$\begin{aligned} x^2+3x+4 &= A(x^2+2x+5)+(Bx^2+Bx+Cx+C) \\ &= (A+B)x^2+(2A+B+C)x+(5A+C) \end{aligned}$$ Collecting like powers of x

So, matching coefficients for each power of x:

Constants: $4 = 5A+C$ $\quad (R_1)$

x: $3 = 2A+B+C$ $\quad (R_2)$

x^2: $1 = A+B$ $\quad (R_3)$

We have three equations in three unknowns. To solve for A, B, and C we start by reducing to two equations in two unknowns; that is, we eliminate one of the constants. We see that, if we subtract (R_2) from (R_1), C will be eliminated, giving:

$$1 = 3A-B \qquad (R_4)$$

Together with (R_3) we now have two equations in two unknowns (A and B). Adding (R_3) to (R_4) eliminates B, giving:

$$1+1 = (3A-B)+(A+B) \Rightarrow 4A = 2$$

or $A = 1/2$. Then substituting into (R_4), we see $1/2+B = 1$ so $B = 1/2$. Finally substituting into (R_1) we obtain $5/2+C = 4$ or $C = 3/2$.

So

$$\begin{aligned} \int \frac{x^2+3x+4}{(x+1)(x^2+2x+5)}\,dx &= \frac{1}{2}\int\left(\frac{1}{x+1}+\frac{x+3}{x^2+2x+5}\right)dx \\ &= \frac{1}{2}\ln|x+1| + \frac{1}{2}\int \frac{x+3}{x^2+2x+5}\,dx. \end{aligned}$$

To integrate the second term, we start, as we did in Example 8, by making a change of variables: i.e., we set $u = x^2+2x+5$, $du = (2x+2)\,dx = 2(x+1)\,dx$. We use this substitution to remove the x from the denominator:

$$\begin{aligned} \int \frac{x+3}{x^2+2x+5}\,dx = \int \frac{(x+1)+2}{x^2+2x+5}\,dx &= \frac{1}{2}\int\frac{du}{u} + \int \frac{2}{x^2+2x+5}\,dx \\ &= \frac{1}{2}\ln|u| + \int \frac{2}{x^2+2x+5}\,dx. \end{aligned}$$ $u = x^2+2x+5$

To integrate the remaining term, make use of the generalized arctan integral formula. To use this formula we need to complete the square in the denominator to put it in the form $(x-b)^2+a^2$:

$$x^2+2x+5 = (x+1)^2+4 = (x+1)^2+2^2$$

So:

$$\int \frac{2}{x^2+2x+5}\,dx = 2\left(\frac{1}{2}\tan^{-1}\left(\frac{x+1}{2}\right)+C\right) \qquad a=2, b=-1$$

And then, assembling all of the pieces

$$\begin{aligned}\int \frac{x+3}{x^2+2x+5}\,dx &= \frac{1}{2}\ln|u| + \int \frac{2}{x^2+2x+5}\,dx \\ &= \frac{1}{2}\ln|x^2+2x+5| + \tan^{-1}\left(\frac{x+1}{2}\right)+2C\end{aligned}$$

So:

$$\begin{aligned}\int \frac{x^2+3x+4}{(x+1)(x^2+2x+5)} &= \frac{1}{2}\ln|x+1| + \frac{1}{2}\int \frac{x+3}{x^2+2x+5}\,dx \\ &= \frac{1}{2}\ln|x+1| + \frac{1}{4}\ln|x^2+2x+5| \\ &\quad + \frac{1}{2}\tan^{-1}\left(\frac{x+1}{2}\right)+C\end{aligned}$$

At this stage, if the constant of integration appeared with a coefficient in front of it (e.g., we had $2C$ instead of C), we could redefine the constant of integration to absorb the coefficient. •

This result can be used to integrate any function of the form $\frac{1}{g(x)}$ where the denominator $g(x)$ is an irreducible polynomial, by first completing the square in the denominator.

EXAMPLE 11 **Distinct Irreducible Quadratic Factors** Evaluate $\displaystyle\int \frac{2x^3-x^2+2x-2}{(x^2+2)(x^2+1)}\,dx$.

Solution The rational function in the integrand is proper. The denominator is already factored, each factor is an irreducible quadratic polynomial, and the two factors are distinct. We can therefore write the integrand as

$$\begin{aligned}\frac{2x^3-x^2+2x-2}{(x^2+2)(x^2+1)} &= \frac{Ax+B}{x^2+2} + \frac{Cx+D}{x^2+1} \\ &= \frac{(Ax+B)(x^2+1)+(Cx+D)(x^2+2)}{(x^2+2)(x^2+1)} \\ &= \frac{Ax^3+Ax+Bx^2+B+Cx^3+2Cx+Dx^2+2D}{(x^2+2)(x^2+1)} \\ &= \frac{(A+C)x^3+(B+D)x^2+(A+2C)x+B+2D}{(x^2+2)(x^2+1)} \qquad \text{Collect like powers of } x\end{aligned}$$

Comparing the denominators on the left- and right-hand sides, we obtain four equations in four unknowns (A, B, C, and D):

$$\begin{aligned}-2 &= B+2D && \text{Constants} && (R_1)\\ 2 &= A+2C && x && (R_2)\\ -1 &= B+D && x^2 && (R_3)\\ 2 &= A+C && x^3 && (R_4)\end{aligned}$$

This is a lot of equations (and unknowns) but we must have courage and press ahead by eliminating unknowns and reducing the number of equations. We notice in this case that Equations (R_1) and (R_3) contain B and D only, while (R_2) and (R_4) contain A and C only. So we first use (R_1) with (R_3) to solve for B and D. Subtracting (R_3) from (R_1) eliminates B:

$$\begin{aligned}-2-(-1) &= B+2D-(B+D) = D\\ -1 &= D\end{aligned}$$

then substituting for D in (R_3) we obtain:

$$-1 = B + (-1) \Rightarrow B = 0$$

Similarly we use (R_2) and (R_4) to solve for A and B. Subtracting (R_4) from (R_2) we obtain:

$$2 - 2 = A + 2C - (A + C)$$
$$0 = C$$

and then by substituting for C in (R_2) we obtain $A = 2$.

So

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

and the integral can be rewritten as:

$$\int \frac{2x^3 - x^2 + 2x - 2}{(x^2+2)(x^2+1)}\,dx = \int \frac{2x}{x^2+2}\,dx - \int \frac{1}{x^2+1}\,dx \qquad \text{Let } u = x^2 + 2$$
$$= \int \frac{du}{u} - \tan^{-1} x + C = \ln|u| - \tan^{-1} x + C$$
$$= \ln|x^2+2| - \tan^{-1} x + C.$$

7.3.5 Summary

We conclude this section by summarizing the two most important cases in which you will encounter partial fraction expansions: when the integrand is a rational function for which the quotient is a polynomial of degree 2 and is either (1) a product of two not necessarily distinct linear factors or (2) an irreducible quadratic polynomial.

The first step is to make sure that the degree of the numerator is less than the degree of the denominator. If not, then we use long division to simplify the integrand.

We will now assume that the degree of the numerator is strictly less than the degree of the denominator (i.e., the integrand is a proper rational function). We write the rational function $f(x)$ as

$$f(x) = \frac{P(x)}{Q(x)}$$

with $Q(x) = ax^2 + bx + c$, $a \neq 0$, and $P(x) = rx + s$. Either $Q(x)$ can be factored into two linear factors, or it is irreducible (i.e., does not have real roots).

Case 1a: $Q(x)$ is a product of two distinct linear factors. In this case, we write

$$Q(x) = a(x - x_1)(x - x_2)$$

where x_1 and x_2 are the two distinct roots of $Q(x)$. We then use the method of partial fractions to simplify the rational function:

$$\frac{P(x)}{Q(x)} = \frac{rx + s}{ax^2 + bx + c} = \frac{A}{x - x_1} + \frac{B}{x - x_2}$$

The constants A and B must now be determined as in Examples 3 or 4.

Case 1b: $Q(x)$ is a product of two identical linear factors. In this case, we write

$$Q(x) = a(x - x_1)^2$$

where x_1 is the root of $Q(x)$. We then use the method of partial fractions to simplify the rational function:

$$\frac{P(x)}{Q(x)} = \frac{rx + s}{ax^2 + bx + c} = \frac{A}{x - x_1} + \frac{B}{(x - x_1)^2}$$

The constants A and B must now be determined as in Example 6.

Case 2: (Optional) $Q(x)$ is an irreducible quadratic polynomial. In this case,

$$Q(x) = ax^2 + bx + c \quad \text{with } b^2 - 4ac < 0$$

and we must complete the square as in Example 10. Doing so then leads to integrals of the form

$$\int \frac{dx}{(x - x_3)^2 + x_4^2} \quad \text{or} \quad \int \frac{x}{(x - x_3)^2 + x_4^2}\, dx$$

For some constants x_3 and x_4

The first integral is given by the generalized arctan integral, whereas the second integral can be evaluated by substitution. (See Examples 9 and 10.)

Section 7.3 Problems

7.3.1

In Problems 1–16, use long division to write $f(x)$ as a sum of a polynomial and a proper rational function.

1. $f(x) = \dfrac{x+2}{x+1}$

2. $f(x) = \dfrac{x^2+1}{x+1}$

3. $f(x) = \dfrac{2x^2+5x-1}{x+2}$

4. $f(x) = -\dfrac{x^2-4x-1}{x-1}$

5. $f(x) = \dfrac{3x^3+5x-2x^2-2}{x^2+1}$

6. $f(x) = \dfrac{x^3-3x^2-15}{x^2+x+3}$

7. $f(x) = \dfrac{x^2+x+1}{x^2+2x+1}$

8. $f(x) = \dfrac{x^3+2x+3}{x+1}$

9. $f(x) = \dfrac{x^3+3x^2+3x+1}{x^2+1}$

10. $f(x) = \dfrac{x^3+1}{x^2+x+1}$

11. $f(x) = \dfrac{x^3}{x^2+x}$

12. $f(x) = \dfrac{x^3+x}{x^2+x}$

13. $f(x) = \dfrac{x^4+1}{x-1}$

14. $f(x) = \dfrac{x^5-1}{x-1}$

15. $f(x) = \dfrac{x+1}{2x+3}$

16. $f(x) = \dfrac{x^2+1}{3x+1}$

7.3.2

In Problems 17–22, write out the partial-fraction decomposition of the function $f(x)$.

17. $f(x) = \dfrac{2x-3}{x(x+1)}$

18. $f(x) = -\dfrac{x+1}{(2x+1)(x-1)}$

19. $f(x) = \dfrac{x+1}{x(x+2)}$

20. $f(x) = \dfrac{2x+1}{(x+1)(x-1)}$

21. $f(x) = \dfrac{16x-6}{(2x-5)(3x+1)}$

22. $f(x) = \dfrac{4x^2-14x-6}{x(x-3)(x+1)}$

In Problems 23–30, write out the partial-fraction decomposition of the function $f(x)$.

23. $f(x) = \dfrac{5x-1}{x^2-1}$

24. $f(x) = \dfrac{9x-7}{2x^2-7x+3}$

25. $f(x) = -\dfrac{10}{3x^2+8x-3}$

26. $f(x) = \dfrac{4x+1}{x^2-3x-10}$

27. $f(x) = \dfrac{x+1}{x^2-2x}$

28. $f(x) = \dfrac{x+1}{x^2-3x+2}$

29. $f(x) = \dfrac{1}{x^3-4x^2+3x}$

30. $f(x) = \dfrac{1}{x^3-2x^2-x+2}$

In Problems 31–36, use partial-fraction decomposition to evaluate the integrals.

31. $\displaystyle\int \frac{1}{x(x-2)}\, dx$

32. $\displaystyle\int \frac{1}{x(2x+1)}\, dx$

33. $\displaystyle\int \frac{1}{(x+1)(x-3)}\, dx$

34. $\displaystyle\int \frac{1}{(x-1)(x+2)}\, dx$

35. $\displaystyle\int \frac{x^2-2x-2}{x(x+2)}\, dx$

36. $\displaystyle\int \frac{4x^2-x-1}{(x+1)(x-3)}\, dx$

7.3.3

In Problems 37–46 use partial fraction decompositions to evaluate each integral.

37. $\displaystyle\int \frac{2x-3}{(x-1)^2}\, dx$

38. $\displaystyle\int \frac{x-1}{(x+1)^2}\, dx$

39. $\displaystyle\int \frac{3x^2-x+1}{x(x-1)^2}\, dx$

40. $\displaystyle\int \frac{4x^2+3x+1}{(x+1)^2(x-1)}\, dx$

41. $\displaystyle\int \frac{x^2-2x-2}{x^2(x+2)} \cdot dx$

42. $\displaystyle\int \frac{4x^2-x-1}{(x+1)^2(x-3)}\, dx$

43. $\displaystyle\int \frac{x-2}{(x-1)^2}\, dx$

44. $\displaystyle\int \frac{2x^2-2x+1}{x^2(x-1)}\, dx$

45. $\displaystyle\int \frac{2x^2+x+1}{x(x+1)^2}\, dx$

46. $\displaystyle\int \frac{2x^3-x-1}{x^2(x+1)^2}\, dx$

In Problems 47–48 we will discuss alternatives to comparing powers of x for finding the coefficients in a partial fraction expansion when the denominator polynomial has a repeated root.

47. Substituting Values of x

(a) Consider the rational function

$$f(x) = \frac{x+3}{(x-1)(x+1)^2}$$

whose partial fraction expression is of the form

$$f(x) = \frac{A}{x-1} + \frac{B}{x+1} + \frac{C}{(x+1)^2}$$

for some set of constants A, B, C that need to be determined. To calculate these constants put all of the terms over a common denominator.

$$\frac{x+3}{(x+1)^2(x-1)} = \frac{A(x+1)^2 + B(x-1)(x+1) + C(x-1)}{(x-1)(x+1)^2}$$

Explain why

$$x+3=A(x+1)^2+B(x-1)(x+1)+C(x-1) \qquad (7.24)$$

(b) One method to calculate A, B, C from (7.24) is to substitute in specific values of x. We showed in this section that some good choices are values that make one or more terms vanish. Show by substituting in $x=1$ and $x=-1$ that $A=1$ and $C=-1$.

(c) Explain why there is no value of x that will make both the $A(x+1)^2$ and $C(x-1)$ terms vanish, without causing the $B(x-1)(x+1)$ term to vanish.

(d) Although we cannot isolate the term in B, we can obtain more equations by substituting different values of x. By letting $x=0$, show that:

$$3=A-B-C$$

and use this equation to calculate B.

(e) Use your answers from (a)–(d) to evaluate

$$\int \frac{x+3}{(x+1)^2(x-1)}\,dx$$

(f) Use the method given in parts (b) and (d) to find the partial fraction expansion of $f(x)=\frac{3x^2-2x-4}{2x^2(1+x)}$

48. Differentiating Both Sides of the Identity

(a) Consider the rational function:

$$f(x)=\frac{x^2-3x-7}{(x-1)(x+2)^2}$$

whose partial fraction expansion must take the form:

$$f(x)=\frac{A}{x-1}+\frac{B}{x+2}+\frac{C}{(x+2)^2}$$

for some constants A, B, and C, all of which we need to determine. To calculate these constants put all of the terms over a common denominator, and explain why

$$x^2-3x-7=A(x+2)^2+B(x-1)(x+2)+C(x-1) \qquad (7.25)$$

(b) Previously we obtained the values of A, B, and C from equations like (7.25) by substituting in specific values of x. Good choices for these values are substitutions that make one or more terms vanish in the right-hand side of (7.25). Show by substituting in two different values of x that $A=-1$ and $C=-1$.

(c) Explain why there is no value of x that will make both the $A(x+2)^2$ and $C(x-1)$ terms vanish, but will not also cause the $B(x-1)(x+2)$ term to vanish.

(d) Since (7.25) is an identity, that is, the two polynomials on the left- and right-hand sides are identical, they must also have the same derivatives. Explain why, as a result:

$$2x-3=2A(x+2)+B(2x+1)+C$$

(e) By choosing $x=-2$ show that $-7=-3B+C$ and calculate B. (Why was this value of x chosen?)

(f) Use your answers to parts (a)–(e) to evaluate

$$\int \frac{x^2-3x-7}{(x-1)(x+2)^2}\,dx.$$

(g) Use the method from parts (b), (d), and (e) to evaluate

$$\int \frac{3x^2+x-3}{x^2(x-1)}\,dx.$$

7.3.4

Find the partial fraction expansion for each of the following functions.

49. $f(x)=\dfrac{x^3-x^2+x-4}{(x^2+1)(x^2+4)}$

50. $f(x)=\dfrac{x^3-3x^2+x-6}{(x^2+2)(x^2+1)}$

51. $f(x)=\dfrac{x^2+2x-1}{(x-1)(x^2+1)}$

52. $f(x)=\dfrac{x^2-x+6}{(x-2)(x^2+4)}$

53. $f(x)=\dfrac{x^3+2x^2+x+1}{x^2(x^2+1)}$

54. $f(x)=\dfrac{x^3-3x}{(x-1)^2(x^2+1)}$

In Problems 55–58, complete the square in the denominator and evaluate the integral.

55. $\displaystyle\int \frac{1}{x^2-2x+2}\,dx$

56. $\displaystyle\int \frac{1}{x^2+4x+5}\,dx$

57. $\displaystyle\int \frac{1}{x^2-4x+13}\,dx$

58. $\displaystyle\int \frac{1}{x^2+2x+5}\,dx$

For Problems 59–68 evaluate each integral.

59. $\displaystyle\int \frac{1}{x^2+9}dx$

60. $\displaystyle\int \frac{1}{x^2+2x+10}\,dx$

61. $\displaystyle\int \frac{x}{x^2+4x+5}\,dx$

62. $\displaystyle\int \frac{x+1}{x^2+1}\,dx$

63. $\displaystyle\int \frac{x+1}{x(x^2+1)}\,dx$

64. $\displaystyle\int \frac{1}{(x+1)(x^2+4)}\,dx$

65. $\displaystyle\int \frac{2x^2+x+5}{x(x^2+2x+5)}\,dx$

66. $\displaystyle\int \frac{2x^2+x}{(x-1)(x^2+x+1)}\,dx$

67. $\displaystyle\int \frac{1}{x^2(x^2+1)}\,dx$

68. $\displaystyle\int \frac{x^2+2x}{(x+1)(x^2+2x+2)}\,dx$

7.3.5

In Problems 69–78, evaluate each integral.

69. $\displaystyle\int \frac{1}{(x-3)(x+2)}\,dx$

70. $\displaystyle\int \frac{2x-1}{(x+4)(x+1)}\,dx$

71. $\displaystyle\int \frac{1}{x^2-9}\,dx$

72. $\displaystyle\int \frac{1}{x^2+16}\,dx$

73. $\displaystyle\int \frac{1}{x^2-x-2}\,dx$

74. $\displaystyle\int \frac{1}{x^2-x+2}\,dx$

75. $\displaystyle\int \frac{x^2+1}{x^2+3x+2}\,dx$

76. $\displaystyle\int \frac{x^3+1}{x^2+3}\,dx$

77. $\displaystyle\int \frac{x^2+4}{x^2-4}\,dx$

78. $\displaystyle\int \frac{x^4+3}{x^2-4x+3}\,dx$

In Problems 79–84, evaluate each definite integral.

79. $\displaystyle\int_3^5 \frac{x-1}{(x+1)(x+2)}\,dx$

80. $\displaystyle\int_3^5 \frac{x}{x-1}\,dx$

81. $\displaystyle\int_0^1 \frac{x}{x^2+1}\,dx$

82. $\displaystyle\int_1^2 \frac{x+1}{x^2+1}\,dx$

83. $\displaystyle\int_2^3 \frac{1}{1-x}\,dx$

84. $\displaystyle\int_2^3 \frac{1}{1-x^2}\,dx$

In Problems 85 and 86, evaluate the definite integral. Hint: First integrate by parts to turn the integrand into a rational function.

85. $\displaystyle\int_0^1 \tan^{-1}x\,dx$

86. $\displaystyle\int_0^1 x\tan^{-1}x\,dx$

7.4 Improper Integrals

In this section, we discuss definite integrals of two types with the following characteristics:

1. One or both limits of integration are infinite; that is, the integration interval is unbounded; or
2. The integrand becomes infinite at one or more points of the interval of integration.

We call such integrals **improper integrals**.

Improper integrals of the first type arise quite commonly in the study of probability distributions. If you plan to study Chapter 12, where probability distributions are discussed, you may wish to work through Section 7.4.1 now. Improper integrals of the second type arise very commonly in physics, in the analysis of how charges or masses interact with each other. Because these applications are not directly relevant in life sciences, we regard the material in Sections 7.4.2 and 7.4.3 to be optional.

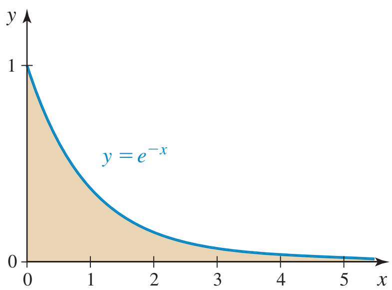


Figure 7.9 The unbounded region between the graph of $y = e^{-x}$ and the x-axis for $x \geq 0$.

7.4.1 Type 1: Unbounded Intervals

Suppose that we wanted to compute the area of the unbounded region below the graph of $f(x) = e^{-x}$ and above the x-axis for $x \geq 0$. (See Figure 7.9.) How would we proceed? We know how to find the area of a region bounded by the graph of a continuous function [here, $f(x) = e^{-x}$] and the x-axis between 0 and z for any value of z, namely,

$$A(z) = \int_0^z e^{-x} dx = -e^{-x}\Big]_0^z = 1 - e^{-z}$$

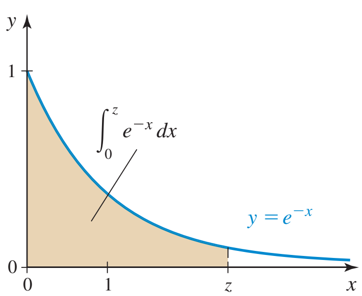


Figure 7.10 The area in Figure 7.9 is approximated by the region $0 \leq x \leq z$.

This is the shaded area in Figure 7.10. If we now let z tend to infinity, we may regard the limiting value (if it exists) as the area of the unbounded region below the graph of $f(x) = e^{-x}$ and above the x-axis for $x \geq 0$ (see Figure 7.9):

$$A = \lim_{z \to \infty} A(z) = \lim_{z \to \infty} (1 - e^{-z}) = 1$$

We write

$$\int_0^\infty e^{-x}\, dx = 1$$

Therefore, for functions that are continuous on unbounded intervals (see Figures 7.11 and 7.12), we define

$$\boxed{\int_a^\infty f(x)\, dx = \lim_{z \to \infty} \int_a^z f(x)\, dx}$$

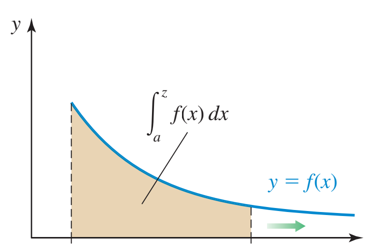


Figure 7.11 The definition of the improper integral $\int_a^\infty f(x)\, dx$ as the limit of $\int_a^z f(x)\, dx$ as $z \to \infty$.

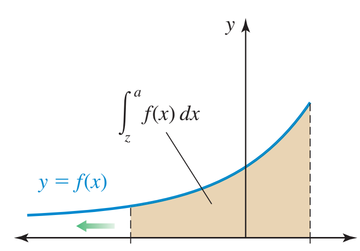


Figure 7.12 The definition of the improper integral $\int_{-\infty}^a f(x)\, dx$ as the limit of $\int_z^a f(x)\, dx$ as $z \to -\infty$.

and

$$\boxed{\int_{-\infty}^{a} f(x)\,dx = \lim_{z\to-\infty}\int_{z}^{a} f(x)\,dx}$$

You might be surprised that the area of an unbounded region can be finite. This need not be the case, and it happens only if the graph of $f(x)$ approaches the x-axis sufficiently fast. We illustrate this property in the next two examples.

EXAMPLE 1 **Finite Area** Compute $\displaystyle\int_1^\infty \frac{1}{x^2}\,dx$.

Solution The function $y = 1/x^2$ is continuous on $[1, \infty)$. We first compute

$$A(z) = \int_1^z \frac{1}{x^2}\,dx = -\frac{1}{x}\bigg]_1^z = 1 - \frac{1}{z} \qquad \text{See Figure 7.13}$$

and then let $z \to \infty$. We find that

$$\lim_{z\to\infty} A(z) = \lim_{z\to\infty}\left(1 - \frac{1}{z}\right) = 1$$

Hence,

$$\int_1^\infty \frac{1}{x^2}\,dx = 1.$$

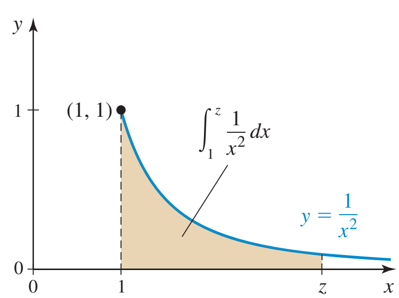


Figure 7.13 The region corresponding to $A(z)$ in Example 1.

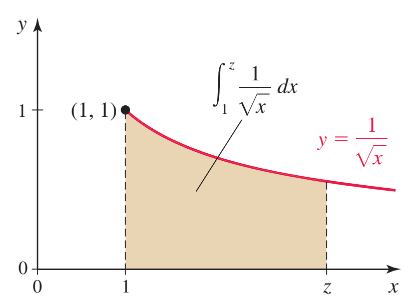


Figure 7.14 The region corresponding to $A(z)$ in Example 2.

EXAMPLE 2 **Infinite Area** Compute $\displaystyle\int_1^\infty \frac{1}{\sqrt{x}}\,dx$.

Solution The function $f(x) = 1/\sqrt{x}$ is continuous on $[1, \infty)$. We first compute

$$A(z) = \int_1^z \frac{1}{\sqrt{x}}\,dx = 2\sqrt{x}\,\Big]_1^z = 2(\sqrt{z} - 1) \qquad \text{See Figure 7.14}$$

and then let $z \to \infty$. We find that

$$\lim_{z\to\infty} A(z) = \lim_{z\to\infty} 2(\sqrt{z} - 1) = \infty$$

Hence, $\int_1^\infty \frac{1}{\sqrt{x}}\,dx$ does not exist.

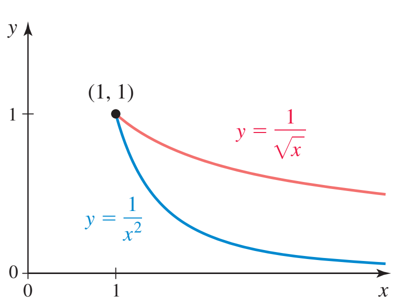


Figure 7.15 The function $y = \frac{1}{x^2}$ approaches the x-axis much faster than the function $y = \frac{1}{\sqrt{x}}$.

Looking back at Examples 1 and 2, we see that, in both cases, the respective integrands approached the x-axis as $x \to \infty$; that is, both

$$\lim_{x\to\infty}\frac{1}{x^2} = 0 \qquad \text{and} \qquad \lim_{x\to\infty}\frac{1}{\sqrt{x}} = 0$$

However, $\frac{1}{x^2}$ approaches the x-axis much faster than $\frac{1}{\sqrt{x}}$, as can be seen from the graphs in Figure 7.15. The exponent of x in the denominator determines how fast the function approaches the x-axis. The area between the graph and the x-axis from $x = 1$ to infinity is finite only if the graph approaches the x-axis fast enough. Indeed, if we tried to compute

$$\int_1^\infty \frac{1}{x^p}\,dx$$

for $0 < p < \infty$, we would find that

$$\int_1^\infty \frac{1}{x^p}\,dx = \begin{cases} \frac{1}{p-1} & \text{for } p > 1 \\ \infty & \text{for } 0 < p \le 1 \end{cases} \qquad y = \tfrac{1}{x^p} \text{ is continuous on } [1, \infty).$$

For $p > 1$, the function $\frac{1}{x^p}$ approaches the x-axis fast enough as $x \to \infty$ for the area under the graph to be finite. (We investigate this integral further in Problem 35.)

We will use the following terminology to indicate whether an improper integral is finite or infinite:

Definition Let $f(x)$ be continuous on the interval $[a, \infty)$. If $\lim_{z\to\infty} \int_a^z f(x)\,dx$ exists and has a finite value, we say that the improper integral $\int_a^\infty f(x)\,dx$ **converges** and define:

$$\int_a^\infty f(x)\,dx = \lim_{z\to\infty} \int_a^z f(x)\,dx.$$

Otherwise, we say that the improper integral **diverges**.

Analogous definitions can be given when the lower limit of integration is infinite.

EXAMPLE 3 **Infinite Lower Limit** Show that the improper integral $\displaystyle\int_{-\infty}^0 \frac{1}{(x-1)^2}\,dx$ converges.

Solution To show that the integral converges, we compute

$$A(z) = \int_z^0 \frac{1}{(x-1)^2}\,dx \quad \text{for } z < 0 \qquad y = \tfrac{1}{(x-1)^2} \text{ is continuous on } (-\infty, 0]$$

and then let $z \to -\infty$. (See Figure 7.16.) We find that

$$\int_z^0 (x-1)^{-2}\,dx = -(x-1)^{-1}\Big]_z^0$$

$$= -\frac{1}{x-1}\Big]_z^0 = -\frac{1}{-1} + \frac{1}{z-1} = 1 + \frac{1}{z-1}$$

and

$$\lim_{z\to-\infty}\left(1 + \frac{1}{z-1}\right) = 1.$$

Therefore,

$$\int_{-\infty}^0 \frac{1}{(x-1)^2}\,dx = 1.$$

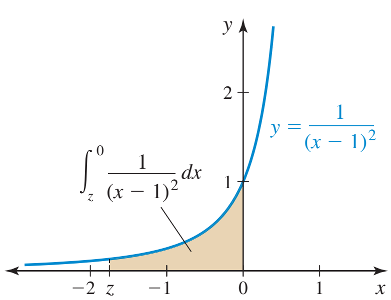


Figure 7.16 The region corresponding to $A(z)$ in Example 3.

We next discuss the case when both limits of integration are infinite.

Assume that $f(x)$ is continuous on $(-\infty, \infty)$. Then

$$\int_{-\infty}^\infty f(x)\,dx = \int_{-\infty}^a f(x)\,dx + \int_a^\infty f(x)\,dx \tag{7.26}$$

where a is a real number. If *both* improper integrals on the right-hand side of (7.26) are convergent, then the improper integral on the left-hand side of (7.26) is the sum of the two limiting values on the right-hand side.

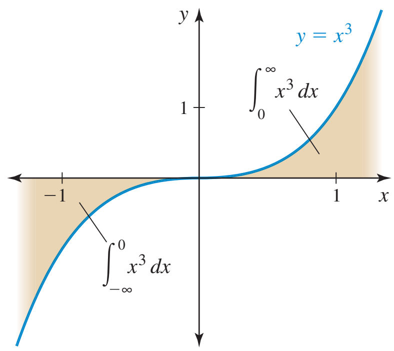


Figure 7.17 The integral $\int_{-\infty}^{\infty} x^3\,dx$ is divergent.

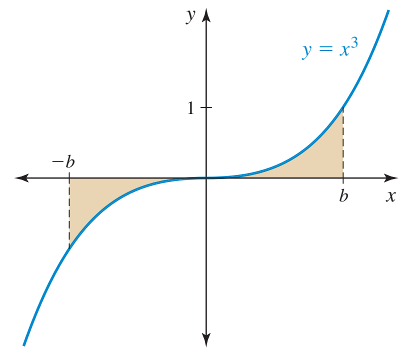


Figure 7.18 Because of symmetry, $\int_{-b}^{b} x^3\,dx = 0$.

Suppose that we wish to compute $\int_{-\infty}^{\infty} x^3\,dx$. We choose a value $a \in (-\infty, \infty)$—for instance, $a = 0$. Then

$$\int_{-\infty}^{\infty} x^3\,dx = \int_{-\infty}^{0} x^3\,dx + \int_{0}^{\infty} x^3\,dx$$

Looking at Figure 7.17, you can see that both improper integrals on the right-hand side are divergent. We check this assertion for the second one: We have

$$\int_0^\infty x^3\,dx = \lim_{z\to\infty} \int_0^z x^3\,dx = \frac{1}{4}x^4\Big]_0^z = \frac{1}{4}\lim_{z\to\infty}(z^4 - 0)$$

which does not exist. Hence, $\int_{-\infty}^{\infty} x^3\,dx$ is divergent.

It is important to realize that the definition of $\int_{-\infty}^{\infty} f(x)\,dx$ is different from that of

$$\lim_{b\to\infty} \int_{-b}^{b} f(x)\,dx$$

We use $f(x) = x^3$ again to illustrate this difference. For any $b > 0$, we find that

$$\int_{-b}^{b} x^3\,dx = \frac{1}{4}x^4\Big]_{-b}^{b} = \frac{1}{4}(b^4 - (-b)^4) = 0 \qquad \text{See Figure 7.18}$$

Therefore, $\lim_{b\to\infty} \int_{-b}^{b} x^3\,dx = 0$. This limit is not the same as $\int_{-\infty}^{\infty} x^3\,dx$.

Looking at (7.26), we see that, in order to evaluate $\int_{-\infty}^{\infty} f(x)\,dx$, we need to split up the integral at some $a \in \mathbf{R}$. There are often natural choices for a; we illustrate this in the next example.

EXAMPLE 4 **Infinite Upper and Lower Limit** Compute $\displaystyle\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx$.

Solution The graph of $f(x) = \frac{1}{1+x^2}$ is shown in Figure 7.19. The function $f(x) = 1/(1+x^2)$ is continuous for all $x \in \mathbf{R}$. It is symmetric about $x = 0$; a good choice for splitting up the integral is therefore $a = 0$. We write

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx = \int_{-\infty}^{0} \frac{1}{1+x^2}\,dx + \int_{0}^{\infty} \frac{1}{1+x^2}\,dx$$

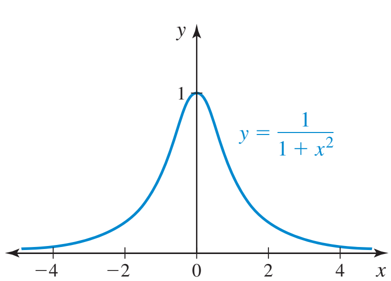


Figure 7.19 The graph of $f(x) = \frac{1}{1+x^2}$ in Example 4.

Now,

$$\begin{aligned}\lim_{z\to\infty} \int_0^z \frac{1}{1+x^2}\,dx &= \lim_{z\to\infty} \left[\tan^{-1} x\right]_0^z \\ &= \lim_{z\to\infty} (\tan^{-1} z - \tan^{-1} 0) = \frac{\pi}{2}\end{aligned}$$

and

$$\lim_{z\to-\infty}\int_z^0 \frac{1}{1+x^2}\,dx = \lim_{z\to-\infty}\left[\tan^{-1}x\right]_z^0$$
$$= \lim_{z\to-\infty}(\tan^{-1}0 - \tan^{-1}z) = \frac{\pi}{2}$$

That $\int_{-\infty}^0 \frac{1}{1+x^2}\,dx = \frac{\pi}{2}$ is expected because of symmetry. The area of the region to the left of the y-axis is equal to the area of the region to the right of the y-axis. Putting things together, we find that

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx = \frac{\pi}{2}+\frac{\pi}{2} = \pi.$$

EXAMPLE 5 **Infinite Upper and Lower Limit** Compute $\displaystyle\int_{-\infty}^{\infty}\frac{x}{1+x^2}\,dx$.

Solution The graph of $f(x) = \frac{x}{1+x^2}$ is shown in Figure 7.20. The function $f(x) = x/(1+x^2)$ is continuous for all $x \in \mathbf{R}$. Because $f(x)$ is an odd function, you might be tempted to say that the signed area to the left of 0 is the negative of the area to the right of 0 and, therefore, the value of the improper integral should be 0. But this is wrong! We choose $a = 0$, and write

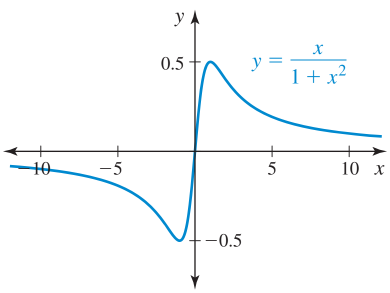


Figure 7.20 The graph of $f(x) = \frac{x}{1+x^2}$ in Example 5.

$$\int_{-\infty}^{\infty}\frac{x}{1+x^2}\,dx = \int_{-\infty}^{0}\frac{x}{1+x^2}\,dx + \int_{0}^{\infty}\frac{x}{1+x^2}\,dx$$

We begin by computing

$$\int_0^z \frac{x}{1+x^2}\,dx$$

Using the substitution $u = 1+x^2$ and $du = 2x\,dx$, we find that

$$\int_0^z \frac{x}{1+x^2}\,dx = \int_1^{1+z^2}\frac{1}{2u}\,du = \frac{1}{2}\ln|u|\Big]_1^{1+z^2}$$
$$= \frac{1}{2}[\ln(1+z^2) - \ln 1] = \frac{1}{2}\ln(1+z^2).$$

Taking the limit as $z \to \infty$, we obtain

$$\int_0^{\infty}\frac{x}{1+x^2}\,dx = \lim_{z\to\infty}\frac{1}{2}\ln(1+z^2) = \infty.$$

Since one of the integrals is already divergent, we conclude that $\int_{-\infty}^{\infty}\frac{x}{1+x^2}\,dx$ is divergent and therefore cannot be equal to 0. This example has an important take-home message: Before we can use symmetry to compute an improper integral, we need to make sure that the integral exists.

7.4.2 Type 2: Unbounded Integrand

So far, when we computed a definite integral, we made sure that the integrand was continuous over the interval of integration so that an antiderivative exists. We will now explain what to do when the integrand becomes infinite at one or more points of the interval. Suppose we wish to integrate

$$\int_0^1 \frac{dx}{\sqrt{x}}$$

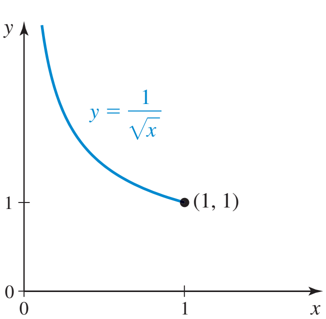


Figure 7.21 The graph of $y = \frac{1}{\sqrt{x}}$.

The graph of $f(x) = \frac{1}{\sqrt{x}}$ is shown in Figure 7.21. We see immediately that $f(x)$ is continuous on $(0, 1]$ and undefined at $x = 0$, and that

$$\lim_{x\to 0^+}\frac{1}{\sqrt{x}} = \infty$$

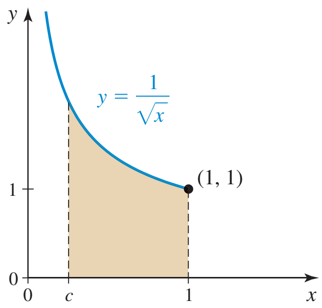


Figure 7.22 The area of the shaded region is $\int_c^1 \frac{1}{\sqrt{x}}\,dx = 2(1-\sqrt{c})$.

Let's choose a number $c \in (0, 1)$ and compute

$$\int_c^1 \frac{dx}{\sqrt{x}} = 2\sqrt{x}\,\Big]_c^1 = 2(1-\sqrt{c})$$ See Figure 7.22

If we now let $c \to 0^+$, we may regard the limiting value (if it exists) as the definite integral $\int_0^1 \frac{1}{\sqrt{x}}\,dx$. That is,

$$\int_0^1 \frac{dx}{\sqrt{x}} = \lim_{c\to 0^+} \int_c^1 \frac{dx}{\sqrt{x}} = \lim_{c\to 0^+} 2(1-\sqrt{c}) = 2$$

If f is continuous on $(a, b]$ and $\lim_{x\to a^+} f(x) = \pm\infty$ (see Figure 7.23), we define

$$\int_a^b f(x)\,dx = \lim_{c\to a^+} \int_c^b f(x)\,dx$$

provided that this limit exists. If the limit exists, we say that the improper integral on the left-hand side **converges**; if the limit does not exist, we say that the integral **diverges**.

Similarly, if f is continuous on $[a, b)$ and $\lim_{x\to b^-} f(x) = \pm\infty$ (see Figure 7.24), we define

$$\int_a^b f(x)\,dx = \lim_{c\to b^-} \int_a^c f(x)\,dx$$

provided that this limit exists.

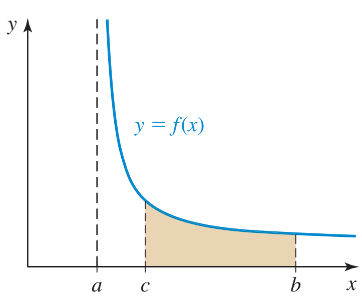


Figure 7.23 The improper integral $\int_a^b f(x)\,dx$ is defined as the limit of $\int_c^b f(x)\,dx$ as $c \to a^+$.

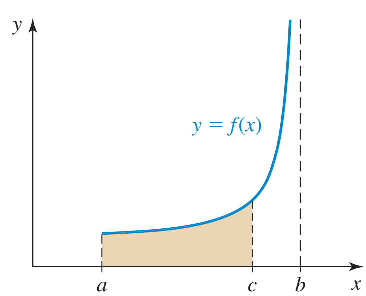


Figure 7.24 The improper integral $\int_a^b f(x)\,dx$ is defined as the limit of $\int_a^c f(x)\,dx$ as $c \to b^-$.

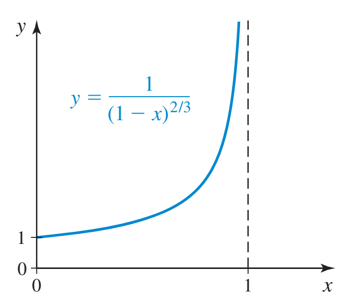


Figure 7.25 The graph of $f(x) = \frac{1}{(x-1)^{2/3}}$, $0 \le x < 1$, in Example 6.

EXAMPLE 6 **Integrand Undefined at Right Endpoint** Compute

$$\int_0^1 \frac{dx}{(x-1)^{2/3}}$$

Solution The graph of $f(x) = \frac{1}{(x-1)^{2/3}}$ is shown in Figure 7.25. From the graph we see that $f(x)$ is continuous on $[0, 1)$ and undefined at $x = 1$, and $\lim_{x\to 1^-} f(x) = \infty$. To compute the integral, we choose a number $c \in (0, 1)$ and compute

$$\int_0^c \frac{dx}{(x-1)^{2/3}}$$ See Figure 7.26

Letting $c \to 1^-$ will then produce the desired integral. That is,

$$\int_0^1 \frac{dx}{(x-1)^{2/3}} = \lim_{c\to 1^-} \int_0^c \frac{dx}{(x-1)^{2/3}}$$

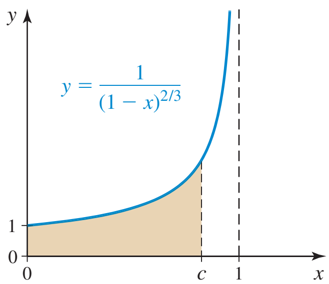


Figure 7.26 The area of the shaded region is $\int_0^c \frac{1}{(x-1)^{2/3}}\,dx$.

We first compute the indefinite integral

$$\int \frac{dx}{(x-1)^{2/3}} = 3(x-1)^{1/3} + C$$

If we set $F(x) = 3(x-1)^{1/3}$, then

$$\begin{aligned}\lim_{c\to 1^-}\int_0^c \frac{dx}{(x-1)^{2/3}} &= \lim_{c\to 1^-}[F(c) - F(0)] \\ &= \lim_{c\to 1^-}\left[3(c-1)^{1/3} - 3(-1)^{1/3}\right] = 3\end{aligned} \tag{7.27}$$

We therefore find that $\int_0^1 \frac{dx}{(x-1)^{2/3}} = 3$. •

EXAMPLE 7 **Integrand Undefined at Left Endpoint** Compute $\displaystyle\int_0^1 \ln x\,dx$.

Solution The graph of $f(x) = \ln x$, $0 < x \le 1$, is shown in Figure 7.27. We see from the graph that $f(x)$ is continuous on $(0, 1]$ and not defined at $x = 0$, and that $\lim_{x\to 0^+} f(x) = -\infty$. To determine the definite integral, we need to compute $\lim_{c\to 0^+}\int_c^1 \ln x\,dx$.

Since $F(x) = x\ln x - x$ is an antiderivative of $f(x) = \ln x$ (see Example 4 in Section 7.2), we find that

$$\begin{aligned}\lim_{c\to 0^+}\int_c^1 \ln x\,dx &= \lim_{c\to 0^+}[F(1) - F(c)] \\ &= \lim_{c\to 0^+}[1\ln 1 - 1 - c\ln c + c].\end{aligned}$$

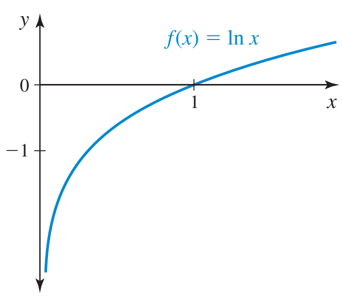


Figure 7.27 The graph of $f(x) = \ln x$, $0 < x \le 1$, in Example 7.

We need to find $\lim_{c\to 0^+} c\ln c$. The limit is of the form $0\cdot\infty$. L'Hôpital's rule yields

$$\begin{aligned}\lim_{c\to 0^+} c\ln c = \lim_{c\to 0^+}\frac{\ln c}{\frac{1}{c}} &= \lim_{c\to 0^+}\frac{\frac{1}{c}}{-\frac{1}{c^2}} \qquad \text{Write in form } \tfrac{\infty}{\infty} \\ &= \lim_{c\to 0^+}\left(-\frac{1}{c}\cdot\frac{c^2}{1}\right) = -\lim_{c\to 0^+} c = 0.\end{aligned}$$

Together with $\lim_{c\to 0^+} c = 0$, we therefore find that

$$\int_0^1 \ln x\,dx = \lim_{c\to 0^+}\int_c^1 \ln x\,dx = -1.$$ •

EXAMPLE 8 **Integrand Discontinuous in Interval** Compute $\displaystyle\int_{-1}^1 \frac{1}{x^2}\,dx$

Solution The function $f(x) = \frac{1}{x^2}$ is not defined at $x = 0$. In fact, it has a vertical asymptote at $x = 0$, since

$$\lim_{x\to 0^-}\frac{1}{x^2} = \infty \qquad \text{and} \qquad \lim_{x\to 0^+}\frac{1}{x^2} = \infty.$$

The graph of $f(x) = \frac{1}{x^2}$, $x \ne 0$, is shown in Figure 7.28. We see that $f(x) = 1/x^2$ is continuous, except at $x = 0$. To deal with this discontinuity, we split the integral at $x = 0$. We write

$$\int_{-1}^1 \frac{1}{x^2}\,dx = \lim_{c\to 0^-}\int_{-1}^c \frac{1}{x^2}\,dx + \lim_{d\to 0^+}\int_d^1 \frac{1}{x^2}\,dx \tag{7.28}$$

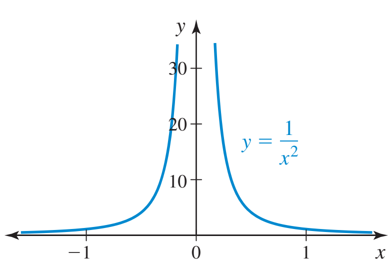


Figure 7.28 The graph of $f(x) = \frac{1}{x^2}$.

(See Figure 7.29.) The function $F(x) = -\frac{1}{x}$ is an antiderivative of $\frac{1}{x^2}$. Therefore,

$$\begin{aligned}\lim_{c\to 0^-}\int_{-1}^c \frac{1}{x^2}\,dx &= \lim_{c\to 0^-}[F(c) - F(-1)] \qquad \text{Evaluate first term in (7.28)} \\ &= \lim_{c\to 0^-}\left[-\frac{1}{c} - 1\right] = \infty.\end{aligned}$$

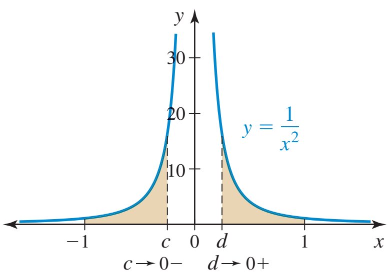


Figure 7.29 The improper integral $\int_{-1}^{1} \frac{1}{x^2}\,dx$.

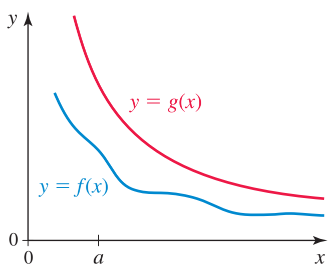


Figure 7.30 The function $g(x)$ lies above the function $f(x)$.

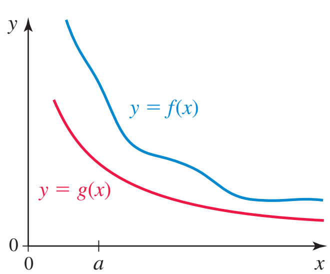


Figure 7.31 The graph of $g(x)$ is below the graph of $f(x)$.

We don't need to calculate the other integral in (7.28) because if the first term is divergent then the whole integral is divergent.

Therefore, $\int_{-1}^{1} \frac{1}{x^2}\,dx$ is divergent. •

7.4.3 A Comparison Result for Improper Integrals

In many cases, it is impossible to evaluate an integral exactly. In dealing with improper integrals, we frequently must know whether the integral converges. We may be able to determine whether or not the integral converges even without being able to evaluate the integral exactly. To show convergence it is often helpful to compare the integrand to simpler integrands that either dominate or are dominated by the integrand of interest. We will explain this idea graphically.

We assume that $f(x) \geq 0$ for $x \geq a$. Suppose we wish to show that $\int_a^\infty f(x)\,dx$ is convergent. Then it is enough to find a function $g(x)$ such that $g(x) \geq f(x)$ for all $x \geq a$ and $\int_a^\infty g(x)\,dx$ is convergent. This is illustrated in Figure 7.30. From the graph we can see that:

$$0 \leq \int_a^\infty f(x)\,dx \leq \int_a^\infty g(x)\,dx$$

If $\int_a^\infty g(x)\,dx < \infty$, it follows that $\int_a^\infty f(x)\,dx$ is convergent, since $\int_a^\infty f(x)\,dx$ must take on a value between 0 and a finite number, given by $\int_a^\infty g(x)\,dx$.

Suppose instead we have an integral $\int_a^\infty f(x)\,dx$ that we suspect is divergent. If we assume that $f(x) \geq 0$ for all $x \geq a$ it is then enough to find a function $g(x)$ such that $0 \leq g(x) \leq f(x)$ for all $x \geq a$ and $\int_a^\infty g(x)\,dx$ is divergent. This is illustrated in Figure 7.31. From the graph we see that

$$\int_a^\infty f(x)\,dx \geq \int_a^\infty g(x)\,dx \geq 0$$

If $\int_a^\infty g(x)\,dx$ is divergent, it follows that $\int_a^\infty f(x)\,dx$ is divergent.

You can see from the preceding discussion that in one case we selected a function that dominated $f(x)$, whereas in the other case we selected a function that was dominated by $f(x)$. This indicates that, before you find a comparison function, you must first guess whether the integral is likely to converge. (With practice, you get better at guessing whether an integral converges or diverges.) Sketching the functions involved can help you convince yourself that you are making the comparison in the right direction. Your comparison function, of course, should be simple enough so that you can integrate it without any problems. We present two examples that illustrate both cases.

EXAMPLE 9 **Convergence** Show that $\displaystyle\int_0^\infty e^{-x^2}\,dx$ is convergent.

Solution This slightly odd looking integral is tremendously important in probability models—we will meet it again in Chapter 12. The function $f(x) = e^{-x^2}$ is continuous and positive for $x \in [0, \infty)$. We cannot compute the antiderivative of $f(x) = e^{-x^2}$ with any of the techniques we have learned in this text. In fact, there is no simple way to express the value of $\int_0^z e^{-x^2}\,dx$ for $z > 0$. (It can be expressed as a sum of infinitely many terms.) But we can still determine whether the integral is convergent. To do so, we write $\int_0^\infty e^{-x^2}\,dx$ as a sum of two integrals and then show that each one is finite. We have

$$\int_0^\infty e^{-x^2}\,dx = \int_0^1 e^{-x^2}\,dx + \int_1^\infty e^{-x^2}\,dx$$

Since $0 < e^{-x^2} \leq 1$, it follows that

$$0 < \int_0^1 e^{-x^2}\,dx \leq \int_0^1 1dx = 1 < \infty$$

To show that $\int_1^{\infty} e^{-x^2}\,dx$ is convergent, we use the fact that e^{-x} is a decreasing function and that if $x \geq 1$, then $x \leq x^2$. It then follows that

$$0 \leq e^{-x^2} \leq e^{-x} \quad \text{for } x \geq 1.$$

Therefore,

$$0 \leq \int_1^{\infty} e^{-x^2}\,dx \leq \int_1^{\infty} e^{-x}\,dx = \lim_{c\to\infty} \left[-e^{-x}\right]_1^c = e^{-1} < \infty$$

Since both integrals are convergent, $\int_0^{\infty} e^{-x^2}\,dx$ is convergent.

Although for $0 < z < \infty$, $\int_0^z e^{-x^2}\,dx$ can be computed only approximately (e.g., using numerical methods of the sort we will discuss in Section 7.5), we can show with very different tools (which we do not cover in this text) that

$$\int_0^{\infty} e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}$$

•

EXAMPLE 10 **Divergence** Show that $\displaystyle\int_1^{\infty} \frac{1}{\sqrt{x+\sqrt{x}}}\,dx$ is divergent.

Solution The function $f(x) = 1/\sqrt{x+\sqrt{x}}$ is continuous on $[1, \infty)$. The integrand looks rather complicated, but since $x + \sqrt{x} \leq x + x$ for $x \geq 1$, it follows that

$$\frac{1}{\sqrt{x+\sqrt{x}}} \geq \frac{1}{\sqrt{2x}} \quad \text{for } x \geq 1$$

Hence,

$$\int_1^{\infty} \frac{1}{\sqrt{x+\sqrt{x}}}\,dx \geq \int_1^{\infty} \frac{1}{\sqrt{2x}}\,dx = \frac{1}{\sqrt{2}} \int_1^{\infty} \frac{1}{\sqrt{x}}\,dx = \infty$$

as shown in Example 2. Therefore, $\int_1^{\infty} \frac{dx}{\sqrt{x+\sqrt{x}}}$ is divergent. •

Section 7.4 Problems

7.4.1, 7.4.2

All the integrals in Problems 1–16 are improper and converge. Explain in each case why the integral is improper, and evaluate each integral.

1. $\displaystyle\int_0^{\infty} 3e^{-x}\,dx$

2. $\displaystyle\int_0^{\infty} xe^{-x}\,dx$

3. $\displaystyle\int_0^{\infty} \frac{2}{1+x^2}\,dx$

4. $\displaystyle\int_e^{\infty} \frac{dx}{x(\ln x)^2}$

5. $\displaystyle\int_{-\infty}^{-1} \frac{1}{1+x^2}\,dx$

6. $\displaystyle\int_1^{\infty} \frac{1}{x^{3/2}}\,dx$

7. $\displaystyle\int_{-\infty}^{\infty} e^{-|x|}\,dx$

8. $\displaystyle\int_{-\infty}^{\infty} xe^{-x^2/2}\,dx$

9. $\displaystyle\int_{-\infty}^{\infty} \frac{x}{(1+x^2)^2}\,dx$

10. $\displaystyle\int_{-\infty}^{\infty} x^3 e^{-x^4}\,dx$

11. $\displaystyle\int_0^{9} \frac{dx}{\sqrt{9-x}}$

12. $\displaystyle\int_0^{1} \frac{dx}{\sqrt{1-x^2}}$ Make the substitution $x = \sin u$.

13. $\displaystyle\int_{-2}^{2} \frac{dx}{\sqrt{4-x^2}}$ Make the substitution $x = 2\sin u$.

14. $\displaystyle\int_{-2}^{0} \frac{dx}{(x+1)^{1/3}}$

15. $\displaystyle\int_{-1}^{1} \ln|x|\,dx$

16. $\displaystyle\int_0^{2} \frac{dx}{(x-1)^{2/5}}$

In Problems 17–30, determine whether each integral is convergent. If the integral is convergent, compute its value.

17. $\displaystyle\int_0^{\infty} \frac{1}{x+1}\,dx$

18. $\displaystyle\int_{-1}^{0} \frac{1}{x+1}\,dx$

19. $\displaystyle\int_1^{\infty} \frac{1}{x^3}\,dx$

20. $\displaystyle\int_1^{\infty} \frac{1}{x^{1/3}}\,dx$

21. $\displaystyle\int_0^{4} \frac{1}{x^4}\,dx$

22. $\displaystyle\int_0^{4} \frac{1}{x^{1/4}}\,dx$

23. $\displaystyle\int_0^{2} \frac{1}{(x-1)^{1/3}}\,dx$

24. $\displaystyle\int_0^{\infty} \frac{1}{\sqrt{x+1}}\,dx$

25. $\displaystyle\int_0^{2} \frac{1}{(x-1)^4}\,dx$

26. $\displaystyle\int_{-1}^{0} \frac{1}{\sqrt{x+1}}\,dx$

27. $\displaystyle\int_e^{\infty} \frac{dx}{x\ln x}$

28. $\displaystyle\int_1^{e} \frac{dx}{x\ln x}$

29. $\displaystyle\int_{-2}^{2} \frac{2x\,dx}{(x^2-1)^{1/3}}$

30. $\displaystyle\int_{-\infty}^{1} \frac{3}{1+x^2}\,dx$

31. Determine whether

$$\int_{-\infty}^{\infty} \frac{1}{x^2 - 1}\, dx$$

is convergent. *Hint*: Use the partial-fraction decomposition

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

32. Although we cannot compute the antiderivative of $f(x) = e^{-x^2/2}$, it can be shown that:

$$\int_{-\infty}^{\infty} e^{-x^2/2}\, dx = \sqrt{2\pi}$$

Use this fact to show that

$$\int_{-\infty}^{\infty} x^2 e^{-x^2/2}\, dx = \sqrt{2\pi}$$

Hint: Write the integrand as $x \cdot (xe^{-x^2/2})$ and use integration by parts.

33. Determine the constant c so that

$$\int_0^{\infty} ce^{-3x}\, dx = 1$$

34. Determine the constant c so that

$$\int_{-\infty}^{\infty} \frac{c}{1+x^2}\, dx = 1$$

35. In this problem, we investigate the integral $\int_1^{\infty} \frac{1}{x^p}\, dx$ for $0 < p < \infty$.

(a) For $z > 1$, set $A(z) = \int_1^z \frac{1}{x^p}\, dx$ and show that

$$A(z) = \frac{1}{1-p}(z^{-p+1} - 1) \quad \text{for } p \neq 1$$

and

$$A(z) = \ln z \quad \text{for } p = 1.$$

(b) Use your results in (a) to show that, for $0 < p \leq 1$,

$$\lim_{z\to\infty} A(z) = \infty$$

(c) Use your results in (a) to show that, for $p > 1$,

$$\lim_{z\to\infty} A(z) = \frac{1}{p-1}$$

36. In this problem, we investigate the integral $\int_0^1 \frac{1}{x^p}\, dx$ for $0 < p < \infty$.

(a) Compute $\int \frac{1}{x^p}\, dx$ for $0 < p < \infty$. (*Hint*: Treat the case where $p = 1$ separately.)

(b) Use your result in (a) to compute $\int_c^1 \frac{1}{x^p}\, dx$ for $0 < c < 1$.

(c) Use your result in (b) to show that $\int_0^1 \frac{1}{x^p}\, dx = \frac{1}{1-p}$ for $0 < p < 1$.

(d) Show that $\int_0^1 \frac{1}{x^p}\, dx$ is divergent for $p \geq 1$.

7.4.3

37. (a) Show that

$$0 \leq e^{-x^2} \leq e^{-x}$$

for $x \geq 1$.

(b) Use your result in (a) to show that

$$\int_1^{\infty} e^{-x^2}\, dx$$

is convergent.

38. (a) Show that

$$0 \leq \frac{1}{\sqrt{1+x^4}} \leq \frac{1}{x^2}$$

for $x > 0$.

(b) Use your result in (a) to show that

$$\int_1^{\infty} \frac{1}{\sqrt{1+x^4}}\, dx$$

is convergent.

39. (a) Show that

$$\frac{1}{\sqrt{1+x^2}} \geq \frac{1}{2x} > 0$$

for $x \geq 1$.

(b) Use your result in (a) to show that

$$\int_1^{\infty} \frac{1}{\sqrt{1+x^2}}\, dx$$

is divergent.

40. (a) Show that

$$0 \leq \frac{1}{\sqrt{x+x^4}} \leq \frac{1}{x^2}$$

for $x > 0$.

(b) Use your result in (a) to show that

$$\int_1^{\infty} \frac{1}{\sqrt{x+x^4}}\, dx$$

is convergent.

In Problems 41–44, find a comparison function for each integrand and determine whether the integral is convergent.

41. $\int_0^{\infty} e^{-x^2/2}\, dx$

42. $\int_1^{\infty} \frac{1}{(1+x^4)^{1/3}}\, dx$

43. $\int_1^{\infty} \frac{1}{\sqrt{1+x}}\, dx$

44. $\int_0^{\infty} \frac{1}{e^x + e^{-x}}\, dx$

45. (a) Show that

$$\lim_{x\to\infty} \frac{\ln x}{\sqrt{x}} = 0$$

(b) Use your result in (a) to show that

$$2\ln x \leq \sqrt{x} \tag{7.29}$$

for sufficiently large x. Use a graphing calculator to determine just how large x must be for (7.29) to hold.

(c) Use your result in (b) to show that

$$\int_0^{\infty} e^{-\sqrt{x}} dx \tag{7.30}$$

converges. Use a graphing calculator to sketch the function $f(x) = e^{-\sqrt{x}}$ together with its comparison function(s), and use your graph to explain how you showed that the integral in (7.30) is convergent.

46. (a) Show that

$$\lim_{x\to\infty} \frac{\ln x}{x} = 0$$

(b) Use your result in (a) to show that, for any $c > 0$,

$$cx \geq \ln x$$

for sufficiently large x.

(c) Use your result in (b) to show that, for any $p > 0$,

$$x^p e^{-x} \le e^{-x/2}$$

provided that x is sufficiently large.

(d) Use your result in (c) to show that, for any $p > 0$,

$$\int_0^\infty x^p e^{-x}\,dx$$

is convergent.

47. **Cellular Response Functions** ***Schwabe and Bruggeman (2014)*** modeled how yeast cells respond to a change in the amount of nutrient available in their environment. Schwabe and Bruggeman found that the time taken by the yeast cells to respond to an increase in the amount of nutrient available in their environment could be modeled by a **Gamma** distributed random variable. Specifically the probability that a cell responds in time t is proportional to $p(t) = t^{a-1}e^{-bt}$, where a and b are both positive constants. It can be shown (see Chapter 12) that the probability a cell responds at all (i.e., in finite time) to the change in environmental conditions is proportional to

$$\int_0^\infty p(t)\,dt$$

(a) Assume $a = 1$; show that the integral $\int_0^\infty p(t)\,dt$ is convergent and find its value.

(b) Now assume $a = 2$; again show that the integral is convergent, and find its value.

(c) If $a = 3/2$, you cannot use integration by parts to find the value of the integral; but you can still show that the integral is convergent using the comparison theorem. Use the integrand from part (b) as a comparison function to show that $\int_0^\infty p(t)\,dt$ still converges.

48. **Forest Fires** The time between forest fires is often modeled using a **Weibull** distribution. According to this distribution the likelihood that a time t elapses between the end of one forest fire and the start of the next one is proportional to $p(t) = t^{k-1}\exp(-t^k)$ where k is a positive constant.

It can be shown using the laws of probability (see Chapter 12) that the probability of a second forest fire starting at any time following the first is proportional to $\int_0^\infty p(t)\,dt$.

(a) Assuming $k = 1$, show that $\int_0^\infty p(t)\,dt$ is convergent and calculate the value of the integral.

(b) Now assume $k = 2$. Show that $\int_0^\infty p(t)\,dt$ is convergent and calculate the value of the integral (*Hint*: you will need to use integration by substitution).

(c) Now adapt your argument from (b) to show that the integral is convergent for any value of k obeying $k \ge 1$.

7.5 Numerical Integration

This section makes use of Σ-notation, which was introduced in Section 2.2. But you can work through the section ignoring the expressions that use Σ-notation, because each expression is also written out in longhand form first.

Some integrals, such as $\int_0^4 e^{-x^2}\,dx$ are impossible to evaluate exactly. That is, there is no simple representation of the antiderivative. In such situations, numerical approximations are needed.

One way to approximate an integral numerically comes directly from our initial approach to the area problem. To solve that problem, we approximated areas by rectangles; that is, we used the Riemann sum approximation. Recall that for f continuous,

$$\int_a^b f(x)\,dx = \lim_{n\to\infty}(f(x_0) + f(x_1) + \cdots + f(x_{n-1}))h$$

$$= \lim_{n\to\infty}\sum_{k=0}^{n-1} f(x_k)h$$

In Chapter 6 we used w for the subinterval length.

where for each n the points $x_0, x_1, x_2, \ldots, x_n$ evenly divide the interval $[a, b]$ into subintervals of length $h = \frac{b-a}{n}$. That is: $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ and $x_k - x_{k-1} = h$ for all k.

The central idea of integration is that we approximate the area under the graph $y = f(x)$ by n rectangular strips. We let the number of strips go to infinity to calculate the integral. But if we evaluate the sum of the strip areas for a finite value of n, we have a numerical approximation to the integral. In this section we will discuss methods based, respectively, on approximating the area under the graph by a set of rectangles (the midpoint rule) or by a set of trapezoids (the trapezoidal rule). Throughout, we assume that the function f is continuous on $[a, b]$.

7.5.1 The Midpoint Rule

In the Riemann sum formula we approximate the area under the curve $y = f(x)$ between $x = x_{k-1}$ and $x = x_k$ by a rectangle of width $h = x_k - x_{k-1}$ and height $f(x_{k-1})$. The rectangle will touch the curve at the left edge of the interval. From Figure 7.32, we can see that this rectangle will underestimate the true area if $f(x)$ is increasing over

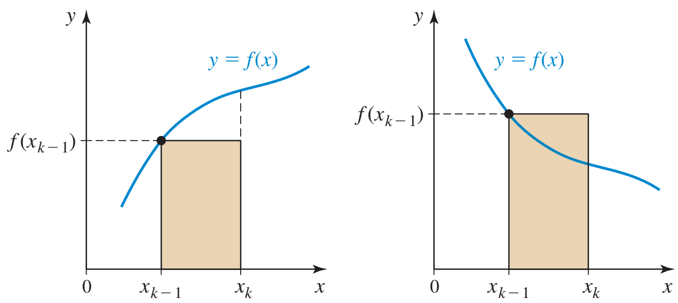


Figure 7.32 Approximating the area under $y = f(x)$ between x_{k-1} and x_k by a rectangle of height $f(x_{k-1})$. The area is underestimated if $f(x)$ is increasing and overestimated if $f(x)$ is decreasing.

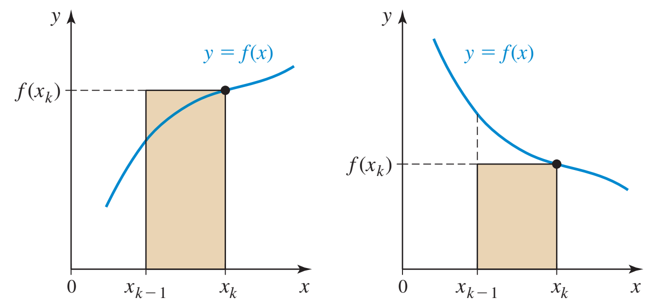


Figure 7.33 Approximating the area under $y = f(x)$ between x_{k-1} and x_k by a rectangle of height $f(x_k)$. The area is overestimated if $f(x)$ is increasing and underestimated if $f(x)$ is decreasing.

the interval $[x_{k-1}, x_k]$ and it will overestimate the area if $f(x)$ is decreasing. Can we approximate the area more accurately? If we instead make the height of the rectangle $f(x_k)$, then it will touch the curve on its right edge (see Figure 7.33). From the figure you can see that we now have the opposite problem: The area is underestimated if $f(x)$ is increasing and overestimated if $f(x)$ is decreasing. But we aren't forced to make the height of the strip equal to one of the end point values of the function. In fact, although the proof is beyond the scope of this book, the most accurate approximation comes from making the height of the strip equal to $f(x)$ at the midpoint of the interval (see Figure 7.34). That is, the height of our approximating rectangle is $f\left(\frac{x_{k-1}+x_k}{2}\right)$ and its area is $f\left(\frac{x_{k-1}+x_k}{2}\right) h$.

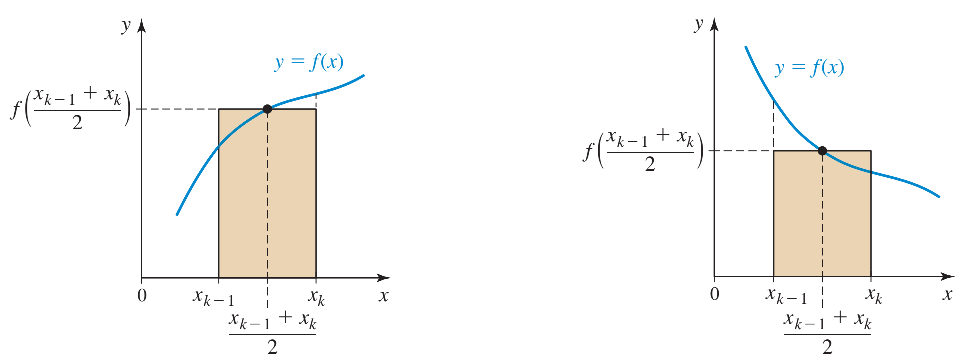


Figure 7.34 The midpoint rule accurately approximates the area under $y = f(x)$ between x_{k-1} and x_k, whether $f(x)$ is increasing or decreasing.

To approximate the area under the entire curve we sum the areas of the rectangular strips (see also Figure 7.35):

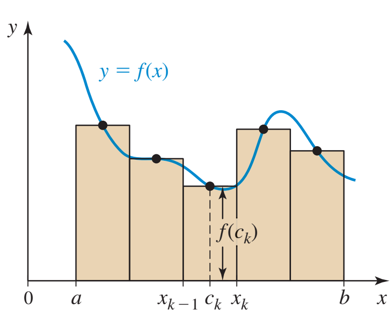


Figure 7.35 Approximating the area under a curve using n rectangular strips and the midpoint rule.

Midpoint Rule Suppose that $f(x)$ is continuous on $[a, b]$ and that $[x_0, x_1, \ldots, x_n]$ divides $[a, b]$ into n subintervals of equal length. We approximate $\int_a^b f(x)\,dx$ by

$$M_n = \frac{b-a}{n}(f(c_1) + f(c_2) + \cdots + f(c_n)) \qquad \text{Interval length is } h = \frac{b-a}{n}$$

$$= \frac{b-a}{n}\sum_{k=1}^{n} f(c_k)$$

where $c_k = \frac{x_{k-1}+x_k}{2}$ is the midpoint of $[x_{k-1}, x_k]$.

In the next example, we choose an integral that we can evaluate exactly, so that we can see how close the approximation is.

EXAMPLE 1 **Midpoint Rule** Use the midpoint rule with $n = 4$ to approximate

$$\int_0^1 x^2\,dx$$

Solution The function $f(x) = x^2$ is continuous on $[0, 1]$. For $n = 4$, we find $h = \frac{b-a}{4} = \frac{1}{4}$, so we obtain four subintervals, each of length $\frac{1}{4}$ (see Figure 7.36), as given in the following table:

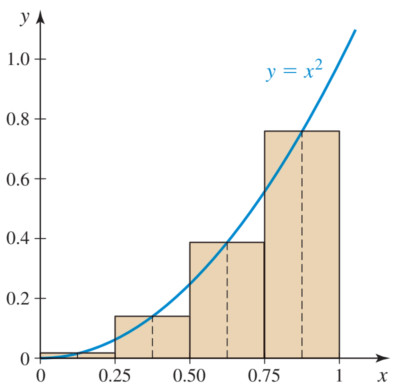


Figure 7.36 The midpoint rule for $\int_0^1 x^2\,dx$ with $n = 4$.

Subinterval $[x_{k-1}, x_k]$	***Midpoint*** $c_k = \frac{1}{2}(x_{k-1} + x_k)$	$f(c_k)$	$f(c_k)h$
$\left[0, \frac{1}{4}\right]$	$\frac{1}{8}$	$\frac{1}{64}$	$\frac{1}{64} \cdot \frac{1}{4}$
$\left[\frac{1}{4}, \frac{1}{2}\right]$	$\frac{3}{8}$	$\frac{9}{64}$	$\frac{9}{64} \cdot \frac{1}{4}$
$\left[\frac{1}{2}, \frac{3}{4}\right]$	$\frac{5}{8}$	$\frac{25}{64}$	$\frac{25}{64} \cdot \frac{1}{4}$
$\left[\frac{3}{4}, 1\right]$	$\frac{7}{8}$	$\frac{49}{64}$	$\frac{49}{64} \cdot \frac{1}{4}$

So the midpoint approximation is:

$$M_4 = \frac{b-a}{4}\sum_{k=1}^{4} f(c_k) = \frac{1}{4}\left(\frac{1}{64} + \frac{9}{64} + \frac{25}{64} + \frac{49}{64}\right) = \frac{1}{4} \cdot \frac{84}{64} = \frac{21}{64} \approx 0.3281$$

We know that $\int_0^1 x^2\,dx = \frac{1}{3}$. Hence, the error is

$$\left|\int_0^1 x^2\,dx - M_4\right| \approx 0.0052$$

A larger value of n would improve the approximation. •

Instead of memorizing the formula for the midpoint rule, it is easier to keep a picture in mind. We illustrate this heuristic in the next example.

EXAMPLE 2 **Midpoint Rule** Use the midpoint rule with $n = 5$ to approximate

$$\int_1^2 \frac{1}{x}\,dx$$

Solution The graph of $f(x) = \frac{1}{x}$, together with the five approximating rectangles, is shown in Figure 7.37. We see that $f(x) = 1/x$ is continuous on $[1, 2]$.

With $n = 5$, the partition of $[1, 2]$ is given by $P = [1, 1.2, 1.4, 1.6, 1.8, 2]$ and the midpoints are 1.1, 1.3, 1.5, 1.7, and 1.9. Since the width of each rectangle is 0.2 and $f(x) = \frac{1}{x}$, the area of the first rectangle is $(0.2)\frac{1}{1.1}$, the area of the second rectangle is $(0.2)\frac{1}{1.3}$, and so on. We thus find that

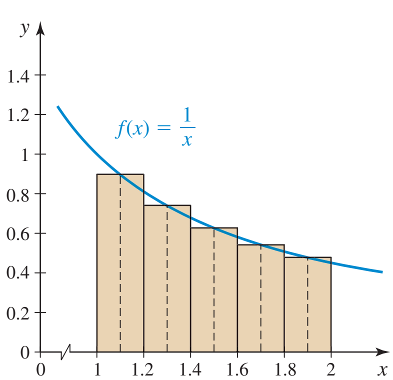


Figure 7.37 The midpoint rule for Example 2.

$$M_5 = (0.2)\left[\frac{1}{1.1} + \frac{1}{1.3} + \frac{1}{1.5} + \frac{1}{1.7} + \frac{1}{1.9}\right] = 0.6919$$

Note that we factored out 0.2, the width of each rectangle, since it is a common factor of the areas of the five rectangles.

We know that

$$\int_1^2 \frac{1}{x}\,dx = \ln x]_1^2 = \ln 2 - \ln 1 = \ln 2$$

Hence, the error in the approximation is

$$\left|\int_1^2 \frac{1}{x}\,dx - M_5\right| = |\ln 2 - 0.6919| = 0.0012$$ •

7.5.2 The Trapezoidal Rule

For many functions we can improve the accuracy of the approximation for the integral if, instead of using rectangular strips, we use trapezoids that have the same height as $f(x)$ at both endpoints of the subinterval. As you can see from Figure 7.38, this approximates the area of the subinterval whether $f(x)$ is increasing or decreasing.

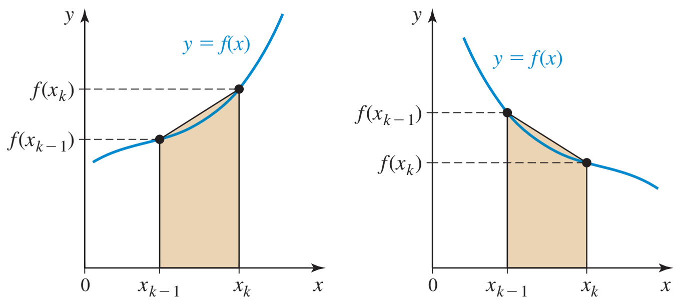


Figure 7.38 Approximating the area under the curve $y = f(x)$ between x_{k-1} and x_k using a trapezoid.

A trapezoid has two heights, the height of its left side and of its right side. If the left side has height d_1, the right side has height d_2, and the width is h (see Figure 7.39), then recall from geometry that the area of the trapezoid is:

$$\text{area } = h\frac{(d_1 + d_2)}{2}$$

In our approximation of the integral, the left height is $f(x_{k-1})$, the right height is $f(x_k)$, and the width is $h = x_k - x_{k-1}$, so the area is:

$$\text{area } = h\frac{(f(x_{k-1}) + f(x_k))}{2}$$

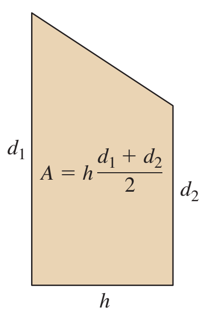


Figure 7.39 The area of a trapezoid.

To approximate the entire area under the curve we split the interval $[a, b]$ into n equal subintervals, that is, divide the interval into subintervals: $a = x_0 < x_1 < x_2 < \cdots < x_n = b$ with $x_k - x_{k-1} = h$. We then approximate the area in each subinterval by a trapezoid (see Figure 7.40), and sum up the (signed) areas of the trapezoids. The area of the kth trapezoid is $h\left(\frac{f(x_{k-1})+f(x_k)}{2}\right)$ and the width of each trapezoid is $h = \frac{b-a}{n}$. So, the sum of the areas of the trapezoids is:

$$T_n = \frac{b-a}{n}\left[\frac{f(x_0) + f(x_1)}{2} + \frac{f(x_1) + f(x_2)}{2} + \frac{f(x_2) + f(x_3)}{2} + \cdots + \frac{f(x_{n-2}) + f(x_{n-1})}{2} + \frac{f(x_{n-1}) + f(x_n)}{2}\right]$$

$$= \frac{b-a}{n}\left[\frac{f(x_0)}{2} + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{f(x_n)}{2}\right]$$

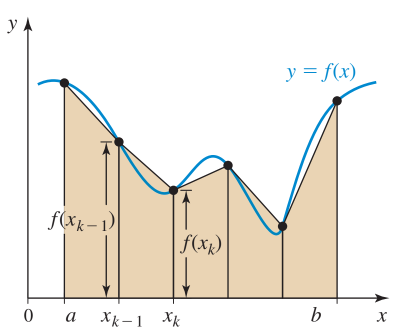


Figure 7.40 The trapezoidal rule.

Trapezoidal Rule Suppose that $f(x)$ is continuous on $[a, b]$ and that $a = x_0 < x_1 < x_2 < \cdots < x_n = b$; divide $[a, b]$ into n subintervals of equal length. Then we can approximate

$$\int_a^b f(x)\,dx$$

by

$$T_n = \frac{b-a}{n}\left[\frac{f(x_0)}{2} + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{f(x_n)}{2}\right]$$

Or using Σ-notation:

$$T_n = \frac{b-a}{n}\left(\frac{f(x_0)}{2} + \sum_{k=1}^{n-1} f(x_k) + \frac{f(x_n)}{2}\right).$$

We will now use the trapezoidal rule to evaluate the same two integrals that we approximated using the midpoint rule in Examples 1 and 2.

EXAMPLE 3 **Trapezoidal Rule** Use the trapezoidal rule with $n = 4$ to approximate

$$\int_0^1 x^2\, dx$$

Solution The function $f(x) = x^2$ is continuous on [0, 1]. As in Example 1, there are four subintervals, each of length $\frac{1}{4}$. (See Figure 7.41.) We find the following:

k	x_k	$f(x_k)$
0	0	0
1	$\frac{1}{4}$	$\frac{1}{16}$
2	$\frac{1}{2}$	$\frac{1}{4}$
3	$\frac{3}{4}$	$\frac{9}{16}$
4	1	1

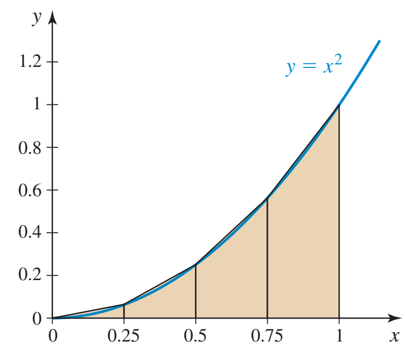


Figure 7.41 The trapezoidal rule for $\int_0^1 x^2\, dx$ with $n = 4$.

The approximation is

$$T_4 = \frac{1}{4}\left[\frac{0}{2} + \frac{1}{16} + \frac{1}{4} + \frac{9}{16} + \frac{1}{2}\right] = 0.34375$$

Since we know that $\int_0^1 x^2\, dx = \frac{1}{3}$, we can compute the error:

$$\left|\int_0^1 x^2\, dx - T_4\right| = 0.0104$$

Note that because $y = x^2$ is concave up, the line segments of the polygon are above the curve $y = x^2$. The area of the trapezoid therefore exceeds the area under the curve $y = x^2$ in each subinterval, and the trapezoidal approximation overestimates the integral. •

EXAMPLE 4 **Trapezoidal Rule** Use the trapezoidal rule with $n = 5$ to approximate

$$\int_1^2 \frac{1}{x}\, dx$$

Solution The situation is illustrated in Figure 7.42. The function $1/x$ is continuous on [1, 2]. With $n = 5$, the partition of [1, 2] is given by $P = [1.0, 1.2, 1.4, 1.6, 1.8, 2.0]$. The base of each trapezoid has length 0.2. Hence,

$$T_5 = (0.2)\left[\frac{1}{2}\cdot\frac{1}{1.0} + \frac{1}{1.2} + \frac{1}{1.4} + \frac{1}{1.6} + \frac{1}{1.8} + \frac{1}{2}\cdot\frac{1}{2.0}\right] = 0.69563$$

Since we know from Example 2 that $\int_1^2 \frac{1}{x}\, dx = \ln 2$, we can compute the error:

$$\left|\int_1^2 \frac{1}{x}\, dx - T_5\right| = 0.00249$$ •

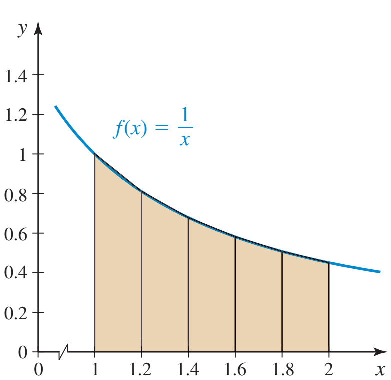


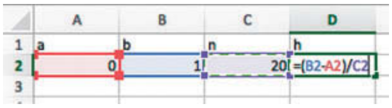
Figure 7.42 The trapezoidal rule for $\int_1^2 \frac{1}{x}\, dx$ with $n = 5$.

7.5.3 Using a Spreadsheet for Numerical Integration

Both the midpoint rule and trapezoid rule for numerical integration rely on evaluating the function $f(x)$ many times, and then summing these values. As we have seen when we studied recurrence equations (in Chapter 2) or solved an equation by an iterative method (in Chapter 3), processes that involve large numbers of repetitive operations can be automated using a spreadsheet.

Let's start by calculating an integral that we already know how to evaluate exactly, namely, $\int_0^1 x^3\,dx$, which we know evaluates to $\frac{1}{4}$.

In our spreadsheet we will start by entering the values of a, b (the two endpoints of our interval) and n (the number of subintervals that we will use in our approximation). We will enter these data in columns A, B; and C. Let's start by labeling the columns. In cell $A1$ enter **a**; in cell $B1$ enter **b**; and in cell $C1$ enter **n**. In our example $a = 0$, $b = 1$; let's suppose we want to use 20 subintervals to evaluate the integral (so $n = 20$). Thus in cells $A2$, $B2$, and $C2$ we enter **0**, **1**, and **20**, respectively.

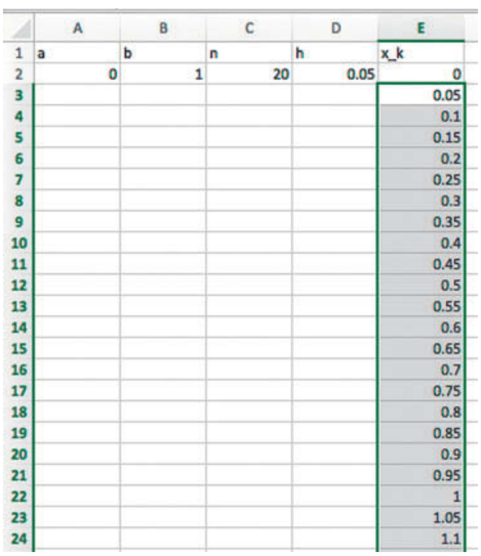


	A	B	C	D
1	a	b	n	h
2	0	1	20	=(B2-A2)/C2
3				

Figure 7.43 Entering the parameters used to evaluate an integral numerically.

We now need to calculate h (the size of our subintervals). Enter **h** in cell D1. Recall that $h = (b - a)/n$. In this case, since $n = 20$, $h = (1 - 0)/20 = 0.05$. But we can use the spreadsheet to calculate h. In the cell $D2$ enter the formula **= (B2 – A2)/C2**. (Your spreadsheet will now appear as in Figure 7.43).

We next find the points x_0, x_1, x_2 that are used to subdivide the interval of integration. We'll store those values in column E. We title the column (in cell *E1*) **x_k**. *E2* will then store the value of x_0; we can copy that over from cell *A2* (that is, enter the formula **= A2** in cell *E2*). Cell $E3$ will store the value of x_1. Since $x_1 = x_0 + h$, enter the formula **= E2 + D\$2** in cell $E3$ (we will explain the \$ sign shortly). To fill in the values of $x_2, x_3, x_4, \ldots, x_n$, we will use Autofill. Select the cell $E3$, click on the small square in the bottom right corner of the cell, and drag this square down many (that is, at least n) rows. Your spreadsheet should now appear in the form shown in Figure 7.44.



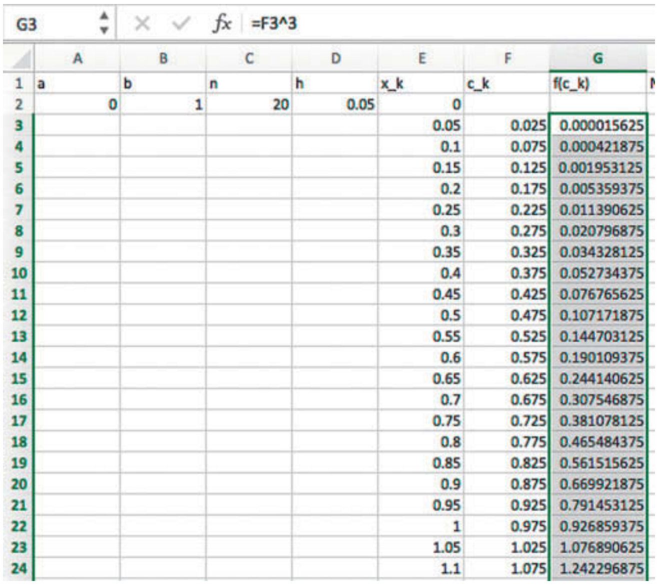
	A	B	C	D	E
1	a	b	n	h	x_k
2	0	1	20	0.05	0
3					0.05
4					0.1
5					0.15
6					0.2
7					0.25
8					0.3
9					0.35
10					0.4
11					0.45
12					0.5
13					0.55
14					0.6
15					0.65
16					0.7
17					0.75
18					0.8
19					0.85
20					0.9
21					0.95
22					1
23					1.05
24					1.1

Figure 7.44 Adding the points $x_0, x_1, x_2, \ldots$, to the spreadsheet.

Note that it is not necessary to AutoFill down exactly n rows—in general column E should contain at least as many points as you are using to approximate the integral; that is, the column of numbers should go up to b at least, but it is not a problem if we go past it (i.e., include numbers larger than b). In the spreadsheet shown in Figure 7.44, you can see that we go up past b (which is included in cell $E22$) to include 1.05 (in cell $E23$), 1.1 (in cell $E24$), and so on.

Why did we add a **\$** sign to the formula in cell *E3*? When AutoFill is used to fill a formula down one row, it automatically increments all rows in the formula by one. So if the formula that we enter in *E3* is **= E2 + D2** and we AutoFill one row down then in cell *E4*, AutoFill will enter the formula **= E3 + D3**. We get x_2 by adding h to x_1, just as we got x_1 by adding h to x_0. So we want the formula in cell $E4$ to calculate $x_1 + h$. The AutoFilled formula is partly correct: the first term is x_1, that is, it is correct to increment the first term in the formula. But the value of h is stored in cell *D2*. So the correct formula to enter in cell $E4$ is **= E3 + D2**. How do we stop AutoFill from incrementing the **D2** term? Adding the \$ symbol to a term in a formula prevents the term from being changed when the formula is AutoFilled to a new cell. So if we enter **= E2 + D\$2** in cell *E3* and we AutoFill one cell down, then the formula AutoFill put in cell *E4* will be **= E3 + D\$2**, which is what we need. Note, however, that if we had AutoFilled right from cell $E4$ to cell $F4$, then the AutoFilled formula would be **= F2 + E\$2**. The \$ symbol in **D\$2** only prevents the row from being changed by AutoFill; it does not affect the column. If we want to also stop the column from being changed, we would need to add another \$ symbol in front of the **D**; i.e., enter the formula **= E2 + \$D\$2** in cell $E3$.

For the midpoint method we need to calculate the midpoints of each interval: $c_k = \frac{1}{2}(x_{k-1} + x_k)$. We will store these values in column F. We title this column **c_k**. We will store the value of c_k next to the value of x_k, so c_1 goes in $F3$, next to x_1 (in $E3$). Since $c_1 = \frac{1}{2}(x_0 + x_1)$, we enter in cell $F3$ the formula **= 0.5 * (E2 + E3)**, since *E2* stores the value of x_0 and *E3* the value of x_1. We then use AutoFill to fill in the value of c_2 in *F4*, c_3 in *F5*. Again you need to fill down at least as far as c_n, but it is OK if we have extra values beyond c_n in column F. Once we have calculated each of the midpoints (i.e., the sequence $\{c_k\}$), we need to calculate $\{f(c_k)\}$. We will enter those values in column G. Title this column (in cell G1) **f(c_k)**. In cell *G3* enter the formula for $f(c_0)$: i.e., **= F3ˆ3**. Then use AutoFill to fill this formula down to the other cells in the column; so, for example, cell *G4* will receive the formula **= F4ˆ3**. After this



G3 =F3^3

	A	B	C	D	E	F	G
1	a	b	n	h	x_k	c_k	f(c_k)
2	0	1	20	0.05	0		
3					0.05	0.025	0.000015625
4					0.1	0.075	0.000421875
5					0.15	0.125	0.001953125
6					0.2	0.175	0.005359375
7					0.25	0.225	0.011390625
8					0.3	0.275	0.020796875
9					0.35	0.325	0.034328125
10					0.4	0.375	0.052734375
11					0.45	0.425	0.076765625
12					0.5	0.475	0.107171875
13					0.55	0.525	0.144703125
14					0.6	0.575	0.190109375
15					0.65	0.625	0.244140625
16					0.7	0.675	0.307546875
17					0.75	0.725	0.381078125
18					0.8	0.775	0.465484375
19					0.85	0.825	0.561515625
20					0.9	0.875	0.669921875
21					0.95	0.925	0.791453125
22					1	0.975	0.926859375
23					1.05	1.025	1.076890625
24					1.1	1.075	1.242296875

Figure 7.45 Using AutoFill to calculate the sequences of midpoints c_k and function values $f(c_k)$.

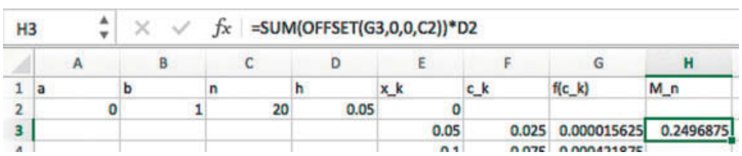
last AutoFill, your spreadsheet should appear as shown in Figure 7.45.

We will now apply the midpoint rule to evaluate the integral. Let's enter the title **M_n** in cell *H1* and use cell *H3* to store the approximate value of the integral. The main difficulty here is that we do not necessarily want to use all the values of $f(c_k)$ in column *G*, since only the values $f(c_1), f(c_2), \ldots, f(c_n)$ are needed in the midpoint rule formula. To sum only the values of $f(c_1), f(c_2), \ldots, f(c_n)$, that is, only the first n entries in column G, we combine the SUM and OFFSET functions.

OFFSET allows us to specify a particular set of cells in the spreadsheet. The way we want to use it is as follows:

OFFSET (G3, 0, 0, C2)

What does this do? The first argument needs to be the first term in our sum. This is $f(c_1)$, stored in cell *G3*. The next two arguments shift the row and column of this first entry. For example, if the second argument were 2, then the first entry in our sum would be shifted 2 rows down to *G5*. We don't need to make any shifts, so we just enter 0 for the second and third arguments. The fourth argument specifies how many terms we want to include. We need to sum n terms (that is, $f(c_1), f(c_2), \ldots$, up to $f(c_n)$), which are respectively stored in cells *G*3, *G*4, *G*5, ..., *G*22 (if $n = 20$). The number of terms that we need to sum (n) is stored in cell *C*2, so we enter **C2** as the last argument. So when OFFSET is called with this set of arguments, we get $f(c_1), f(c_2), \ldots, f(c_{20})$ as required.



H3 =SUM(OFFSET(G3,0,0,C2))*D2

	A	B	C	D	E	F	G	H
1	a	b	n	h	x_k	c_k	f(c_k)	M_n
2	0	1	20	0.05	0			
3					0.05	0.025	0.000015625	0.2496875

Figure 7.46 Using the OFFSET function to sum only the terms $f(c_1) + f(c_2) + \cdots + f(c_{20})$ to calculate the midpoint approximation to the integral.

Now since $M_n = h(f(c_1) + \cdots + f(c_n))$, we need to sum these terms and multiply by h. To do this we enter the formula

= SUM(OFFSET(G3, 0, 0, C2))*D2

in cell *H3*.

When we do this, the spreadsheet returns the numerical approximation to the integral: $M_{20} = 0.2496875$ (to 7 decimal places) see Figure 7.46. This approximation is close to the actual value of the integral:

$$\int_0^1 x^3\,dx = \frac{1}{4}x^4\Big]_0^1 = \frac{1}{4}.$$

The advantage of the spreadsheet approach, once it is set up, is that we can quickly change the interval of integration (i.e., a or b), or the number of subintervals (n), by changing the numbers in cells A2, B2, or C2. If we were evaluating M_n by hand we would have to recalculate all of the values c_n, and $f(c_n)$, effectively restarting our calculation from scratch. For example, suppose we want to calculate M_{40}, that is, approximate the integral using 40, rather than 20, subintervals. Then we simply replace the entry in (*C*2) by 40. For this to work, column *E* must contain enough entries to calculate $x_0, x_1, x_2, \ldots, x_{40}$; if it doesn't, then you will have to AutoFill the formula down extra cells until the entries in column *C* go at least as far as $x_{40} = 1$. Then the entry for M_n in cell *H2* is automatically updated to 0.2499219. We can calculate the error by comparing cell *H2* with the exact value of the integral, which is $\frac{1}{4} = 0.25$. Doubling the number of intervals from 20 to 40 decreased the error from $|0.2496875 - 0.25| \approx 3.111 \times 10^{-4}$ to $|0.2499219 - 0.25| \approx 7.8 \times 10^{-5}$ i.e., reduces the error by a factor of almost exactly 4. In fact it is generally true for the midpoint rule that, if the number of subintervals is doubled, then the error in the approximation will decrease by a factor of 4, as we shall discuss in the next subsection.

We can modify our previous spreadsheet to calculate an integral using the trapezoidal rule. To keep things interesting, let's evaluate an integral that we cannot calculate using any of our previous rules for integration.

EXAMPLE 5 Use the trapezoidal rule with (a) $n = 20$, (b) $n = 50$ subintervals to approximate $\int_0^1 e^{-x^2}\,dx$.

Solution This odd-looking function will show up throughout our study of probability in Chapter 12. Its antiderivative cannot be expressed in terms of familiar functions like exponentials or polynomials, and integrals like the one given above can only be evaluated numerically. Unlike the midpoint rule, we only need the endpoints of each subinterval (i.e., the points $x_0, x_1, \ldots, x_n$) and not also their midpoints (i.e., the points $c_1, c_2, \ldots, c_n$). So we set up the spreadsheet just as for the midpoint rule (that is, carry out the steps up to Figure 7.49) but we do not need to use column F to store the sequence $\{c_k\}$. Instead we will store the values of $f(x_k)$ there. That is, in cell *F1*, enter the title $f(x_k)$ and in cell *F2* enter the formula **= EXP(-1*E2^2)**, which calculates $f(x_0) = \exp(-x_0^2)$. Then AutoFill down the rest of column F.

We will use column G to calculate the numerical approximation to the integral. In cell *G1* enter **T_n**. Then cell *G2* will store the formula for T_n. Recall that to calculate T_n we need to sum $n+1$ terms:

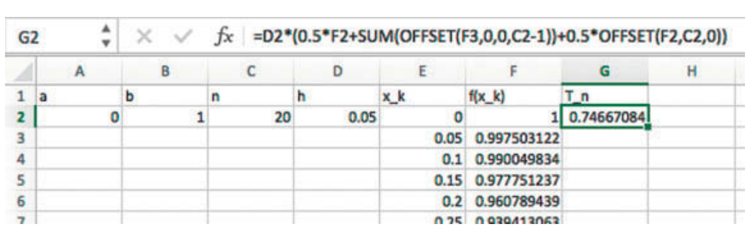
$$T_n = h\left(\frac{1}{2}f(x_0) + \overbrace{(f(x_1) + f(x_2) + \cdots + f(x_{n-1}))}^{n-1 \text{ terms}} + \frac{1}{2}f(x_n)\right)$$

Let's identify where to find those terms in the spreadsheet. h is stored in cell *D2*, $f(x_0)$ is stored in cell *F2*. To calculate $f(x_1) + f(x_2) + \cdots + f(x_{n-1})$ we need to sum cells *F3*, *F4*,$\ldots$, *F21* (assuming $n = 20$). Since the number of cells we need to sum depends on n, we use the OFFSET command, just as we did when applying the midpoint rule, since it allows us to sum sequences with variable numbers of terms. Recall that **OFFSET(F3,0,0,C2–1)** extracts a total of **C2–1** terms (i.e., $n-1$ terms) starting with cell *F3* and working downwards. Finally we need $f(x_n)$. If $n = 20$ this value is stored in cell *F22*, but its location will vary depending on the value of n. Again the **OFFSET** function comes to our aid. We need the cell that is n cells below *F2* (which stores $f(x_0)$). That cell is given by **OFFSET (F2,C2,0)** since the second argument of **OFFSET** shifts down a certain number of cells from the first argument (here it shifts down **C2**$= n$ cells from *F2*). So we can write the formula for T_n in the form

$$= \overbrace{\mathbf{D2}}^{h} * \Big(\overbrace{\mathbf{0.5 * F2}}^{\frac{1}{2}f(x_0)} + \overbrace{\mathbf{SUM(OFFSET(F3, 0, 0, C2 - 1)}}^{(f(x_1)+f(x_2)+\cdots+f(x_{n-1}))}$$

$$+ \underbrace{\mathbf{0.5 * OFFSET(F2,C2,0)}}_{\frac{1}{2}f(x_n)}\Big)$$

and enter in cell *G2*. After entering the formula your spreadsheet should appear as in Figure 7.47. We find $T_{20} = 0.74667084$ to 8 decimal places.


G2 =D2*(0.5*F2+SUM(OFFSET(F3,0,0,C2-1))+0.5*OFFSET(F2,C2,0))

	A	B	C	D	E	F	G	H
1	a	b	n	h	x_k	f(x_k)	T_n	
2	0	1	20	0.05	0	1	0.74667084	
3					0.05	0.997503122		
4					0.1	0.990049834		
5					0.15	0.977751237		
6					0.2	0.960789439		
7					0.25	0.939413063		

Figure 7.47 Using a spreadsheet to approximate $\int_0^1 e^{-x2}dx$ by the trapezoidal rule with $n = 20$ subintervals.

(b) To recalculate the integral using $n = 50$ terms, we need only enter 50 in cell *C2*. (You should check though that the terms in column E go all the way to $x_{50} = 1$. If they don't, use AutoFill to add more terms in columns E and F.) The spreadsheet now returns in cell *G2*, $T_{50} = 0.746800$ (to 6 decimal places). Note that this agrees to

3 decimal places with our answer from part (a). For an integral like this one, in which we have no exact answer to compare with, convergence in our numerical answer is our main way of checking the accuracy of our numerical approximation. •

7.5.4 Estimating Error in a Numerical Integration

In Sections 7.5.1 and 7.5.2 we calculated the error in our numerical approximation by comparing the numerical answer to the exact value of the integral. But the real value of numerical methods is that they can be used even when it is impossible to calculate an integral exactly.

In these situations we know from Section 7.5.3 that using more subintervals gives a more accurate approximation to the integral. But suppose we want to calculate the value of the integral to some desired accuracy (e.g., 10^{-3}); how many subintervals should we use? Fortunately there are methods for estimating the accuracy of the numerical approximation that don't require us to know the exact value of the integral. We will state these results without proof. The proof is outside the scope of this book, but can actually be derived using just the techniques from the next section. You can use the results without needing to know how they are proved.

Theorem Error Bound for the Midpoint Rule Suppose that $|f''(x)| \leq K$ for all $x \in [a, b]$. Then the error in the midpoint rule is at most

$$\left| \int_a^b f(x)\, dx - M_n \right| \leq K \frac{(b-a)^3}{24n^2}$$

Let's start by checking this bound for Examples 1 and 2 from Section 7.5.1. In the first example, $f(x) = x^2$, and therefore $f''(x) = 2$. Hence, with $n = 4$, the error is at most

$$\left| M_4 - \int_0^1 x^2 dx \right| \leq 2 \frac{(1-0)^3}{24(4^2)} \approx 0.0052$$

This is in fact the error that we obtained.

In the second example, $f(x) = \frac{1}{x}$. Since $f'(x) = -\frac{1}{x^2}$ and $f''(x) = \frac{2}{x^3}$, it follows that

$$|f''(x)| = \left| \frac{2}{x^3} \right| \leq 2 \quad \text{for } 1 \leq x \leq 2$$

Hence, with $n = 5$, the error is at most

$$\left| M_5 - \int_1^2 \frac{1}{x} dx \right| \leq 2 \frac{(2-1)^3}{24(5^2)} = 0.0033$$

The actual error was in fact smaller, only 0.0012. The actual error can be quite a bit smaller than the theoretical error bound, which is the worst-case scenario, but it will never be larger. One of the most important features of this error bound is that it tells us how changing the number of subintervals affects the error. Since the maximum error is proportional to $1/n^2$, if we double the number of subintervals from n to $2n$, then the maximum error will decrease from $\frac{K(b-a)^3}{24n^2}$ to $\frac{K(b-a)^3}{24(2n)^2} = \frac{1}{4} \cdot \frac{K(b-a)^3}{24n^2}$, that is, the maximum error will decrease by a factor of $\frac{1}{2^2} = \frac{1}{4}$. Similarly, if we use three times as many subintervals, the error will decrease by a factor of $\frac{1}{3^2} = \frac{1}{9}$, while using four times as many subintervals will decrease the error by a factor of $\frac{1}{4^2} = \frac{1}{16}$.

We may use the error bound result to find the number of subintervals required to obtain a certain accuracy. For instance, if we want to numerically approximate $\int_0^1 x^2\, dx$

so that the error is at most 10^{-4}, then we must choose n so that

$$K\frac{(b-a)^3}{24n^2} \le 10^{-4}$$

$$2 \cdot \frac{1}{24n^2} \le 10^{-4} \qquad f''(x) = 2 \Rightarrow K = 2$$

$$\frac{1}{12} \cdot 10^4 \le n^2$$

$$28.9 \le n$$

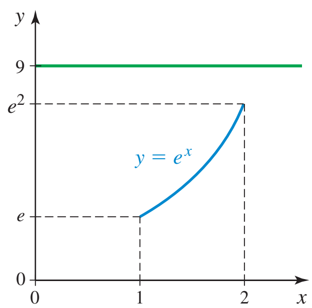


Figure 7.48 An upper bound on $|e^x|$ over $[1, 2]$ is 9.

That is, $n = 29$ would suffice to produce an error of at most 10^{-4}.

Finding a value for K in the estimate is not always easy. A graph of $f''(x)$ over the interval of interest can facilitate finding a bound on the second derivative. We need not find the best possible bound. For instance, if we wanted to integrate $f(x) = e^x$ over the interval $[1, 2]$, we would need to find a bound on $f''(x) = e^x$ over the interval $[1, 2]$. Since $|e^x| \le e^2 \approx 7.39$ over that interval, we could use, for instance, $K = e^2$. (See Figure 7.48.) However, any value of K for which $|f''(x)| \le K$ over $[a, b]$ would also work. For example, we could take $K = 9$, since $|e^x| \le 9$ for $1 \le x \le 2$.

A similar theoretical error bound can also be derived for the trapezoidal rule.

Error Bound for the Trapezoidal Rule Suppose that $|f''(x)| \le K$ for all $x \in [a, b]$. Then the error in the trapezoidal rule is at most

$$\left|\int_a^b f(x)\,dx - T_n\right| \le K\frac{(b-a)^3}{12n^2}$$

In Example 3, since $f(x) = x^2$, it follows that $f''(x) = 2$ and hence $K = 2$. The error is therefore bounded by

$$2\frac{1}{12(4^2)} = 0.0104$$

which is the same as the actual error, we computed in Example 3.

In Example 4, since $f(x) = 1/x$, we have $|f''(x)| = 2/x^3 \le 2$ for $1 \le x \le 2$ (as in Example 2). Hence, with $n = 5$, the error bound is at most

$$2\frac{(2-1)^3}{12(5^2)} = 0.0067$$

The actual error was in fact smaller, only 0.00249. As with the midpoint rule, the theoretical error can be quite a bit larger than the actual error.

Notice that when the error from approximating an integral using the midpoint rule is compared with the error from the trapezoidal rule, the upper bound on error from the midpoint rule is half that for the trapezoidal rule. That comparison is supported by the four examples that we worked through in Sections 7.5.1 and 7.5.2. For example, $|M_4 - \int_0^1 x^2\,dx|$ (the error from using the midpoint rule with four subintervals) was 0.0052, but the error from using the trapezoidal rule with the same number of intervals was twice as large because $|T_4 - \int_0^1 x^2\,dx| = 0.0104$.

Although the theoretical error bound for the trapezoidal rule has different coefficients from the theoretical error bound for the midpoint rule, the error is predicted to scale in the same way with n, the number of subintervals. That is, if we double the number of subintervals, then our predicted error will decrease by a factor of $\frac{1}{4}$. Just as with the midpoint rule error bound, we can use the bound to find the number of subintervals required to obtain a desired accuracy, and you will consider several instances of this type of calculation in the problems following this section.

Section 7.5 Problems

7.5.1, 7.5.2

In Problems 1–4, use the midpoint rule to approximate each integral with the specified value of n.

1. $\int_1^2 x^2\,dx, n = 4$

2. $\int_{-1}^0 (x+1)^2\,dx, n = 5$

3. $\int_0^1 \exp(x^2)\,dx, n = 3$

4. $\int_0^1 \sin(x^2)\,dx, n = 4$

In Problems 5–8, use the midpoint rule to approximate each integral with the specified value of n. Compare your approximation with the exact value.

5. $\int_1^2 \frac{1}{x}\,dx, n = 4$

6. $\int_0^1 (e^{2x} - 1)\,dx, n = 4$

7. $\int_0^4 x^3\,dx, n = 4$

8. $\int_1^3 \frac{2}{\sqrt{x}}\,dx, n = 5$

In Problems 9–12, use the trapezoidal rule to approximate each integral with the specified value of n.

9. $\int_1^2 \sqrt{x^2+1}\,dx, n = 4$

10. $\int_{-1}^0 \sin(x^2)\,dx, n = 5$

11. $\int_0^1 \exp(\sqrt{x})\,dx, n = 3$

12. $\int_0^1 \sin(\sqrt{x})\,dx, n = 4$

In Problems 13–16, use the trapezoidal rule to approximate each integral with the specified value of n. Compare your approximation with the exact value.

13. $\int_1^3 x^3\,dx, n = 5$

14. $\int_{-1}^1 (1 - e^{-x})\,dx, n = 4$

15. $\int_0^2 \sqrt{x}\,dx, n = 4$

16. $\int_1^2 \frac{1}{x}\,dx, n = 5$

7.5.3

T ***In Problems 17–22 use a spreadsheet to approximate each of the following integrals using the midpoint rule with each of the specified values of n.***

17. $\int_0^1 x^2\,dx$

(a) $n = 10$

(b) $n = 20$

18. $\int_{-1}^1 \sqrt{x+1}\,dx$

(a) $n = 10$

(b) $n = 30$.

19. $\int_0^1 \sin(x^2)\,dx$

(a) $n = 10$

(b) $n = 20$.

20. $\int_0^1 e^{-\sqrt{x}}\,dx$

(a) $n = 10$

(b) $n = 40$.

21. $\int_0^\pi e^{-x}\cos x\,dx$

(a) $n = 20$

(b) $n = 40$.

22. $\int_2^4 \frac{1}{\ln x}\,dx$

(a) $n = 20$

(b) $n = 50$.

T ***In Problems 23–28 use a spreadsheet to approximate each of the following integrals using the trapezoidal rule with each of the specified values of n.***

23. $\int_1^4 x^3\,dx$

(a) $n = 10$

(b) $n = 20$.

24. $\int_0^1 (x^2+1)^{1/3}\,dx$

(a) $n = 10$

(b) $n = 30$.

25. $\int_0^\pi x\sin x\,dx$

(a) $n = 10$

(b) $n = 20$.

26. $\int_1^\pi \frac{\cos x}{x}\,dx$

(a) $n = 15$

(b) $n = 30$.

27. $\int_1^5 \frac{e^{-x}}{x}\,dx$

(a) $n = 20$

(b) $n = 40$.

28. $\int_0^1 \exp(\cos x)\,dx$

(a) $n = 20$

(b) $n = 50$.

7.5.4

In Problems 29–36, use the theoretical error bound to determine how large n should be. [Hint: In each case, find the second derivative of the integrand you can use a graphing calculator to find an upper bound on $|f''(x)|$.]

29. How large should n be so that the midpoint rule approximation of

$$\int_0^2 x^2\,dx$$

is accurate to within 10^{-4}?

30. How large should n be so that the midpoint rule approximation of

$$\int_1^2 \frac{1}{x}\,dx$$

is accurate to within 10^{-3}?

31. How large should n be so that the midpoint rule approximation of

$$\int_0^2 e^{-x^2/2}\,dx$$

is accurate to within 10^{-4}?

32. How large should n be so that the midpoint rule approximation of

$$\int_2^4 \frac{1}{\ln t}\,dt$$

is accurate to within 10^{-3}?

33. How large should n be so that the trapezoidal rule approximation of

$$\int_0^1 e^{-x}\,dx$$

is accurate to within 10^{-5}?

34. How large should n be so that the trapezoidal rule approximation of

$$\int_0^\pi \sin x\,dx$$

is accurate to within 10^{-4}?

35. How large should n be so that the trapezoidal rule approximation of

$$\int_1^2 \frac{e^t}{t}\,dt$$

is accurate to within 10^{-4}?

36. How large should n be so that the trapezoidal rule approximation of

$$\int_0^1 e^{-x^2}\,dx$$

is accurate to within 10^{-3}?

37. (a) Show graphically that, for $n = 5$, the trapezoidal rule overestimates, and the midpoint rule underestimates,

$$\int_0^1 x^3\,dx$$

In each case, compute the approximate value of the integral and compare it with the exact value.

(b) The result in (a) has to do with the fact that $y = x^3$ is concave up on $[0, 1]$. To generalize that result to functions with this concavity property, we assume that the function $f(x)$ is continuous, nonnegative, and concave up on the interval $[a, b]$. Denote by M_n the midpoint rule approximation, and by T_n the trapezoidal rule approximation, of $\int_a^b f(x)\,dx$. Explain in words why

$$M_n \le \int_a^b f(x)\,dx \le T_n$$

(c) If we assume that $f(x)$ is continuous, nonnegative, and concave *down* on $[a, b]$, then

$$M_n \ge \int_a^b f(x)\,dx \ge T_n$$

Explain why this is so. Use this result to give an upper and a lower bound on

$$\int_0^1 \sqrt{x}\,dx$$

when $n = 4$ in the approximation.

T **38.** Using a spreadsheet, approximate the integral $\int_0^1 \sqrt{x}\,dx$ using the midpoint rule and:

(a) $n = 5$ subintervals, **(b)** $n = 10$ subintervals, **(c)** $n = 20$ subintervals, **(d)** $n = 50$ subintervals.

(e) What is the exact value of the integral?

(f) By comparing your answers from (a)–(d) with the exact answer, calculate the error $L_n = |M_n - \int_0^1 \sqrt{x}\,dx|$, and make a plot of L_n against n using your data.

(g) By plotting $\log L_n$ against $\log n$, show how your data support the claim that the error decreases proportional to $1/n^2$.

T **39.** Using a spreadsheet, approximate the integral $\int_0^2 e^{-x}\,dx$ using the trapezoidal rule and:

(a) $n = 5$ subintervals, **(b)** $n = 10$ subintervals, **(c)** $n = 20$ subintervals, **(d)** $n = 50$ subintervals.

(e) What is the exact value of the integral?

(f) By comparing your answers from (a)–(d) with the exact answer, calculate the error $L_n = |T_n - \int_0^2 e^{-2x}\,dx|$, and make a plot of L_n against n using your data.

(g) By plotting $\log L_n$ against $\log n$, show how your data support the claim that the error decreases proportional to $1/n^2$.

7.6 The Taylor Approximation

In Section 4.11 we discussed how to use a function derivative to approximate the function by a linear function. This is a useful trick, for example, to find solutions of equations (see Section 5.8). The trapezoid rule in Section 7.5 also relies on approximating a function by a series of linear functions to numerically approximate its integral. It is reasonable to ask, however, whether we can approximate the function even more accurately. More concretely, polynomial functions are easy to integrate and manipulate. Approximating a function by a linear function is equivalent to approximating it by a first order polynomial. In this section we will provide a procedure for approximating a function by polynomials of higher degree.

7.6.1 Taylor Polynomials

In Section 4.11 we discussed how to linearize a function about a given point. This discussion led to the linear, or tangent, approximation. We found the following:

> The linear approximation of $f(x)$ at $x = a$ is $f(x) \approx L(x)$ where
>
> $$L(x) = f(a) + f'(a)(x - a)$$

As an example, we look at

$$f(x) = e^x$$

and approximate this function by its linearization at $x = 0$. We find that

$$L(x) = f(0) + f'(0)x = 1 + x \tag{7.31}$$

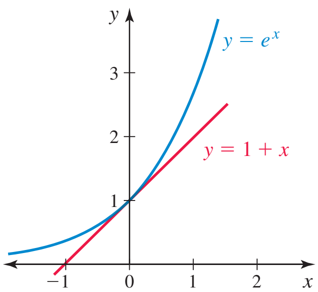


Figure 7.49 The graph of $y = e^x$ and its linear approximation at 0.

since $f'(x) = e^x$ and $f(0) = f'(0) = 1$. To see how close the approximation is, we graph both $f(x)$ and its approximation (Figure 7.49). The approximation is quite good as long as x is close to 0. The figure suggests that it gets gradually worse as we move away from 0. We derived the approximation for $f(x)$ geometrically, but we can be a little more systematic. Suppose we want to approximate $f(x)$ by a linear function $L(x)$. Then we must have

$$L(x) = mx + b$$

where m and b are both constant coefficients. Since there are two unknown coefficients, we must impose no constraints on $L(x)$. If we want to approximate $f(x)$ in the neighborhood $x = 0$, we make the first constraint that $L(0) = f(0)$, i.e., L and f give the same value at $x = 0$. Equivalently

$$m \cdot 0 + b = f(0) \quad \text{or} \quad b = f(0) \qquad \text{$f(0)$ is known, m and b are unknown}$$

To solve for coefficient m we need another constraint. Remembering that m represents the slope (or gradient) of $L(x)$, we make our second constraint that the gradients of L and f should agree at $x = 0$, i.e.,

$$\begin{aligned} L'(0) &= f'(0) \\ m &= f'(0). \qquad L'(x) = m \end{aligned}$$

Hence,

$$\begin{aligned} f(x) \approx L(x) &= b + mx \\ &= f(0) + f'(0)x \end{aligned}$$

and this is exactly our linear approximation formula when $a = 0$.

To improve the approximation, we may wish to use an approximating function whose higher-order derivatives also agree with those of $f(x)$ at $x = 0$. The function $L(x)$ is a polynomial of degree 1. To improve the approximation, we will continue to work with polynomials, but require that the function and its first n derivatives at $x = 0$ agree with those of the polynomial. To be able to match up the first n derivatives, the polynomial must be of degree n. A polynomial of degree n can be written as

$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \tag{7.32}$$

and there are $(n+1)$ unknown coefficients, $a_0, a_1, \ldots, a_n$ to solve for.

If we want to approximate $f(x)$ at $x = 0$, then we require that

$$\begin{aligned} f(0) &= P_n(0) \\ f'(0) &= P_n'(0) \\ f''(0) &= P_n''(0) \\ &\vdots \\ f^{(n)}(0) &= P_n^{(n)}(0) \end{aligned} \tag{7.33}$$

Now,

$$\begin{aligned} P_n(0) &= a_0 + a_1 x + \cdots + a_n x^n \big|_{x=0} = a_0 \\ P_n'(0) &= a_1 + 2a_2 x + 3a_3 x^2 + \cdots + n a_n x^{n-1} \big|_{x=0} = a_1 \\ P_n''(0) &= 2a_2 + (3)(2)a_3 x + \cdots + n(n-1)a_n x^{n-2} \big|_{x=0} = 2a_2 \\ P_n'''(0) &= (3)(2)a_3 + (4)(3)(2)a_4 x + \cdots + n(n-1)(n-2)a_n x^{n-3} \big|_{x=0} \\ &= (3)(2)a_3 \\ &\vdots \\ P_n^{(n)}(0) &= n(n-1)(n-2)\cdots(3)(2)(1)a_n \big|_{x=0} \\ &= n(n-1)(n-2)\cdots(3)(2)(1)a_n \end{aligned}$$

Factorial Notation If $k > 0$, define

$$k! = k(k-1)(k-2)\cdots(3)(2)(1)$$

and define $0! = 1$.

$k!$ is read "k factorial." Using the factorial notation, the derivatives of $P_n(x)$ can be written as:

$$P_n(0) = a_0, \quad P_n'(0) = a_1, \quad P_n''(0) = 2a_2, \quad P_n'''(0) = 3!a_3, \quad \ldots,$$
$$P_n^{(k)}(0) = k!a_k, \quad \ldots, \quad P_n^{(n)}(0) = n!a_n$$

We then substitute for the derivatives of P_n in (7.33) to solve for $a_0, a_1, \ldots, a_n$.

$$\begin{aligned} a_0 &= P_n(0) = f(0) \\ a_1 &= P_n'(0) = f'(0) \\ a_2 &= \frac{1}{2}P_n''(0) = \frac{1}{2!}f''(0) \\ a_3 &= \frac{1}{3!}P_n'''(0) = \frac{1}{3!}f'''(0) \\ &\vdots \\ a_n &= \frac{1}{n(n-1)\cdots 2\cdot 1}P_n^{(n)}(0) = \frac{1}{n!}f^{(n)}(0) \end{aligned} \tag{7.34}$$

A polynomial of degree n of the form (7.32) and whose coefficients satisfy (7.34) is called a **Taylor polynomial** of degree n. We summarize this definition as follows:

Definition The **Taylor polynomial of degree *n*** about $x = 0$ for the function $f(x)$ is given by

$$P_n(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \cdots + \frac{f^{(n)}(0)}{n!}x^n$$

EXAMPLE 1 Compute the Taylor polynomial of degree 3 about $x = 0$ for the function $f(x) = e^x$.

Solution To find the Taylor polynomial of degree 3, we need the first three derivatives of $f(x)$ at $x = 0$. We have

$$\begin{aligned} f(x) &= e^x, &\quad \text{so } f(0) &= 1 \\ f'(x) &= e^x, &\quad \text{so } f'(0) &= 1 \\ f''(x) &= e^x, &\quad \text{so } f''(0) &= 1 \\ f'''(x) &= e^x, &\quad \text{so } f'''(0) &= 1 \end{aligned}$$

Therefore,

$$P_3(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$

since $2! = (2)(1) = 2$ and $3! = (3)(2)(1) = 6$.

Our claim was that this polynomial would provide a better approximation to e^x than the linearization $1 + x$. We check this claim by evaluating e^x, $L(x)$, and $P_3(x)$ at a few values. The results are summarized in the following table:

x	e^x	$1+x$	$1+x+\frac{x^2}{2}+\frac{x^3}{6}$
-1	0.36788	0	0.3333
-0.1	0.90484	0.9	0.9048
0	1	1	1.0000
0.1	1.1052	1.1	1.1052
1	2.7183	2	2.6667

We see from the table that the third-degree Taylor polynomial provides a better approximation. Indeed, for x sufficiently close to 0, the values of $f(x)$ and $P_3(x)$ are very close. For instance,

$$f(0.1) = 1.105170918 \qquad \text{and} \qquad P_3(0.1) = 1.105166667$$

That is, their first five digits are identical. The error of approximation is

$$|f(0.1) - P_3(0.1)| = 4.25 \times 10^{-6}$$

which is quite small. By contrast the error from approximating e^x by $L(x)$ is:

$$|f(0.1) - L(0.1)| = 0.0052.$$

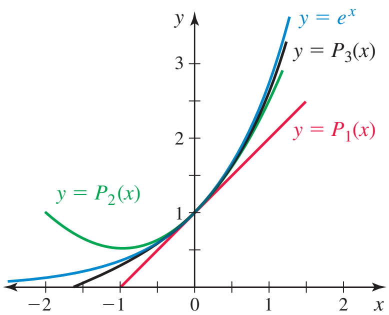


Figure 7.50 The graph of $y = e^x$ and the first three Taylor polynomials.

In Figure 7.50, we display the graphs of $f(x)$ and the Taylor polynomials $P_1(x)$, $P_2(x)$, and $P_3(x)$. We see from the graphs that the approximation is good only as long as x is close to 0. We also see that increasing the degree of the Taylor polynomial improves the approximation. •

When we look at Example 1, we find that the successive Taylor polynomials for $f(x) = e^x$ about $x = 0$ are

$$P_0(x) = 1$$
$$P_1(x) = 1 + x$$
$$P_2(x) = 1 + x + \frac{x^2}{2!}$$
$$P_3(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}$$

$P_1(x)$ is the same linear approximation $L(x)$ that we found in (7.31). There appears to be a pattern to the terms in the Taylor series approximation, and we might be tempted to guess the form of $P_n(x)$ for an arbitrary n. Our guess would be

$$P_n(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$$

and this is indeed the case.

You might wonder why we bother to find an approximation for the function $f(x) = e^x$. To compute $f(1) = e$, for instance, it seems a lot easier simply to use a calculator. But how does the calculator calculate e^x? In fact it makes use of Taylor series (and other approximations).

We now find the Taylor polynomial of degree n for other important functions.

EXAMPLE 2 Compute the Taylor polynomial of degree n about $x = 0$ for the function $f(x) = \sin x$.

Solution We begin by computing successive derivatives of $f(x) = \sin x$ at $x = 0$:

$$\begin{aligned}
f(x) &= \sin x & &\text{so} & f(0) &= 0\\
f'(x) &= \cos x & &\text{so} & f'(0) &= 1\\
f''(x) &= -\sin x & &\text{so} & f''(0) &= 0\\
f'''(x) &= -\cos x & &\text{so} & f'''(0) &= -1\\
f^{(4)}(x) &= \sin x & &\text{so} & f^{(4)}(0) &= 0
\end{aligned}$$

Since $f^{(4)}(x) = f(x)$, we find that $f^{(5)}(x) = f'(x)$, $f^{(6)}(x) = f''(x)$, and so on. We also conclude that all even derivatives are equal to 0 at $x = 0$ and that the odd derivatives alternate between 1 and -1 at $x = 0$. We find that

$$\begin{aligned}
P_1(x) &= P_2(x) = x \\
P_3(x) &= P_4(x) = x - \frac{x^3}{3!} \\
P_5(x) &= P_6(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} \\
P_7(x) &= P_8(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}
\end{aligned} \tag{7.35}$$

and so on. To find the Taylor polynomial of degree n, we must find out how to write the last term. Note that the sign in front of successive terms alternates between plus and minus. To account for this alternating sign, we introduce the factor

$$(-1)^n = \begin{cases} 1 & \text{if } n \text{ is even} \\ -1 & \text{if } n \text{ is odd} \end{cases}$$

An odd number can be written as $2n+1$ for any integer n. For a term of the form $\pm\frac{x^k}{k!}$ with k odd, we write

$$(-1)^n \frac{x^{2n+1}}{(2n+1)!} \tag{7.36}$$

where n is an integer. Inserting successive values of n into (7.36), we find the following:

n	0	1	2	3
$(-1)^n \frac{x^{2n+1}}{(2n+1)!}$	x	$-\frac{x^3}{3!}$	$\frac{x^5}{5!}$	$-\frac{x^7}{7!}$

We see from the table that the term (7.36) produces the successive terms in the Taylor polynomial for $f(x) = \sin x$ that we calculated in (7.35). The Taylor polynomial of degree $2n+1$ is thus

$$P_{2n+1}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

EXAMPLE 3 Compute the Taylor polynomial of degree n about $x = 0$ for the function $f(x) = \frac{1}{1-x}$, $x \neq 1$.

Solution We begin by computing successive derivatives of $f(x) = \frac{1}{1-x}$ at $x = 0$.

$$\begin{aligned}
f(x) &= \frac{1}{1-x}, && \text{so } f(0) = 1 \\
f'(x) &= \frac{1}{(1-x)^2}, && \text{so } f'(0) = 1 \\
f''(x) &= \frac{2}{(1-x)^3}, && \text{so } f''(0) = 2 = 2! \\
f'''(x) &= \frac{(2)(3)}{(1-x)^4}, && \text{so } f'''(0) = (2)(3) = 3! \\
f^{(4)}(x) &= \frac{(2)(3)(4)}{(1-x)^5}, && \text{so } f^{(4)}(0) = (2)(3)(4) = 4!
\end{aligned}$$

and so on. Continuing in this way, we find that

$$f^{(k)}(x) = \frac{(2)(3)(4)\cdots(k)}{(1-x)^{k+1}} = \frac{k!}{(1-x)^{k+1}} \quad \text{so } f^{(k)}(0) = k!$$

For the Taylor polynomial of degree n about $x = 0$, we obtain

$$\begin{aligned} P_n(x) &= 1 + x + \frac{2!}{2!}x^2 + \frac{3!}{3!}x^3 + \frac{4!}{4!}x^4 + \frac{5!}{5!}x^5 + \cdots + \frac{n!}{n!}x^n \\ &= 1 + x + x^2 + x^3 + \cdots + x^n \end{aligned}$$

Taylor approximations are widely used in biology. Here is an example that is already familiar to us.

EXAMPLE 4 Denote the size of a population at time t by $N(t)$. A differential equation modeling the growth of this population may take the form:

$$\frac{dN}{dt} = f(N) \qquad \text{with } f(0) = 0$$

Find the linear and the quadratic approximation of $f(N)$ about $N = 0$.

Solution We impose $f(0) = 0$ so that when $N = 0$, $\frac{dN}{dt} = 0$. If members are only added by reproduction, it is impossible for an initially zero-sized population to grow. The linear approximation of $f(N)$ about $N = 0$ is the Taylor polynomial of degree 1:

$$P_1(N) = \underbrace{f(0)}_{=0} + f'(0)N$$

If we set $r = f'(0)$, then the first-order approximation of this growth model is

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

which is the equation that describes exponential growth.

The quadratic approximation of $f(N)$ about $N = 0$ is the Taylor polynomial of degree 2:

$$P_2(N) = \underbrace{f(0)}_{=0} + \underbrace{f'(0)}_{r}N + \frac{f''(0)}{2}N^2$$

Factoring rN yields:

$$P_2(N) = rN\left[1 + \frac{f''(0)}{2r}N\right]$$

If we set $K = -\frac{2r}{f''(0)}$, then the second-order approximation of the growth model is

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

which is the equation that describes logistic growth if K and r are positive. In either approximation, $r = f'(0)$ is the intrinsic rate of growth, i.e., the limiting rate of growth if N is very small, so the population is not resource limited.

7.6.2 The Taylor Polynomial about $x = a$

Thus far, we have considered Taylor polynomials about $x = 0$. Because Taylor polynomials typically are good approximations only close to the point of approximation, it is useful to have approximations about points other than $x = 0$. We have already done this for linear approximations. For instance, the tangent-line approximation of $f(x)$ at $x = a$ is

$$f(x) \approx L(x) = f(a) + f'(a)(x - a) \tag{7.37}$$

Note that $L(a) = f(a)$ and $L'(a) = f'(a)$. That is, the linear approximation and the original function, together with their first derivatives, agree at $x = a$. If we want to approximate $f(x)$ at $x = a$ by a polynomial of degree n, we then require that the polynomial and the original function, together with their first n derivatives, agree at $x = a$. This leads us to a polynomial of the form

$$P_n(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + \cdots + c_n(x-a)^n \tag{7.38}$$

Comparing (7.38) and (7.37), we conclude that $c_0 = f(a)$ and $c_1 = f'(a)$. To find the remaining coefficients, we proceed as in the case $a = 0$. That is, we differentiate $f(x)$ and $P_n(x)$ and require that their first n derivatives agree at $x = a$. We then arrive at the following formula:

The **Taylor polynomial of degree n about $x = a$** of the function $f(x)$ is

$$P_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

EXAMPLE 5 Find the Taylor polynomial of degree 3 for $f(x) = \ln x$ at $x = 1$.

Solution We need to evaluate $f(x)$ and its first three derivatives at $x = 1$. We find that

$$\begin{aligned}
f(x) &= \ln x, & &\text{so } f(1) = 0\\
f'(x) &= \frac{1}{x}, & &\text{so } f'(1) = 1\\
f''(x) &= -\frac{1}{x^2}, & &\text{so } f''(1) = -1\\
f'''(x) &= \frac{2}{x^3}, & &\text{so } f'''(1) = 2
\end{aligned}$$

Using the definition of the Taylor polynomial, we get

$$\begin{aligned}
P_3(x) &= 0 + (1)(x-1) + \frac{(-1)}{2!}(x-1)^2 + \frac{2}{3!}(x-1)^3\\
&= (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3
\end{aligned}$$

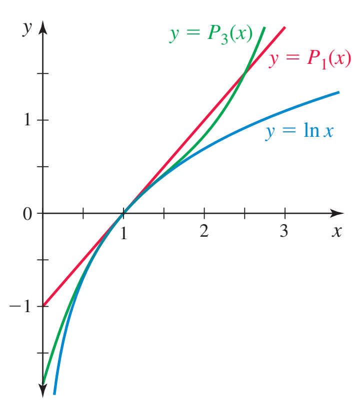


Figure 7.51 The graph of $y = \ln x$, the linear approximation, and the Taylor polynomial of degree 3.

Figure 7.51 shows $f(x)$, the linear approximation $P_1(x) = x - 1$, and $P_3(x)$. We see that the approximation is good when x is close to 1 and that the approximation $P_3(x)$ is better than the linear approximation. •

7.6.3 How Accurate Is the Approximation?

We saw in Example 1 that the approximation improved when the degree of the polynomial was higher. We will now investigate how accurate the Taylor approximation is. We can assess the accuracy of the approximation directly for the function in Example 3. The material in this subsection is quite technical. Although approximation of functions by Taylor series is immeasurably important for solving differential equations with a computer, and for analyzing them theoretically, for many applications it is quite enough to appreciate that the more terms there are in a Taylor series approximation, the more accurate it becomes. For this reason we encourage readers studying this material for the first time to skip this subsection.

In Example 3, we showed that the Taylor polynomial of degree n about $x = 0$ for $f(x) = \frac{1}{1-x}$, $x \neq 1$, is

$$P_n(x) = 1 + x + x^2 + x^3 + \cdots + x^{n-1} + x^n \tag{7.39}$$

We claimed that adding more terms to a Taylor series approximation often makes it more accurate. Is that always the case? We want to calculate the error (i.e., the difference between $f(x)$ and its Taylor series approximation).

There is a nice "trick" we may use to find an expression for the error in the approximation we derived in Example 3. Note that

$$xP_n(x) = x + x^2 + x^3 + \cdots + x^n + x^{n+1} \tag{7.40}$$

Subtracting (7.40) from (7.39), we find that

$$\begin{aligned} P_n(x) - xP_n(x) &= (1 + x + x^2 + \cdots + x^n) - (x + x^2 + \cdots + x^n + x^{n+1}) \\ &= 1 - x^{n+1} \qquad \text{Canceling terms in pairs.} \end{aligned}$$

That is,

$$(1 - x)P_n(x) = 1 - x^{n+1}$$

or

$$P_n(x) = \frac{1 - x^{n+1}}{1 - x} = \frac{1}{1 - x} - \frac{x^{n+1}}{1 - x} = f(x) - \frac{x^{n+1}}{1 - x}$$

provided that $x \neq 1$. We therefore conclude that

$$|f(x) - P_n(x)| = \left|\frac{x^{n+1}}{1 - x}\right|$$

We can interpret the term $x^{n+1}/(1 - x)$ as the error of approximation. Since

$$\lim_{n\to\infty} \left|\frac{x^{n+1}}{1 - x}\right| = \begin{cases} \infty & \text{if } |x| > 1 \\ 0 & \text{if } |x| < 1 \end{cases}$$

it follows that the error of approximation can be made small only when $|x| < 1$. For $|x| > 1$, the error of approximation increases with increasing n. [When $x = 1$, the function $f(x)$ is not defined.]

In general, it is not straightforward to obtain error estimates. In its general form, the error is given as an integral, which is why we wanted until this chapter to introduce Taylor series. Let's first look at the error terms $P_0(x)$ and $P_1(x)$ before stating the error term for arbitrary n.

Using part II of the Fundamental Theorem of Calculus, we find that

$$\int_a^x f'(t)\,dt = f(x) - f(a)$$

or by rearranging:

$$f(x) = f(a) + \int_a^x f'(t)\,dt.$$

Since $f(a) = P_0(x)$, we can interpret $\int_a^x f'(t)\,dt$ as the error term in the Taylor approximation of $f(x)$ about $x = a$ when $n = 0$.

We can use integration by parts to obtain the next-higher approximation:

$$f(x) - f(a) = \int_a^x f'(t)\,dt = \int_a^x 1 \cdot f'(t)\,dt$$

We set $u' = 1$ with $u = -(x - t)$ and $v = f'(t)$ with $v' = f''(t)$. [Writing $u = -(x - t)$ turns out to be a more convenient antiderivative of $u' = 1$ than $u = t$, as you will see shortly.] We obtain

$$\begin{aligned} f(x) - f(a) = \int_a^x f'(t)\,dt &= -(x - t)f'(t)]_a^x + \int_a^x (x - t)f''(t)\,dt \\ &= (x - a)f'(a) + \int_a^x (x - t)f''(t)\,dt \end{aligned}$$

That is,

$$f(x) = f(a) + f'(a)(x - a) + \int_a^x (x - t)f''(t)\,dt$$

The expression $f(a)+f'(a)(x-a)$ is the linear approximation (i.e. $P_1(x)$); the integral can then be considered as the error term.

Continuing in this way, we find the general formula:

Taylor's Formula Suppose that $f : I \to \mathbf{R}$, where I is an interval, $a \in I$, and f and its first $n+1$ derivatives are continuous at $a \in I$. Then, for $x, a \in I$,

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n+1}(x)$$

where $R_{n+1}(x)$ represents the error in approximating $f(x)$ by $P_n(x)$. It is equal to

$$R_{n+1}(x) = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\, dt$$

We will now examine the error term in Taylor's formula more closely. The error term is given in integral form, and it is often very difficult to evaluate the integral. We will first look at the case $n = 0$; that is, we approximate $f(x)$ by the constant function $f(a)$. The error term is then $R_{n+1}(x)$ when $n = 0$; that is,

$$R_1(x) = \int_a^x f'(t)\, dt$$

If we set

$$K = \begin{bmatrix} \text{largest value of } |f'(t)| \\ \text{for } t \text{ between } a \text{ and } x \end{bmatrix}$$

then because $|f'(t)| \le K$, we can estimate $R_1(x)$ using the comparison results we derived to estimate integrals in Section 6.1.4.

$$|R_1(x)| \le K|x-a| \qquad |\textstyle\int_a^x K\,dt| = K|x-a|$$

Before we give the corresponding results for $R_{n+1}(x)$, we look at one example that illustrates how to find K.

EXAMPLE 6 Estimate the error in the approximation of $f(x) = e^x$ by $P_0(x)$ about $x = 0$ on the interval $[0, 1]$.

Solution Since $f(0) = 1$, it follows that $P_0(x) = 1$ so:

$$f(x) = 1 + R_1(x)$$

where

$$R_1(x) = \int_0^x f'(t)\, dt = x f'(c)$$

for some c between 0 and x. Because $f'(t) = e^t$, the largest value of $|f'(t)|$ in the interval $[0, 1]$, namely, $|f'(1)| = e$, occurs when $t = 1$. Since we want to find an approximation of $f(x) = e^x$ for $x \in [0, 1]$, we should not use e in our error estimate, as e is one of the values we want to estimate. Instead, we use $|f'(t)| \le 3$ for $t \in [0, 1]$. We thus have

$$|R_1(x)| \le 3x \quad \text{for } x \in [0, 1]$$

Again for general n, although we may typically be unable to calculate $R_{n+1}(x)$ exactly, we can bound it using the comparison results that we used to estimate integrals in Section 6.1.4.

Let K be the largest value of $|f^{(n+1)}(t)|$ for $a \le t \le x$. Then the largest value taken by the integrand in the term R_{n+1} is:

$$|\overbrace{(x-t)^n}^{|(x-t)^n| \le |x-a|^n} \cdot \overbrace{f^{(n+1)}(t)}^{|f^{(n+1)}(t)| \le K}| \le |x-a|^n K$$

That is:

$$-|x-a|^n K \le (x-t)^n f^{(n+1)}(t) \le |x-a|^n K$$

so the integrand may be bounded above and below by constants that are independent of t, and we can use these bounds to bound the integral also.

Namely:

$$|R_{n+1}(x)| \le \frac{K|x-a|^{n+1}}{(n+1)!}$$

$\int_a^x K|x-a|^n\,dx = K|x-a|^{n+1}$ assuming $x > a$.

We will use this inequality in the next example to determine in advance what degree of Taylor polynomial will allow us to achieve a given accuracy.

EXAMPLE 7 Suppose that $f(x) = e^x$. What degree of Taylor polynomial about $x = 0$ will allow us to approximate $f(1)$ so that the error is less than 10^{-5}?

Solution In Example 1, we found that, for any $n \ge 1$,

$$f^{(n+1)}(t) = e^t$$

We need to find out how large $f^{(n+1)}(t)$ can get for $t \in [0, 1]$. We obtain

$$|f^{(n+1)}(t)| = e^t \le e \quad \text{for } 0 \le t \le 1$$

As in Example 6, instead of using e as a bound, we use a slightly larger value, namely, 3. Therefore,

$$|R_{n+1}(1)| \le \frac{3|1|^{n+1}}{(n+1)!} = \frac{3}{(n+1)!} \tag{7.41}$$

We want the error to be less than 10^{-5}; that is, we want

$$|R_{n+1}(1)| < 10^{-5}$$

Inserting different values of n shows that

$$\frac{3}{8!} = 7.44 \times 10^{-5} \qquad \text{and} \qquad \frac{3}{9!} = 8.27 \times 10^{-6}$$

That is, when $n = 8$,

$$|R_{n+1}(1)| = |R_9(1)| \le 8.27 \times 10^{-6} < 10^{-5}$$

Because the estimate of the error is greater than 10^{-5} when $n = 7$, we conclude that a polynomial of degree 8 would certainly give us the desired accuracy, whereas a polynomial of degree 7 might not. We can easily check this; we find that

$$1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots + \frac{1}{7!} = 2.71825396825$$

$$1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots + \frac{1}{8!} = 2.71827876984$$

Comparing these with $e = 2.71828182845\ldots$, we see that the error is equal to 2.79×10^{-5} when $n = 7$ and 3.06×10^{-6} when $n = 8$. The error that we computed for Taylor's formula is a worst-case scenario; that is, the true error can be (and typically is) smaller than the error bound.

We have already seen one example in which a Taylor polynomial was useful only for values close to the point at which we approximated the function, regardless of n, the degree of the polynomial. In some situations, the error in the approximation cannot be made small for *any* value close to the point of approximation, regardless of n. One such example is the continuous function

$$f(x) = \begin{cases} e^{-1/x} & \text{for } x > 0 \\ 0 & \text{for } x \le 0 \end{cases}$$

which is used, for instance, to describe the height of a tree as a function of age. We can show that $f^{(k)}(0) = 0$ for *all* $k \geq 1$, which implies that a Taylor polynomial of degree n about $x = 0$ is

$$P_n(x) = 0$$

for all n. This example clearly shows that it will not help to increase n; the approximation just will not improve.

When we use Taylor polynomials to approximate functions, it is important to know for which values of x the approximation can be made arbitrarily close by choosing n large.

Following are a few of the most important functions, together with their Taylor polynomials about $x = 0$ and the range of x values over which the approximation can be made arbitrarily close by choosing n large enough:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + R_{n+1}(x), \qquad -\infty < x < \infty$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + R_{n+1}(x), \quad -\infty < x < \infty$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + R_{n+1}(x), \qquad -\infty < x < \infty$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \cdots + (-1)^{n+1}\frac{x^n}{n} + R_{n+1}(x), \qquad -1 < x \leq 1$$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \cdots + x^n + R_{n+1}(x), \qquad -1 < x < 1$$

Section 7.6 Problems

7.6.1

In Problems 1–5, find the linear approximation of $f(x)$ at $x = 0$.

1. $f(x) = e^{x+1}$

2. $f(x) = \sin(x+1)$

3. $f(x) = \dfrac{1}{1+x}$

4. $f(x) = x^4$

5. $f(x) = \tan x$.

In Problems 6–10, compute the Taylor polynomial of degree n about $x = 0$ for each function.

6. $f(x) = \dfrac{1}{1+x}, n = 4$

7. $f(x) = (1+x)^3, n = 5$

8. $f(x) = e^{-x}, n = 3$

9. $f(x) = x^5, n = 6$

10. $f(x) = \sqrt{1+x}, n = 3$

In Problems 11–16, compute the Taylor polynomial of degree n about $x = 0$ for each function and compare the value of the function at the indicated point with the value of the corresponding Taylor polynomial.

11. $f(x) = \sqrt{1+x}, n = 3, x = 0.1$

12. $f(x) = \dfrac{1}{1+x}, n = 3, x = 0.1$

13. $f(x) = \sin x, n = 5, x = 1$

14. $f(x) = e^{2x}, n = 4, x = 0.3$

15. $f(x) = \tan x, n = 2, x = 0.1$

16. $f(x) = \ln(1+x), n = 3, x = 0.1$

17. (a) Find the Taylor polynomial of degree 3 about $x = 0$ for $f(x) = \sin x$.

(b) Use your result in (a) to explain why

$$\lim_{x \to 0} \frac{\sin x}{x} = 1$$

18. (a) Find the Taylor polynomial of degree 2 about $x = 0$ for $f(x) = \cos x$.

(b) Use your result in (a) to explain why

$$\lim_{x \to 0} \frac{\cos x - 1}{x} = 0$$

7.6.2

In Problems 19–23, compute the Taylor polynomial of degree n about a and compare the value of the approximation with the value of the function at the given point x.

19. $f(x) = \sqrt{x}, a = 1, n = 3; x = 2$

20. $f(x) = \ln x, a = 1, n = 3; x = 2$

21. $f(x) = \cos x, a = \dfrac{\pi}{2}, n = 3; x = \dfrac{\pi}{3}$

22. $f(x) = x^{1/5}, a = 1, n = 3; x = 0.9$

23. $f(x) = e^x, a = 2, n = 3; x = 2.1$

24. Show that

$$x^4 \approx a^4 + 4a^3(x-a)$$

for x close to a.

25. Show that, for positive constants r and K,

$$rN\left(1 - \frac{N}{K}\right) \approx rN$$

for N close to 0.

26. **(a)** Show that, for positive constants a and k,

$$f(R) = \frac{aR}{k+R} \approx \frac{a}{k}R$$

for R close to 0.

(b) Show that, for positive constants a and k,

$$f(R) = \frac{aR}{k+R} \approx \frac{a}{2} + \frac{a}{4k}(R-k)$$

for R close to k.

7.6.3

In Problems 27–30, use the following form of the error term

$$|R_{n+1}(x)| \le \frac{K|x|^{n+1}}{(n+1)!}$$

where K = largest value of $|f^{(n+1)}(t)|$ for $0 \le t \le x$, to determine in advance the degree of Taylor polynomial at $a = 0$ that would achieve the indicated accuracy in the interval $[0, x]$. (Do not compute the Taylor polynomial.)

27. $f(x) = e^x$, $x = 2$, error $< 10^{-3}$

28. $f(x) = \cos x$, $x = 1$, error $< 10^{-2}$

29. $f(x) = 1/(1+x)$, $x = 0.2$, error $< 10^{-2}$

30. $f(x) = \ln(1+x)$, $x = 0.1$, error $< 10^{-2}$

31. It can be shown that the Taylor polynomial for $f(x) = (1+x)^\alpha$ about $x = 0$, with α a positive constant, converges for $x \in (-1, 1)$. Show that

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \cdots + R_{n+1}(x)$$

32. We can show that the Taylor polynomial for $f(x) = \tan^{-1} x$ about $x = 0$ converges for $|x| \le 1$.

(a) Show that the following is true:

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + R_{n+1}(x)$$

(b) Explain why the following holds:

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

(This series converges very slowly, as you would see if you used it to approximate π.)

7.7 Tables of Integrals

At one time, tables of indefinite integrals were useful aids for evaluating integrals. In using a table of integrals, it is still necessary to bring the integrand of interest into a form that is listed in the table—and there are also many integrals that simply cannot be evaluated exactly and must be evaluated numerically. However, mathematical software like Matlab, Mathematica, or Maple now can take up much of the work of evaluating these integrals, making these tables less useful to students today. Nevertheless, it is often helpful to know the antiderivative of a particular function, so we will give a very brief list of indefinite integrals and explain how to use such tables.

I. Basic Functions.

1. $\displaystyle\int x^n dx = \frac{1}{n+1}x^{n+1} + C, n \neq -1$

2. $\displaystyle\int \frac{1}{x}\, dx = \ln|x| + C$

3. $\displaystyle\int e^x\, dx = e^x + C$

4. $\displaystyle\int a^x\, dx = \frac{a^x}{\ln a} + C$ with $a > 0, a \neq 1$

5. $\displaystyle\int \ln x\, dx = x \ln x - x + C$

6. $\displaystyle\int \sin x\, dx = -\cos x + C$

7. $\displaystyle\int \cos x\, dx = \sin x + C$

8. $\displaystyle\int \tan x\, dx = -\ln|\cos x| + C$

II. Rational Functions.

9. $\displaystyle\int \frac{1}{ax+b}\, dx = \frac{1}{a}\ln|ax+b| + C$

10. $\displaystyle\int \frac{x}{ax+b}\, dx = \frac{x}{a} - \frac{b}{a^2}\ln|ax+b| + C$

11. $\displaystyle\int \frac{x}{(ax+b)^2}\,dx = \frac{b}{a^2(ax+b)} + \frac{1}{a^2}\ln|ax+b| + C$

12. $\displaystyle\int \frac{x}{ax^2+bx+c}\,dx = \frac{1}{2a}\ln|ax^2+bx+c| - \frac{b}{2a}\int \frac{1}{ax^2+bx+c}\,dx$

13. $\displaystyle\int \frac{1}{a^2+x^2}\,dx = \frac{1}{a}\arctan\frac{x}{a} + C$

14. $\displaystyle\int \frac{1}{a^2-x^2}\,dx = \frac{1}{2a}\ln\left|\frac{x+a}{x-a}\right| + C$

III. Integrands Involving $\sqrt{a^2+x^2}$, $\sqrt{a^2-x^2}$, or $\sqrt{x^2-a^2}$.

15. $\displaystyle\int \frac{1}{\sqrt{a^2-x^2}}\,dx = \arcsin\frac{x}{a} + C$

16. $\displaystyle\int \frac{1}{\sqrt{x^2\pm a^2}}\,dx = \ln|x+\sqrt{x^2\pm a^2}| + C$

17. $\displaystyle\int \sqrt{a^2\pm x^2}\,dx = \frac{1}{2}\left(x\sqrt{a^2\pm x^2} + a^2\int \frac{1}{\sqrt{a^2\pm x^2}}\,dx\right)$

18. $\displaystyle\int \sqrt{x^2-a^2}\,dx = \frac{1}{2}\left(x\sqrt{x^2-a^2} - a^2\int \frac{1}{\sqrt{x^2-a^2}}\,dx\right)$

IV. Integrands Involving Trigonometric Functions.

19. $\displaystyle\int \sin(ax)\,dx = -\frac{1}{a}\cos(ax) + C$

20. $\displaystyle\int \sin^2(ax)\,dx = \frac{1}{2}x - \frac{1}{4a}\sin(2ax) + C$

21. $\displaystyle\int \sin(ax)\sin(bx)\,dx = \frac{\sin(a-b)x}{2(a-b)} - \frac{\sin(a+b)x}{2(a+b)} + C$, for $a^2 \neq b^2$

22. $\displaystyle\int \cos(ax)\,dx = \frac{1}{a}\sin(ax) + C$

23. $\displaystyle\int \cos^2(ax)\,dx = \frac{1}{2}x + \frac{1}{4a}\sin(2ax) + C$

24. $\displaystyle\int \cos(ax)\cos(bx)\,dx = \frac{\sin(a-b)x}{2(a-b)} + \frac{\sin(a+b)x}{2(a+b)} + C$, for $a^2 \neq b^2$

25. $\displaystyle\int \sin(ax)\cos(ax)\,dx = \frac{1}{2a}\sin^2(ax) + C$

26. $\displaystyle\int \sin(ax)\cos(bx)\,dx = -\frac{\cos(a+b)x}{2(a+b)} - \frac{\cos(a-b)x}{2(a-b)} + C$, for $a^2 \neq b^2$

V. Integrands Involving Exponential Functions.

27. $\displaystyle\int e^{ax}\,dx = \frac{1}{a}e^{ax} + C$

28. $\displaystyle\int xe^{ax}\,dx = \frac{e^{ax}}{a^2}(ax-1) + C$

29. $\displaystyle\int x^n e^{ax}\,dx = \frac{1}{a}x^n e^{ax} - \frac{n}{a}\int x^{n-1}e^{ax}\,dx$

30. $\displaystyle\int e^{ax}\sin(bx)\,dx = \frac{e^{ax}}{a^2+b^2}(a\sin(bx) - b\cos(bx)) + C$

31. $\displaystyle\int e^{ax}\cos(bx)\,dx = \frac{e^{ax}}{a^2+b^2}(a\cos(bx) + b\sin(bx)) + C$

VI. Integrands Involving Logarithmic Functions.

32. $\displaystyle\int \ln x\,dx = x\ln x - x + C$

33. $\displaystyle\int (\ln x)^2\,dx = x(\ln x)^2 - 2x\ln x + 2x + C$

34. $\displaystyle\int x^m \ln x\,dx = x^{m+1}\left[\frac{\ln x}{m+1} - \frac{1}{(m+1)^2}\right] + C,\ m \neq -1$

35. $\displaystyle\int \frac{\ln x}{x}\,dx = \frac{(\ln x)^2}{2} + C$

36. $\displaystyle\int \frac{1}{x\ln x}\,dx = \ln(\ln x) + C$

37. $\displaystyle\int \sin(\ln x)\,dx = \frac{x}{2}(\sin(\ln x) - \cos(\ln x)) + C$

38. $\displaystyle\int \cos(\ln x)\,dx = \frac{x}{2}(\sin(\ln x) + \cos(\ln x)) + C$

We will now illustrate how to use the preceding table. We begin with examples that fit one of the listed integrals exactly.

EXAMPLE 1 **Square Root** Find $\int \sqrt{3 - x^2}\,dx$.

Solution The integrand involves $\sqrt{a^2 - x^2}$ and is of the form III.17 with $a^2 = 3$. Hence,

$$\int \sqrt{3 - x^2}\,dx = \frac{1}{2}\left(x\sqrt{3 - x^2} + 3\int \frac{1}{\sqrt{3 - x^2}}\,dx\right)$$

To evaluate $\int \frac{1}{\sqrt{3-x^2}}\,dx$, we use III.15 with $a^2 = 3$ and find that

$$\int \frac{1}{\sqrt{3 - x^2}}\,dx = \arcsin\frac{x}{\sqrt{3}} + C$$

Thus,

$$\int \sqrt{3 - x^2}\,dx = \frac{1}{2}\left(x\sqrt{3 - x^2} + 3\arcsin\frac{x}{\sqrt{3}}\right) + C$$

•

EXAMPLE 2 **Trigonometric Function** Find $\int \sin(3x)\cos(4x)\,dx$.

Solution The integrand involves trigonometric functions, and we can find it in IV.26 with $a = 3$ and $b = 4$. Hence,

$$\int \sin(3x)\cos(4x)\,dx = -\frac{\cos(7x)}{14} - \frac{\cos(-x)}{(2)(-1)} + C$$

Since $\cos(-x) = \cos x$, this simplifies to

$$\int \sin(3x)\cos(4x)\,dx = -\frac{\cos(7x)}{14} + \frac{\cos x}{2} + C$$

•

EXAMPLE 3 **Exponential Function** Find $\int x^2 e^{3x}\,dx$.

Solution This integrand is of the form V.29 with $n = 2$ and $a = 3$. Hence,

$$\int x^2 e^{3x}\,dx = \frac{1}{3}x^2 e^{3x} - \frac{2}{3}\int x e^{3x}\,dx$$

We now use V.28 to continue the evaluation of the integral and find that

$$\int x e^{3x}\,dx = \frac{e^{3x}}{9}(3x - 1) + C$$

Thus,

$$\int x^2 e^{3x}\,dx = \frac{1}{3}x^2 e^{3x} - \frac{2}{3}\left[\frac{e^{3x}}{9}(3x - 1)\right] + C$$

•

Thus far, each of our examples exactly matched one of the integrals in our table. Often, this will not be the case, and the integrand must be manipulated until it matches one of the integrals in the table. Among the manipulations that are used are expansions, long division, completion of the square, and substitution. We give a few examples to illustrate.

EXAMPLE 4 **Exponential Function** Find $\int e^{2x}\sin(3x-1)\,dx$.

Solution This integrand looks similar to V.30. If we use the substitution

$$u = 3x - 1 \quad \text{with } dx = \frac{1}{3}\,du \text{ and } 2x = \frac{2}{3}(u+1)$$

then the integrand can be transformed so that it matches V.30 exactly, and we have

$$\begin{aligned}
\int e^{2x}\sin(3x-1)\,dx &= \int e^{2(u+1)/3}(\sin u)\frac{1}{3}\,du \\
&= \frac{e^{2/3}}{3}\int e^{2u/3}\sin u\,du \\
&= \frac{e^{2/3}}{3}\frac{e^{2u/3}}{\frac{4}{9}+1}\left[\frac{2}{3}\sin u - \cos u\right] + C \\
&= \frac{e^{2/3}}{3\cdot\frac{13}{9}}e^{2(3x-1)/3}\left[\frac{2}{3}\sin(3x-1) - \cos(3x-1)\right] + C \\
&= \frac{3}{13}e^{2x}\left[\frac{2}{3}\sin(3x-1) - \cos(3x-1)\right] + C
\end{aligned}$$

•

EXAMPLE 5 **Rational Function** Find $\displaystyle\int \frac{x^2}{9+x^2}\,dx$.

Solution The integrand is a rational function; we can use long division to simplify it:

$$\frac{x^2}{9+x^2} = 1 - \frac{9}{9+x^2}$$

Then, using II.13 with $a = 3$, we obtain

$$\begin{aligned}
\int \frac{x^2}{9+x^2}\,dx &= \int dx - 9\int \frac{1}{9+x^2}\,dx \\
&= x - 9\left(\frac{1}{3}\arctan\frac{x}{3}\right) + C \\
&= x - 3\arctan\frac{x}{3} + C
\end{aligned}$$

•

EXAMPLE 6 **Rational Function** Find $\displaystyle\int \frac{1}{x^2-2x-3}\,dx$.

Solution The first step is to complete the square in the denominator:

$$\begin{aligned}
\frac{1}{x^2-2x-3} &= \frac{1}{(x^2-2x+1)-1-3} \\
&= \frac{1}{(x-1)^2-4}
\end{aligned}$$

Then, using the substitution $u = x - 1$ with $du = dx$, we find that

$$\int \frac{dx}{(x-1)^2-4} = \int \frac{du}{u^2-4} = -\int \frac{du}{4-u^2}$$

which is of the form II.14 with $a = 2$. Therefore,

$$\begin{aligned}
\int \frac{1}{x^2-2x-3}\,dx &= -\int \frac{du}{4-u^2} = -\frac{1}{4}\ln\left|\frac{u+2}{u-2}\right| + C \\
&= -\frac{1}{4}\ln\left|\frac{x+1}{x-3}\right| + C
\end{aligned}$$

•

Section 7.7 Problems

In Problems 1–8, use the Table of Integrals to compute each integral.

1. $\displaystyle\int \frac{x}{2x+3}\,dx$ **2.** $\displaystyle\int \frac{dx}{4+x^2}$

3. $\displaystyle\int \sqrt{x^2+16}\,dx$ **4.** $\displaystyle\int \sin(2x)\cos(2x)\,dx$

5. $\displaystyle\int_0^1 x^3 e^{2x}\,dx$ **6.** $\displaystyle\int_0^{\pi/2} e^{-x}\cos(x)\,dx$

7. $\displaystyle\int_1^e x^2\ln x\,dx$ **8.** $\displaystyle\int_e^{e^2} \frac{\ln x}{x}dx$

In Problems 9–22, use the Table of Integrals to compute each integral after manipulating the integrand in a suitable way.

9. $\displaystyle\int_0^{\pi/2} e^x\cos\left(x-\frac{\pi}{6}\right)\,dx$ **10.** $\displaystyle\int_1^2 x\ln(x+3)\,dx$

11. $\displaystyle\int (x^2-1)e^{-x/2}\,dx$ **12.** $\displaystyle\int (x+1)^2 e^{-2x}\,dx$

13. $\displaystyle\int \cos^2(5x-3)\,dx$ **14.** $\displaystyle\int \frac{x^2}{x^2+4x+1}\,dx$

15. $\displaystyle\int \sqrt{x^2+2x+2}\,dx$ **16.** $\displaystyle\int \frac{1}{\sqrt{16-9x^2}}\,dx$

17. $\displaystyle\int e^{2x+1}\sin\left(\frac{\pi}{2}x\right)\,dx$ **18.** $\displaystyle\int (x-1)^2 e^{2x}\,dx$

19. $\displaystyle\int_2^4 \frac{\ln\sqrt{x}}{x}\,dx$ **20.** $\displaystyle\int_1^e (x+2)^2\ln x\,dx$

21. $\displaystyle\int \sin(\ln(3x))\,dx$ **22.** $\displaystyle\int \frac{3}{x^2-4x+5}\,dx$

Chapter 7 Review

Key Terms

Discuss the following definitions and concepts:

1. The substitution rule for indefinite integrals
2. The substitution rule for definite integrals
3. Integration by parts
4. The "trick" of "multiplying by 1"
5. Partial-fraction decomposition
6. Proper rational function
7. Long division of polynomials
8. Irreducible quadratic factor
9. Improper integral
10. Integration when the interval is unbounded
11. Integration when the integrand is discontinuous
12. Convergence and divergence of improper integrals
13. Comparison results for improper integrals
14. Numerical integration: midpoint and trapezoidal rule
15. Using a spreadsheet to numerically estimate an integral
16. Error bounds for the midpoint and the trapezoidal rule
17. Linear approximation
18. Taylor polynomial of degree n
19. Taylor's formula
20. Using tables of integrals for integration

Review Problems

In Problems 1–30, evaluate each indefinite integral.

1. $\displaystyle\int 2x^2(1+x^3)^2\,dx$ **2.** $\displaystyle\int \frac{\cos x}{1+\sin^2 x}\,dx$

3. $\displaystyle\int xe^{-x}\,dx$ **4.** $\displaystyle\int \frac{x\ln(1+x^2)}{1+x^2}\,dx$

5. $\displaystyle\int (1+\sqrt{x})^{1/3}\,dx$ **6.** $\displaystyle\int x\sqrt{x+3}\,dx$

7. $\displaystyle\int x\cos x\,dx$ **8.** $\displaystyle\int \tan x\sec^2 x\,dx$

9. $\displaystyle\int x\ln x\,dx$ **10.** $\displaystyle\int x\ln(1+x^2)\,dx$

11. $\displaystyle\int \cos x\exp(\sin x)\,dx$ **12.** $\displaystyle\int \sqrt{x}\ln\sqrt{x}\,dx$

13. $\displaystyle\int \frac{1}{9-x^2}\,dx$ **14.** $\displaystyle\int \frac{1}{9+x^2}\,dx$

15. $\displaystyle\int \tan x\,dx$ **16.** $\displaystyle\int \tan^{-1} x\,dx$

17. $\displaystyle\int e^x\sin 2x\,dx$ **18.** $\displaystyle\int x\sin(x^2+1)\,dx$

19. $\displaystyle\int x^3\sin(x^2+1)\,dx$ **20.** $\displaystyle\int e^{2x}\ln(e^x+1)\,dx$

21. $\displaystyle\int \sin^2 x\,dx$ **22.** $\displaystyle\int \sin x\cos x\,dx$

23. $\displaystyle\int \frac{1}{x(x-1)}\,dx$ **24.** $\displaystyle\int \frac{1}{(x+1)(x-2)}\,dx$

25. $\displaystyle\int \frac{x}{x+2}\,dx$ **26.** $\displaystyle\int \frac{1}{x+5}\,dx$

27. $\displaystyle\int \frac{x}{x^2+4}\,dx$ **28.** $\displaystyle\int \frac{1}{x^2+4}\,dx$

29. $\displaystyle\int \frac{(x+1)^2}{x-1}\,dx$ **30.** $\displaystyle\int \frac{x+1}{x^2-4}dx$

In Problems 31–42, evaluate each definite integral.

31. $\int_0^1 x\sqrt{x^2+1}\,dx$ **32.** $\int_0^{\pi} x\sin x\,dx$

33. $\int_0^1 xe^{-x^2/2}\,dx$ **34.** $\int_1^2 \ln x\,dx$

35. $\int_0^2 \frac{1}{4+x^2}\,dx$ **36.** $\int_0^1 \frac{1}{x^2+2x+1}\,dx$

37. $\int_2^6 \frac{1}{\sqrt{x-2}}\,dx$ **38.** $\int_0^2 \frac{1}{x^2+5x+6}\,dx$

39. $\int_0^1 x\ln x\,dx$ **40.** $\int_0^1 \frac{x}{x^2-2x-3}\,dx$

41. $\int_0^{\pi/4} e^{\cos x}\sin x\,dx$ **42.** $\int_0^{\pi/4} x\sin(2x)\,dx$

In Problems 43–50, calculate each improper integral. (Some integrals may be divergent.)

43. $\int_0^{\infty} \frac{1}{9+x^2}\,dx$ **44.** $\int_0^{\infty} \frac{1}{x^2+3}\,dx$

45. $\int_0^{\infty} \frac{1}{x+3}\,dx$ **46.** $\int_{-1}^{0} \frac{1}{x^2+5x+6}\,dx$

47. $\int_0^1 \frac{1}{x^2}\,dx$ **48.** $\int_1^{\infty} \frac{1}{x^2}\,dx$

49. $\int_0^1 \frac{1}{\sqrt{x}}\,dx$ **50.** $\int_1^{\infty} \frac{1}{\sqrt{x}}\,dx$

In Problems 51–54, use (a) the midpoint rule and (b) the trapezoidal rule to approximate each integral with the specified value of n.

51. $\int_2^4 (x^2-4)^{1/3}\,dx, n=4$ **52.** $\int_1^2 \sqrt{(x^3-1)}\,dx, n=4$

53. $\int_0^1 \exp(-x^2)\,dx, n=5$ **54.** $\int_0^1 \sin(x^2+1)\,dx, n=4$

In Problems 55–58, find the Taylor polynomial of degree n about x = a for each function.

55. $f(x)=\sin(2x), a=0, n=3$

56. $f(x)=e^{-x^2/2}, a=0, n=3$

57. $f(x)=\ln x, a=1, n=3$

58. $f(x)=\frac{1}{x-3}, a=4, n=4$

59. Random Events: Swimming Bacterium An important tool when studying random events is the normal distribution, which we shall study in Chapter 12. The normal distribution is used to describe random processes like diffusion. Unless they are following chemical cues freely swimming bacteria follow random paths. If a bacterium is released at a point $x = 0$, then some time later the probability that it is at a point x may be proportional to:

$$f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right);$$

$x < 0$ corresponds to swimming leftward and $x > 0$ corresponds to swimming rightward. You may assume (it can be proven, but not using the techniques in this chapter) that $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

(a) An important quantity to study is the average distance that the bacterium travels from the origin. In Chapter 12 we will show that the average distance of travel is $\bar{x} = \int_{-\infty}^{\infty} x f(x)\,dx$.

Show that $\bar{x} = 0$.

(b) Your answer from (a) can be intuitively interpreted that the bacterium is equally likely to swim leftward as rightward, so on average its displacement is 0. It may make more sense, then, to measure the total distance that the bacterium travels. This is given on average by

$$d = \int_{-\infty}^{\infty} |x| f(x)\,dx.$$

Explain why

$$d = \frac{1}{\sqrt{2\pi}}\int_0^{\infty} xe^{-x^2/2}\,dx - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} xe^{-x^2/2}\,dx$$

and calculate d.

(c) The main use of $f(x)$, we shall learn in Chapter 12, is to calculate the total probability of finding the bacterium within a certain interval of distances from its starting position. For example, the probability the bacterium is somewhere between $x = -1$ and $x = 1$ is:

$$p = \int_{-1}^{1} f(x)\,dx$$

This integral cannot be evaluated analytically; however, using the trapezoidal rule with $n = 5$ subintervals, estimate the probability p.

60. Random Events: Modeling Forest Fires One model that is often used to model the time elapsed between forest fires is the gamma distribution.

According to the Gamma distribution, the probability that a time t elapses between the end of one forest fire and the start of another forest fire is proportional to

$$f(t) = t^a\exp(-bt)$$

In this question we will assume for definiteness that $a = 1$ and $b = 1$ (in practice, a and b are coefficients that can be used to fit the model to different forests and climates).

(a) Show that the improper integral $\int_0^{\infty} f(t)\,dt$ is convergent and calculate its value.

(b) It can be derived using the methods in Chapter 12 that the average time between forest fires is given by:

$$\bar{t} = \frac{\int_0^{\infty} t f(t)\,dt}{\int_0^{\infty} f(t)\,dt}.$$

Calculate $\bar{t}$.

(c) One useful calculation that can be performed with $f(t)$ is to estimate the probability that the next forest fire will occur by a certain time t, e.g., $t = 2$. It can be shown that this probability is given by:

$$p = \frac{\int_0^{2} f(t)\,dt}{\int_0^{\infty} f(t)\,dt}$$

Calculate p.

61. Cost of Gene Substitution *(Adapted from* ***Roughgarden, 1996****)* Over time populations continuously adapt to their environment. One term of adaptation occurs when a new gene emerges by mutation. If this gene makes the individuals who carry it "fitter" or better adapted to their environment, then the gene may spread through the population, as the descendants of those individuals outcompete or breed with individuals that do not have the gene. As the gene spreads through the population, the average fitness of the population increases. We denote by $f_{\text{avg}}(t)$ the average fitness of the population at time t, by $f_{\text{avg}}(0)$ the average fitness of the population at time 0 (when the mutation arose), and by K the final value of the average fitness after the mutation has spread through the population. ***Haldane (1957)*** suggested measuring the *cost of evolution* (also called the *cost of gene substitution*) by the cumulative difference between the current and the final fitness—that is, by

$$\int_0^\infty (K - f_{\text{avg}}(t))\, dt$$

In Figure 7.52, shade the region whose area is equal to the cost of gene substitution.

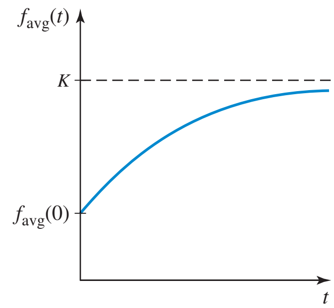
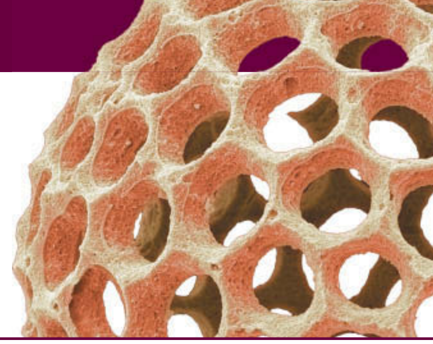


Figure 7.52 The cost of gene substitution. See Problem 61.



CHAPTER

8

Differential Equations

The focus of this chapter is on solving and analyzing differential equations. Specifically, we will learn how to

- use the method of separation of variables to solve separable differential equations;
- find equilibria and determine their stability graphically and analytically;
- describe the behavior of solutions of differential equations, starting from different initial conditions;
- derive differential equations to model many different biological systems;
- use the method of integrating factors to solve linear first order differential equations;
- use compartment models to analyze biological systems with multiple interacting components.

In Chapter 5 we showed how mathematical models of population growth and of the passage of medication through the human body often take the form of differential equations. For example, suppose that after t hours, the size of a population of cells is $N(t)$. If half of the cells divide every hour and if cell death can be ignored, then we may model the growth of the population by a differential equation:

$$\frac{dN}{dt} = \frac{1}{2}N \tag{8.1}$$

Previously we have shown that if the population size at time $t = 0$ is 50 (that is, $N(0) = 50$) then the function $N(t) = 50e^{t/2}$ solves (i.e., satisfies) the differential equation and matches the initial population size. But is there a way to derive the solution from the differential equation directly? In Chapter 6 we learned that integration may be thought of as being the inverse of differentiation. In Section 8.1 we will describe the most important use of this result: solving differential equations by integrating them.

We call (8.1) a **first order differential equation** because it includes the first order derivative dN/dt, but no higher order derivatives (like d^2N/dt^2, d^3N/dt^3, etc.). The techniques introduced in this chapter will enable you to solve first order differential equations. In Section 8.1 we will learn how to solve differential equations of the form:

$$\frac{dN}{dt} = f(t)g(N) \tag{8.2}$$

in which the right-hand side of the equation can be written as the product of two functions, f and g, $f(t)$ is a function of t only, and $g(N)$ is a function of N only. Equation (8.1) is an example of this kind of equation; we may set $g(N) = \frac{1}{2}N$ (a function of N) and $f(t) = 1$ (a function of t).

Using the method from Section 8.1, we will be able to solve many different mathematical models. For some types of differential equations, however, it is possible to get qualitative information about the solution (e.g., the shape of the graph of $N(t)$ against t or the limit of $N(t)$ as $t \to \infty$) without solving the equation, as we shall learn in Section 8.2. In Section 8.3 we will derive and analyze differential equation models for many different biological systems, either by solving the equations, or by analyzing them using the methods from Section 8.2.

In Section 8.4 we will go on to learn how to solve equations of the form:

$$\frac{dN}{dt} + a(t)N = b(t)$$

where $a(t)$ and $b(t)$ are functions of t only. An example of this type of equation is:

$$\frac{dN}{dt} + \frac{N}{t} = t^2 + 1$$

in which $a(t) = 1/t$ and $b(t) = t^2 + 1$. In addition to learning both of these techniques, an important skill for solving differential equations is to recognize which of the two types an equation belongs to, and in Section 8.4 you will practice identifying the types of different equations.

The method of Section 8.4 will enable us to solve a type of biological model called a two-compartment model. These models are useful to understand how matter moves through biological systems (e.g., how a cell exchanges salts and ions with its surroundings).

8.1 Solving Separable Differential Equations

Let's return to the growth model in (8.1):

$$\frac{dN}{dt} = \frac{1}{2}N, \ t \geq 0 \tag{8.3}$$

We are interested in finding a function $N(t)$ that satisfies (8.3). Such a function is called a **solution** of the differential equation. We already know that, with $N(0) = 50$

$$N(t) = 50e^{t/2}, \ t \geq 0$$

is a solution of (8.3). To confirm that the function $N(t)$ is a solution, we differentiate $N(t)$:

$$\frac{dN}{dt} = \frac{1}{2}\underbrace{50e^{t/2}}_{N(t)} = \frac{1}{2}N(t).$$

Let's recall how a differential equation like (8.3) might arise. (8.3) can be written in the form:

$$\frac{1}{N}\frac{dN}{dt} = \frac{1}{2}$$

So for this population, the per capita growth rate is constant. Put another way, a fixed fraction of organisms within the population will die in one unit of time, and a fixed fraction of organisms will reproduce in one unit of time. In many bacteria and fungi the rate of reproduction varies over the course of a day—reproduction is slowest at night when temperatures are lowest, and then the reproductive rate climbs during the day as temperatures increase again. Thus, the per capita growth rate may oscillate over time, as illustrated in Figure 8.1.

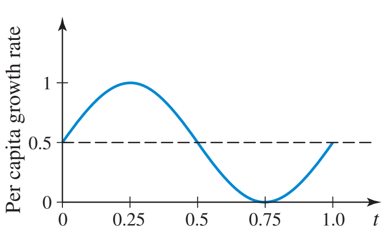


Figure 8.1 Per capita growth rate in Equation (8.4).

We can modify our original differential equation (8.3) to include these oscillations:

$$\frac{dN}{dt} = \frac{1}{2}(1 + \sin(2\pi t))\,N, \ t \geq 0 \tag{8.4}$$

In this section, we will learn how to solve differential equations like (8.4). To find the solution, we must integrate. We begin with a general method for solving **separable differential equations**; that is, equations of the form:

$$\frac{dy}{dt} = f(t)g(y) \tag{8.5}$$

where $f(t)$ is a function of t only and $g(y)$ is a function of y only.

We divide both sides of (8.5) by $g(y)$ [assuming that $g(y) \neq 0$]:

$$\frac{1}{g(y)}\frac{dy}{dt} = f(t)$$

Then, integrating both sides with respect to t, we find that

$$\int \frac{1}{g(y)}\frac{dy}{dt}\,dt = \int f(t)\,dt.$$

We can use the rule for integration by substitution to convert the left-hand side of this integration from an integral over t to an integral over y. We get:

$$\int \frac{1}{g(y)}\,dy = \int f(t)\,dt \qquad dy = \tfrac{dy}{dt}\,dt \tag{8.6}$$

Evaluating the two integrals in (8.6) and rearranging terms then gives y as a function of t.

In summary, to solve a separable differential equation, separate the variables t and y so that one side of the equation depends only on y and the other side only on t and then integrate both sides with respect to t. A good mnemonic for using this method is to treat dy/dt as if it were an ordinary ratio; then multiply both sides of the equation by dt when separating variables:

$$\frac{1}{g(y)}dy = f(t)\,dt.$$

Then we may integrate both sides to obtain Equation (8.6).

The method of separating the variables t and y works because the right-hand side of (8.5) can be separated into the product of two functions, $f(t)g(y)$, which gave this type of differential equation its name. Note that when we divided (8.5) by $g(y)$, we had to be careful, since $g(y)$ might be 0 for some values of y. We will address this problem in Subsection 8.1.2. Before solving general equations of form (8.5), we will first consider the special cases where either $g(y) = 1$ or $f(t) = 1$.

8.1.1 Pure-Time Differential Equations

If the rate of change of a function depends only on time, we call the resulting differential equation a **pure-time differential equation**. Such a differential equation is of the form

$$\frac{dy}{dt} = f(t), \;\; t \in I \tag{8.7}$$

where I is an interval and t represents time. This equation is a special case of (8.5) in which $g(y) = 1$, and Equation (8.6) can then be rewritten as

$$y = \int f(t)\,dt \qquad g(y) = 1 \Rightarrow \int \tfrac{dy}{g(y)} = \int 1\,dy = y \tag{8.8}$$

We previously solved equations like (8.7) in Section 5.10. Our derivation above shows that separation of variables produces the same answer. Namely, if $F(t)$ is an anti-derivative of $f(t)$, then (8.8) implies that:

$$y(t) = F(t) + C$$

where C is any constant. To determine C, we must, in addition to (8.7), have an initial condition on $y(t)$. If the initial condition on $y(t)$ is that $y(0) = y_0$, then we can write the solution to the initial value problem as a definite integral

$$y(t) = \int_0^t f(s)ds + y_0 \qquad \text{Since } \textstyle\int_0^0 f(s)ds = 0,\; y(0) = y_0 \tag{8.9}$$

EXAMPLE 1 Suppose that the volume $V(t)$ of a cell at time t changes according to

$$\frac{dV}{dt} = \sin t \quad \text{with } V(0) = 3$$

Find $V(t)$.

Solution We can find $V(t)$ directly from (8.9), but it is not necessary to memorize the formula; we can solve the differential equation by separating variables:

$$\int dV = \int \sin t \, dt$$

$$V = -\cos t + C$$

Applying the initial condition $V(0) = 3$ we obtain:

$$3 = -1 + C$$

$$4 = C$$

So the volume of the cell at time t is:

$$V(t) = 4 - \cos t$$

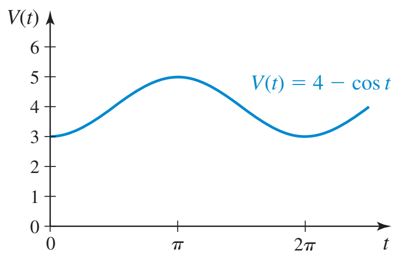


Figure 8.2 The solution $V(t) = 4 - \cos t$ in Example 1.

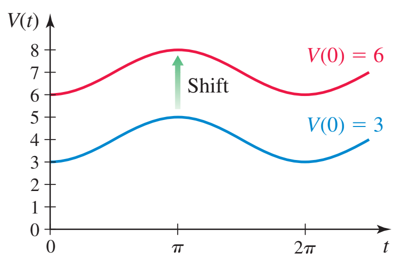


Figure 8.3 The function $V(t) = 7 - \cos t$ solves the differential equation in Example 1 with $V(0) = 6$. The solution can be obtained from the solution to Example 1 by shifting upward by 3.

See Figure 8.2 for a graph of $V(t)$. If we changed the initial condition in Example 1, the graph of the new solution could be obtained from the old solution by shifting the old solution vertically to satisfy the new initial condition. (See Figure 8.3.) •

8.1.2 Autonomous Differential Equations

A different important simplification of (8.5) comes from setting $f(t) = 1$. Then:

$$\frac{dy}{dt} = g(y) \tag{8.10}$$

These equations are called **autonomous differential equations**.

To interpret the biological meaning of *autonomous*, let's return to the growth model

$$\frac{dN}{dt} = \frac{1}{2}N \tag{8.11}$$

We will show very shortly that the general solution of (8.11) is

$$N(t) = Ce^{t/2} \tag{8.12}$$

where C is a constant that can be determined if the population size is known at one time. Suppose we conduct an experiment in which we follow a population over time, and suppose the population satisfies (8.11) with $N(0) = 20$. Using (8.12), we find that $N(0) = C = 20$. Then the size of the population at time t is given by

$$N(t) = 20e^{t/2} \tag{8.12a}$$

If we repeat the experiment at, say, time $t = 2$ with the exact same initial population size, then, everything else being equal, the population evolves in exactly the same way as the one starting at $t = 0$. Returning to (8.12) but now setting $N(2) = 20$, we find that $N(2) = Ce = 20$, or $C = 20e^{-1}$. The size of the population is then given by

$$N(t) = 20e^{-1}e^{t/2} = 20e^{(t-2)/2}. \tag{8.12b}$$

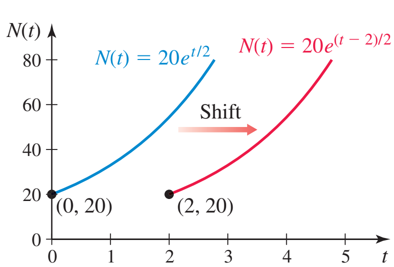


Figure 8.4 If $N(2) = 20$, then the graph of the solution $N(t) = 20e^{t/2}$ is shifted to the new starting point: $(t, N) = (2, 20)$.

The graph of this solution can be obtained from the graph of (8.12a), by shifting the old graph 2 units to the right to the new starting point $(t, N) = (2, 20)$ (see Figure 8.4).

This means that all populations starting with $N = 20$ grow in the same way, regardless of when we start the experiment. We can understand this biologically: If the growth conditions do not depend explicitly on time, the experiment should yield the same outcome regardless of when the experiment is performed.

We can solve (8.10) by separation of variables. We divide both sides of (8.10) by $g(y)$ and integrating both sides with respect to t.

$$\int \frac{1}{g(y)}\frac{dy}{dt}dt = \int dt$$

$$\Rightarrow \int \frac{dy}{g(y)} = \int dt \qquad \text{Special case of (8.6) with } f(t) = 1$$

Before we turn to biological applications, we give an example in which we see how to solve an autonomous differential equation and how to use the initial condition.

EXAMPLE 2 Solve $\frac{dy}{dx} = 2 - 3y$, where $y(1) = 1$.

Solution The independent variable is now x, but we can still separate the variables and integrate just as we did in Equation (8.6):

$$\int \frac{dy}{2-3y} = \int dx \tag{8.13}$$

We need to assume $2 - 3y \neq 0$ to divide by $2 - 3y$. We will discuss what to do if $2 - 3y = 0$ below. Since an antiderivative of $\frac{1}{2-3y}$ is $-\frac{1}{3}\ln|2-3y|$, we find that

$$-\frac{1}{3}\ln|2-3y| = x + C_1$$

We want to solve this equation for y (i.e., write y as function of x).

$$\ln|2-3y| = -3x - 3C_1$$

$$|2-3y| = e^{-3x-3C_1} \qquad \text{Exponentiate both sides}$$

$$|2-3y| = e^{-3C_1}e^{-3x} \qquad \text{Split exponential}$$

$$2-3y = \pm e^{-3C_1}e^{-3x} \qquad \text{Remove absolute value signs}$$

C_1 is an arbitrary constant, so $\pm e^{-3C_1}$ is also an arbitrary constant. We may define a new constant, $C = \pm e^{C_1}$, allowing us to write the solution in a more readable form:

$$2 - 3y = Ce^{-3x} \tag{8.14}$$

$$y = \frac{2}{3} - \frac{C}{3}e^{-3x}$$

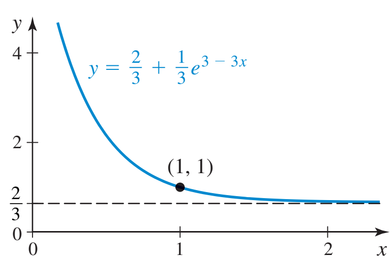


Figure 8.5 The solution to Example 2.

For any value of C the function $y = \frac{2}{3} - \frac{C}{3}e^{-3x}$ satisfies the differential equation. To determine C, we use the initial $y(1) = 1$. That is,

$$1 = \frac{2}{3} - \frac{C}{3}e^{-3}, \qquad \text{or} \qquad C = -e^3.$$

Hence,

$$y = \frac{2}{3} + \frac{1}{3}e^{3-3x}$$

(See Figure 8.5.) •

To obtain (8.13) we divided by $2-3y$, which we are only allowed to do if $2-3y \neq 0$ (i.e., if $y \neq 2/3$). Fortunately we see from the solution that $y \neq 2/3$ for all x. But if $y = 2/3$, then according to the differential equation $dy/dx = 0$, so y is a constant. In other words, if $y = 2/3$ initially, then $y(x) = 2/3$ for all x. This solution can be obtained from our general solution (8.14) by setting $C = 0$.

We now turn to two biological applications.

EXAMPLE 3 **Exponential Population Growth** We have previously modeled the growth of a population by assuming that the per capita growth rate of a population is constant

If the number of organisms is $N(t)$, then:

$$r = \frac{1}{N}\frac{dN}{dt}$$

is a constant. When $r > 0$, this model represents a growing population. When $r < 0$, the size of the population decreases. We can rearrange the terms in the formula for r into a differential equation.

$$\frac{dN}{dt} = rN. \tag{8.15}$$

To solve this differential equation we must also know the initial population size $N(0) = N_0$. (8.1) is a special case of this differential equation, in which $r = 1/2$.

We solve (8.15) by separating variables:

$$\int \frac{1}{N} dN = \int r\, dt \qquad \text{Assume } N \neq 0$$
$$\ln |N| = rt + C_1$$
$$|N| = e^{C_1} e^{rt}$$
$$N = \pm e^{C_1} e^{rt} = Ce^{rt} \qquad \text{Define a new constant } C = \pm e^{C_1}$$

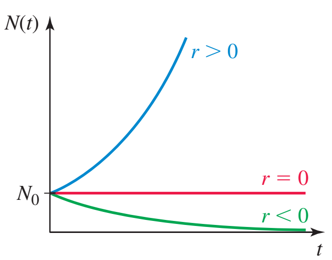


Figure 8.6 Solution curves for $dN/dt = rN$.

Our initial condition is $N(0) = N_0$. But if $N(t) = Ce^{rt}$ then $N(0) = C$, so imposing the initial condition gives $C = N_0$, or

$$N(t) = N_0 e^{rt} \tag{8.16}$$

Equation (8.16) shows that the population size grows exponentially when $r > 0$. When $r < 0$, the population size decreases exponentially. When $r = 0$, the population size stays constant.

We show solution curves of $N(t) = N_0 e^{rt}$ for $r > 0$, $r = 0$, and $r < 0$ in Figure 8.6. Exponential growth (or decay) is one of the most important growth phenomena in biology. You should therefore memorize both the differential equation (8.15) and its solution (8.16). •

In Example 3 when $r > 0$, the population size grows without bound ($\lim_{t\to\infty} N(t) = \infty$). Exponential growth cannot continue indefinitely in any real population. For example, if a small population of bacteria starts growing in a flask, then while the bacteria have plenty of resources, the population will grow exponentially. Eventually, however, the bacteria will run out of resources, and the population growth will slow. We will discuss one model for this **density dependent growth** in Example 6.

The type of growth in Example 3 is referred to as **Malthusian growth**, named after Thomas Malthus (1766–1834), a British clergyman and economist. Malthus warned about the consequences of unrestricted growth on the welfare of humans. He argued that while populations grow exponentially, food production can grow only linearly. He concluded that, since exponential growth ultimately overtakes linear growth, populations would eventually experience starvation (see Problem 62.)

Recall from Section 4.6 that, when $r < 0$, (8.15) has the same form as the differential equation that describes radioactive decay. $N(t)$ would then represent the amount of radioactive material left at time t. We will revisit this application in Problem 22.

EXAMPLE 4 **Restricted Growth: von Bertalanffy's Equation** Some species of fish show indefinite growth, that is, they continue to grow over their entire lifetime. However, the fish grow more slowly as they age. One model for fish growth is von Bertalanffy's equation, which models the length of a fish $L(t)$, at age t, using a differential equation:

$$\frac{dL}{dt} = k(L_\infty - L) \tag{8.17}$$

where k and L_∞ are both positive constants. Assuming that at age $t = 0$, the fish has length L_0, solve this initial value problem.

Solution Assuming that $L < L_\infty$, the right-hand side of this equation is positive, so $dL/dt > 0$ (the fish is growing). When or if L ever reaches L_∞, then $dL/dt = 0$, so growth will stop. The solution will depend on the constants k and L_∞, as well as the initial length, L_0. We solve the equation by separating variables.

$$\int \frac{dL}{L_\infty - L} = \int k\, dt$$

Hence,

$$-\ln|L_\infty - L| = kt + C_1$$

$$|L_\infty - L| = e^{-C_1}e^{-kt} \qquad \times(-1) \text{ and exponentiate both sides.}$$

$$L_\infty - L = Ce^{-kt} \qquad \text{Let: } C = \pm e^{-C_1}.$$

We solve for C by setting $t = 0$ on both sides

$$L_\infty - L_0 = C \qquad L(0) = L_0$$

So, substituting for C, we obtain the solution:

$$L_\infty - L(t) = (L_\infty - L_0)e^{-kt}$$

or

$$L(t) = L_\infty - (L_\infty - L_0)e^{-kt} \tag{8.18}$$

(See Figure 8.7.)

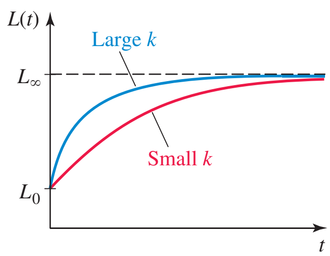


Figure 8.7 The solution of von Bertalanffy's equation.

Since $\lim_{t\to\infty} L(t) = L_\infty$ the parameter L_∞ denotes the **asymptotic length** of the fish. According to the model, if $L_0 < L_\infty$ the fish will grow over its lifetime, approaching $L = L_\infty$ asymptotically (that is, getting closer and closer to a length of L_∞ but never reaching it). k represents the rate of growth; between two fish with the same values of L_0 and L_∞, the one with the larger value of k will approach its asymptotic length quicker (see Figure 8.7).

We now consider an important type of autonomous differential equation, in which the function $g(y)$ is a quadratic polynomial.

EXAMPLE 5 Solve

$$\frac{dy}{dt} = 2(y-1)(y+2) \quad \text{with } y(0) = 2$$

Solution Separation of variables yields

$$\int \frac{dy}{(y-1)(y+2)} = \int 2\, dt. \tag{8.19}$$

We use partial fractions to integrate the left-hand side. There are (unknown) constants A and B for which

$$\frac{1}{(y-1)(y+2)} = \frac{A}{y-1} + \frac{B}{y+2} \qquad \text{for all } y$$

$$= \frac{A(y+2) + B(y-1)}{(y-1)(y+2)}$$

Comparing numerators we have: $1 = A(y+2) + B(y-1)$. As in Section 7.3 we find A and B by substituting specific values of y into this equation:

$$y = -2 \quad \Rightarrow \quad 1 = -3B$$

$$y = 1 \quad \Rightarrow \quad 1 = 3A$$

Thus, $A = \frac{1}{3}$ and $B = -\frac{1}{3}$. So:

$$\frac{1}{3}\int\left(\frac{1}{y-1} - \frac{1}{y+2}\right)dy = \int 2\,dt$$

$$\frac{1}{3}\left[\ln|y-1| - \ln|y+2|\right] = 2t + C_1$$

So

$$\ln\left|\frac{y-1}{y+2}\right| = 6t + 3C_1 \qquad \text{Combining logarithms}$$

$$\left|\frac{y-1}{y+2}\right| = e^{3C_1}e^{6t} \qquad \text{Exponentiating}$$

$$\frac{y-1}{y+2} = \pm e^{3C_1}e^{6t} \qquad \text{Removing absolute values}$$

$$\frac{y-1}{y+2} = Ce^{6t} \qquad \text{Define } C = \pm e^{3C_1}$$

We can solve for C using the initial condition:

$$\frac{1}{4} = C \qquad \text{when } t = 0,\ y = 2, \text{ so } \tfrac{y-1}{y+2} = \tfrac{1}{4}$$

The solution is therefore

$$\frac{y-1}{y+2} = \frac{1}{4}e^{6t}$$

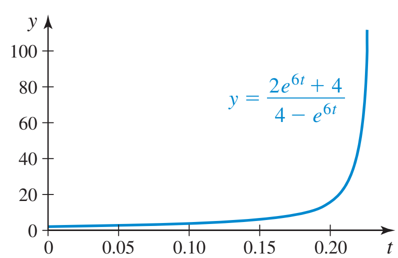


Figure 8.8 The solution for Example 5.

If we want the solution in the form $y = f(t)$, we must solve for y:

$$y - 1 = (y+2)\frac{1}{4}e^{6t}$$

$$y\left(1 - \frac{1}{4}e^{6t}\right) = \frac{1}{2}e^{6t} + 1$$

$$y = \frac{\frac{1}{2}e^{6t} + 1}{1 - \frac{1}{4}e^{6t}} = \frac{2e^{6t}+4}{4-e^{6t}} \qquad \text{Isolate terms in } y$$

See Figure 8.8 for a graph of this solution.

Why do we stress the need to study equations where $g(y)$ is a quadratic polynomial? A very important model of this type is the **logistic equation**, which can be used to analyze density dependent growth of populations. The logistic equation was introduced around 1835 by Pierre François Verhulst, and it modifies Malthus' equation from Example 3 to account for the finite **carrying capacity** of the environment that the organisms are growing in.

Malthus' equation assumes that a population has a constant per capita rate of growth (i.e., $\frac{1}{N}\frac{dN}{dt} = r$ for some constant r). In a population with density dependent growth the per capita rate will depend on the population size, N. That is,

$$\frac{1}{N}\frac{dN}{dt} = R(N)$$

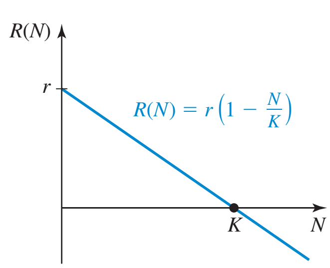


Figure 8.9 The per capita growth rate in the logistic equation is a linearly decreasing function of population size.

for some function $R(N)$. Different growth models (we will explore others in Section 8.2) are associated with different functions $R(N)$. For the logistic equation we assume

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

where r and K are both positive coefficients ($R(N)$ is graphed in Figure 8.9). Why is this form chosen? As you can see from the graph, per capita growth is positive if $N < K$ and is negative if $N > K$. Populations smaller than K will therefore grow, while populations larger than K will decrease. If $N = K$, then $\frac{dN}{dt} = 0$ (the population stays steady): K therefore represents the carrying capacity of the population environment. If $N > K$ then there are not enough resources for all of the organisms present, so the population will shrink. Also from Figure 8.9, we see that $R(0) = r$. This means that r is the limiting per capita growth rate when N is very small. If N is small (specifically, if N is much smaller than K), then $\frac{1}{N}\frac{dN}{dt} \approx r$, so the population will grow exponentially with per capita growth rate r until overcrowding starts to affect growth. The function $R(N)$ is the simplest function (a straight line) that incorporates both of the required features: $R(0) = r$ and $R(N)$ changes sign at $N = K$.

EXAMPLE 6 **The Logistic Equation for Population Growth** Solve the logistic equation

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) \text{ with initial condition } N(0) = N_0 \tag{8.20}$$

where $r > 0$, $K > 0$, and $N_0 \geq 0$ are all constants.

Solution The constants r, K, and N_0 allow the logistic equation to be fit to different species and environments. It is important, therefore, to be able to solve the equation in this general form, without specifying the values of these coefficients.

To solve (8.20) we separate variables and integrate:

$$\int \frac{dN}{N(1 - N/K)} = \int r\,dt,$$

provided $N \neq 0$ and $N \neq K$.

To evaluate the integral on the left-hand side, we use the partial fraction expansion:

$$\frac{1}{N(1 - N/K)} = \frac{A}{N} + \frac{B}{1 - N/K}$$

$$1 = A(1 - N/K) + BN. \qquad \times N(1 - N/K) \text{ on both sides} \tag{8.21}$$

To find A and B, substitute specific values into both sides of (8.21):

$$N = K \quad \Rightarrow \quad 1 = BK \quad \Rightarrow \quad B = 1/K$$

$$N = 0 \quad \Rightarrow \quad 1 = A \quad \Rightarrow \quad A = 1.$$

So our integrals become:

$$\int \left(\frac{1}{N} + \frac{1}{K(1 - N/K)}\right) dN = \int r\,dt$$

$$\ln|N| - \ln|1 - N/K| = rt + C_1$$

$$\ln\left|\frac{N}{1 - N/K}\right| = rt + C_1 \qquad \text{Combining logarithms}$$

$$\frac{N}{1 - N/K} = \pm e^{rt}e^{C_1} = Ce^{rt} \qquad \text{Defining } C = \pm e^{C_1} \tag{8.22}$$

To solve for C we can apply the initial condition:

$$\frac{N_0}{1 - N_0/K} = C. \qquad N(0) = N_0$$

But the calculations will be a little easier if we leave C in the equation for the time being. Rearrange (8.22) as:

$$N\left(1 + \frac{Ce^{rt}}{K}\right) = Ce^{rt} \qquad \times(1 - N/K) \text{ and then isolate terms in } N$$

$$\Rightarrow N = \frac{Ce^{rt}}{1 + Ce^{rt}/K} = \frac{C}{e^{-rt} + C/K} \qquad \text{Solve for } N$$

$$\Rightarrow N = \frac{K}{\frac{Ke^{-rt}}{C} + 1} \qquad \text{Multiply numerator and denominator by } \frac{K}{C}$$

Now since $C = \frac{N_0}{1-N_0/K}$, $\frac{K}{C} = \frac{K-N_0}{N_0}$, so assuming $N_0 \neq 0$ we obtain,

$$N(t) = \frac{K}{1 + \left(\frac{K-N_0}{N_0}\right)e^{-rt}} = \frac{K}{1 + \left(\frac{K}{N_0} - 1\right)e^{-rt}} \tag{8.23}$$

Our derivation requires that $N \neq 0$ and $N \neq K$, so to complete our solution we observe that if $N_0 = 0$, then $N(t) = 0$ for all t, while if $N_0 = K$, then $N(t) = K$ for all t. •

The formula for $N(t)$ in (8.23) is a somewhat complicated function, so we spend a little time determining the shape of its graph, using the steps from Section 5.6. We may check that

$$N(0) = \frac{K}{1 + \left(\frac{K}{N_0} - 1\right)} = \frac{K}{\frac{K}{N_0}} = N_0,$$

so our solution does satisfy the initial conditions.

Then notice that since $r > 0$, $e^{-rt} \to 0$ as $t \to \infty$. So

$$\lim_{t\to\infty} N(t) = \frac{K}{1 + \left(\frac{K}{N_0} - 1\right)\cdot 0} = K.$$

That is, so long as $N_0 \neq 0$, the population size will converge to K as $t \to \infty$.

If $0 < N < K$, then

$$\frac{dN}{dt} = rN(1 - N/K) > 0 \qquad r > 0, N > 0, 1 - N/K > 0$$

while if $N > K$ then $\frac{dN}{dt} = rN(1 - N/K) < 0$. $\quad r > 0, N > 0, 1 - N/K < 0$

So if $N_0 < K$ then $N(t)$ will increase monotonically until it converges to a horizontal asymptote $N = K$, while if $N_0 > K$, then $N(t)$ will decrease monotonically, approaching the horizontal asymptote $N = K$. To complete the information needed to sketch the function $N(t)$, we only need to know whether it is concave up or concave down. We may obtain the curvature from the differential equation. (Recall that $N(t)$ is concave up if $d^2N/dt^2 > 0$ and concave down if $d^2N/dt^2 < 0$.)

$$\begin{aligned}\frac{d^2N}{dt^2} &= \frac{d}{dt}\left(\frac{dN}{dt}\right) = \frac{d}{dt}\left(rN(1 - N/K)\right)\\ &= r\frac{dN}{dt} - \frac{2rN}{K}\frac{dN}{dt} \qquad \frac{d}{dt}(N^2) = 2N\frac{dN}{dt} \text{ using implicit differentiation}\\ &= r\left(1 - \frac{2N}{K}\right)\frac{dN}{dt}\end{aligned}$$

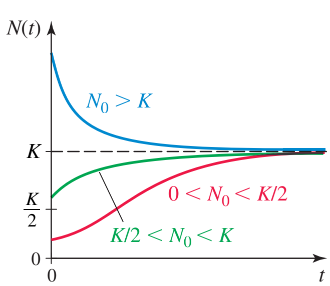


Figure 8.10 Solution curves of the logistic equation $\frac{dN}{dt} = rN(1 - N/K)$ for different initial values N_0.

If $0 < N < K$ then $\frac{dN}{dt} > 0$ (the solution increases), so $\frac{d^2N}{dt^2}$ is positive if $\left(1 - \frac{2N}{K}\right) > 0$ and negative if $\left(1 - \frac{2N}{K}\right) < 0$; that is, $\frac{d^2N}{dt^2} > 0$ if $N < K/2$ and $\frac{d^2N}{dt^2} < 0$ if $K/2 < N < K$. So if the initial condition $0 < N_0 < K/2$, then the solution initially curves upward until the population reaches $N = K/2$; once $N(t)$ crosses $K/2$ the solution curves downward (but continues to grow) approaching a horizontal asymptote $N = K$. The point at which $N(t) = K/2$ is an inflection point. If $K/2 < N_0 < N$, then $N(t)$ grows toward $N = K$ but always curves downward. If $N > K$, then $\left(1 - \frac{2N}{K}\right) < 0$, and $\frac{dN}{dt} < 0$ so $\frac{d^2N}{dt^2} > 0$. So if $N_0 > K$, then the solution curve is both monotonic decreasing and concave upward. Figure 8.10 shows some representative solution curves.

8.1.3 General Separable Equations

Although many important equations from biology are either of pure-time or autonomous forms, we can combine the techniques from 8.1.1 and 8.1.2 to solve general equations of the form: $\frac{dN}{dt} = f(t)g(N)$ (i.e., where the right-hand side includes both functions of the independent and dependent variables). As an example we first consider the problem introduced at the beginning of this chapter.

EXAMPLE 7 **Circadian Rhythm** The rate of division of some cells varies over the course of one day. One model that incorporates this effect is to have the per capita growth rate of the population be a function of the time of day. Let t represent time in fractions of one complete day, so $t = 0$ is the beginning of the first day and $t = 1$ is the end of the day. For example:

$$\frac{dN}{dt} = \frac{1}{2}(1 + \sin 2\pi t)N, \quad t \geq 0$$

Assuming $N(0) = N_0$, solve for $N(t)$.

Solution This is a separable equation with $f(t) = \frac{1}{2}(1 + \sin 2\pi t)$ and $g(N) = N$. Separate variables:

$$\frac{1}{N}\frac{dN}{dt} = \frac{1}{2}(1 + \sin 2\pi t) \qquad \text{Functions of } N \text{ on left, function of } t \text{ on right}$$

$$\int \frac{1}{N}dN = \int \frac{1}{2}(1 + \sin 2\pi t)dt \qquad \text{Integrate with respect to } t \text{ and use } \tfrac{dN}{dt}dt = dN$$

$$\ln|N| = \frac{t}{2} - \frac{1}{4\pi}\cos 2\pi t + C_1$$

$$N(t) = C\exp\left(\frac{t}{2} - \frac{1}{4\pi}\cos 2\pi t\right) \qquad \text{Define } C = \pm e^{C_1}$$

To find the value of C, apply the initial condition:

$$N_0 = C\exp\left(\frac{0}{2} - \frac{1}{4\pi}\cos 0\right) = Ce^{-1/4\pi}$$

$$\Rightarrow C = N_0 e^{1/4\pi}$$

so

$$N(t) = N_0 e^{1/4\pi}\exp\left(\frac{t}{2} - \frac{1}{4\pi}\cos 2\pi t\right) = N_0\exp\left(\frac{t}{2} + \frac{1}{4\pi}(1 - \cos 2\pi t)\right).$$

An important application of equations of this type is **allometry**. *Allometry* is the study of how different parts of an organism (e.g., the size of two different organs, or parts) grow differently as the organism grows. We denote by $L_1(t)$ and $L_2(t)$ the respective sizes of two different parts of an individual of age t. We say that L_1 and L_2 are related through an allometric law if their *specific* (or *relative*) growth rates are proportional—that is, if

$$\frac{1}{L_1}\frac{dL_1}{dt} = k\frac{1}{L_2}\frac{dL_2}{dt} \qquad \text{Relative growth rate} = \text{growth rate}/\text{current size.} \tag{8.24}$$

for some constant k. If the constant k is equal to 1, then the growth is called **isometric**; otherwise it is called **allometric**. Integrating both sides of (8.24), we find that

$$\int \frac{dL_1}{L_1} = k\int \frac{dL_2}{L_2} \qquad \tfrac{dL_1}{dt}dt = dL_1$$

or

$$\ln|L_1| = k\ln|L_2| + C_1$$

Solving for L_1, we obtain

$$L_1 = CL_2^k \tag{8.25}$$

where $C = \pm e^{C_1}$. (Since L_1 and L_2 are typically positive, the constant C will typically be positive.) If $k = 1$ (isometric growth) then $L_1 \propto L_2$, so the two organs maintain the same relative size during growth. More generally, the relationship between L_1 and L_2 is a power law.

Section 8.1 Problems

8.1.1

In Problems 1–8, solve each pure-time differential equation.

1. $\dfrac{dy}{dt} = t + \sin t$, where $y(0) = 0$.

2. $\dfrac{dy}{dt} = e^{-3t}$, where $y(0) = 10$.

3. $\dfrac{dy}{dx} = \dfrac{1}{x}$, where $y(1) = 0$.

4. $\dfrac{dy}{dx} = \dfrac{1}{1-x^2}$, where $y(0) = 0$.

5. $\dfrac{dx}{dt} = \dfrac{1}{1-t}$, where $x(0) = 2$.

6. $\dfrac{dx}{dt} = \sin(2\pi(t+3))$, where $x(3) = 1$.

7. $\dfrac{ds}{dt} = \sqrt{t+1}$, where $s(0) = 1$.

8. $\dfrac{dh}{dt} = 4 - 16t^2$, where $h(1) = 0$.

9. Suppose that the volume $V(t)$ of a cell at time t changes according to

$$\frac{dV}{dt} = 1 + \cos t \quad \text{with } V(0) = 5$$

Find $V(t)$.

10. Suppose that the amount of phosphorus in a lake at time t, denoted by $P(t)$, follows the equation

$$\frac{dP}{dt} = 3t + 1 \quad \text{with } P(0) = 0$$

Find the amount of phosphorus at time $t = 10$.

11. Drug Absorption For a drug with zeroth order elimination kinetics, a constant amount of drug is removed from the blood per unit time. So the amount of drug in a patient's blood obeys a differential equation

$$\frac{dM}{dt} = -k_0$$

where $k_0 > 0$ is a constant. Assuming $M(0) = M_0$, find (in terms of M_0 and k_0) the time at which the level of drug will drop to 0.

12. Drug Absorption Drug enters a patient's blood by being absorbed from the gut. Assume that the drug enters the patient's blood at a rate that depends on time as ce^{-rt} where c and r are positive constants (the rationale for this formula will be discussed in Section 8.4) and the drug is eliminated at constant rate k. So:

$$\frac{dM}{dt} = ce^{-rt} - k.$$

(a) Assuming $c > k$ and $M(0) = 0$ (there is no drug present in the patient's blood at the start of the experiment), solve this differential equation.

(b) Suppose $k > 0$. What does your solution predict will happen to $M(t)$ as $t \to \infty$? Does your answer make sense? (In reality, drug can only be removed at a constant rate until all drug is removed from the blood. That is, the rate of elimination will be k if $M > 0$ and 0 once M drops to 0.)

(c) Assume $k = 0$ (i.e., this drug is never eliminated from blood, or is eliminated so slowly that elimination can be neglected). Show that $\lim_{t\to\infty} M(t) = \frac{c}{r}$.

(d) Suppose that $k = 0$ and you measure the following data for $M(t)$ as a function of t:

t	$M(t)$
0	0
1	1
2	1.5

Find parameters c and r that fit your model to these data.

8.1.2

In Problems 13–18, solve each autonomous differential equation.

13. $\dfrac{dy}{dt} = 2y$, where $y(0) = 2$

14. $\dfrac{dy}{dt} = 2(1-y)$, where $y(0) = 0$

15. $\dfrac{dx}{dt} = -2x$, where $x(1) = 3$

16. $\dfrac{dx}{dt} = 1 - 3x$, where $x(1) = -2$

17. $\dfrac{dh}{ds} = 2h + 1$, where $h(0) = 4$

18. $\dfrac{dN}{dt} = 5 - N$, where $N(2) = 2$

19. Suppose that a population, whose size at time t is denoted by $N(t)$, grows according to

$$\frac{dN}{dt} = 0.3N \quad \text{with } N(0) = 20$$

Solve this differential equation, and find the size of the population at time $t = 5$.

20. Suppose that you follow the size of a population over time. When you plot the size of the population versus time on a semilog plot (i.e., the horizontal axis, representing time, is on a linear scale, whereas the vertical axis, representing the size of the population, is on a logarithmic scale) you find that your data fit a straight line that intercepts the vertical axis at 1 (on the log scale) and has slope -0.43. Find a differential equation that relates the growth rate of the population at time t to the size of the population at time t.

21. Suppose that a population, whose size at time t is denoted by $N(t)$, grows according to

$$\frac{1}{N}\frac{dN}{dt} = r \tag{8.26}$$

where r is a constant.

(a) Solve (8.26).

(b) Transform your solution in (a) appropriately so that the resulting graph is a straight line. How can you determine the constant r from your graph?

(c) Suppose now that, over time, you followed a population which grew according to (8.26) with some unknown reproductive rate r. Describe how you would determine r from your data.

22. **Radioactive Decay** Assume that $W(t)$ denotes the amount of radioactive material in a substance at time t. Radioactive decay is described by the differential equation

$$\frac{dW}{dt} = -\lambda W(t) \quad \text{with } W(0) = W_0 \tag{8.27}$$

where λ is a positive constant called the *decay constant*.

(a) Solve (8.27).

(b) Assume that $W(0) = 123$ g and $W(5) = 20$ g and that time is measured in minutes. Find the decay constant λ and determine the half-life of the radioactive substance. (Remember that the half-life of the substance is the time taken for $W(t)$ to decrease to half of its initial value.)

23. **Drug Absorption** A drug has first order elimination kinetics, meaning that a fixed fraction of drug is eliminated from the body in each unit of time. So if no further drug is absorbed into the patient's blood after time $t = 0$, the amount of drug in their blood will decay with time according to:

$$\frac{dM}{dt} = -k_1 M$$

where $k_1 > 0$ is the fraction of drug eliminated in one unit of time.

(a) Assuming $M(0) = M_0$, solve the differential equation.

(b) According to your model, does $M(t)$ ever reach 0?

(c) Given that $M_0 = 10$ and $k_1 = 2$, calculate the time at which $M(t)$ drops to $M = 1$.

24. **Fish Growth** Denote by $L(t)$ the length of a fish at time t, and assume that the fish grows according to von Bertalanffy's equation

$$\frac{dL}{dt} = k(34 - L(t)) \quad \text{with } L(0) = 2$$

(a) Solve the differential equation.

(b) Use your solution in (a) to determine k under the assumption that $L(4) = 10$. Sketch the graph of $L(t)$ for this value of k.

(c) Find the length of the fish when $t = 10$.

(d) Find the asymptotic length of the fish; that is, find $\lim_{t\to\infty} L(t)$.

25. **Fish Growth** Denote by $L(t)$ the length of a certain fish at time t, and assume that this fish grows according to von Bertalanffy's equation

$$\frac{dL}{dt} = k(L_\infty - L(t)) \quad \text{with } L(0) = 1 \tag{8.28}$$

where k and L_∞ are positive constants. It is known that the asymptotic length is equal to 123 in. and that it takes the fish 27 months to reach half its asymptotic length.

(a) Use this information to determine the constants k and L_∞ in (8.28). [*Hint*: Solve (8.28).]

(b) Determine the length of the fish after 10 months.

(c) How long will it take until the fish reaches 90% of its asymptotic length?

26. **Insulin Pump** A diabetic patient receives insulin at constant rate from an implanted insulin pump. Insulin has first order elimination kinetics, so the amount of insulin in the blood will obey a differential equation:

$$\frac{dM}{dt} = a - k_1 M$$

where $a > 0$ is the rate at which insulin is released into their blood by the pump and k_1 is the fraction of insulin removed from the blood in one unit of time

(a) Assuming $M(0) = 0$ (i.e., there is no insulin present in the patient's blood at time $t = 0$), solve the differential equation to find $M(t)$ as a function of t.

(b) Find the limit of $M(t)$ as $t \to \infty$.

(c) Assume that a (the rate of release from the pump) is 1 IU/hr and $M(t) \to 0.2$ IU as $t \to \infty$. Calculate, k_1, the rate of insulin elimination.

27. **Amnesia During Surgery** During surgery a patient receives midazolam, a sedative, to produce amnesia (memory loss) and ensure they do not remember the surgery. ***Holazo et al. (1988)*** studied the rate at which midazolam is eliminated from a patient's body. They gave healthy volunteers one injection of the drug at time $t = 0$, and then measured the rate at which it disappeared from each volunteer's blood. If no further drug is added after time $t = 0$, then the concentration will obey the differential equation:

$$\frac{dC}{dt} = -k_1 C$$

where k_1 is the fraction of drug eliminated in one hour.

(a) If the concentration at $t = 0$ is C_0, solve for $C(t)$.

(b) Holazo et al. found the following data:

t	$C(t)$
1	90
4	34

where t is measured in hours, and $C(t)$ is measured in ng of midazolam per milliliter of blood. From these data estimate the rate of elimination k_1.

(c) During surgery midazolam may be infused continuously by intravenous line, at some constant rate r per milliliter of blood. Then the concentration must obey a differential equation.

$$\frac{dC}{dt} = r - k_1 C.$$

Assuming that $C(0) = 0$, find $C(t)$ as a function of t.

(d) Calculate $\lim_{t\to\infty} C(t)$ as a function of r and k_1.

(e) Using the value of k_1 from part (b), at what rate, t, must midazolam be infused into the patient's blood to maintain a constant concentration of 130 ng/milliliter?

28. Let $N(t)$ denote the size of a population at time t. Assume that the population exhibits exponential growth.

(a) If you plot $\log N(t)$ versus t, what kind of graph do you get?

(b) Find a differential equation that describes the growth of this population and sketch possible solution curves.

29. Use the partial-fraction method to solve

$$\frac{dy}{dx} = y(1 + y)$$

where $y(0) = 2$.

30. Use the partial-fraction method to solve

$$\frac{dy}{dx} = y(1 - y)$$

where $y(0) = 2$.

31. Use the partial-fraction method to solve

$$\frac{dy}{dx} = y(y-2)$$

where $y(0) = 1$.

32. Use the partial-fraction method to solve

$$\frac{dy}{dx} = (y+1)(y-2)$$

where $y(0) = 0$.

33. Use the partial-fraction method to solve

$$\frac{dy}{dt} = 2y\left(1 - \frac{y}{3}\right)$$

where $y(1) = 5$.

34. Use the partial-fraction method to solve

$$\frac{dy}{dt} = \frac{1}{2}y^2 - 2y$$

where $y(0) = -3$.

In Problems 35–38, solve each differential equation.

35. $\frac{dy}{dx} = y(1+y)$

36. $\frac{dy}{dx} = (1+y)^2$

37. $\frac{dy}{dx} = (1+y)^3$

38. $\frac{dy}{dx} = (1-y^2)$

39. (a) Use partial fractions to show that

$$\int \frac{du}{u^2 - a^2} = \frac{1}{2a} \ln \left| \frac{u-a}{u+a} \right| + C$$

(b) Use your result in (a) to find a solution of

$$\frac{dy}{dx} = y^2 - 4$$

with initial conditions **(i)** $y(0) = 0$, **(ii)** $y(0) = 2$, and **(iii)** $y(0) = 4$.

40. Find a solution of

$$\frac{dy}{dx} = y^2 + 4$$

with initial conditions $y(0) = 2$.

41. Suppose that the size of a population at time t is denoted by $N(t)$ and that $N(t)$ satisfies the logistic equation

$$\frac{dN}{dt} = 0.34N\left(1 - \frac{N}{200}\right) \quad \text{with } N(0) = 50$$

Solve this differential equation, and determine the size of the population in the long run; that is, find $\lim_{t\to\infty} N(t)$.

42. Assume that a population, whose size is denoted by $N(t)$, grows according to the logistic equation. Find the limiting growth rate for small N (i.e., find the constant r) if the carrying capacity is 100, $N(0) = 10$, and $N(1) = 30$.

43. Let $N(t)$ denote the size of a population at time t. Assume that the population grows according to the logistic equation. Assume also that the limiting growth rate for small N is 5 and that the carrying capacity is 50.

(a) Find a differential equation that describes the growth of this population.

(b) Without solving the differential equation in (a), sketch solution curves of $N(t)$ as a function of t when **(i)** $N(0) = 10$, **(ii)** $N(0) = 40$, and **(iii)** $N(0) = 50$.

44. Logistic growth is described by the differential equation

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

We showed in Example 6 that the solution of this differential equation with initial condition $N(0) = N_0$ is given by

$$N(t) = \frac{K}{1 + (\frac{K}{N_0} - 1)e^{-rt}}. \tag{8.29}$$

(a) Show that

$$r = \frac{1}{t}\ln\left(\frac{K - N_0}{N_0}\right) + \frac{1}{t}\ln\left(\frac{N(t)}{K - N(t)}\right) \tag{8.30}$$

by solving (8.29) for r.

(b) Equation (8.30) can be used to estimate r. Suppose we are studying a population that grows according to the logistic equation and find that $N(0) = 10$, $N(5) = 22$, $N(100) = 30$, and $N(200) = 30$. Estimate r. (*Hint*: First estimate K from the behavior of the solution for large t.)

45. Population Genetics Population genetics is the study of how the frequency of particular traits changes within a population over time. We are studying a gene that comes in two alleles (i.e., variants) A and a. The A allele makes individuals reproduce a little faster than the a allele. So we expect the A alleles to take over the population with time. Suppose that a proportion p of all individuals within the population carry the A allele (with the remaining proportion, $1-p$, carrying the a allele). If the A allele boosts reproduction rate by an amount s it can be shown under some assumptions that the proportion of A-allele individuals obeys a differential equation

$$\frac{dp}{dt} = \frac{1}{2}s\,p(1-p) \tag{8.31}$$

(a) Use separation of variables and partial fractions to find the solution of (8.31), assuming $p(0) = p_0$.

(b) Show that if $p_0 \neq 0$, then $\lim_{t\to\infty} p(t) = 1$. Explain why this behavior makes sense biologically.

(c) Suppose $p_0 = 0.1$ and $s = 0.01$; how long will take until $p(t) = 0.5$?

8.1.3

In Problems 46–54, solve each differential equation with the given initial condition.

46. $\frac{dy}{dx} = 2\frac{y}{x}$, with $y(1) = 1$.

47. $\frac{dy}{dx} = \frac{x+1}{y}$, with $y(0) = 2$.

48. $\frac{dy}{dx} = \frac{xy}{x+1}$, with $y(0) = 1$.

49. $\frac{dy}{dx} = (y+1)e^{-x}$, with $y(0) = 2$.

50. $\frac{dy}{dx} = \frac{y^2}{x}$, with $y(1) = 1$.

51. $\frac{dy}{dx} = \frac{y+1}{x-1}$, with $y(2) = 5$.

52. $\frac{du}{dt} = \frac{\sin t}{u+1}$, with $u(0) = 3$.

53. $\frac{dy}{dt} = \frac{t}{y}$, with $y(0) = 1$.

54. $\frac{dx}{dy} = \frac{1}{2}\frac{x}{y}$, with $x(3) = 2$.

In Problems 55–60 you will need to solve differential equations by separation of variables. In these problems it will not always be possible to solve explicitly for y in terms of t; instead your solution may take the form of an implicit function relating the two variables.

55. $\frac{dy}{dt} = \frac{y^2 + y}{t - 1}$ where $y(0) = 1$.

56. $\frac{dy}{dt} = \frac{yt}{\ln y}$ where $y(1) = e$.

57. $\frac{dy}{dt} = \frac{t + 1}{y + y^2}$ where $y(0) = 1$.

58. $\frac{dy}{dt} = \frac{t^2 + 1}{\cos y + \sin y}$ where $y(0) = 0$.

59. $\frac{dy}{dt} = \sqrt{\frac{t + 1}{y + 1}}$ where $y(0) = 1$.

60. $\frac{dy}{dt} = \frac{t + 1}{ty + ty^3}$ where $y(1) = 1$.

61. Circadian Rhythm The per capita growth rate of a population of cells varies over the course of a day. Assume that time t is measured in hours and

$$\frac{dN}{dt} = 2\left(1 - \cos\frac{2\pi t}{24}\right)N$$

if $N(0) = 5$, find the number of cells after one day (that is, find $N(24)$).

62. Malthusian Population Growth This problem addresses Malthus's concerns, which were discussed in Example 3. Assume that a population size grows exponentially according to

$$N(t) = 1000e^t$$

and the food supply grows linearly according to

$$F(t) = 3t$$

(a) Write a differential equation for each of $N(t)$ and $F(t)$.

(b) Does exponential growth eventually overtake linear growth? Explain.

63. Bite Strength in Carnivores Bite strength varies as animals grow, which may mean that the animal's diet must change. ***Christiansen and Adolfsson (2005)*** studied the relationship between the strength of animal teeth with skull size in carnivores from the cat and dog families. They found that tooth strength S, and skull length L, were related in a power law:

$$S = CL^{2.85}$$

where C is some constant. Find the relationship between the relative rates of growth of S and L (i.e., between $\frac{1}{S}\frac{dS}{dt}$ and $\frac{1}{L}\frac{dL}{dt}$).

64. Homeostasis ***Sterner and Elser (2002)*** studied the relationship between the amount of nitrogen in an animal's body and the amount of nitrogen present in the food that it eats. Many animals maintain homeostasis (balance), that is, they control their own nitrogen content. As the amount of nitrogen present in their food increases, the amount of nitrogen in the animal's body increases more slowly. If the amount of nitrogen in the animal is N and the amount of nitrogen in its food is F, Sterner and Elser argue that:

$$\frac{1}{N}\frac{dN}{dt} = \frac{\sigma}{F}\frac{dF}{dt}$$

where σ is a constant.

(a) Show that if $\sigma = 1$, then $N \propto F$; that is, the nitrogen content of the animal increases in proportion to its food. This is called **absence of homeostasis**.

(b) If $\sigma = 0$, then N is a constant, independent of F. This is called **homeostasis** (the animal maintains a balanced amount of nitrogen, independent of its food).

(c) Show that if $0 < \sigma < 1$, then, if F doubles, N also increases but by a factor less than 2.

8.2 Equilibria and Their Stability

In Subsection 8.1.2, we learned how to solve autonomous differential equations. Once we solve a differential equation we may draw the graph of the solution. For instance, logistic growth can be modeled using the differential equation

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) \tag{8.32}$$

The solution of this differential equation was derived in Section 8.1 (see Equation (8.23)) and graphed in Figure 8.10. In particular we saw that so long as $N(0) \neq 0$ the population size converges to K as $t \to \infty$. We identified K as the carrying capacity of the organism's habitat. In this section we will show that this behavior could have been predicted without solving the differential equation. In particular, if $N = K$ then the right-hand side of (8.32) vanishes. This is no coincidence. If $N(t)$ converges to any constant, as $t \to \infty$, then $\frac{dN}{dt}$ must converge to 0, since the curve approaches a horizontal asymptote (i.e., a flat line) and the gradient of this flat line is 0. Since $dN/dt = 0$, the right-hand side of the equation must also vanish. In fact there are two values of N for which the right-hand side of (8.32) vanishes: $N = K$ and $N = 0$. Why don't solutions converge to 0 as $t \to \infty$? To answer this question we will need to introduce the concept of **stability**.

8.2.1 Equilibrium Points

We will consider autonomous differential equations of the form

$$\frac{dy}{dt} = g(y) \tag{8.33}$$

The point $\hat{y}$ is said to be an **equilibrium** of the differential equation if $g(\hat{y}) = 0$. Some books use the term **fixed point** instead of equilibrium. We will always refer to these points as equilibria in this book, but you should be aware of the alternate term.

Why are equilibria important? If $y(t)$ solves the differential equation $\frac{dy}{dt} = g(y)$ and $y(0) = \hat{y}$, then because $\frac{dy}{dt} = g(\hat{y}) = 0$, $y(t) = \hat{y}$ for all t (i.e., if a solution starts at one of the equilibria of the differential equation, then it will remain at that equilibrium). In other words $y(t) = \hat{y}$ is a constant solution of the differential equation.

EXAMPLE 1 Find all equilibria of the differential equations:

(a) $\frac{dy}{dt} = 2y$ **(b)** $\frac{dy}{dt} = y^3 - 3y^2 + 2y$ **(c)** $\frac{du}{dt} = \sin u$

Solution In all cases the equilibria are the zeros of the function on the right-hand side of the differential equation.

(a) $2\hat{y} = 0 \quad \Rightarrow \quad \hat{y} = 0$

(b) $\hat{y}^3 - 3\hat{y}^2 + 2\hat{y} = 0 \quad \Rightarrow \quad \hat{y}(\hat{y}^2 - 3\hat{y} + 2) = 0$

$\hat{y}(\hat{y} - 1)(\hat{y} - 2) = 0$ $\quad$ $\hat{y} = 0$ is a root

so $\hat{y} = 0,\ 1$ or 2. $\quad$ Factorize the quadratic

(c) $\sin \hat{u} = 0 \Rightarrow$ the equilibria are: $\hat{u} = 0, \pm\pi, \pm 2\pi, \ldots$ (alternatively: $\hat{u} = K\pi$ where $K \in \mathbf{Z}$.) •

EXAMPLE 2 **Tumor Growth** ***Laird (1969)*** showed that the growth of many different kinds of solid tumors in mice, rats, and rabbits could be described by Gompertz's equation, which predicts that the number of cells in the tumor, N, increases with time t, according to the differential equation:

$$\frac{dN}{dt} = aN \ln\left(\frac{b}{N}\right) \quad , \quad N > 0 \tag{8.34}$$

Here a and b are both positive coefficients. Find the possible equilibria of the number of cells.

Solution The function on the right-hand side of the differential equation is:

$$g(N) = aN \ln\left(\frac{b}{N}\right)$$

$g(N)$ can only be equal to zero if $N = 0$ or if $\ln\left(\frac{b}{N}\right) = 0$, since one of the two factors that make up $g(N)$ must equal 0 to make $g(N)$ equal 0. $N = 0$ is not in the domain for which $g(N)$ is defined. But $\ln\left(\frac{b}{N}\right) = 0$ if $N = b$, and this will in general be within the domain. In fact, we will show in subsection 8.2.4 that all solutions of (8.34) converge to $N = b$. •

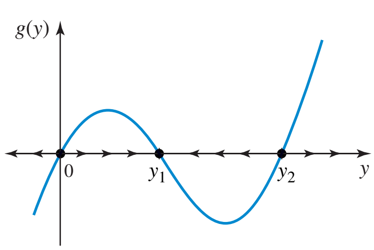


Figure 8.11 Vector field plot for the differential equation $dy/dt = g(y)$.

8.2.2 Graphical Approach to Finding Equilibria

Suppose that $g(y)$ is of the form given in Figure 8.11. To find the equilibria of $dy/dt = g(y)$, we must find all points $y = \hat{y}$ for which $g(y) = 0$. Graphically, this means that if we graph $g(y)$ as a function of y, then the equilibria are the points of intersection of $g(y)$ with the horizontal axis, which is the y-axis in this case, since y is the independent variable. We see that, for this choice of $g(y)$, the equilibria are at $y = 0$, y_1, and y_2.

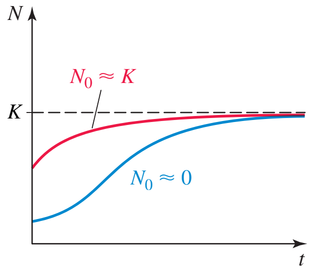


Figure 8.12 Solution curves of the logistic equation. If $N_0 \approx 0$, the population will grow away from $N = 0$ as t increases. If $N_0 \approx K$, the population will approach K as t increases.

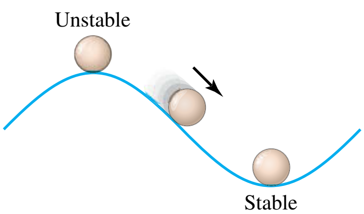


Figure 8.13 A ball rolls around on a landscape containing hills, and valleys. Equilibria occur both at the top of hills, and at the bottom of valleys.

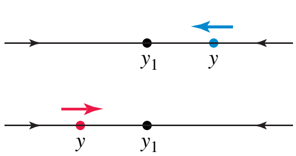


Figure 8.14 If y is perturbed from the stable equilibrium $y = y_1$ then $y(t)$ will either decrease or increase to bring $y(t)$ back toward y_1.

8.2.3 Stability of Equilibrium Points

Although the logistic equation $dN/dt = rN(1 - N/K)$ has two equilibria, at $N = 0$ and at $N = K$, the equilibria differ in the following fundamental respect. If the population starts at either equilibrium, it will remain there (i.e., $N = 0$ and $N = K$ are both valid constant solutions of the differential equation). But if the population starts with a size close to 0, but not exactly equal to 0, it will grow further from 0 as t increases, whereas if the population starts close to K, it will approach K as $t \to \infty$ (see Figure 8.12).

We say that the equilibrium point $N = 0$ is **unstable**, while the equilibrium point $N = K$ is **stable**. Whether an equilibrium is stable or unstable is determined by the behavior of the solution when the solution is **perturbed** away from the equilibrium, meaning that it doesn't start at the equilibrium but at a value close to the equilibrium. If it returns to the equilibrium, the equilibrium is stable; while if it diverges from the equilibrium, then the equilibrium is unstable.

A useful analogy is to think of a ball rolling around on a landscape that contains hills and valleys (see Figure 8.13). The ball can be balanced and at rest when it is either at the top of a hill, or at the bottom of a valley. But if it is perturbed slightly from the top of a hill, it will roll down the hill away from its rest position—these equilibria are unstable. If it starts at the bottom of a valley, and is perturbed from the bottom of a valley, it will return to its starting position—these equilibria are stable.

We will present two methods for identifying whether an equilibrium is stable or unstable. One is based on the graph of the function $g(y)$; the other requires calculating $g'(y)$. You should be comfortable using both approaches.

Suppose we are studying the differential equation

$$\frac{dy}{dt} = g(y)$$

where the function $g(y)$ is graphed in Figure 8.11. The equilibria of this differential equation are points $\hat{y}$ at which $g(\hat{y}) = 0$. On the graph these correspond to points at which $g(y)$ crosses the horizontal axis. There are three such points: 0, y_1, and y_2.

We can also use the graph in Figure 8.11 to identify the values of y for which $y(t)$ increases with t and the values for which $y(t)$ decreases with t. (Remember, since the equation is autonomous, the value of $\frac{dy}{dt}$ depends only on y.) If $g(y) > 0$, then according to the differential equation $\frac{dy}{dt} > 0$ (i.e., $y(t)$ is an increasing function of t). Conversely, if $g(y) < 0$ (so $\frac{dy}{dt} < 0$), then $y(t)$ is a decreasing function of t. We can label the horizontal axis of Figure 8.11 with direction arrows to show the direction in which $y(t)$ travels as t increases, putting a rightward arrow to show where $y(t)$ is increasing, and a leftward arrow to show where y is decreasing. For the function $g(y)$ in the figure, this means that we add rightward arrows in the intervals $(0, y_1)$, and (y_2, ∞) and we add leftward arrows in the intervals $(-\infty, 0)$ and (y_1, y_2). This plot, which now includes the direction in which y travels with time, is called a **vector field plot** of the differential equation $\frac{dy}{dt} = g(y)$.

We may use the vector field plot to determine which of the equilibria are stable and which are unstable. Consider, for example, the equilibrium $y = y_1$. Recall that stability of an equilibrium is determined from the behavior of the solution when it is slightly perturbed away from the equilibrium. Suppose that we start with an initial condition that is slightly above y_1 (i.e., right of it on the graph); then, since the arrows are leftward, $y(t)$ will *decrease* back toward y_1. Similarly, if we solve the differential equation with an initial condition that is slightly smaller than y_1 (i.e., left of it on the graph), then the arrows show that $y(t)$ will *increase* back toward y_1. In either case the solution will tend to return to y_1 if it is started from an initial condition that is perturbed a small distance from y_1 (see Figure 8.14). Therefore y_1 is a **stable** equilibrium of the differential equation.

Let's apply the same arguments to study the stability of the point y_2. If y starts above (right of) y_2, then it will follow the rightward arrows (i.e., increase) and move further from y_2. If y starts below (left of) y_2 then it will follow the leftward arrows (i.e., decease), again moving further from y_2. Thus, however y is perturbed from y_2, it

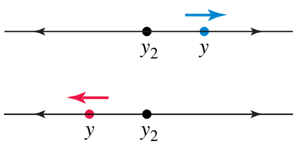


Figure 8.15 If y is perturbed from the unstable equilibrium y_2, then $y(t)$ will increase or decrease to travel further from y_2.

will tend to move further from y_2 over time, as summarized in Figure 8.15. So y_2 is an **unstable** equilibrium of the differential equation.

Our reasoning here did not require much knowledge of $g(y)$; we only needed to know the direction of arrows on either side of each equilibrium. We can examine the final equilibrium $\hat{y} = 0$ in exactly the same way. The arrows are leftward left of $y = 0$, and rightward right of $\hat{y} = 0$, so if $y(t)$ starts from an initial condition on either side of 0, it will tend to travel away from $y = 0$; the equilibrium $\hat{y} = 0$ is therefore unstable.

Graphical Criteria for Stability
If $g(y) > 0$ left of $\hat{y}$ and $g(y) < 0$ right of $\hat{y}$, then $\hat{y}$ is a stable equilibrium.
If $g(y) < 0$ left of $\hat{y}$ and $g(y) > 0$ right of $\hat{y}$, then $\hat{y}$ is an unstable equilibrium.

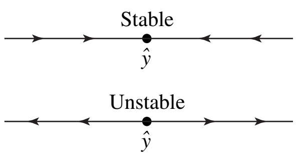


Figure 8.16 Arrows showing direction in which $y(t)$ moves near stable and unstable equilibria.

In terms of the vector field plot, stable equilibria have arrows pointing toward the equilibrium, while for unstable equilibria, the arrows point away from the equilibrium (see Figure 8.16).

An alternative way of writing the graphical criteria for stability may be easier to apply in some circumstances. Notice that if $g(\hat{y}) = 0$ and $g(y) > 0$ left of $\hat{y}$ and $g(y) < 0$ right of $\hat{y}$ (i.e., $\hat{y}$ is a stable equilibrium according to the graphical criteria), then $g(y)$ must be a decreasing function of y at $y = \hat{y}$ (since it goes from positive values to negative values as y passes from values smaller than $\hat{y}$ to values larger than $\hat{y}$). We can make similar arguments for unstable fixed points.

If $g(\hat{y}) = 0$ and $g(y)$ is decreasing at $y = \hat{y}$, then $\hat{y}$ is a stable equilibrium.
If $g(\hat{y}) = 0$ and $g(y)$ is increasing at $y = \hat{y}$, then $\hat{y}$ is an unstable equilibrium.

If g is differentiable then we may turn this criteria into one based on the derivative $g'(\hat{y})$ because $g'(\hat{y}) < 0$ implies that g is decreasing at $y = \hat{y}$, and $g'(\hat{y}) > 0$ implies that g is increasing at $y = \hat{y}$.

Derivative-Based Criterion for Stability of Equilibria
If $g(\hat{y}) = 0$ and $g'(\hat{y}) < 0$, then $\hat{y}$ is a stable equilibrium.
If $g(\hat{y}) = 0$ and $g'(\hat{y}) > 0$, then $\hat{y}$ is an unstable equilibrium.

The derivative-based criterion says nothing about the stability of an equilibrium if $g'(\hat{y}) = 0$, since the function g may be increasing, decreasing, or have a local extremum if $g'(\hat{y}) = 0$. We will discuss the last possibility after some examples.

EXAMPLE 3 Identify the equilibria of the following differential equations and determine whether they are stable or unstable.

(a) $\dfrac{dy}{dt} = y^2 - y$ **(b)** $\dfrac{dy}{dx} = y^3$

Solution Notice that the independent variables are different for the two equations. But since the right-hand side is a function of y (the dependent variable) in both cases, they are both autonomous differential equations. We plot these functions $g(y)$ for (a) and (b) in Figures 8.17 and 8.18.

(a) $g(y) = y(y-1)$ so the equilibria are at $y = 0$ and $y = 1$. In the vector field plot, the arrows point in toward $y = 0$, so $y = 0$ is stable. Arrows point out from $y = 1$, so $y = 1$ is unstable. In this case we could also have used the derivative test since

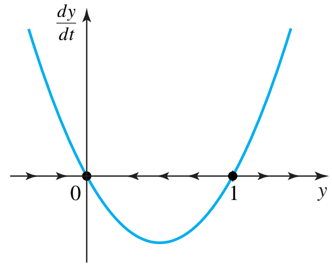


Figure 8.17 Vector field plot of $\frac{dy}{dt} = y^2 - y$ for Example 3(a).

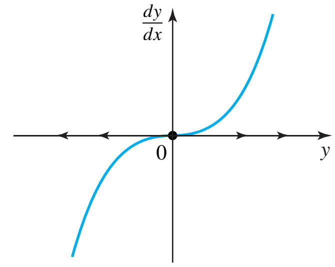


Figure 8.18 Vector field plot of $\frac{dy}{dx} = y^3$ for Example 3(b).

$g'(y) = 2y - 1$, so $g'(0) = -1$, and $g'(1) = 1$. Since $g'(0) < 0$, $y = 0$ is stable. Since $g'(1) > 0$, $y = 1$ is unstable.

(b) The only zero of $g(y) = y^3$ is $y = 0$. In the vector field plot the arrows point away from $y = 0$, so $y = 0$ is an unstable equilibrium, Although $g(y)$ is an increasing function of y at $y = 0$, $g'(0) = 0$, so the derivative criterion for stability cannot be used here. •

EXAMPLE 4 **The Logistic Equation** The logistic equation models the growth or decay of a population living in a habitat with carrying capacity K using a differential equation:

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

where $r, K > 0$ are constants. Find the stable and unstable equilibria of this differential equation.

Solution The right-hand side of the differential equation is $g(N) = rN(1 - N/K)$ and is already written in factorized form, so we can read off the roots $N = 0$ and $N = K$. These roots are the equilibria of the differential equation. To determine stability we make a vector field plot (Figure 8.19). To draw the graph of $g(N)$ notice that $g(N)$ is a quadratic polynomial in N (i.e., a parabola). For large N, $g(N) \approx -rN^2/K$ (i.e., $g(N)$ can be approximated by the largest degree term; see Section 3.3), so $\lim_{N\to\infty} g(N) = -\infty$ and $\lim_{N\to-\infty} g(N) = -\infty$, meaning that the parabola bends downward for $N > K$, and $g(N) > 0$ for $0 < N < K$.

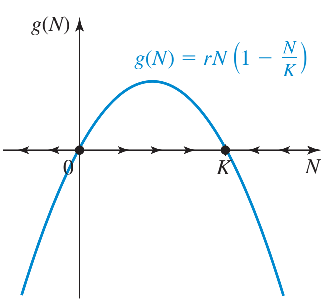


Figure 8.19 Vector field plot for Example 4.

The arrows in the vector field plot (Figure 8.19) point toward the equilibrium at $N = K$, so this equilibrium is stable; while arrows point away from the equilibrium at $N = 0$, so this equilibrium is unstable. •

Our graphical classification of stability of equilibria agrees with our direct solution of the logistic equation in Example 6 of Section 8.1. In the discussion following that example we showed that solutions tend to converge toward $N = K$ as $t \to \infty$, and tend to diverge away from $N = 0$.

EXAMPLE 5 **The Allee Effect** If a population is too small, then it may be outcompeted by other species present in the same habitat, or individuals may be forced to leave in search of mates outside the habitat. So, rather than growing exponentially, small populations may decline into extinction, an effect known as the Allee effect. We can modify the logistic equation to capture this effect by introducing another threshold, a, that represents the minimum size a population must exceed to avoid decline. (We assume that $0 < a < K$, where K is the carrying capacity.)

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

The only difference between this equation and the logistic equation is the extra factor of $(N - a)$ on the right-hand side.

Find the equilibria of the logistic equation with Allee effect and determine their stability.

Solution The right-hand side of the differential equation is already written in factorized form, so we may read off the values of the equilibria: $N = 0$, a, and K. To determine stability we refer to the vector field plot in Figure 8.20. The arrows in the vector field plot point toward the equilibria $N = 0$ and $N = K$, making these equilibria stable. The arrows point away from the equilibrium $N = a$, making $N = a$ unstable. In Example 9 we will discuss how the vector field plot represents the Allee effect. •

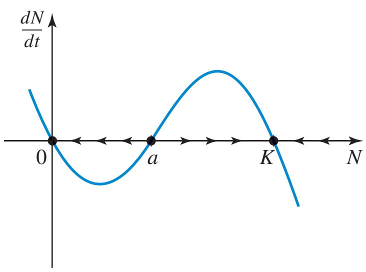


Figure 8.20 Vector field plot for Example 5.

The graphical approach can be used to find the number of equilibria, and to determine their stability, even if the function $g(y)$ is too complicated to graph directly, provided that $g(y)$ can be written as the difference between two functions that can be graphed individually.

Suppose

$$\frac{dy}{dt} = f(y) - h(y) \qquad g(y) = f(y) - h(y)$$

where $f(y)$ and $h(y)$ can both be graphed individually. Where the graphs for $f(y)$ and $h(y)$ intersect, $f(y) = h(y)$, so $g(y) = f(y) - h(y) = 0$. So the points of intersection of the two graphs are the equilibria of the differential equation. To create the vector field plot (i.e., determine whether dy/dt is positive or negative) we compare the graphs. On any interval in which $f(y) > h(y)$ (i.e., the graph of $f(y)$ is above the graph of $h(y)$), $g(y) > 0$, so we draw an arrow to the right on the horizontal axis. On any interval where $f(y) < h(y)$ (the graph of $f(y)$ is below $h(y)$), $g(y) < 0$, so we draw an arrow to the left on the horizontal axis.

EXAMPLE 6 Consider the differential equation

$$\frac{dy}{dt} = e^{-y} - y.$$

Determine how many equilibria the differential equation has and whether they are stable or unstable.

Solution To find the equilibria, we would need to solve the equation $g(y) = e^{-y} - y = 0$. This equation cannot be solved analytically, although it can be solved using a numerical method like bisection (from Section 3.5) or Newton's method (from Section 5.8). However, the graphical method allows us to find the number of equilibria and determine their stability without needing to know precisely where they are.

We cannot graph $g(y) = e^{-y} - y$ directly, but we can decompose it into two functions: $f(y) = e^{-y}$ and $h(y) = y$, whose graphs we should be familiar with (see Figure 8.21). We see from the graphs that the two functions intersect at one point, $\hat{y}$, and this point will be an equilibrium of the differential equation. In fact, since $f(y)$ is a decreasing function of y, and $h(y)$ is an increasing function of y, there can be only one intersection point. To create the vector field plot we observe that if $y < \hat{y}$ (left of the intersection point), $f(y) > h(y)$, so $g(y) = f(y) - h(y) > 0$, and we draw right pointing arrows on the horizontal axis to represent the fact that $y(t)$ is increasing. For $y > \hat{y}$ (i.e., right of the point of intersection), $f(y) < h(y)$, so $g(y) < 0$, and we draw left pointing arrows on the horizontal axis. Since the arrows point toward $\hat{y}$ on both sides, the equilibrium point is stable. •

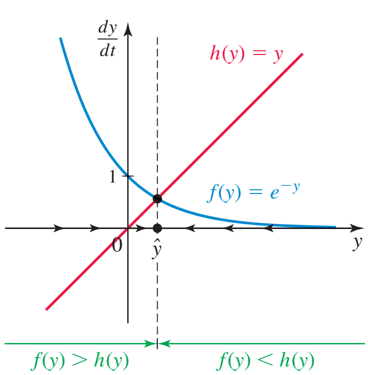


Figure 8.21 From the graphs of $f(y) = e^{-y}$ and $h(y) = y$ we can create the vector field plot for $dy/dt = f(y) - h(y)$.

The graphical approach is particularly helpful when analyzing how the locations and number of equilibria are affected by unknown coefficients that may appear in the differential equation.

EXAMPLE 7 **Sustainable Harvesting** A lake contains a managed population of fish. In the absence of any fishing, the population grows or decreases following the logistic equation with carrying capacity K, and limiting growth rate, r, as $N \to 0$. That is, absent fishing:

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

Suppose time, t, is measured in months and fishing removes fish from the lake at a rate of H fish per month. So, if fishing takes place, then

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) - H \tag{8.35}$$

where $H > 0$ is a constant.

Successful management of the fish stocks in the lake means finding a value of H for which the differential equation has a stable equilibrium with $N > 0$. Show that this stable equilibrium exists, provided $H < \frac{1}{4}rK$.

Solution Since the right-hand side of (8.35) is a quadratic polynomial in N, we could find any equilibria by solving the polynomial using the quadratic formula. However, the solution is a complicated function of r, K, and H, and it is hard to see what role the different parameters play. Instead we use the graphical approach, splitting $g(N) = rN(1-N/K)-H$ into functions $f(N) = rN(1-N/K)$ and $h(N) = H$. We sketched the function $f(N)$ already for Example 4. The function $h(N)$ is just a horizontal line. Two possible scenarios can occur, depending on the value of H. If H is small, then there are two points of intersection (N_1 and N_2). If H is large, then the curves do not intersect (see Figure 8.22).

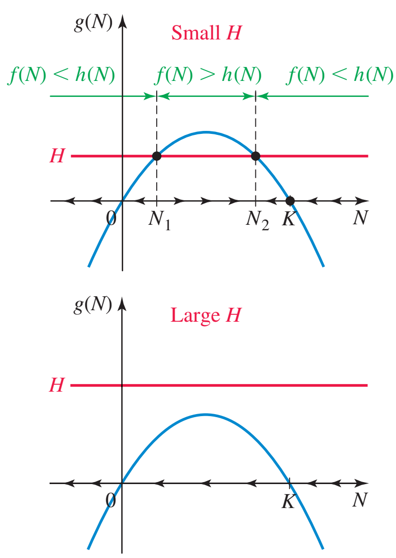


Figure 8.22 Vector field plot for $\frac{dN}{dt} = rN(1 - N/K) - H$

If H is small we can determine the stability of the two equilibria from the vector field plot. If $N_1 < N < N_2$ (i.e., between the equilibria), $f(N) > h(N)$, meaning that $dN/dt > 0$, so arrows are rightward. If $N < N_1$, or $N > N_2$ (i.e., left of N_1, or right of N_2 on the graph), $f(N) < h(N)$, meaning that $dN/dt < 0$, so the arrows point leftward. The arrows therefore point toward the equilibrium N_2 and away from N_1, making N_2 stable and N_1 unstable (Figure 8.22).

The equilibrium N_2 represents a sustainable state for the lake, since we expect the fish population to stay at this value, and not be strongly affected by perturbations. But this stable equilibrium does not exist if H is too large. If H is too large, then $f(N) < h(N)$ everywhere, so $dN/dt < 0$ for all values of N. What is the threshold value for H at which the two equilibria disappear? From the plot we can see that the curves stop intersecting if H exceeds the height of the parabola, $f(N)$ (i.e., if H exceeds the value f takes at its local maximum). To find the local maximum calculate

$$f'(N) = r(1 - 2N/K).$$

$f'(N) = 0$ when $N = K/2$ (from the plot we know that $f(N)$ has one local maximum and no local minima, so there is no need to check the type of extremum point). At $N = K/2$ the height of the function is:

$$f(K/2) = \frac{rK}{4}.$$

So, if $H < \frac{rK}{4}$, there is a stable equilibrium. If $H > rK/4$, then there is no stable equilibrium (in fact, $dN/dt < 0$ for all values of N, meaning that the fish population decreases over time).

If $H > rK/4$ in the above example, then our model predicts that $N(t)$ will decrease indefinitely (i.e., $N(t) \to -\infty$ as $t \to \infty$). This is an unrealistic prediction because $N(t)$ should never drop below 0 in a real population! The problem is that in our model we assume that the same number of fish is removed from the lake each month, independent of the size of the fish population, even when the fish population

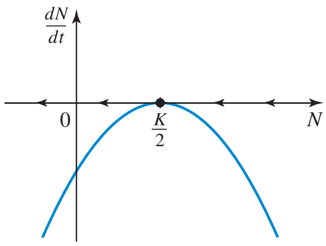


Figure 8.23 $\frac{dN}{dt} = rN(1-N/K)-K/4$ has a semi-stable equilibrium at $N = K/2$

is 0. In Problem 91 you will analyze a different fishing model, in which a fraction of fish is removed from the fish each month. In that model $N(t)$ will not drop below 0.

What happens if $H = rK/4$ in the above example? In that case $\frac{dN}{dt} < 0$ if $N \neq K/2$, and $\frac{dN}{dt} = 0$ if $N = K/2$. The vector field plot is shown in Figure 8.23. So the arrows showing whether $N(t)$ is increasing or decreasing point toward $N = K/2$ for $N > K/2$ and away from $N = K/2$ for $N < K/2$. Thus $N = K/2$ is stable to perturbations that increase N, but unstable to perturbations that decrease N. We call this kind of equation **semi-stable**: It is stable in one direction but unstable in the other.

Definition An equilibrium $\hat{y}$ is **semi-stable** if $y(t)$ returns to $\hat{y}$ if perturbed in one direction from $\hat{y}$ and diverges from $\hat{y}$ if perturbed in the other direction from $\hat{y}$.

Criterion for Semi-Stability If $g(\hat{y}) = 0$ and $\hat{y}$ is a local minimum or local maximum of $g(y)$, then $\hat{y}$ is a semi-stable equilibrium.

8.2.4 Sketching Solutions Using the Vector Field Plot

Using the information in the vector field plot, we can sketch the solutions $y(t)$ as a function of time t. To do this we make use of an important result.

Monotonicity of Solutions of Autonomous Differential Equations If $y(t)$ solves the differential equation $\frac{dy}{dt} = g(y)$, then $y(t)$ must be either constant or monotonic increasing or monotonic decreasing.

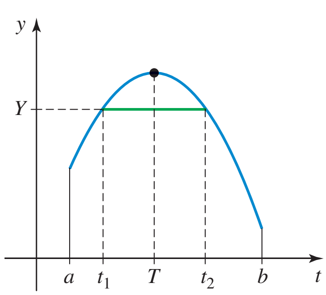


Figure 8.24 A sketch of a solution to $\frac{dy}{dt} = g(y)$ that has both increasing and decreasing intervals.

This result means that if $y(t)$ is initially increasing (i.e., if our initial condition is $y(0) = y_0$, and $g(y_0) > 0$, so that $\frac{dy}{dt}$ starts positive) then $y(t)$ will always be increasing; that is, so long as the solution exists, $\frac{dy}{dt}$ will remain positive. Similarly if $\frac{dy}{dt}$ starts off negative then it will remain negative so long as the solution exists.

We will justify, but not rigorously prove, the monotonicity result. Suppose that a solution of the differential equation $\frac{dy}{dt} = g(y)$ does not obey the monotonicity result. For example, it may start off increasing; but then switch to decreasing (the following argument can be adapted for solutions that start off decreasing but then switch to increasing; see Problem 67). For definiteness assume that this switch happens at a time $t = T$; that is, $dy/dt > 0$ over the interval $[0, T)$ and $\frac{dy}{dt} < 0$ for some interval (T, b) where $b > T$. So, $t = T$ must then be a local maximum of $y(t)$, and the solution locally will look like the sketch in Figure 8.24.

Since $y(t)$ must pass through the same values climbing up to $y(T)$ as those it passes through when it decreases again, there must be some pair of times t_1, and t_2 with $t_1 < T < t_2$, at which $y(t)$ takes the same value, say $y(t_1) = y(t_2) = Y$. However, since $t_1 < T$, $\frac{dy}{dt} > 0$ for $t = t_1$, and since $t_2 > T$, $\frac{dy}{dt} < 0$ for $t = t_2$. But $\frac{dy}{dt} = g(y)$, so $\frac{dy}{dt} \geq 0$ for $t = t_1$ requires $g(Y) \geq 0$, whereas $\frac{dy}{dt} < 0$ for $t = t_2$ requires $g(Y) \leq 0$. The only way that we may have both $g(y) \geq 0$ and $g(y) \leq 0$ is if $g(y) = 0$. In that case $y = Y$ is an equilibrium of the differential equation, meaning that $y(t) = Y$ for all $t > t_1$. This contradicts the assumption that $y(t)$ starts decreasing at $t = T$.

What else can we say about solutions? Suppose $y(t)$ is an increasing solution of the differential equation. Then as $t \to \infty$, $y(t)$ may either increase indefinitely (i.e., $y(t) \to +\infty$) or it must converge to a constant. That is $\lim_{t\to\infty} y(t) = \hat{y}$ for some constant $\hat{y}$. If $y(t)$ converges to a constant as $t \to \infty$, then $\frac{dy}{dt} \to 0$ and $g(y) \to g(\hat{y})$. For the two sides of the equation to be equal we must have $g(\hat{y}) = 0$, which implies that $\hat{y}$ is an equilibrium of the differential equation.

Let's put all of this information together.

Rules for Sketching the Solutions of a Differential Equation.

1. If $\frac{dy}{dt} = 0$ initially then $y(t)$ is a constant.
2. If $\frac{dy}{dt} > 0$ initially, $y(t)$ will grow to $+\infty$, or converge to the first equilibrium it meets.
3. If $\frac{dy}{dt} < 0$ initially, $y(t)$ will decrease to $-\infty$, or converge to the first equilibrium it meets.

Let's use these rules to sketch the solutions for some differential equations.

EXAMPLE 8 The Gompertz law for tumor growth predicts that the number of cells in a tumor, N, changes with time t, according to $dN/dt = g(N)$ where

$$g(N) = \begin{cases} aN\ln(b/N) & \text{if } N > 0 \\ 0 & \text{if } N = 0. \end{cases}$$

Use l'Hôpital's rule to show that $g(N)$ is continuous from the right at $N = 0$

where a, b are positive coefficients that depend on the type of tumor.

By drawing the vector field plot of the differential equation, sketch possible solutions.

Solution Previously we showed that $\hat{N} = b$ is an equilibrium of this differential equation. Adding $N = 0$ to the domain of the differential equation adds the additional equilibrium $\hat{N} = 0$. To draw the vector field plot we need to know where $g(N)$ is positive and where it is negative. Since $aN > 0$ for all $N > 0$, the sign of $g(N)$ is determined by the sign of $\ln(b/N)$.

$$\ln(b/N) > 0 \text{ if } N < b \quad \text{and} \quad \ln(b/N) < 0 \text{ if } N > b$$

$\ln x > 0$ if and only if $x > 1$

The vector field plot for the differential equation is shown in Figure 8.25. Both $N = 0$ and $N = b$ are constant solutions of the differential equation if $N(0) \in (0, b)$ (that is, the initial tumor size starts somewhere between 0 and b). Then we see from the vector field plot the solution is initially growing. It must then continue to grow, but it cannot grow past $N = b$, since $dN/dt < 0$ for $N > b$. Thus $N(t)$ will converge to b as $t \to \infty$. Similarly if $N(0) > b$ (i.e., if the starting tumor size is right of b on the vector field plot, then $N(t)$ will decrease until it again converges to b). We sketch the solutions in Figure 8.26.

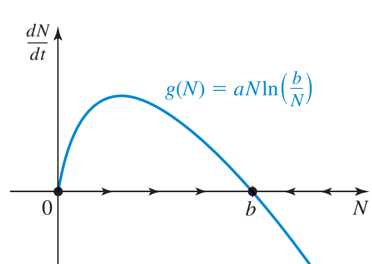


Figure 8.25 Vector field plot for the Gompertz tumor growth model.

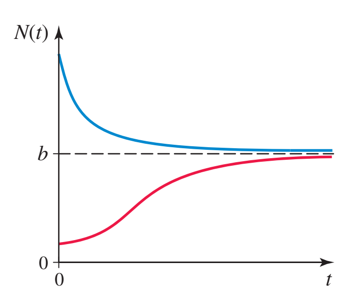


Figure 8.26 Solutions of the Gompertz tumor growth model.

EXAMPLE 9 **The Allee Effect** In Example 5 we modified the logistic equation to include the Allee effect; the tendency of small populations to shrink due to emigration and competition with other populations. The population size is modeled using a differential equation:

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

Sketch the possible solutions of this differential equation.

Solution Again we start from the vector field plot that was plotted in Figure 8.20. There are three possible constant solutions: $N = 0$, a, and K. If $0 < N(0) < a$, then from the vector field plot we see that $N(t)$ starts decreasing. It will continue to decrease, converging to 0 as $t \to \infty$. If $a < N(0) < K$, then $N(t)$ will increase, converging to $N = K$. Finally, if $N(0) > K$, then $N(t)$ will decrease, again converging to $N = K$. This behavior is consistent with the Allee effect as it was described in Example 5: Populations smaller than a decline to extinction, and populations larger than a converge to the habitat's carrying capacity. A sketch of the solutions is given in Figure 8.27. •

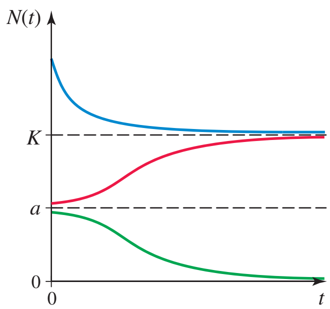


Figure 8.27 Solutions of the logistic equation with Allee effect.

8.2.5 Behavior Near an Equilibrium

The graphical method allows all equilibria of a differential equation to be identified and classified as either stable or unstable. We also showed that the stability of an equilibrium $\hat{y}$ can also be determined from the sign of $g'(\hat{y})$. What the graphical method cannot tell us is how quickly the solution converges back to $\hat{y}$, or diverges away from $\hat{y}$, if perturbed from $\hat{y}$. Put another way, are some equilibria more stable than others? To determine the rate of convergence (or divergence) we analyze the behavior of the differential equation if y is slightly perturbed from $\hat{y}$.

Consider a differential equation

$$\frac{dy}{dt} = g(y) \tag{8.36}$$

and an equilibrium $\hat{y}$. We consider a small perturbation of the solution away from the equilibrium $\hat{y}$; we express this perturbation as

$$y(t) = \hat{y} + Y(t)$$

where $Y(t)$ is small and may be either positive or negative. The function $Y(t)$ measures how far the solution has been perturbed from the equilibrium. The perturbed initial condition is $y(0) = \hat{y} + Y(0)$. Then substituting for $y(t)$:

$$\frac{dy}{dt} = \frac{d}{dt}(\hat{y} + Y) = \frac{dY}{dt} \qquad \hat{y} \text{ is a constant, so } \frac{d\hat{y}}{dt} = 0.$$

Substituting for y into both sides of (8.36), we find that:

$$\frac{dY}{dt} = g(\hat{y} + Y)$$

If Y is sufficiently small, we can approximate $g(\hat{y}+Y)$ by its linear approximation. The linear approximation of $g(y)$ close to $y = \hat{y}$ is given by

$$g(y) \approx g(\hat{y}) + (y - \hat{y})g'(\hat{y}) = (y - \hat{y})g'(\hat{y}) \qquad g(\hat{y}) = 0$$

Or, in terms of the perturbation variable, Y:

$$g(\hat{y} + Y) \approx Yg'(\hat{y})$$

If we set $\lambda = g'(\hat{y})$ then the linear approximation of (8.36) takes the form:

$$\frac{dY}{dt} = \lambda Y$$

This equation has the solution

$$Y(t) = Y(0)e^{\lambda t} \tag{8.37}$$

The value $Y(0)$ represents our initial perturbation; that is, how far the solution starts from the equilibrium. According to Equation (8.37), if $Y(0)$ is non-zero, then $Y(t)$ will grow exponentially (i.e., diverge from $\hat{y}$ if $\lambda > 0$), or will decay exponentially (i.e., converge back to $\hat{y}$, if $\lambda < 0$). Since we have defined $\lambda = g'(\hat{y})$, this analysis reproduces the derivative test (including the ambiguity of what must occur if $\lambda = 0$). However it also tells us the role that $g'(\hat{y})$ plays in controlling how quickly the solution converges back to $\hat{y}$. If $g'(\hat{y}) < 0$, then perturbations decay more quickly for larger values of $|g'(\hat{y})|$. If $g'(\hat{y}) > 0$, then perturbations grow faster for larger values of $g(\hat{y})$. Because the parameter λ controls the behavior of small perturbations, we need a name for it: We will call it the **eigenvalue** of the equilibrium. The main importance of the eigenvalue method for determining stability of equilibria is that it can also be used to study systems of differential equations, a problem that we will study at length in Chapter 11.

EXAMPLE 10 Find the equilibria and associated eigenvalues for each of the following differential equations. Then use the eigenvalue to determine whether each equilibrium is stable or unstable.

(a) $\dfrac{dy}{dt} = y(1-y)$ **(b)** $\dfrac{dz}{dt} = \frac{1}{z^2} - \frac{1}{z},\ z > 0$ **(c)** $\dfrac{dy}{dx} = \ln y - 2,\ y > 0$

Solution **(a)** The function $g(y) = y(1-y)$ has roots at $\hat{y} = 0$ and $\hat{y} = 1$. $g'(y) = 1 - 2y$, so the corresponding eigenvalues are $g'(0) = 1$ and $g'(1) = -1$. Since $g'(0) > 0$, $\hat{y} = 0$ is unstable, while $g'(1) < 0$ implies that $\hat{y} = 1$ is stable.

(b) The function $g(z) = \frac{1}{z^2} - \frac{1}{z}$ has roots where:

$$\frac{1}{z^2} - \frac{1}{z} = 0 \Rightarrow 1 - z = 0 \qquad \times \text{ both sides by } z^2$$

$$z = 1.$$

$$g'(z) = -\frac{2}{z^3} + \frac{1}{z^2}$$

so the eigenvalue for this equilibrium is: $g'(1) = -2 + 1 = -1$. Since $g'(1) < 0$, the equilibrium $\hat{z} = 1$ is stable.

(c) The function $g(y) = \ln y - 2$ has roots when $\ln y - 2 = 0$ (i.e., the equilibrium is $\hat{y} = e^2$). $g'(y) = \frac{1}{y}$, so the eigenvalue is $g'(e^2) = e^{-2}$. Since $g'(e^2) > 0$, the equilibrium $\hat{y} = e^2$ is unstable. •

Section 8.2 Problems

8.2.1

Find the equilibria of the following differential equations.

1. $\dfrac{dy}{dt} = y(y^2 - 1)$

2. $\dfrac{dy}{dt} = y^3 + y$

3. $\dfrac{dx}{dt} = x^2 - 3x + 2$

4. $\dfrac{dx}{dt} = 6 + 5x + x^2$

5. $\dfrac{dy}{dt} = \dfrac{y-2}{y+1}$

6. $\dfrac{dy}{dt} = \dfrac{y-1}{y^2+1}$

7. $\dfrac{dx}{dt} = x^8 - 1$

8. $\dfrac{dy}{dt} = y^{1/3} - 1$

9. $\dfrac{dN}{dt} = Ne^{-N}$

10. $\dfrac{dN}{dt} = N \ln N,\ N > 0$

11. $\dfrac{dN}{dt} = \sin N$

12. $\dfrac{dN}{dt} = N \cos 2N$

8.2.2, 8.2.3

For Problems 13–28 make vector field plots of each of the differential equations. Find any equilibria of each differential equation and use your vector field plot to classify whether each equilibrium is stable or unstable.

13. $\dfrac{dy}{dt} = y - 1$

14. $\dfrac{dy}{dt} = 2 - y$

15. $\dfrac{dy}{dt} = 4 - y^2$

16. $\dfrac{dy}{dt} = y(y-2)$

17. $\dfrac{dy}{dt} = y^2 - y$

18. $\dfrac{dy}{dt} = y^2 - 2y - 8$

19. $\dfrac{dx}{dt} = x - x^3$

20. $\dfrac{dx}{dt} = x^5 - x$

21. $\dfrac{dN}{dt} = N \ln(2/N),\ N > 0$

22. $\dfrac{dN}{dt} = N^3 e^{-N}$

23. $\frac{dx}{dt} = \frac{x^2 - x}{x^2 + 1}$

24. $\frac{dx}{dt} = \frac{x + 1}{x - 1}, x \neq 1$

25. $\frac{dx}{dt} = \frac{x}{x + 1}, x \neq -1$

26. $\frac{dx}{dt} = \frac{x + 1}{x}, x \neq 0$

27. $\frac{dS}{dt} = \frac{1}{S^3} - \frac{1}{S}, S > 0$

28. $\frac{dS}{dt} = \frac{1}{S} - \frac{1}{S^5}, S > 0$

In Problems 29–38, by breaking down each equation into two parts that you can sketch, determine how many equilibria each differential equation has, and classify them as stable or unstable. You do not need to determine the location of the equilibria.

29. $\frac{dy}{dt} = e^y - (1 - y)$

30. $\frac{dy}{dt} = \ln y - e^{-y} \quad y > 0$

31. $\frac{dx}{dt} = \frac{1}{x} - \frac{x}{x + 1} \quad x > 0$

32. $\frac{dx}{dt} = 3e^{-x^2} - x^2$

33. $\frac{dx}{dt} = \frac{1}{2} - \frac{x^2}{x^2 + 1}$

34. $\frac{dx}{dt} = x^2 - \frac{1}{x + 1} \quad x \neq -1$

35. $\frac{dN}{dt} = N^2 - N + 1 \quad N > 0$

36. $\frac{dN}{dt} = 1 - N - N^3$

37. $\frac{dy}{dx} = y + y^5 - 1$

38. $\frac{dy}{dx} = y^4 + y^3 - 1$

In Problems 39–48 you should treat h as a constant. For what values of h (if any) does each equation have equilibria? Use a graphical argument to show which of the equilibria (if any) are stable.

39. $\frac{dy}{dt} = y(1 - y) - h$

40. $\frac{dy}{dt} = y - h$

41. $\frac{dy}{dx} = y^2 - h$

42. $\frac{dy}{dx} = (y - 1)(y + 3) - h$

43. $\frac{dy}{dx} = (y - 2)(y + 4) + h$

44. $\frac{dy}{dx} = y^3 - y - h$

45. $\frac{dx}{dt} = x^2 - hx$

46. $\frac{dx}{dt} = x^3 - hx$

47. $\frac{dx}{dt} = x(x^2 - 1) - h$

48. $\frac{dx}{dt} = hx - x^3$

For Problems 49–56 determine whether the equilibrium at $x = 0$ is stable, unstable, or semi-stable.

49. $\frac{dx}{dt} = x^3$

50. $\frac{dx}{dt} = -x^5$

51. $\frac{dx}{dt} = x^4$

52. $\frac{dx}{dt} = x^3 - x^5$

53. $\frac{dx}{dt} = x^3 + x^4$

54. $\frac{dx}{dt} = x^2 - x^3$

55. $\frac{dx}{dt} = \frac{x^3}{x - 1}$

56. $\frac{dx}{dt} = x^3 e^{-x}$

8.2.4

For Problems 57–66 draw the vector field plot of the differential equation. Then, using the given initial conditions, sketch the solutions (i.e., draw a graph showing the dependent variable as a function of the independent variable).

57. $\frac{dy}{dt} = 3y - 2$
(a) $y(0) = 2$, **(b)** $y(0) = 0$.

58. $\frac{dy}{dt} = 1 - y$
(a) $y(0) = 2$, **(b)** $y(0) = -1$.

59. $\frac{dy}{dt} = y(1 - y)$
(a) $y(0) = 0$, **(b)** $y(0) = 1/2$, **(c)** $y(0) = 1/4$, **(d)** $y(0) = 2$.

60. $\frac{dy}{dt} = y^2 - 1$
(a) $y(0) = -1$, **(b)** $y(0) = -1/2$, **(c)** $y(0) = 1/2$, **(d)** $y(0) = 2$.

61. $\frac{dy}{dt} = (y + 3)(1 - y)$
(a) $y(0) = -1$, **(b)** $y(0) = -1/2$, **(c)** $y(0) = -2$, **(d)** $y(0) = 2$.

62. $\frac{dy}{dt} = (y + 1)(y + 3)$
(a) $y(0) = -3/2$, **(b)** $y(0) = -5/2$, **(c)** $y(0) = 0$, **(d)** $y(0) = -5$.

63. $\frac{dN}{dt} = N(N - 1)(5 - N)$
(a) $N(0) = 1$, **(b)** $N(0) = 1/2$, **(c)** $N(0) = 3/2$, **(d)** $N(0) = 7$.

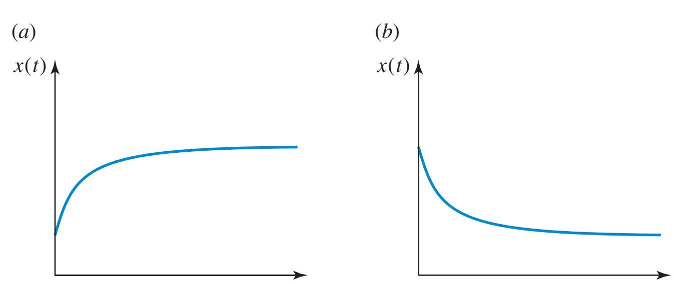
64. $\frac{dN}{dt} = (N - 1)(N + 1)(N - 4)$
(a) $N(0) = 0$, **(b)** $N(0) = 2$, **(c)** $N(0) = 6$, **(d)** $N(0) = -2$.

65. $\frac{dy}{dx} = (y - 1)(y - 2)(y - 5)$
(a) $y(0) = 0$, **(b)** $y(0) = 4$, **(c)** $y(0) = 3/2$, **(d)** $y(0) = 6$.

66. $\frac{dy}{dx} = -y - y^3$
(a) $y(0) = 0$, **(b)** $y(0) = 1$, **(c)** $y(0) = 2$, **(d)** $y(0) = 3$.

67. **Monotonicity of Solutions** One of the key ideas for sketching solutions from vector field plots is that a solution curve must be monotonic; that is, $x(t)$ is either increasing or decreasing or constant but cannot switch from one behavior to another. We showed that a solution $x(t)$ could not start by increasing and then switch to decreasing. Suppose that $x(t)$ is a solution of the differential equation $\frac{dx}{dt} = g(x)$ and that $x(t)$ starts off decreasing with time. Show that $x(t)$ cannot switch to increasing.

68. **Monotonicity of Solutions** Figure 8.28 shows the graphs of some functions $x(t)$. Which of these functions could **not** arise as solutions of a differential equation $dx/dt = f(x)$ for some continuous function $f(x)$?



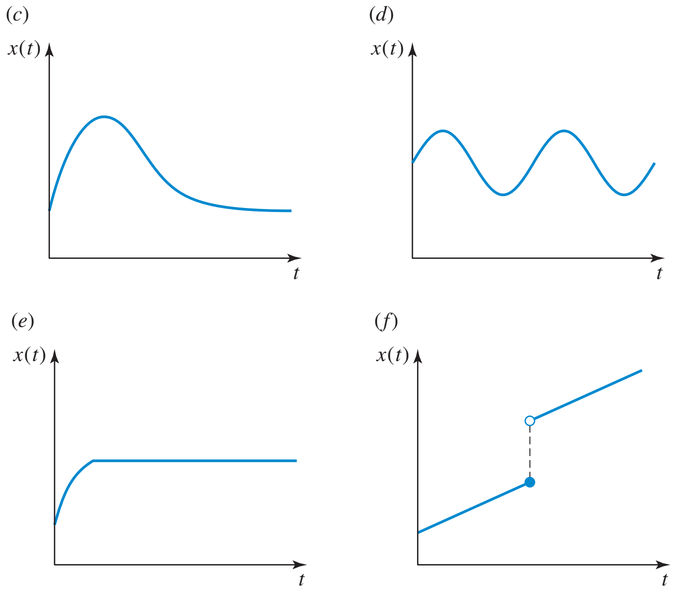


Figure 8.28 Functions $x(t)$ for Problem 68.

69. Curvature of Solutions In Section 5.6 we showed how knowing the curvature of a function (that is, whether it is concave up or concave down) can assist in drawing the graph of the function. Suppose that $x(t)$ solves the differential equation $dx/dt = f(x)$, where $f(x)$ is a differential function and $f'(x)$ is continuous.

(a) By applying the chain rule, show that

$$\frac{d^2x}{dt^2} = f'(x)f(x)$$

(b) If $f(x) > 0$, explain why a solution $x(t)$ is concave up if $f'(x) > 0$ and concave down if $f'(x) < 0$.

(c) What about if $f(x) < 0$? When is the solution $x(t)$ concave up and when is it concave down?

(d) Using the results of (b) and (c), show on the vector field plot in Figure 8.29 the intervals on which $x(t)$ will be concave up and the intervals on which $x(t)$ will be concave down.

(e) What happens to $x(t)$ when x passes through one of the local extrema of $f(x)$?

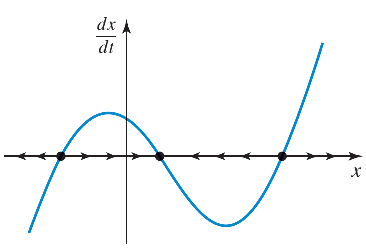


Figure 8.29 Vector field plot for $\frac{dx}{dt} = f(x)$ in Problem 69(d).

70. Logistic Equation with Allee Effect

You must solve Problem 69 before solving this problem.

The vector field plot for a model of density dependent growth of a population is shown in Figure 8.30.

(a) Label on the plot the intervals on which $N(t)$ is concave up.

(b) Label also the intervals on which $N(t)$ is concave down.

(c) What happens when $N = b$ or $N = c$?

(d) Suppose that $N(0) \in (a, c)$. Explain why the solution $N(t)$ will look like the sketch in Figure 8.31.

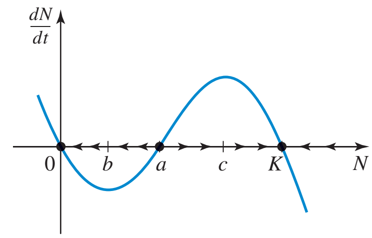


Figure 8.30 Vector field plot for a population modeling density dependent growth and the Allee effect (see Problem 70). Local extrema b and c are shown as well as equilibria $N = 0, a, K$.

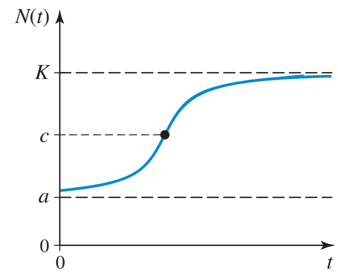


Figure 8.31 Sketch of the solution if $N(0) \in (a, c)$

(e) Draw sketches of the solution if (i) $N(0) \in (0, b)$, (ii) $N(0) \in (b, a)$, (iii) $N(0) \in (c, K)$, (iv) $N(0) \in (K, \infty)$

In Problems 71–76 you are given graphs of the function $f(y)$ for a differential equation $dy/dt = f(y)$. (See Figure 8.32.) You are also given initial conditions y_0 or y_1 (shown on the plot). For each graph make a sketch of the solution $y(t)$ against t for (a) $y(0) = y_0$, (b) $y(0) = y_1$. If you have solved Problem 69, your sketch can show where $y(t)$ is concave up and where $y(t)$ is concave down.

71. The function shown in 8.32a.

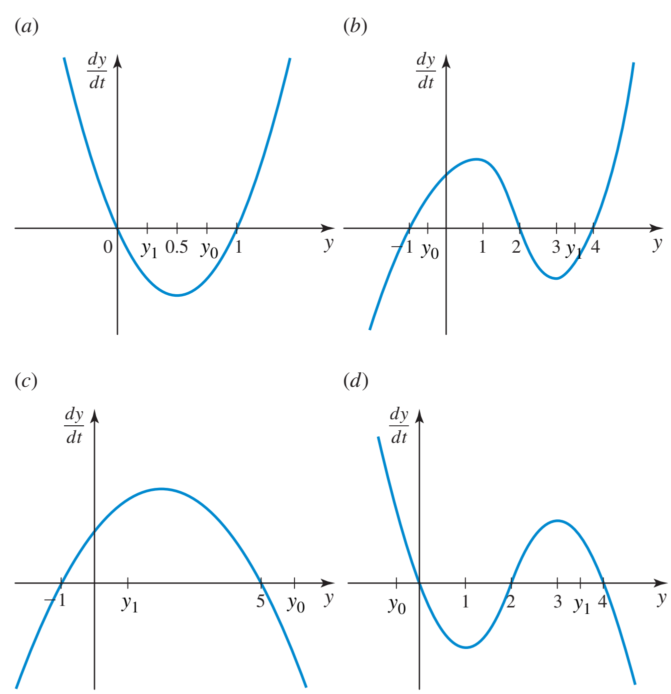
72. The function shown in 8.32b.

73. The function shown in 8.32c.

74. The function shown in 8.32d.

75. The function shown in 8.32e.

76. The function shown in 8.32f.



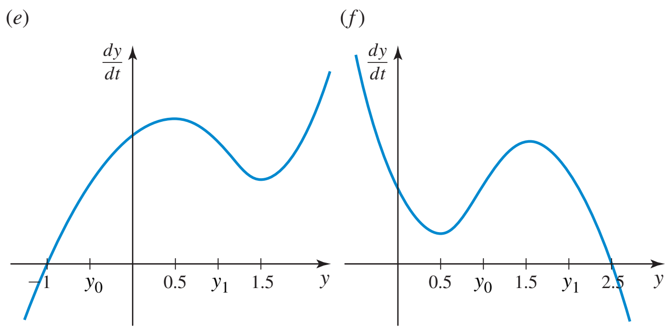


Figure 8.32 Graphs of $f(y)$ for Problems 71–76. The initial conditions $y(0) = y_0$ or $y(0) = y_1$ are marked on each graph.

8.2.5

For Problems 77–88 find all equilibria, and, by calculating the eigenvalue of the differential equation, determine which equilibria are stable and which are unstable.

77. $\frac{dy}{dt} = 2 - 3y$ **78.** $\frac{dy}{dt} = y - 2$

79. $\frac{dy}{dt} = y(2-y)(y-3)$ **80.** $\frac{dy}{dt} = y(y-1)(y-2)$

81. $\frac{dN}{dt} = N \ln\left(\frac{2}{N}\right) \qquad N > 0$

82. $\frac{dN}{dt} = \frac{N-1}{N+1} \qquad N \geq 0$

83. $\frac{dy}{dx} = \frac{y^2 - y}{y^2 + 1}$ **84.** $\frac{dy}{dx} = \frac{1}{y^3} - \frac{1}{y}, \quad y > 0$

85. $\frac{dx}{dt} = xe^{-x}$ **86.** $\frac{dx}{dt} = e^{-x} - e^{-2x}$

87. $\frac{dx}{dt} = hx - x^2$, where h is a constant and

(a) $h > 0$, **(b)** $h < 0$

88. $dx/dt = hx - x^3$, where h is a constant and

(a) $h > 0$, **(b)** $h < 0$

89. Effect of Predation on Population Growth Suppose that $N(t)$ denotes the size of a population at time t. The population evolves according to the logistic equation, but in addition, predation reduces the size of the population so that the rate of change is given by

$$\frac{dN}{dt} = g(N) \tag{8.38}$$

where

$$g(N) = N\left(1 - \frac{N}{50}\right) - \frac{9N}{5+N}.$$

The first term on the right-hand side describes the logistic growth; the second term describes the effect of predation.

(a) Make the vector field plot for this differential equation.

(b) Find all equilibria of (8.38).

(c) Use your vector field plot in (a) to determine the stability of the equilibria you found in (b).

(d) Repeat your analysis from part (c) but now use the method of eigenvalues to determine the stability of the equilibria you found in (b).

90. Sustainable Harvesting Suppose that a fish population evolves according to the logistic equation and that a fixed number of fish per unit time are removed. That is,

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

Assume that $r = 2$ and $K = 1000$.

(a) Find possible equilibria, and discuss their stability when $H = 100$.

(b) What is the maximal harvesting rate that maintains a positive population size?

91. Sustainable Harvesting Suppose that a fish population evolves according to a logistic equation and that fish are harvested at a rate proportional to the population size. If $N(t)$ denotes the population size at time t, then

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

Assume that $r = 2$ and $K = 1000$.

(a) Find possible equilibria and use the graphical approach to discuss their stability, when $h = 0.1$.

(b) Show that if $h < r = 2$, then there is a nontrivial equilibrium. Find the equilibrium.

(c) Use **(i)** the eigenvalue approach and **(ii)** the graphical approach to analyze the stability of the equilibrium you found in (b).

92. Growth of a Tumor Gompertz's equation can be used to model the growth of some solid tumors. It predicts that the number of cells $N(t)$ in a tumor will grow over time according to:

$$\frac{dN}{dt} = g(N) \tag{8.39}$$

where

$$g(N) = \begin{cases} 0 \text{ if } N = 0 \\ aN\ln(b/N) \quad \text{if } N > 0 \end{cases}$$

where a and $b > 0$ are constants

(a) Find all of the equilibria of (8.39).

(b) Can you use the eigenvalues of g to calculate the stability of each equilibrium? If not, why not?

93. Logistic Equilibrium with Allee Effect A population whose growth is affected by the Allee effect is modeled using the differential equation:

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

where r, a, k are all positive constants and $a < K$.

The equilibria of this equation are $N = 0$, $N = a$, and $N = K$. Use the eigenvalue method to classify whether each of these equilibria is stable or unstable.

8.3 Differential Equation Models

In sections 8.1.2 and 8.2 we learned methods to solve autonomous differential equations, and to identify features of their solutions (e.g., equilibria) without solving the differential equation. We will now show how these methods can be used to study many different mathematical models.

8.3.1 Compartment Models

This example is adapted from ***DeAngelis (1992)***. A common question that many models set out to answer is how matter, nutrients, energy, or drugs move through a system like the human body. One way to model this movement is to treat the system as a series of compartments; for example, we could treat the blood as one compartment and organs as separate compartments, and model the flows of a drug between those compartments.

Let's start with a physically motivated example. A single compartment of fixed volume, V, of water contains a concentration $C(t)$ of a particular solute (e.g., salt). New water is added to the volume at a rate q. That is, in one unit of time a volume q flows into the compartment. At the same time water also flows out of the compartment. Let's assume that the rate of outflow is also q (that is, a volume, q, of water is removed in one unit of time; see Figure 8.33). We denote by $C(t)$ the concentration of the solution in the compartment at time t. Then the total mass of the solute in the compartment is $C(t)V$, where V is the volume of the compartment. For instance, if the concentration of the solution is 2 grams per liter and the volume of the compartment is 10 liters, then the total mass of the solute in the compartment is $2\ \text{g liter}^{-1} \times 10$ liters, which is equal to 20 g. Since the inflow and outflow rates are the same, the total volume of water in the compartment will remain constant with time. However, if the concentration of solute in the water flowing into the compartment is different from the concentration flowing out of the compartment, then $C(t)$ can change over time.

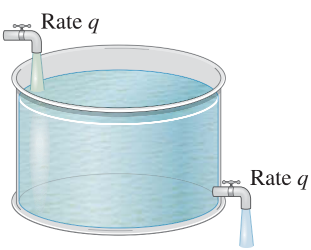


Figure 8.33 Inflows and outflows of water and solute in a single compartment.

The concentration in the outflow should be the same as the concentration in the compartment, $C(t)$, provided the solute is mixed evenly through the compartment. Suppose that the concentration in the inflow is a constant, C_I. Like q, C_I is an arbitrary constant that allows us to tailor the model to different problems. To calculate $C(t)$ we determine the rate of change of solute in the compartment.

$$\begin{array}{c}\text{Rate of change of}\\ \text{amount of solute}\\ \text{in compartment}\end{array} = \begin{array}{c}\text{Rate at which solute}\\ \text{flows in}\end{array} - \begin{array}{c}\text{Rate at which solute}\\ \text{flows out}\end{array}$$

In one unit of time a volume q of water flows in, so an amount qC_I of solute flows in: The *rate* at which solute flows in is qC_I. Similarly the rate at which solute flows out is $qC(t)$. On the left-hand side, since the total amount of solute in the compartment is $C(t) \cdot V$, the rate of change is $\frac{d}{dt}(C(t) \cdot V)$. Putting these ingredients together, we obtain:

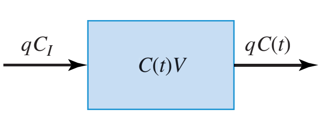


Figure 8.34 Flow diagram for the single-compartment model.

$$\frac{d}{dt}(CV) = qC_I - qC(t)$$

$$V\frac{dC}{dt} = q(C_I - C) \qquad V \text{ is a constant}$$

Figure 8.34 is a diagram showing the rates of flow of solute into and out of the compartment.

To analyze this equation we rearrange it into:

$$\frac{dC}{dt} = \frac{q}{V}(C_I - C) \tag{8.40}$$

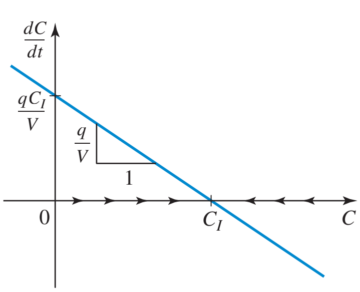


Figure 8.35 Vector field plot for the single compartment model.

and draw the vector field plot for this differential equation (Figure 8.35). Since C represents the concentration of solute we only need to consider $C \geq 0$. The graph of this equation is a straight line with slope $(-q/V)$ that meets the horizontal axis at $C = C_I$.

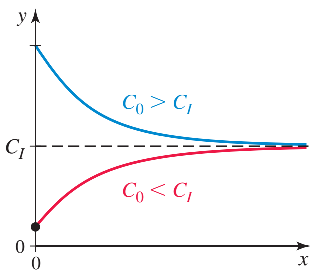


Figure 8.36 The solution curves for the single-compartment model for different values of C_0.

The equilibrium concentration is $C = C_I$, and all arrows point toward $C = C_I$, so whatever the starting concentration is, $C(t)$ will converge to C_I over time. Some possible solutions are shown in Figure 8.36 for different values of the initial concentration, C_0.

If $C(0) = C_0$, then we can solve the differential equation using separation of variables. We will just give the solution here (you will derive the solution in Problem 1).

$$C(t) = C_I - (C_I - C_0)e^{-qt/V}$$

Since $e^{-qt/V} \to 0$ as $t \to \infty$ the solution confirms that $\lim_{t\to\infty} C(t) = C_I$.

What role do the parameters q and V play? Notice from the exact solution that the exponential decay part of the solution has the form $e^{-qt/V}$. From the form of the exponential we see that C converges to C_I more rapidly if q/V is large, that is, if either q is large (fast inflow and outflow) or V is small (i.e., the compartment is small). In fact we do not need to solve the differential equation to see how q and V affect the rate at which $C(t)$ converges to C_I. Remember that, in general, near the equilibrium the solution decays with decay rate $e^{\lambda t}$ where λ is the eigenvalue. If $g(C) = \frac{q}{V}(C_I - C)$, then for Equation (8.40) $\lambda = g'(C_I) = -\frac{q}{V}$.

8.3.2 An Ecological Model

The habitats of many organisms are fragmented: Fungi, for example, may only live on certain hosts, or may only feed on logs that are spread through a forest. Ants in the same forest have a limited number of potential nest sites, depending on the availability of food and building materials and also the danger of flooding, predators, and so on. Thus the ants or fungi in a forest are broken down into small subpopulations; for example, the fungi on a single log, or the ants in a single nest. The number of organisms at each of these sites will typically vary greatly over time; each subpopulation may last only a few years before going extinct. Yet the total number of ants or fungi in the forest may remain quite stable over time: When an ant nest goes extinct, the site that it occupied is quickly recolonized by ants from nearby sites. We will study a mathematical model that was created by ***Levins (1969)*** to explain why under some conditions the total number of organisms in a particular habitat can be stable, though individual subpopulations don't last very long.

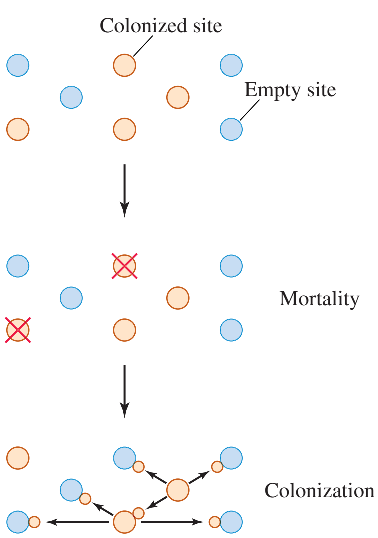


Figure 8.37 The number of occupied sites changes due to mortality and colonization.

In Levins' model, we start by imagining that the habitat is divided into a large number of sites that could contain subpopulations of the organism (e.g., potential nest sites). We will keep track of how the proportion of sites, that are occupied, p, changes with time, t. The number of occupied sites changes due to two processes.

Mortality: The subpopulation occupying a particular site may die out. Assume that in one unit of time, a fraction m of the occupied sites die out (the sites themselves remain but they cease to be occupied). The constant m is called the **mortality rate**.

Colonization: Each subpopulation sends out propagules to nearby sites. A propagule could be, for example, a fungal spore, or an insect scout that is sent out to find new potential nest sites and to colonize those sites. Suppose that in one unit of time each subpopulation sends out c propagules; these propagules are sent out indiscriminately, landing at sites that are already occupied as well as sites that are not. The constant c is called the colonization rate. The two steps, mortality and colonization, are illustrated in Figure 8.37.

Suppose that there are N sites that could be occupied by subpopulations, and that at a time t, a proportion $p(t)$ of the sites are occupied. Then the rate of change of the number of occupied sites can be written as a word equation:

$$\begin{matrix}\text{Rate of change} \\ \text{of occupied sites}\end{matrix} = -\begin{matrix}\text{Rate of loss by} \\ \text{mortality}\end{matrix} + \begin{matrix}\text{Rate of colonization} \\ \text{of new sites}\end{matrix}$$

If a proportion of p sites are occupied, the number of subpopulations is Np. In one unit of time, a fraction m of subpopulations will die out. The total number of subpopulations that die out is therefore:

$$\begin{matrix}\text{Rate of loss} \\ \text{by mortality}\end{matrix} = \begin{matrix}\text{Number of} \\ \text{subpopulations}\end{matrix} \times \begin{matrix}\text{Fraction that} \\ \text{die}\end{matrix} = Np \times m = mNp$$

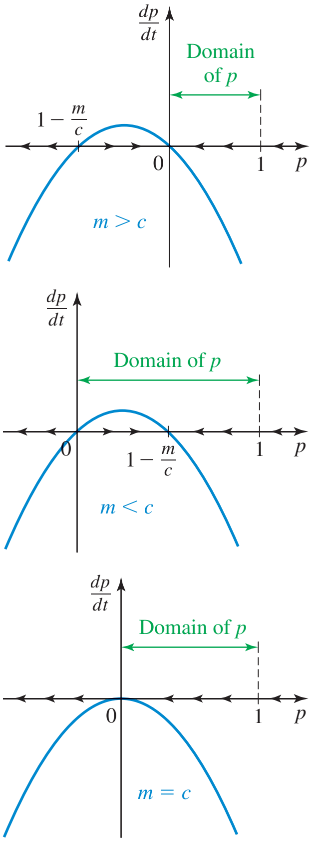


Figure 8.38 The vector field plot for the Levins' model of population growth in a patchy environment depends on whether $m \geq c$ or $m < c$.

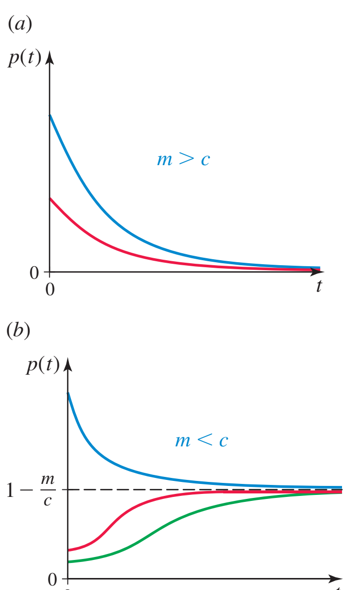


Figure 8.39 (a) If $m \geq c$, then $p(t)$ converges to 0. (b) If $m < c$, if $p(0) \neq 0$, then $p(t)$ converges to $1 - m/c$.

The rate of colonization is slightly harder to calculate. If in one unit of time each occupied site sends out c propagules, then since there are Np occupied sites, in one unit of time a total of cNp propagules (number of occupied sites $\times$ number of propagules per site) will be sent out. Only propagules that land on unoccupied sites will start new subpopulations. Since a fraction p of sites are occupied, a fraction p of propagules sent out will land on already occupied sites, and a fraction $1 - p$ will land on unoccupied sites and start new subpopulations. So the rate of colonization is:

$$\begin{array}{c}\text{Rate of}\\ \text{colonization}\end{array} = \begin{array}{c}\text{Number of propagules}\\ \text{released in}\\ \text{unit time}\end{array} \times \begin{array}{c}\text{Fraction of}\\ \text{propagules that}\\ \text{start new subpopulations}\end{array} = cNp(1-p)$$

So, putting everything together, we obtain:

$$\frac{d}{dt}(Np) = -mNp + cNp(1-p)$$

$$\frac{dp}{dt} = -mp + cp(1-p) \qquad \text{Divide by } N \qquad (8.41)$$

Although (8.41) can be solved using a separation of variables, we will not do that here. Instead we will use the methods from Section 8.2 to analyze how the solution depends on the parameters m and c.

Let $g(p) = -mp + cp(1-p)$. Then the equilibria of (8.41) correspond to values $\hat{p}$ for which $g(\hat{p}) = 0$, that is,

$$\hat{p}((c-m) - c\hat{p}) = 0 \qquad \text{Factoring } g(\hat{p})$$

This equation has two solutions, $\hat{p} = 0$ and $\hat{p} = 1 - m/c$. However, p represents the proportion of sites that are occupied, and this fraction must be between 0 and 1. Although the equilibrium $\hat{p} = 0$ always lies in this interval, $p = 1 - m/c \geq 0$ only if $m/c \leq 1$, that is, if $m \leq c$ (since $m \geq 0$ and $c \geq 0$, we certainly have $\hat{p} \leq 1$).

Thus there are three cases: If $m > c$, then there is only one equilibrium in $[0, 1]$. If $m < c$, then there are two distinct equilibria in $[0, 1]$. There is also a borderline case $m = c$ in which the two roots of $g(p) = 0$ coincide at $\hat{p} = 0$.

We draw the vector field plots corresponding to each of the cases in Figure 8.38. In the plots we have drawn $g(p)$ over a large enough interval to show both roots, but you should remember that the domain of p is always $[0, 1]$.

From the vector field plots we see that if $m \geq c$, then all solutions converge to 0, whatever the initial condition on p (remember $0 \leq p(0) \leq 1$). But if $m < c$, then the equilibrium $\hat{p} = 1 - m/c$ is stable, and $\hat{p} = 0$ is unstable. Thus $p(t) = 0$ and $p(t) = 1 - m/c$ are constant solutions. But if $p(0) > 0$, then $p(t)$ will either grow or decrease monotonically until it converges to $1 - m/c$. We show the different behaviors of $p(t)$ in these two cases in Figure 8.39.

We can interpret the two types of behaviors biologically. When $m \geq c$, mortality removes subpopulations faster than they can be replaced by colonization, so the subpopulations die out faster than sites can be recolonized; eventually all subpopulations die out. If $m < c$, then although subpopulations are continuously lost, the sites are then recolonized, maintaining a stable number of subpopulations. At the stable equilibrium the proportion of sites that are occupied is $\hat{p} = 1 - m/c$. This is always less than 1 (i.e., colonization cannot keep all sites filled), but is closer to 1 if m/c is small; that is, if either the mortality rate, m, is small, or the colonization rate, c, is large.

8.3.3 Modeling a Chemical Reaction

Now we will revisit the mass action laws that were introduced in Section 4.2. Suppose we want to build a mathematical model for a chemical reaction in which two reactants (A and B) combine to form a product (C).

$$A + B \to C$$

In Section 4.2 we argued that this reaction will proceed until either A or B is completely used up. What if the reaction is reversible; that is, C can also break down spontaneously back into A and B? We represent the reaction and the two rates (for the forward and backward reactions) by:

$$A + B \underset{k_C}{\overset{k_{AB}}{\rightleftarrows}} C$$

where k_{AB} is the rate constant for the reaction between A and B, and k_C is the rate at which C breaks down.

Suppose that the initial amount of A present is a (this could be a total amount; e.g., measured in moles, or concentration) and the amount of B initially is b. Let x denote the amount of C present at time t. Then we derive a mathematical equation for the rate of change of x from the word equation:

$$\begin{array}{c}\text{Rate of change}\\ \text{of } x\end{array} = \begin{array}{c}\text{Rate at which}\\ C \text{ is produced}\\ \text{by } A+B\to C\end{array} - \begin{array}{c}\text{Rate at which } C \text{ is consumed}\\ \text{by } C \to A+B\end{array}$$

The $A + B \to C$ reaction produces C at rate $k_{AB}[A][B]$ (remember from Section 4.2 that $[A]$ denotes the amount of A present). But to produce an amount x of C, the same amount of A must be consumed, so $[A] = a-x$. Similarly $[B] = b-x$. So C is produced at a rate $k_{AB}(a-x)(b-x)$.

The reverse reaction consumes C at a rate $k_C x$. So

$$\frac{dx}{dt} = k_{AB}(a-x)(b-x) - k_C x.$$

To find the equilibria of this differential equation we can set $g(x) = k_{AB}(a-x)(b-x) - k_C x$ and use the quadratic formula to find the roots of $g(x) = 0$. However, even without knowing the precise values of the roots we can use a vector field plot to determine how many equilibria there are and to determine their stability. First, though, we need to establish the domain of x for this plot. Since $[A]$, $[B]$, $[C]$ all represent amounts of chemicals, they must all be non-negative; this constraint means that we must have $a - x \geq 0$, $b - x \geq 0$, and $x \geq 0$. To satisfy all of these constraints we must have:

$$0 \leq x \leq \min(a, b)$$

which generalizes the condition that we derived in Section 1.3, Example 5.

To identify equilibria in this interval we will use the method from Section 8.2.3 of separately plotting the two parts of $g(x)$ (i.e., make plots of $f(x) = k_{AB}(a-x)(b-x)$ and $h(x) = k_C x$ and look for points of intersection). For our plot we will assume that $a < b$ (in Problem 25, you will investigate what modifications are necessary if $a > b$ or $a = b$).

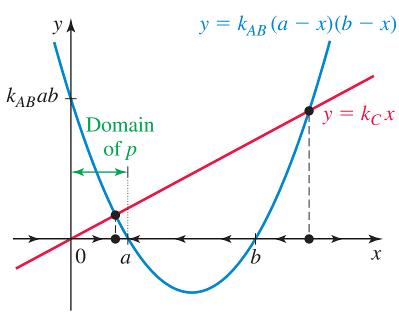


Figure 8.40 The curves $y = k_{AB}(a-x)(b-x)$ and $y = k_C x$ have two points of intersection, including one stable equilibrium with $0 < x < a$.

The function $f(x) = k_{AB}(a-x)(b-x)$ has roots $x = a$ and $x = b$, and $\lim_{x\to\pm\infty} f(x) = +\infty$, since the largest degree term is $k_{AB}x^2$ and $k_{AB} > 0$. So the function is a concave up parabola. $f(x)$ is a decreasing function of x over the interval $0 \leq x \leq a$, starting at $k_{AB}ab$ when $x = 0$ and decreasing to 0 when $x = a$. By contrast, $h(x) = k_C x$ is an increasing function of x on the same interval starting at $h(0) = 0$ and increasing to $h(a) = k_C a$. Thus, the two functions must intersect at exactly one point in the interval $0 < x < a$, meaning that there is a single equilibrium value $\hat{x}$. The vector field plot of the differential equation is given in Figure 8.40. From the vector field plot we see that the equilibrium is stable; and in fact, whatever the initial amount of x is, $x(t) \to \hat{x}$ as $t \to \infty$. A sketch of these solutions are shown in Figure 8.41.

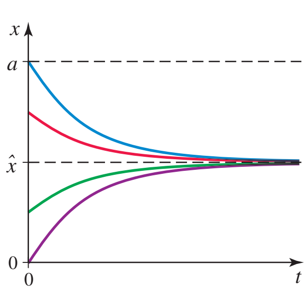


Figure 8.41 Concentration for C over time starting with different initial conditions.

The graphical method also allows us to determine how the equilibrium value, $\hat{x}$, varies as the parameters in the differential equation are changed. We will focus on the effect of varying the rate constants k_{AB} and k_C. Increasing k_{AB} dilates the curve $y = f(x)$ in the vertical direction (Figure 8.42). Since increasing k_{AB} increases $f(x)$ for all $x \in [0, a)$, the point of intersection, $\hat{x}$, between $y = f(x)$ and $y = h(x)$ moves rightward toward $x = a$ (see Figure 8.42). The change in $\hat{x}$ makes sense chemically, because as k_{AB} increases, C is produced at faster and faster rates. But the rate of decay of C back into A and B is left unaffected. So the concentration of C, at which rate of production of C equals decay, increases.

What about changing k_C? If we increase k_C, then the curve $y = k_C x$ gets steeper. It therefore intersects $y = k_{AB}(a-x)(b-x)$ at smaller and smaller values of x

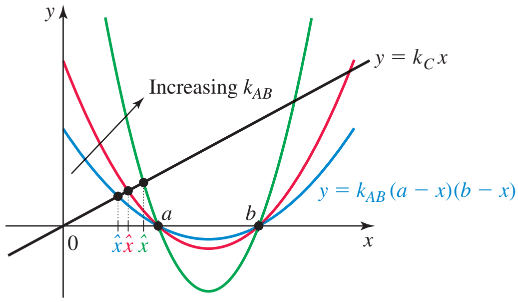


Figure 8.42 Increasing k_{AB} increases $\hat{x}$, the equilibrium amount of C.

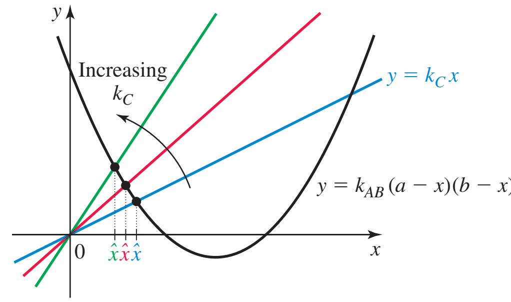


Figure 8.43 Increasing k_C decreases the equilibrium amount of C.

(see Figure 8.43), that is, the equilibrium value of x gets smaller and smaller. Again there is a chemical interpretation of this behavior: Increasing k_C increases the rate at which C decomposes back into A and B. Since the rate of production of C is unaffected, the equilibrium amount of C decreases.

8.3.4 The Evolution of Cooperation

Many organisms cooperate to perform tasks that they could not achieve working individually. For example, the cells in the human body cooperate to process food, fight off infection, and move the body around. An ant colony may contain millions of ants that work together to raise young, gather food, and defend the nest. Cooperation can also occur between different species. The health of a plant depends on **symbiotic** interactions between the plant and bacteria and fungi that the plant harbors in its root system. Despite the frequency with which it occurs, cooperation is also a paradoxical biological process. When two organisms cooperate they both contribute to a common good, but these contributions may be costly to each. For example, if a worker ant in a colony devotes her life to building the nest or defending it, she loses the opportunity to have offspring of their own. In many cooperative arrangements, the cooperating organisms are related (e.g., ants in a nest may all be sisters, born to the same queen); and so contributing to the common good may indirectly benefit each organism's genes, since these genes are shared by the organisms that they cooperate with. However, in cases where the two partners are not related we could imagine one of the organisms allowing the other to contribute to the common good but not doing so itself. That way, the organism that cheats (i.e., chooses not to cooperate) receives the benefit of cooperation but does not pay any cost. The cooperative organism pays all of the cost of cooperating.

One method for modeling these kinds of interactions is through **evolutionary game theory**. The model that we will give in this section is adapted from ***Nowak (2006)***. Imagine a population of organisms playing what evolutionary biologists call a **snowdrift game**. In this game there are two types of organisms: **cooperators** and **cheaters**. When two cooperators meet they cooperate; that is, the two organisms work together so each pays a cost $c/2$ and receives a benefit b. When a cooperator meets a cheater, then the cooperator contributes but the cheater doesn't. In that case the cooperating organism pays a cost c, and the cheating organism pays no cost. So long as one organism cooperates, the organisms still receive a benefit b.

Bio Info • The name **snowdrift game** is used because biologists consider this form of cooperation to be analogous to two people driving on a snowy street who find themselves blocked by the same snowdrift. The benefit of shoveling through the snowdrift to each driver (i.e., getting home) is b independent of whether that driver helped to clear the snowdrift. The total cost of clearing the snowdrift so that each may drive home is c. If one person clears the snowdrift unassisted, that individual pays the entire cost, c, but if both work together they split the cost equally between them (i.e., each pays cost $c/2$). In this scenario a cooperator is one who shovels the snowdrift. A cheater, on the other hand, sits in their car, listening to the radio, and staying warm, while the other driver shovels. You shouldn't imagine literal snowdrifts, however. Consider, for example, a plant that interacts with a fungus or bacterium in the soil. The soil fungus or bacterium can cooperate by providing the plant with nitrogen that the plant itself cannot extract from the soil. The plant cooperates by sharing with the bacterium sugars that it creates by photosynthesis. A cheating bacterium would take sugars, but not provide any nitrogen in return, while a cheating plant might take nitrogen but not provide any sugars in return.

Thus we may model the interaction by using a **pay-off matrix** that represents the net benefit each organism gets from the interaction. If we think of one organism as a player in a game, and the other organism as its opponent, then the entries in the table tell us what net return (benefit minus cost) the *player* receives from the interaction. The opponent's pay-off is not represented in the pay-off matrix.

		Opponent	
		Cooperate	Cheat
Player	Cooperate	$b - c/2$	$b - c$
	Cheat	b	0

As we can see from the table, if the opponent is a cooperator, then a player who cheats will receive a higher pay-off than a player who cooperates. But if the opponent is a cheater, then depending on whether $b > c$ or $b < c$, a cooperator may receive higher pay-off than a cheater. **Evolutionary game theory** models the effects of these different strategies on the numbers of organisms playing the game.

Imagine that we have a (large) population of N organisms, a fraction x of which are cooperators, and a fraction y of which are cheaters. In each unit of time, each organism (whether it is a cooperator or a cheater) interacts with n other organisms, which are randomly chosen (that is, a cooperator cannot choose to interact only with other cooperators; see Problem 35 for a discussion of what can occur if cooperators are able to choose whom to interact with). From each interaction it receives a payoff that can be calculated from the payoff matrix. For each organism the rate of births is proportional to the net payoff from all of its interactions with other organisms. The offspring of a cooperator organism are all cooperators, and the offspring of a cheater are all cheaters. However, we also assume that the habitat that the organisms share is at its carrying capacity. This means that for each birth that occurs, another organism must die to maintain the total size of the population at N. All organisms (cheaters and cooperators) are equally likely to die. Assume that the death rate is m for both types of organism. Figure 8.44 summarizes how the population changes from step to step.

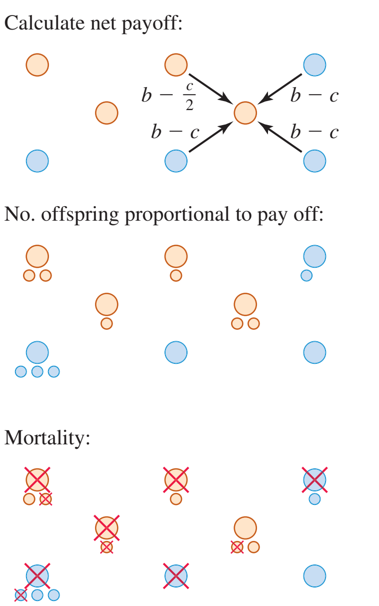


Figure 8.44 Evolutionary snowdrift game. A population contains both cheaters (blue) and cooperators (red). Each organism plays a snowdrift game with n other organisms receiving a pay-off from each game. The organism has a number of offspring proportional to its total pay-off. Organisms then die to keep population size constant.

So we can write down word equations for the rate at which the cooperator and cheater populations grow or decrease.

$$\begin{array}{c} \text{Rate of change} \\ \text{of number of} \\ \text{cooperators} \end{array} = \begin{array}{c} \text{Rate of} \\ \text{cooperator} \\ \text{births} \end{array} - \begin{array}{c} \text{Rate of} \\ \text{cooperator} \\ \text{deaths} \end{array}$$

Since x denotes the proportion of organisms that are cooperators and y represents the proportion that are cheaters, the total number of cooperators is Nx, and the total number of cheaters is Ny.

The rate of change of the number of cooperators is:

$$\frac{d}{dt}(Nx)$$

Since a fraction m of all organisms die in one unit of time, the rate of cooperator deaths is $m\times$ number of cooperators $= mNx$.

Thus:

$$\frac{d(Nx)}{dt} = \begin{array}{c} \text{Rate of cooperator} \\ \text{births} \end{array} - mNx \tag{8.42a}$$

Similarly the word equation for the rate of change of the cheater population is:

$$\begin{array}{c} \text{Rate of change} \\ \text{of cheater} \\ \text{population} \end{array} = \begin{array}{c} \text{Rate of} \\ \text{cheater} \\ \text{births} \end{array} - \begin{array}{c} \text{Rate of} \\ \text{cheater} \\ \text{deaths} \end{array}$$

which can be rewritten as:

$$\frac{d}{dt}(Ny) = \text{Rate of cheater births} - mNy. \tag{8.42b}$$

To complete our derivation we need mathematical expressions for the birth rates that appear in Equations (8.42a) and (8.42b). To derive these expressions let's first consider the average payoff that a cooperator receives in one unit of time. In that time it interacts with n other organisms. On average, since a fraction x of those organisms are cooperators and y are cheaters, a single cooperator will interact with nx cooperators and ny cheaters. Its payoff from each cooperator is $b - c/2$, and its payoff from each cheater is $b - c$. So the total payoff from all n interactions is:

$$\begin{array}{c} \text{Payoff} \\ \text{to cooperator} \end{array} = \begin{array}{c} \text{Number of cooperators} \\ \text{interacted with} \end{array} \times \begin{array}{c} \text{Payoff from} \\ \text{interacting} \\ \text{with cooperator} \end{array} + \begin{array}{c} \text{Number of cheaters} \\ \text{interacted with} \end{array} \times \begin{array}{c} \text{Payoff from} \\ \text{interacting with} \\ \text{a cheater} \end{array}$$

$$= nx \cdot (b - c/2) + ny \cdot (b - c)$$

The number of offspring this individual will have is proportional to its total payoff. Let the constant of proportionality be k. Then

$$\begin{aligned} \text{Number of offspring per cooperator} &= k \times \text{ total payoff} \\ &= k \cdot n\,(x(b - c/2) + y(b - c)) \end{aligned}$$

So:

$$\begin{array}{c} \text{Rate of} \\ \text{cooperator births} \end{array} = \text{Number of cooperators} \times \begin{array}{c} \text{Number of offspring per} \\ \text{cooperator} \end{array}$$

$$= Nx \cdot kn\,(x(b - c/2) + y(b - c))$$

We can calculate similarly the birth rate of cheaters, by first calculating the total payoff to each cheater in one unit of time.

$$\begin{array}{c} \text{Total payoff} \\ \text{to cheater} \end{array} = \begin{array}{c} \text{Number of cooperators} \\ \text{interacted with} \end{array} \times \begin{array}{c} \text{Payoff from} \\ \text{interacting} \\ \text{with a cooperator} \end{array} + \begin{array}{c} \text{Number of cheaters} \\ \text{interacted with} \end{array} \times \begin{array}{c} \text{Payoff from} \\ \text{interaction} \\ \text{with a cheater} \end{array}$$

$$= nx \cdot b + ny \cdot 0$$

$$= nx \cdot b$$

Hence

$$\begin{aligned} \text{Rate of cheater births } &= \text{Number of cheaters } \times \text{Birth rate per cheater} \\ &= Ny \cdot knxb \end{aligned}$$

Birth rate per cheater $= k\times$ Payoff to each cheater

Equations (8.42a) and (8.42b) can now be written as mathematical formulas:

$$\frac{d(Nx)}{dt} = Nxkn\,(x(b - c/2) + y(b - c)) - mNx$$

and

$$\frac{d(Ny)}{dt} = Nyknxb - mNy$$

or on rearranging some of the terms:

$$\frac{dx}{dt} = knx\,(x(b - c/2) + y(b - c)) - mx \tag{8.43}$$

N is a constant, and can be cancelled from both sides.

$$\frac{dy}{dt} = knyxb - my \tag{8.44}$$

The mortality rate, m, appears not to have been specified. But we can use Equations (8.43) and (8.44) to calculate m. The total rate of births must match the total rate of deaths, to keep the population size constant. Alternatively, since each organism can only be one of the two types (cheater or cooperator), we must have $x + y = 1$. Thus

$$\frac{dx}{dt} + \frac{dy}{dt} = \frac{d}{dt}(x + y) = \frac{d}{dt}(1) = 0.$$

And substituting in our expressions for dx/dt and dy/dt from (8.43) and (8.44) into this equation, we obtain:

$$knx\,(x(b - c/2) + y(b - c)) + knyxb - m(x + y) = 0$$

or

$$m(x + y) = knx\,(x(b - c/2) + y(2b - c)) \qquad \text{Isolating term in } m$$

so:

$$m = knx(b - c/2)(x + 2y) = knx(b - c/2)(1 + y) \qquad x + y = 1, \text{ and factorize out } (b - c/2)$$

From this equation we see that it is essential that $b - c/2 \geq 0$. Otherwise, $m < 0$ (i.e., we have negative mortality rate). The biological rationale for this inequality is that if $b - c/2 < 0$, then cooperators always receive negative payoffs, even from interacting with other cooperators. There is clearly no incentive in this case to cooperate.

Thus, substituting for m in (8.43) and (8.44), we obtain

$$\frac{dx}{dt} = knx\Big(x(b - c/2) + y(b - c) - (b - c/2)x(1 + y)\Big)$$
$$\frac{dy}{dt} = kny\Big(bx - (b - c/2)x(1 + y)\Big)$$

or

$$\frac{dx}{dt} = knxy\,(-(b - c/2)x + (b - c)) \tag{8.45}$$

and

$$\frac{dy}{dt} = knxy\left(\frac{c}{2} - (b - c/2)y\right) \tag{8.46}$$

Initially it may appear that we have to solve two differential equations; one for $x(t)$ and one for $y(t)$. But if we know x, then we can obtain y as $1 - x$, without needing to solve a second differential equation. So we only need to analyze one of the equations. Let's study the equation for $x(t)$ (i.e., (8.45)), which can be written as $\frac{dx}{dt} = g(x)$, where

$$g(x) = knx(1 - x)\Big((b - c) - (b - c/2)x\Big) \text{ and } 0 \leq x \leq 1 \qquad \text{Set } y = 1 - x \text{ in (8.45).}$$

This is a single autonomous differential equation, so we can analyze it using the methods from Section 8.2. To find the equilibria observe $g(\hat{x}) = 0$ if $\hat{x} = 0$, or $\hat{x} = 1$, or $\hat{x} = \frac{(b-c)}{(b-c/2)}$, assuming that $b \neq c/2$. You will analyze the case $b = c/2$ in Problem 31. We will assume that $b > c/2$. The first two roots certainly lie in the interval $0 \leq x \leq 1$. Whether the other root does or doesn't depends on the values of b and c. In order for the third root to lie in the interval we must have

$$0 \leq \frac{b - c}{b - c/2} \leq 1 \tag{8.47}$$

To analyze these inequalities we must multiply all sides by $b - c/2$. Since we are assuming that $b > c/2$, we do not need to reverse any of our inequalities if we multiply all three parts of the inequality by $b - c/2$. So (8.47) is satisfied if

$$0 \leq \quad b - c \quad \leq b - c/2$$

The inequality $b - c \leq b - c/2$ is automatically satisfied if $c > 0$. Thus the equilibrium $x = \frac{b-c}{b-c/2}$ lies in $[0, 1]$, provided $b \geq c$. Hence, if $b > c$, there are three equilibria in $[0, 1]$, at $x = 0$, $x = 1$, and $x = \frac{b-c}{b-c/2}$. If $\frac{c}{2} < b < c$, then there are two equilibria in $[0, 1]$, at $x = 0$ and $x = 1$.

Which of these equilibria are stable? We will use the derivative test to decide:

$$
\begin{aligned}
g'(x) &= kn(1-x)\Big((b-c)-(b-c/2)x\Big) \\
&\quad - knx\Big((b-c)-(b-c/2)x\Big) \\
&\quad - kn(b-c/2)x(1-x) \qquad \text{Using product rule}
\end{aligned}
$$

So:

$$
\begin{aligned}
g'(0) &= kn(b-c) \qquad g'(0) > 0 \text{ if } b > c \\
g'(1) &= knc/2 \qquad g'(1) > 0
\end{aligned}
$$

and

$$
\begin{aligned}
g'\left(\frac{b-c}{b-c/2}\right) &= -kn(b-c/2)\left(\frac{b-c}{b-c/2}\right)\left(1-\frac{b-c}{b-c/2}\right) \\
&= -\frac{kn(b-c)c/2}{b-c/2} \qquad 1-\tfrac{b-c}{b-c/2} = \tfrac{c/2}{b-c/2} \\
&= -\frac{knc}{2}\left(\frac{b-c}{b-c/2}\right)
\end{aligned}
$$

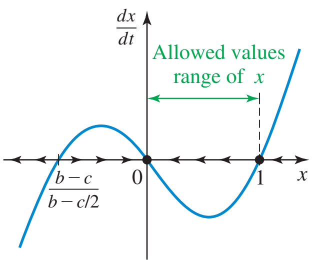


Figure 8.45 The vector field plot for the snowdrift game if $c/2 < b < c$.

So, if $c/2 < b < c$, then there are two equilibria in the interval at $\hat{x} = 0$ and $\hat{x} = 1$. The equilibrium $\hat{x} = 0$ is stable (since $g'(0) < 0$) and $\hat{x} = 1$ is unstable (since $g'(1) > 0$). The vector field plot is shown as Figure 8.45, and possible solutions are drawn in Figure 8.46.

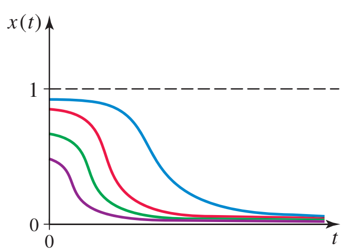


Figure 8.46 If $c/2 < b < c$, then unless $x(0) = 1$, $x(t) \to 0$ as $t \to \infty$.

Conversely, if $b > c$, there are three equilibria in the interval at $\hat{x} = 0$, $\hat{x} = \frac{b-c}{b-c/2}$, and $\hat{x} = 1$. The equilibria $\hat{x} = 0$ and $\hat{x} = 1$ are both unstable, and $\hat{x} = \frac{b-c}{b-c/2}$ is stable. The vector field plot of the differential equation is shown in Figure 8.47 and possible solutions in Figure 8.48.

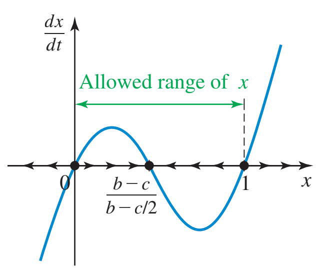


Figure 8.47 The vector field plot for the snowdrift game if $b > c$.

Can we interpret these different behaviors biologically? First we have imposed a condition that $b > c/2$. This inequality ensures that, when two cooperators interact, their payoff $(b - c/2)$ is larger than the payoff received by two cheaters that interact (0). Unless this condition is met, there is no incentive to cooperate. If, in addition, $b < c$, when a cooperator meets a cheater, the cooperator receives a negative payoff $(b - c < 0)$, while the cheater receives a positive payoff. The ability of cheaters to take advantage of cooperators means that according to our model, cheaters will always eventually take over the entire population since $x(t) \to 0$ as $t \to \infty$, meaning that the proportion of cheaters, $y(t) \to 1$. Cooperation can be exploited by cheaters who enjoy the benefit of cooperation without paying the cost, so cooperators are eventually driven extinct. The one exception to this is if $x(0) = 1$, in which case $x(t) = 1$ for all t. If a population starts with no cheaters present $(y(0) = 0)$, then no cheaters will ever appear (since the only way a cheater can be added is when another cheater reproduces). Our analysis shows that even if the starting number of cheaters is very small, they will still eventually drive the cooperators extinct.

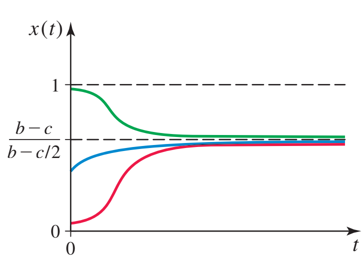


Figure 8.48 If $b > c$, unless $x(0) = 0$ or $x(0) = 1$, all solutions of the snowdrift game converge to the stable fixed point $x(t) = \frac{b-c}{b-c/2}$

If, on the other hand, $b > c$, then when a cooperator and a cheater interact, the cooperator still receives a positive payoff $(b - c > 0)$, although the payoff to the cheater is larger. In an initially mixed population (that is, provided $x(0) \neq 0$ or 1), neither cheaters nor cooperators go extinct, but instead achieve some equilibrium $\hat{x} = \frac{b-c}{b-c/2}$. Notice that for larger values of b (in particular if $b \gg c$), $\hat{x}$ will be close to 1 (i.e., there will be a higher proportion of cooperators in the population's stable equilibrium).

8.3.5 Epidemic Model

Mathematical models can be used to predict disease outbreaks. Models for the spread of disease can provide critical information for efforts to control a disease. For example ***Fisman, Khoo, and Tuite (2014)*** built a mathematical model to predict the rate of growth of an Ebola epidemic in West Africa, and to show that the control measures in place at that time were not enough to stop the disease from continuing to spread.

We will study disease spread in depth in Chapter 11. In this subsection we will introduce a type of model that can be used to predict the spread of a rapidly evolving disease, like the common cold. Colds are caused by several different kinds of viruses that spread by touch, or through droplets that are produced when a person sneezes.

For our model we will consider how a cold spreads through a population of N people. We divide people into two classes: susceptible individuals (who don't currently have a cold) and infected individuals (who are currently sick with a cold). If a susceptible individual catches a cold, they are moved into the infected population. If an infected individual fights off their cold, then they re-enter the susceptible population. Let $S(t)$ represent the number of susceptible people at time t and $I(t)$ the number of infected people at that time. We write down word equations for the rate of change of $S(t)$ and $I(t)$ as follows:

$$\begin{matrix}\text{Rate of change of}\\ \text{no. susceptibles}\end{matrix} = \begin{matrix}\text{Rate at which}\\ \text{infected individuals}\\ \text{recover}\end{matrix} - \begin{matrix}\text{Rate at which}\\ \text{susceptible individuals}\\ \text{are infected}\end{matrix}$$

$$\begin{matrix}\text{Rate of change of}\\ \text{no. infected}\end{matrix} = \begin{matrix}\text{Rate at which}\\ \text{susceptibles}\\ \text{are infected}\end{matrix} - \begin{matrix}\text{Rate at which}\\ \text{infected individuals}\\ \text{recover}\end{matrix}$$

We must now derive mathematical expressions for the rates appearing in these equations. The left-hand sides are respectively dS/dt and dI/dt.

At time t there are $I(t)$ infected individuals. Let's assume that in one unit of time a fraction c of these individuals recover. Then:

$$\begin{matrix}\text{Rate at which}\\ \text{infected individuals}\\ \text{recover}\end{matrix} = \begin{matrix}\text{No. infected}\\ \text{individuals}\end{matrix} \times \begin{matrix}\text{Fraction that}\\ \text{recover}\end{matrix} = I(t) \cdot c$$

We call the coefficient c the **recovery rate**. It depends on the disease that is being studied. For example, the recovery rate for the common cold is typically around $c = 0.3$/day.

The model for infection is a little more involved. Assume that the disease can only be transmitted by direct contact between individuals (e.g., by handshakes, or when they are close enough that one inhales the droplets produced by the other's sneezes). Suppose that in one unit of time each individual comes into contact with b other individuals. We assume that the likelihood that a susceptible individual will become infected is proportional to the number of infected individuals they come into contact with. Of the b individuals, each susceptible individual contacts, a fraction $\frac{I(t)}{N}$ will be infected, and a fraction $\frac{S(t)}{N}$ will not. So;

$$\begin{matrix}\text{Likelihood that}\\ \text{susceptible individual}\\ \text{gets infected}\end{matrix} = \begin{matrix}k\times\text{ No.}\\ \text{infected individuals}\\ \text{contacted in unit time}\end{matrix} = k \times \frac{bI}{N}$$

where k is a constant of proportionality. So total rate of infection is given by

$$\begin{matrix}\text{Rate at which susceptibles}\\ \text{are infected}\end{matrix} = \text{No. susceptibles} \times \begin{matrix}\text{Likelihood susceptible}\\ \text{individual is infected}\\ \text{in one unit of time}\end{matrix}$$

$$= S(t) \times kb\frac{I(t)}{N} = \frac{kb}{N}S(t)I(t).$$

Putting these ingredients together, we obtain differential equations

$$\begin{aligned}\frac{dS}{dt} &= cI - \frac{kb}{N}SI \\ \frac{dI}{dt} &= \frac{kb}{N}SI - cI\end{aligned} \tag{8.48}$$

In these equations c, k, b, and N are all positive constants that allow the model to be used to represent different diseases and different populations.

At first look the equations in (8.48) appear very different from the equations we have previously analyzed because there are two dependent variables (S and I). We cannot solve the differential equation for $I(t)$ unless we know $S(t)$, and we cannot solve the differential equation for $S(t)$ unless we know $I(t)$. How can we proceed? A general method for analyzing systems of differential equations like (8.48) will be introduced in Chapter 11. But in this case we note that all individuals are either susceptible or infected. So:

$$S(t) + I(t) = N \tag{8.49}$$

Hence, if we know $I(t)$ at any time then we can calculate $S(t)$ from (8.49). Using (8.49) to substitute for $S(t)$ in the second equation from (8.48), we obtain:

$$\frac{dI}{dt} = \frac{kb}{N}(N - I)I - cI$$

Or:

$$\frac{dI}{dt} = (kb - c)I - \frac{kb}{N}I^2 \tag{8.50}$$

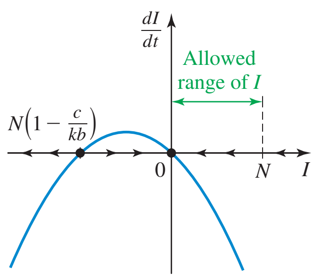


Figure 8.49 Vector field plot for (8.50) if $kb < c$.

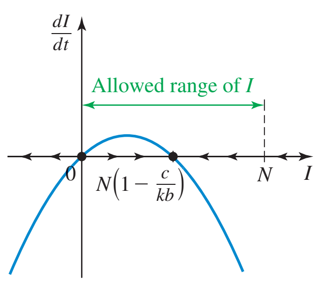


Figure 8.50 Vector field plot for (8.50) if $kb > c$.

This is a single autonomous differential equation so we can analyze the solutions using the methods from Section 8.2. There are two potential equilibria: at $\hat{I} = 0$ and at $\hat{I} = \frac{(kb-c)N}{kb} = N\left(1 - \frac{c}{kb}\right)$. We call these values of $\hat{I}$ potential equilibria because in addition to satisfying (8.50), $I(t)$ must satisfy inequalities.

$$0 \leq I(t) \leq N$$

$\hat{I} = 0$ certainly satisfies these inequalities. The second equilibrium, $\hat{I} = N\left(1 - \frac{c}{kb}\right)$ is certainly less than N (c, k, and b are all positive numbers). But if $kb < c$, then $\hat{I} < 0$, while if $kb > c$, then $\hat{I} > 0$. There are therefore two cases to consider.

If $kb < c$, then the only equilibrium in $[0, N]$ is $\hat{I} = 0$; the vector field plot for Equation (8.50) then shows that $\hat{I} = 0$ is a stable equilibrium (see Figure 8.49). So no matter what the initial number of infected individuals, $I(t) \to 0$ as $t \to \infty$; that is, the disease runs its course and disappears.

If, on the other hand, $kb > c$, then there are two equilibria to be considered. $\hat{I} = 0$ and $\hat{I} = \left(1 - \frac{c}{kb}\right)N$. The equilibrium $\hat{I} = 0$ is unstable, while $\hat{I} = (1 - \frac{c}{kb})N$ is stable. (see the vector field plot in Figure 8.50). $\hat{I} = 0$ is an equilibrium, because if there are initially no infected individuals (i.e., the disease is not present in the population), then no individuals will ever get sick, according to our model. However, our model now predicts that unless $I(0) = 0$, $I(t) \to \left(1 - \frac{c}{kb}\right)N$ as $t \to \infty$. That is, the disease persists in the population at some stable level. In this case we say that the disease has become **endemic**.

We can interpret the condition for determining whether a disease becomes endemic or disappears biologically. A disease will become endemic if $kb > c$, that is, if either k is large, or b is large, or c is small; in other words, if each individual has many contacts with others per unit time (large b), if these contacts have a high chance of passing the disease on (large k), or if the disease takes a long time to recover from (small c).

Section 8.3 Problems

8.3.1

1. In Section 8.3.1 we introduced single-compartment models for the motion of matter through a single tank of water. We derived an equation:

$$\frac{dC}{dt} = \frac{q}{V}(C_I - C) \tag{8.51}$$

for the concentration of solute in the tank, $C(t)$. We analyzed this equation graphically. Now let's solve the equation to confirm our analysis.

Assume that $C(0) = C_0$

(a) Solve (8.51) and use your solution to show that $C(t) \to C_I$ as $t \to \infty$, for any value of C_0.

(b) Explain how your solution from part (a) predicts that larger values of q/V lead to faster convergence of $C(t)$ to C_I, and smaller values of q/V lead to slower convergence.

2. Assume the single-compartment model defined in Section 8.3.1; that is, denote the concentration of the solute at time t by $C(t)$, and assume that

$$\frac{dC}{dt} = 3(20 - C(t)) \quad \text{for } t \geq 0 \tag{8.52}$$

(a) Solve (8.52) when $C(0) = 5$.

(b) Find $\lim_{t\to\infty} C(t)$.

(c) Use your answer in (a) to determine t so that $C(t) = 10$.

3. Use the single-compartment model defined in Section 8.3.1; that is, denote the concentration of the solution at time t by $C(t)$, and assume that the concentration of the incoming solution is 3 g liter^{-1} and the rate at which mass enters is 0.2 liter s^{-1}. Assume, further, that the volume of the compartment $V = 400$ liters.

(a) Find the differential equation for the rate of change of the concentration at time t.

(b) Find all equilibria of the differential equation and discuss their stability.

(c) Solve the differential equation in (a) when $C(0) = 0$, and find $\lim_{t\to\infty} C(t)$.

4. Suppose that a tank holds 1000 liters of water, and 2 kg of salt is poured into the tank.

(a) Compute the concentration of salt in g liter^{-1}.

(b) Assume now that you want to reduce the salt concentration. One method would be to remove a certain amount of the salt water from the tank and then replace it by pure water. How much salt water do you have to replace by pure water to obtain a salt concentration of 1 g liter^{-1}?

(c) Another method for reducing the salt concentration would be to hook up an overflow pipe and pump pure water into the tank. That way, the salt concentration would be gradually reduced. Assume that you have the choice of two pumps, one that pumps water at a rate of 1 liter s^{-1}, the other at a rate of 2 liter s^{-1}. For each pump, find out how long it would take to reduce the salt concentration from the original concentration to 1 g liter^{-1}. (Note that the rate at which water enters the tank is equal to the rate at which water leaves the tank.)

(d) Show that, whichever pump you use in part (c), you need more pure water if you use the pump method than if you follow the method in (b). Can you explain why?

5. Osmosis Through a Cell Membrane A cell constantly gains or loses small molecules to its environment because the small molecules are able to diffuse through the cell membrane. We will build a model for this process.

Suppose a molecule is present in the cell at a concentration $C(t)$, and present in its environment at a concentration C_∞ (you may assume C_∞ is a constant). One model for the diffusion of molecules across the cell membrane is that the rate at which molecules travel through the membrane is proportional to the difference in concentration between the cell and its surroundings. That is:

$$\begin{array}{c}\text{Rate at which}\\ \text{molecules flow out}\\ \text{of cell}\end{array} = k(C - C_\infty)$$

The constant k is known as the **permeability of the membrane**; $k > 0$, and k depends on the surface area of the cell and the chemistry of the membrane, as well as the type of molecule.

(a) Starting with a word equation for the amount of small molecules in the cell, show, if the cell volume is V, then:

$$\frac{dC}{dt} = -\frac{k}{V}(C - C_\infty) \tag{8.53}$$

(b) Find the equilibrium of (8.53) and use a graphical analysis to determine whether it is stable or unstable.

(c) Suppose that the molecule we are studying is produced within the cell. The cell produces the molecule at a rate r; that is, a quantity r is produced (added to the cell) in unit time. Explain why the differential equation for the concentration of molecules in the cell should be modified to:

$$\frac{dC}{dt} = -\frac{k}{V}(C - C_\infty) + \frac{r}{V} \tag{8.54}$$

(d) Analyze Equation (8.54) to find the equilibrium value of the cell concentration. Is this equilibrium stable or unstable? You may use a graphical argument or calculate the eigenvalue to determine the equilibrium's stability.

6. Chemostat A chemostat is a device that can be used to maintain a constant concentration of a chemical in a chamber.

Consider a chemostat consisting of a chamber of volume V and containing a concentration $C(t)$ of the chemical.

(a) Initially we will neglect inflows and outflows in the chamber. The chemical breaks down at a fractional rate p; that is, a proportion p of the chemical contained in the chamber is broken down in unit time. Explain why the concentration of the chemical must obey a differential equation

$$\frac{dC}{dt} = -pC \tag{8.55}$$

(b) By analyzing (8.55) determine the long-term behavior of $C(t)$; that is, find $\lim_{t\to\infty} C(t)$. (You can do this using the methods from Section 8.2. There is no need to solve the differential equation.)

(c) To maintain the concentration $C(t)$ of the chemical, at some desired concentration, fresh chemical is continuously added to the chamber. This is accomplished by adding fluid containing the chemical to the chamber continually at a rate q, and removing fluid from the chamber at the same rate. Show that if the concentration of chemical in the fluid being added to the chamber (the "inflow") is C_I, then:

$$\frac{dC}{dt} = \frac{q}{V}(C_I - C) - pC \tag{8.56}$$

(d) By analyzing (8.56) find the equilibrium concentration of chemical in the chamber as a function of q, V, C_I, and p. Determine whether the equilibrium is stable or unstable.

(e) Suppose $p = 0.2$/hr, $q = 1$ ml/hr, and $V = 10$ ml. If we want to maintain $C(t)$ at 5 g/liter, what should the concentration C_I of chemical in the inflow be?

7. The stability of the equilibrium concentration in a single compartment is often quantified using T_R, which is called the *time to return to equilibrium*. Suppose that the equilibrium concentration is C_0. Then, to measure T_R, perturb the concentration slightly, from C_0 to $C_0 + C_1$. Then T_R is defined to be the time that the tank takes for the perturbation to drop to a factor $\frac{1}{e}$ of its initial value (i.e., for $C(t)$ to drop from $C_0 + C_1$ to $C_0 + \frac{C_1}{e}$).

If the single compartment obeys the single-compartment differential Equation (8.45):

(a) Show that $T_R = V/q$.

(b) Suppose, instead of defining T_R by the time taken for the concentration to drop to $C_0 + C_1/e$, we choose a fraction p, $(p < 1)$, and define T_R to be the time taken for the concentration to drop to $C_0 + pC_1$. Calculate T_R in terms of V, q, and p.

8. Insulin Pump Insulin pumps treat patients with type I diabetes by releasing insulin continuously into the fat in the patient's stomach or thigh. We will develop a model for the transport of insulin from the site where it is released by the pump, by treating the fat as a compartment in a single-compartment model. Let's suppose that the pump releases insulin at a constant rate, r (r is the amount added in one unit of time).

(a) Explain why, if insulin is not transported from the site of release, the amount of insulin at the site of release, $a(t)$, will obey a differential equation:

$$\frac{da}{dt} = r.$$

(b) From the fat, the insulin enters the patient's bloodstream. Suppose that a fraction p of the insulin present in the patient's fat enters the blood in unit time. Explain why:

$$\frac{da}{dt} = r - pa.$$

(c) Find the equilibrium from the differential equation in part (b) and determine whether this equilibrium is stable or unstable.

9. **Freeway Engineering** Compartment models are used to model the flow of traffic between different roads, by treating each road as a compartment. As an example, consider how the number of cars on a freeway on-ramp, $N(t)$, changes with time. For a simplified model let's assume that cars join the on-ramp at a constant rate q (that is, q cars join the on-ramp in one unit of time). Cars then leave the on-ramp by entering the freeway itself. Assume that a fraction f of the cars on the on-ramp enter the freeway in one unit of time.

(a) Derive a differential equation for $N(t)$. Your differential equation will include the unknown constants f and q.

(b) Analyze your model from part (a) to find the equilibrium number of cars on the on-ramp, and determine whether this equilibrium is stable or unstable.

(c) Suppose that the maximum capacity of the on-ramp is 90 cars, and the rate at which cars flow onto the on-ramp is $q = 60$ cars per min. Find the value of f that is needed to keep N below the on-ramp's capacity.

10. **Filling Box Model** In our compartment model we assumed that inflows and outflows are matched at q to keep the volume of water in the tank constant. It's often useful when modeling, for example, the flow of pollutant into a pristine environment, to consider what can occur if the inflows and outflows do not match.

Let's assume that the tank initially contains a volume V_0 of water. Water flows into the tank at rate q_{in}, and out of the tank at rate q_{out}. (You may assume $q_{\text{in}} > q_{\text{out}}$.) Suppose that the water flowing into the tank contains a concentration C_I of solute. As usual we write $C(t)$ for the concentration in the tank.

(a) Show that the concentration in the tank can be modeled using a differential equation:

$$\frac{d}{dt}(CV) = q_{\text{in}}C_I - q_{\text{out}}C$$

(b) Previously we were able to treat V as a constant. Now V changes with time. Derive a formula for $V(t)$.

(c) By substituting your formula for $V(t)$ into (a), derive a differential equation for $C(t)$.

(d) In general we cannot analyze the behavior of the solution $C(t)$ using techniques from Section 8.2. Why not?

(e) Let's assume $C_I = 0$. Then show that your equation from (c) can be written as:

$$\frac{dC}{dt} = \frac{-q_{\text{in}}C}{V_0 + (q_{\text{in}} - q_{\text{out}})t} \tag{8.57}$$

(f) Assume some definite values for the constants in (8.57): $q_{\text{in}} = 2$, $q_{\text{out}} = 1$, and $V_0 = 20$. Assuming $C(0) = 1$, solve (8.57) to find $C(t)$. Show that $\lim_{t\to\infty} C(t) = 0$.

8.3.2

11. **Levins Model** Denote by $p = p(t)$ the fraction of occupied sites in the patchy habitat model, and assume that

$$\frac{dp}{dt} = 2p(1-p) - p \quad \text{for } t \geq 0 \tag{8.58}$$

(a) Set $g(p) = 2p(1-p) - p$. Graph $g(p)$ for $p \in [0, 1]$.

(b) Find all equilibria of (8.58) that are in $[0, 1]$. Use your graph from (a) to determine their stability.

(c) Now use the eigenvalue approach to analyze the stability of the equilibria that you found in (b).

12. **Levins Model** Denote by $p = p(t)$ the fraction of occupied sites in the patchy habitat model, and assume that

$$\frac{dp}{dt} = 0.5p(1-p) - 1.5p \quad \text{for } t \geq 0 \tag{8.59}$$

(a) Set $g(p) = 0.5p(1-p) - 1.5p$. Graph $g(p)$ for $p \in [0, 1]$.

(b) Find all equilibria of (8.59) that are in $[0, 1]$. Use your graph in (a) to determine their stability.

(c) Use the eigenvalue approach to analyze the stability of the equilibria that you found in (b).

Subpopulation Interactions in Patchy Habitats

To derive our model for patchy habitat we assumed that a fixed fraction, m, of occupied sites became extinct in each unit of time. Often, however the survival of the population at a site depends on the number of subpopulations in the surrounding sites. If different subpopulations compete for limited resources, then the per site mortality rate may not be a constant, but may increase with p because, as p increases, competition between subpopulations increases. In questions 13 and 14 we will study the effect of different models for competition between subpopulations.

13. The term p^2 describes the density-dependent extinction of patches; that is, the per-patch extinction rate is p, and a fraction p of patches are occupied, resulting in patches going extinct at a total rate of p^2. The colonization of vacant patches is the same as in the Levins model. Then the fraction of occupied patches obeys a differential equation:

$$\frac{dp}{dt} = cp(1-p) - p^2$$

where $c > 0$.

(a) Show that there are two possible equilibrium values for p in $[0, 1]$ (which you should calculate) and determine their stability.

(b) Does the patch model always predict a nontrivial equilibrium when $c > 0$? Contrast with what we found for the Levins model in Section 8.3.2.

14. Assume that the per site extinction rate is Mp, and recolonization is unaffected by competition between subpopulations. Then our model for the proportion of occupied sites becomes

$$\frac{dp}{dt} = cp(1-p) - Mp^2$$

where $M > 0$ and $c > 0$ are constants.

(a) Show that there are two possible equilibrium values for p in $[0, 1]$ for any values of c and M. You do not need to find the values of both equilibria; instead, follow the method used in Section 8.2 and graph on the same axes the functions $f(p) = cp(1-p)$ and $h(p) = Mp^2$.

(b) Which of the two equilibria from part (a) is stable?

(c) Contrast your answer (about the existence of equilibria) with the analysis from Section 8.3.2.

Competition between Subpopulations

In Problems 13 and 14 we assumed that the per colony extinction rate was proportional to p. This means that the per colony extinction rate goes to 0 for small p. This may not be realistic—subpopulations may still go extinct even if they are not competing among themselves. One way to model this is to say that the per colony extinction rate is a function $m(p)$ of p. In Problems 15 and 16 we will assume that $m(p) = a + bp$ for some constants $a, b > 0$. That is, the extinction rate increases with p because of competition between subpopulations, but $m(p)$ does not vanish as $p \to 0$.

Then our model for proportion of occupied sites must be modified to:

$$\frac{dp}{dt} = cp(1-p) - (a+bp)p \tag{8.60}$$

where c, a, b are all positive constants.

15. Assuming that the subpopulations obey the differential Equation (8.60) and the coefficients are $a = 1, b = 2$, but c is allowed to take any value:

(a) Find the equilibrium values of p (your answer will depend on the unknown coefficient c).

(b) What are the conditions on c for p to have a nontrivial equilibrium, that is, an equilibrium in which $p \in (0, 1]$?

(c) Show that if your condition from (b) is met, then the nontrivial equilibrium is also stable.

16. In this question we will analyze (8.60) using a graphical argument.

(a) Assuming that a, b, c are all positive, draw the graphs of $y = c(1-p)$ and $y = a+bp$ to show that the differential Equation (8.60) has an equilibrium between 0 and 1 if $a < c$. [*Hint*: If the graphs intersect, then $c(1-p) - (a+bp) = 0$.]

(b) Show additionally that if $a < c$, then the nontrivial equilibrium is stable. [*Hint*: if $c(1-p) > a+bp$ and $p > 0$, then $cp(1-p) > (a+bp)p$.]

Cooperation between Subpopulations

Interactions between different subpopulations need not be competitive. In fact, different subpopulations may share resources, and the presence of many subpopulations may provide a pool of genetic diversity that helps the population of organisms to react to changing conditions. We will model cooperation between subpopulations by again assuming that the extinction rate depends on p, but now $m(p) = a - bp$, where a and b are both positive constants. So $m(p)$ decreases as p increases. Our model for the number of subpopulations then becomes:

$$\frac{dp}{dt} = cp(1-p) - (a-bp)p \tag{8.61}$$

We will analyze this model in Problems 17 and 18.

17. Assume that the number of subpopulations obeys (8.61) with $a = 2$, $b = 1$, and c some unknown (positive) constant.

(a) Find the equilibrium values of p (your answer will depend on the constant c). You may assume $c > 1$.

(b) What is the condition on c for p to have a nontrivial equilibrium (i.e., an equilibrium in which $\hat{p} \in (0, 1]$)?

(c) Show that if your condition from (b) is met, then the nontrivial equilibrium is also stable.

18. Assuming that the number of subpopulations obeys (8.61), we will analyze this model graphically.

(a) Explain why we would expect $a \geq b$. [*Hint*: What would having negative $m(p)$ imply?]

(b) Assuming that a, b, c are all positive, show by drawing the graphs of $y = c(1-p)$ and $y = a - bp$ that the differential Equation (8.60) has an equilibrium between 0 and 1 if $a < c$. [*Hint*: If the graphs intersect, then $c(1-p) - (a - bp) = 0$.]

(c) Show additionally that if $a < c$, then the nontrivial equilibrium is stable. [*Hint*: If $c(1-p) > a - bp$ and $p > 0$, then $cp(1-p) > (a-bp)p$.]

Habitat Destruction

To study the effects of habitat destruction on a single species, we modify the Levins model in the following way: We assume that a fraction D of patches is permanently destroyed. Consequently, only patches that are vacant and undestroyed can be successfully colonized. A fraction $1 - p(t) - D$ of patches is both vacant and undestroyed where $p(t)$ is the fraction of occupied patches. Then:

$$\frac{dp}{dt} = cp(1-p-D) - mp \tag{8.62}$$

19. (a) Explain in words the meaning of the different terms in (8.62).

(b) Assume that $m = 0.2$, $c = 2$, and $D = 0.2$. Show that (8.62) predicts a nontrivial equilibrium value for $p(t)$ and that this equilibrium is stable.

20. Assume that a patchy habitat that has been partly destroyed obeys Equation (8.62) with c, D, m all positive constants.

(a) Show that there are two possible equilibria: the trivial equilibrium $\hat{p}_1 = 0$ and the nontrivial equilibrium $\hat{p}_2 = 1 - D - \frac{m}{c}$. Sketch the graph of $\hat{p}_2$ as a function of D.

(b) Assume that $m < c$ such that the nontrivial equilibrium is stable when $D = 0$. Find a condition for D such that the nontrivial equilibrium is between 0 and 1, and investigate the stability of both the nontrivial equilibrium and the trivial equilibrium under that condition.

(c) Assume that the condition that you derived in is met. Show that when the system is in equilibrium, the fraction of patches that are vacant and undestroyed—that is, the sites that are *available* for colonization—is $1 - D - p$ and that this available fraction is independent of D. Show that the **effective colonization rate** in equilibrium—that is, c times the fraction of available patches—is equal to the mortality rate. This equality shows that the effective birth rate of new colonies balances their mortality rate at equilibrium.

8.3.3

A reversible chemical reaction between chemicals A and B produces a product C: $A + B \rightleftharpoons C$. We modeled this reaction in Section 8.3.3 using a differential equation for the amount of C produced:

$$\frac{dx}{dt} = k_{AB}(a-x)(b-x) - k_C x \tag{8.63}$$

Here $x(t)$ is the amount of C at time t, a is the initial amount of chemical A, b is the initial amount of B, and k_{AB} and k_C are respectively the rate constants for the reaction that creates C and for the decay of C back into A and B.

21. Explain what each term in (8.63) represents and how the equation is derived.

For Problems 22–24 find the equilibrium value of x, and use a perturbation analysis to determine the stability of the equilibrium of (8.63).

22. $k_{AB} = 2, k_C = 2, a = 2, b = 2$.

23. $k_{AB} = 2, k_C = 1, a = 2, b = 3$.

24. $k_{AB} = 1, k_C = 2, a = 3, b = 2$.

25. To show that the differential equation (8.63) always has a stable equilibrium between $x = 0$ and $x = \min(a, b)$ we assumed that a and b were different (in fact that $a < b$). Show by redrawing Figure 8.40 that the result still holds if (a) $a > b$ and (b) if $a = b$.

26. Reactions in a Chemostat A chemostat is a device that can be used to maintain a chemical at a particular concentration. Assume that the reaction $A + B \rightleftharpoons C$ takes place in a chemostat that maintains A and B at constant concentrations a and b respectively (that is, the concentrations do not change over time).

(a) Explain why the concentration $x(t)$ of C now obeys a differential equation:

$$\frac{dx}{dt} = k_{AB}ab - k_C x \tag{8.64}$$

(b) Find the equilibrium for x predicted by Equation (8.64).

(c) Is the equilibrium that you found in part (b) stable or unstable?

27. Reactions in a Chemostat A chemostat is a device that can be used to maintain a chemical at a particular concentration. Assume that the reaction $A + B \rightleftharpoons C$ takes place in a chemostat that maintains A at a constant concentration a. The chemical B has initial concentration b and is depleted by the reaction.

(a) Explain why the concentration $x(t)$ of C now obeys a differential equation:

$$dx/dt = k_{AB}a(b - x) - k_C x \tag{8.65}$$

(b) Find the equilibrium for x predicted by Equation (8.65).

(c) Is the equilibrium that you found in part (b) stable or unstable?

28. An Irreversible Reaction In an irreversible reaction A and B combine to produce C, but C cannot disassociate back into A and B. We can represent this reaction symbolically by $A + B \to C$. This is equivalent to setting $k_C = 0$ in our original differential equation model. So the concentration, $x(t)$, of C obeys a differential equation:

$$\frac{dx}{dt} = k_{AB}(a - x)(b - x) \tag{8.66}$$

(a) Find the equilibrium for x predicted by Equation (8.66).

(b) Is the equilibrium that you found in part (b) stable or unstable?

29. Temperature in a Chemical Reaction The rate constants k_{AB} and k_C in the chemical reaction we are modeling depend on temperature: Many reactions speed up at higher temperatures. Both k_{AB} and k_C will be affected by a temperature increase. Suppose that the reaction is run at a higher temperature that doubles both k_{AB} and k_C. Show that the final concentration of C will remain the same, despite the temperature increase.

30. Autocatalytic Reactions An autocatalytic reaction is one in which chemical C is involved in its own production, for example,

$$A + C \rightleftharpoons 2C$$

That is, one molecule of A and one molecule of C react to create two molecules of C. Suppose that the reaction occurs in a chemostat that maintains the concentration of A at a.

(a) If the concentration of C is $x(t)$, explain why we can model this process using an equation:

$$\frac{dx}{dt} = k_{AC}ax - k_C x^2 \tag{8.67}$$

where you should explain what the two terms in this equation represent.

(b) Find the equilibrium for x predicted by Equation (8.67).

(c) Is the equilibrium that you found in part (b) stable or unstable?

8.3.4

To derive the model for the growth or decline of the population of cooperators interacting in a snowdrift game, we modeled the proportion of cooperators using a model.

$$\frac{dx}{dt} = knx(1 - x)\Big(-(b - c/2)x + (b - c)\Big) \tag{8.68}$$

where $b > 0$ represents the benefit of interaction if one player is a cooperator and $c > 0$ is the cost of cooperation.

31. In Section 8.3.4 we analyzed Equation (8.68) if $b > c/2$. Determine the equilibria and what their stability is if $b = c/2$.

Assuming that $x(t)$ is modeled by Equation (8.68) in Problems 32–34, you should locate the equilibria and find which equilibria are stable for each of the following parameter values. Draw a vector field plot for each of the three problems.

32. $k = 1, n = 1, b = 2, c = 1$

33. $k = 1, n = 1, b = 3, c = 4$

34. $k = 1, n = 1, b = 4, c = 4$

35. Greenbeard Genes We showed in Section 8.3.4 that if $b < c$, then cooperators will be eventually outcompeted by cheaters. One mechanism that may allow cooperators to persist under these conditions is the greenbeard gene. Richard Dawkins coined this name (see ***Dawkins, 2006***) to describe how, if the genes that are responsible for cooperation also mark cooperators in some way (e.g., by giving each cooperator a bright green beard), then cooperators can make sure that they interact only with other cooperators (and thus cheaters interact only with other cheaters).

In this case the proportion, $x(t)$, of cooperators in the population will obey a differential equation.

$$\frac{dx}{dt} = (b - c/2)knx(1 - x)$$

and again the proportion of cheaters, $y(t)$, can be obtained from $y = 1 - x$.

(a) Show that if $b > c/2$, then under the greenbeard gene model $x = 1$ is a stable equilibrium and $x = 0$ is unstable.

(b) What are the equilibria and their stability if $b < c/2$?

(c) Explain your answers from part (a) and (b) biologically in terms of the relative costs and benefits of cooperation.

(d) What happens if $b = c/2$? Again explain your answer biologically.

We discussed the snowdrift model as one example of how organisms may interact. In Problems 36–40 we will consider an alternate model for interaction.

A Hawk-Dove game

Bio Info • In the Hawk-Dove game we assume that when two organisms interact they compete for some resource (e.g., territory). Only one organism can win the competition, and the benefit to this organism is b. But the competition may also leave one organism injured. There are two possible strategies that organisms may adopt when interacting with each other. *Hawks* always fight for the resource while *Doves* always back down from a fight. When two doves meet, both back down, and the resource is shared equally between them (i.e., each receives benefit $b/2$). When a hawk meets a dove, the dove surrenders the contested resource; then the hawk automatically gets the benefit (b), while the dove gets nothing. When two hawks meet, they fight: The victor will receive the benefit. But there is a cost to losing, since the loser may be hurt. Let's call this cost c. Since a hawk does not know in advance whether they will win or lose the fight, on average costs and benefits will be evenly split (i.e., on average a hawk's benefit from fighting another hawk is $\frac{1}{2}(b-c)$). We can summarize the results of the competition in a payoff matrix:

		Opponent	
		Hawk	Dove
Player	Hawk	$\frac{1}{2}(b-c)$	b
	Dove	0	$b/2$

If we model the effect of each interaction upon the proportions of hawks $x(t)$, and doves $y(t)$, in the population we may derive a differential equation

$$\frac{dx}{dt} = knx(1-x)\left(\frac{b}{2} - \frac{cx}{2}\right) \tag{8.69}$$

Just as for the snowdrift game model, we can then calculate y, the proportion of doves, from the equation $y = 1 - x$.

In Problems 36–38 we will analyze the population dynamics that are predicted by (8.69), for different values of b and c.

36. Show that if $b > c$ (that is, the maximum benefit of fighting exceeds the maximum cost), then the only equilibria for (8.69) within $0 \le x \le 1$ are $x = 0$ and $x = 1$. Which equilibrium is stable and which is unstable?

37. Show that if $b < c$, then there are three equilibria for (8.69) with $0 \le x \le 1$. What are these equilibria and which one(s) are stable?

38. Suppose $b = c$; then $\frac{dx}{dt} = \frac{knb}{2}x(1-x)^2$. Assuming $k = 1$, $n = 1$, and $b = 1$, sketch the vector field plot for x.

Assuming that $x(t)$ is modeled by Equation (8.69), in Problems 39–40, you should find the equilibria and determine which equilibria are stable for each of the following parameter values. Draw a vector field plot for each problem.

39. $k = 1, n = 1, b = 2, c = 1$

40. $k = 1, n = 1, b = 2, c = 3$

8.3.5

To model the spread of a disease in a population of size N we derived a differential equation model:

$$\frac{dI}{dt} = (kb - c)I - \frac{kb}{N}I^2 \tag{8.70}$$

where $I(t)$ is the number of infected individuals at time t, and k, b, and c are all positive coefficients.

Assuming that $I(t)$ is modeled by Equation (8.70), in Problems 41–44, you should locate the equilibria of the model, and find which of these equilibria are stable. Draw a vector field plot for each problem.

41. $k = 1, b = 1, c = 0.5, N = 50$.

42. $k = 1, b = 1, c = 0.5, N = 200$.

43. $k = 2, b = 2, c = 1, N = 100$.

44. $k = 2, b = 2, c = 4, N = 100$.

45. In this question we will interpret the *recovery rate*, c, that appears in the model. Assume that a population of infected individuals is quarantined (that is, they are unable to transmit the disease to others, or to catch it again once they recover).

(a) Explain why under these assumptions we expect:

$$\frac{dI}{dt} = -cI \tag{8.71}$$

(b) Assuming $I(0) = I_0$, find $I(t)$ by solving (8.71).

(c) How long will it take for the number of infected individuals to decrease from I_0 to $I_0/2$?

(d) Assume that it takes 7 days for the number of infected individuals to decrease from 50 to 25. Calculate the recovery rate c for this disease.

46. **Quarantining** Quarantining is an effective way to prevent diseases from spreading. Infectious individuals are told to stay at home to avoid spreading the disease. However, when a person is in the early stages of a disease, they may not realize they are ill and they then spread the disease to others.

Suppose that a fraction p of infectious individuals continue to spread the disease.

(a) Explain why, if a person contacts b individuals in a unit of time, then a fraction $\left(\frac{pI}{S+pI}\right)$ will be infectious.

(b) Show that the differential equation for the number of infectious individuals needs to be modified to

$$\frac{dI}{dt} = kb\frac{pIS}{(S+pI)} - cI.$$

(c) Use the relationship $S = N - I$ to rewrite your equation from part (b) in terms of I only (i.e., to eliminate $S(t)$ from the equation).

(d) Assume that $p = 1/2$. Analyze the differential equation from (c) to find its equilibria, and determine which are stable.

(e) Under what conditions on k, b, and c will the disease be endemic? (You may continue to assume $p = 1/2$). Compare this condition with the one that we derived in Section 8.3.5.

47. **Handwashing** One way to control the spread of a disease is to run public health programs that educate people on how to limit their exposure to the disease. For example, frequent handwashing can prevent people from picking up a virus after touching surfaces that it may live on.

(a) Explain why in our model such efforts to control the disease would affect the value of the parameter k, but would not affect b or c.

(b) Suppose that for a particular disease $c = 0.3$/day, and $b = 10$/day. What value must k remain below to prevent the disease from becoming endemic?

48. **Public Health** In our analysis of the spread of a disease we showed that a disease will become endemic if $kb > c$. Thinking about the meanings of the coefficients k, b, and c, discuss how the following public health measures can help to prevent a disease from becoming endemic. In particular, explain which coefficients in the model are affected by each measure, and whether the measure increases or decreases the coefficients it affects:

(a) Quarantining sick people (i.e., requiring that sick people stay at home).

(b) Encouraging frequent hand-washing.

(c) Educating people to cover their mouths and noses when they sneeze.

(d) Providing medications to people with the disease.

8.4 Integrating Factors and Two-Compartment Models

In Section 8.1 we learned the method of separation of variables for solving differential equations of the form:

$$\frac{dN}{dt} = f(t)g(N) \tag{8.72}$$

In Section 8.2 we learned how to analyze the solutions of these kinds of differential equations graphically as in the case where $f(t) = 1$ (i.e., the right-hand side of (8.72) is a function of N only). As we saw in Section 8.3 many differential equations that arise as models of biological phenomena are of this form. However, not all equations are of this form, and it is important to recognize when separation of variables can be used and when it cannot be used. For example, the equation:

$$\frac{dN}{dt} = N + t \tag{8.73}$$

cannot be solved by separation of variables—the right-hand side is not the product of a function of N with a function of t. To see why, try pulling out the t (i.e., writing $N + t = t(1 + N/t)$). Then we have one factor, $f(t) = t$, that is truly a function of t only, but the other factor is a function of both N and t, $g(N, t) = 1 + N/t$.

In this section we will learn a technique that can be used to solve any equation of the form:

$$\frac{dN}{dt} + a(t)N = b(t) \tag{8.74}$$

where $a(t)$ and $b(t)$ are both functions of t only. Equation (8.73) is of this form because it can be written as

$$\frac{dN}{dt} - N = t$$

which is of the correct form if we set $a(t) = -1$ and $b(t) = t$ in (8.74). The method that we will use is called integrating factors, because it involves multiplying (8.74) by a new function (a factor) in order to turn both sides of the equation into functions that we can integrate with respect to t. Equations of the form (8.74) are needed for the study of two-compartment models—these are differential equations that are used throughout life sciences, but particularly to study how medications or drugs move through the body, and we will present this important application at the end of this section.

8.4.1 Integrating Factors

Let's start with a specific example, namely the differential equation from (8.73), with an initial condition added

$$\frac{dN}{dt} - N = t \qquad N(0) = 0$$

To solve this differential equation we will start with an unintuitive step: We will multiply both sides of the differential equation by e^{-t}

$$\frac{dN}{dt}e^{-t} - Ne^{-t} = te^{-t}$$

Although it looks like we have, needlessly, made our differential equation even more complicated, we notice that the left-hand side can now be rewritten to put the equation in the form

$$\frac{d}{dt}(Ne^{-t}) = te^{-t}$$

$\frac{d}{dt}(Ne^{-t}) = \frac{dN}{dt} \cdot e^{-t} + N \cdot \frac{d}{dt}(e^{-t})$
$= \frac{dN}{dt} \cdot e^{-t} - N \cdot e^{-t}$

This simplification seems like magic, but it is part of the integrating factor method. We can then integrate both sides of the equation with respect to t:

$$Ne^{-t} = \int te^{-t} dt$$ Fundamental theorem of calculus on left-hand side.

$$= -te^{-t} + \int e^{-t} dt$$ Integrating by parts with $u = t$, $v' = e^{-t}$.

$$= -te^{-t} - e^{-t} + C$$

Thus:

$$N(t) = -(t+1) + Ce^{t}$$

To calculate C we apply the initial condition:

$$0 = -1 + C \Rightarrow C = 1$$ $e^0 = 1$

so

$$N(t) = -(t+1) + e^{t}.$$

Now let's develop a more general version of the integrating factor method. Suppose that we are trying to solve the general form of Equation (8.74). In general we will assume that $a(t)$ is continuous and $b(t)$ is differentiable for all of the values of t for which a solution of the equation is sought. Because $a(t)$ is continuous, it has an antiderivative; that is, there is a function $A(t)$ for which $A'(t) = a(t)$. Multiply both sides of (8.74) by $e^{A(t)}$. For the example studied previously, $a(t) = -1$, so $A(t) = -t$ and $e^{A(t)} = e^{-t}$, which is the factor that we multiplied (8.73) by. The factor $e^{A(t)}$ is known as the **integrating factor** of the equation. (In this section we will use IF as an abbreviation for integrating factor in the help text.) Then (8.74) becomes

$$e^{A(t)} \frac{dN}{dt} + a(t)e^{A(t)} N = e^{A(t)} b(t) \tag{8.75}$$

But we recognize that the left-hand side of (8.75) can be rewritten as $\frac{d}{dt}(N(t)e^{A(t)})$, so

$$\frac{d}{dt}(Ne^{A(t)}) = e^{A(t)} b(t)$$ $\frac{d}{dt}(Ne^{A(t)}) = \frac{dN}{dt} e^{A(t)} + N \frac{d}{dt}(e^{A(t)})$

Or on integrating both sides with t:

$$Ne^{A(t)} = \int e^{A(t)} b(t) dt$$

$$N(t) = e^{-A(t)} \int e^{A(t)} b(t) dt \tag{8.76}$$

Note that the differential equation is solved only in the sense that we have converted the problem of calculating $N(t)$ to the problem of finding the integral $\int e^{A(t)} b(t) dt$. Since $b(t)$ is differentiable, $b(t)$ is certainly continuous and so is $A(t)$, so the integral exists, but you may not be able to write it in terms of functions that you know. It is not necessary to memorize Equation (8.76); in general, you can memorize the sequence of steps that led to the result and perform them for any equation of the form (8.74).

EXAMPLE 1 Solve the differential equation:

$$\frac{dy}{dt} = k(y - a), \quad y(0) = 0 \tag{8.77}$$

using the method of integrating factors.

Solution The function on the right-hand side of this differential equation is a function of the dependent variable, y, only; that is, the differential equation is autonomous. It can therefore be solved by separation of variables. However, (8.77) can also be rearranged into the form of (8.74):

$$\frac{dy}{dt} - ky = -ka, \quad y(0) = 0. \qquad a(t) = -k,\ b(t) = -ka$$

We may then solve the equation using integrating factors. For this equation, $\int a(t)dt = -kt$, so:

$$e^{-kt}\frac{dy}{dt} - ke^{-kt}y = -kae^{-kt} \qquad \text{IF is } e^{-kt}$$

$$\frac{d}{dt}(e^{-kt}y) = -kae^{-kt}$$

$$e^{-kt}y = \int (-kae^{-kt})\,dt \qquad \text{Integrate both sides with respect to } t$$

$$= ae^{-kt} + C \qquad \text{We only need one constant of integration}$$

so $y(t) = a + Ce^{kt}$. We calculate the constant of integration, C, by applying the initial conditions

$$0 = a + C \qquad \text{Set } t = 0$$

$$\Rightarrow y(t) = a - ae^{kt}$$

•

EXAMPLE 2 Solve the differential equation

$$\frac{dy}{dx} + \frac{y}{x} = \frac{1}{x^2}, \quad \text{for } x > 0$$

with initial condition $y(1) = 1$.

Solution This equation is of the form (8.74), with x as the independent variable and y as the dependent variable. $a(x) = \frac{1}{x}$ and $b(x) = \frac{1}{x^2}$. Since $\int a(x)dx = \ln x$, the integrating factor is $e^{\ln x} = x$. So

$$x\frac{dy}{dx} + y = \frac{1}{x} \qquad \text{Multiply both sides by } x.$$

$$\frac{d}{dx}(xy) = \frac{1}{x}$$

$$xy = \int \frac{1}{x}dx$$

$$= \ln x + C$$

So

$$y = \frac{\ln x}{x} + \frac{C}{x}$$

We calculate the constant of integration by applying our initial conditions

$$1 = 0 + C \qquad \text{Substitute } x = 1,\ y = 1$$

so

$$y(x) = \frac{\ln x}{x} + \frac{1}{x}$$

•

Solving differential equations using integrating factors often requires us to use integration methods from Chapter 7.

EXAMPLE 3 Solve the differential equation:

$$\frac{dy}{dt} + \left(\frac{2t}{1+t^2}\right)y = t \tag{8.78}$$

with initial condition $y(0) = 0$.

Solution Here $a(t) = \frac{2t}{1+t^2}$. To calculate the integrating factor we need to find the antiderivative of $a(t)$:

$$A(t) = \int \frac{2t}{1+t^2} dt$$

We can evaluate this integral by the method of substitution. Let $u = 1 + t^2$. Then $du = 2t dt$ and $A(t) = \int \frac{du}{u} = \ln u = \ln(t^2 + 1)$.

So, the integrating factor is $e^{A(t)} = e^{\ln(t^2+1)} = t^2 + 1$, and

$$(t^2+1)\frac{dy}{dt} + 2ty = t(t^2+1)$$

$$\frac{d}{dt}\left((t^2+1)y\right) = t(t^2+1) \qquad \text{Multiply both sides by the IF.}$$

$$(t^2+1)y = \int t(t^2+1)dt = \frac{1}{4}t^4 + \frac{1}{2}t^2 + C \qquad \text{Integrate both sides with } t.$$

Or:

$$y(t) = \frac{\frac{1}{4}t^4 + \frac{1}{2}t^2}{t^2+1} + \frac{C}{t^2+1}$$

Calculate the constant C from the initial condition:

$$0 = 0 + C \qquad t = 0, y = 0$$

So

$$y(t) = \frac{t^2}{4}\frac{(t^2+2)}{(t^2+1)} \qquad \text{Factorize the numerator.}$$

One of the most important skills for solving differential equations is to recognize which equations are separable and which are in a form where integrating factors can be used (remember that in some cases, such as Example 1 in this section, both methods of solution can be used).

EXAMPLE 4 Find the general solution of the following equations:

(a) $\frac{dN}{dt} = 3Nt + t^3$ **(b)** $\frac{dy}{dx} = y^2x + x$ **(c)** $\frac{dx}{dt} = x + t - xt - 1$

Solution **(a)** The right-hand side is not separable. If we try to separate out the function of N, we get $3Nt + t^3 = N(3t + t^3/N)$, but the second factor contains both N and t, rather than being a function of t alone. But we can rewrite the equation in the form of (8.74):

$$\frac{dN}{dt} - 3tN = t^3 \qquad a(t) = -3t, b(t) = t^3$$

So the integrating factor is $\exp\left(\int(-3t)dt\right) = e^{-3t^2/2}$:

$$e^{-3t^2/2}\frac{dN}{dt} - 3te^{-3t^2/2}N = t^3e^{-3t^2/2}$$

$$\frac{d}{dt}(e^{-3t^2/2}N) = t^3e^{-3t^2/2}$$

$$e^{-3t^2/2}N = \int t^3e^{-3t^2/2}dt \qquad \text{Integrate both sides with } t.$$

To evaluate the integral we make a change of variables: $s = 3t^2/2$, $ds = 3tdt$:

$$e^{-3t^2/2}N = \frac{2}{9}\int se^{-s}ds \qquad t^3dt = \frac{t^2}{3}3tdt = \frac{2}{9}s\,ds$$

$$= \frac{2}{9}\left(-se^{-s} + \int e^{-s}ds\right) \qquad \text{Integrate by parts } \frac{dv}{ds} = e^{-s}, u = s$$

$$= \frac{2}{9}(-se^{-s} - e^{-s} + C)$$

$$= -\frac{2}{9}\left(\frac{3t^2}{2} + 1\right)e^{-3t^2/2} + C_1 \qquad \text{Define a new constant } C_1 = 2C/9 \text{ and substitute } s = 3t^2/2$$

So:

$$N(t) = -\frac{2}{9}\left(\frac{3t^2}{2} + 1\right) + C_1 e^{3t^2/2}$$

(b) This differential equation is separable; the right-hand side can be written as:

$$y^2x + x = \overbrace{(y^2+1)}^{g(y)}\overbrace{x}^{f(x)}$$

So we can separate variables:

$$\frac{1}{y^2+1}\frac{dy}{dx} = x$$

$$\int \frac{1}{y^2+1}dy = \int x dx \qquad \text{Integrate both sides with } x.$$

$$\tan^{-1}(y) = \frac{1}{2}x^2 + C \qquad \text{Use Table 6.1 in Section 6.2.3.}$$

$$y = \tan\left(\frac{1}{2}x^2 + C\right)$$

(c) The right-hand side is factorizable.

$$x + t - xt - 1 = (x-1)(1-t)$$

So, we can separate variables

$$\frac{1}{x-1}\frac{dx}{dt} = (1-t)$$

$$\int \frac{1}{x-1}dx = \int (1-t)dt \qquad \text{Integrate both sides with } t.$$

$$\ln|x-1| = t - \frac{t^2}{2} + C$$

$$x - 1 = \pm e^C \exp\left(t - \frac{t^2}{2}\right)$$

So, $x(t) = 1 + C_1 \exp(t - t^2/2)$ if we define a new constant $C_1 = \pm e^C$.

This equation can also be written in the form of (8.74) by pulling the terms in x to the left-hand side:

$$\frac{dx}{dt} + x(t-1) = t - 1.$$

It can then be solved using integrating factors (see Problem 9). •

8.4.2 Two-Compartment Models

Solution of equations by integrating factors is particularly useful to solve two-compartment problems. The flow of a drug through a body is often modeled by treating the body as two linked compartments. One compartment might represent the gut (e.g., the intestines), and the other compartment, the blood. Pills containing the drug enter the gut, and from there the drug passes into the blood, where it is used by the body.

Let's start the analysis of two-compartment models by adapting the single-compartment model from Section 8.3.1. Imagine we have two tanks, of volume V_1 and V_2, that contain both water and solute. Suppose that the concentration in the first tank is C_1, and in the second the concentration is C_2 (both $C_1(t)$ and $C_2(t)$ will vary with time). Water flows into the first tank at a rate q. There is also an outflow, also q, of water from the first tank. (Remember, q represents the volume of water being added to or removed from the tank in one unit of time.) However, unlike the scenario described in Section 8.3.1, the water flowing out of the first tank is not lost, but flows directly into the second tank. Water must then flow out of the second tank, also at rate

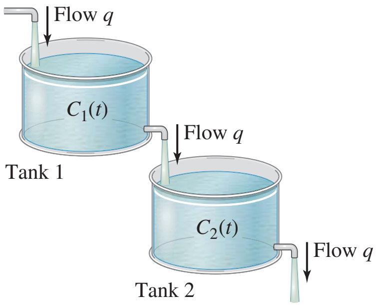


Figure 8.51 Flows of water between two tanks.

q, to keep the volume of water in this tank constant. Figure 8.51 represents the flows of water into and out of each tank.

Suppose that the concentration of solute in the water flowing into tank 1 is C_∞. We also assume, as we did in Section 8.3.1, that the solute in each tank is stirred up well enough to mix through all of the water in the tank. Then the concentration of solute in the flow coming out of tank 1 will be $C_1(t)$, while the concentration in the water coming out of tank 2 will be $C_2(t)$.

We will derive differential equations for the concentration of solute in the two tanks by starting with the word equations:

$$\begin{array}{c}\text{Rate of change of}\\ \text{amount of solute}\\ \text{in tank 1}\end{array} = \begin{array}{c}\text{Rate at which}\\ \text{solute flows into}\\ \text{tank 1}\end{array} - \begin{array}{c}\text{Rate at which}\\ \text{solute flows out}\\ \text{of tank 1}\end{array}$$

Since a volume q of water enters tank 1 in one unit of time, and the concentration of solute in this inflow is C_∞, the total amount of solute flowing into tank 1 in one unit of time is (volume $\times$ concentration) $= qC_\infty$. Similarly, the rate at which solute flows out of the tank is (volume outflow in one unit of time) $\times$ (concentration) $= qC_1$.

So

$$\frac{d}{dt}(C_1V_1) = qC_\infty - qC_1$$

Total amount of solute in tank 1 is $C_1 \times V_1$.

We can write down a similar equation for the solute in tank 2.

$$\frac{d}{dt}(C_2V_2) = qC_1 - qC_2$$

Rate of solute inflow into tank 2 is qC_1.

Because V_1 and V_2 are both constants, we can rewrite these equations in the form:

$$\frac{dC_1}{dt} = \frac{q}{V_1}(C_\infty - C_1) \tag{8.79}$$

and

$$\frac{dC_2}{dt} = \frac{q}{V_2}(C_1 - C_2) \tag{8.80}$$

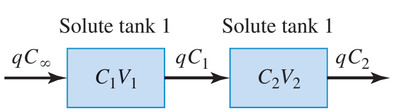


Figure 8.52 Diagram showing flows of solute in the two-compartment model.

We therefore have two differential equations to solve: one for C_1 and one for C_2. But to solve the differential equation for $C_2(t)$ we need to know the value of $C_1(t)$. This is an example of a general problem in math modeling—differential equations with **coupled** variables (i.e., where the terms in one differential equation change with time due to a differential equation of their own). We will develop a general theory for solving this kind of problem in Chapter 11. But in Equation (8.79) the concentration of solute in tank 1 (i.e., C_1) does not depend on the concentration in tank 2. This is made clear if we draw a diagram of the flows of solute in this system (Figure 8.52). If we consider only flows into and out of tank 1, these flows depend only on the concentration C_1 (and on C_∞, which represents the constant concentration of solute in the inflow to tank 1). So we can calculate the concentration $C_1(t)$ without knowing $C_2(t)$ (i.e., solve Equation (8.79) separately from (8.80)).

Equation (8.79) is of a form where we can use either integrating factors or separation of variables to solve the equation. The full solution that accounts for arbitrary initial conditions on C_1 and C_2 is messy and difficult to interpret. So we will solve the system of differential equations in the special case where $C_1(0) = C_2(0) = 0$ (that is, neither tank initially contains any solute). So (8.79) becomes

$$\frac{dC_1}{dt} = \frac{q}{V_1}(C_\infty - C_1) \quad , \quad C_1(0) = 0$$

We solved this equation originally in Section 8.3.1. But the equation is the same as Example 1 in this section, if we rename the variables $y \to C_1$, $a \to C_\infty$, and $k \to -\frac{q}{V_1}$. From Example 1 we may read off the solution:

$$C_1(t) = C_\infty - C_\infty e^{-qt/V_1} = C_\infty(1 - e^{-qt/V_1})$$

To calculate $C_2(t)$ we must substitute our solution for $C_1(t)$ into (8.80):

$$\frac{dC_2}{dt} = \frac{q}{V_2}C_\infty\left(1 - e^{-qt/V_1}\right) - \frac{qC_2}{V_2}. \tag{8.81}$$

Equation (8.81) is not separable, but it can be solved by integrating factors if it is first rewritten as

$$\frac{dC_2}{dt} + \frac{qC_2}{V_2} = \frac{qC_\infty}{V_2}\left(1 - e^{-qt/V_1}\right) \qquad a(t) = q/V_2, b(t) = \tfrac{qC_\infty}{V_2}\left(1 - e^{-qt/V_1}\right).$$

Then:

$$e^{qt/V_2}\frac{dC_2}{dt} + \frac{q}{V_2}e^{qt/V_2}C_2 = \frac{qC_\infty}{V_2}\left(e^{qt/V_2} - e^{qt(1/V_2-1/V_1)}\right) \qquad \text{IF} = e^{qt/V_2}$$

$$\frac{d}{dt}\left(C_2 e^{qt/V_2}\right) =$$

$$C_2 e^{qt/V_2} = \frac{qC_\infty}{V_2}\int \left(e^{qt/V_2} - e^{qt(1/V_2-1/V_1)}\right) dt$$

$$= \frac{qC_\infty}{V_2}\left(\frac{V_2}{q}e^{qt/V_2} - \frac{V_1V_2}{q(V_1-V_2)}e^{qt(1/V_2-1/V_1)}\right) + C \qquad \text{Assume } V_1 \neq V_2;\ (1/V_2 - 1/V_1)^{-1} = \tfrac{V_1V_2}{V_1-V_2}.$$

In our derivation we have assumed that $V_1 \neq V_2$ (i.e., the tanks have different sizes).

Hence:

$$C_2(t) = C_\infty - C_\infty \frac{V_1}{V_1 - V_2}e^{-qt/V_2} + Ce^{-qt/V_2}$$

To calculate the constant of integration, C, apply the initial conditions:

$$0 = C_\infty - C_\infty\left(\frac{V_1}{V_1-V_2}\right) + C. \qquad C_2(0) = 0$$

We find $C = -C_\infty + \frac{V_1}{V_1-V_2}C_\infty$ and so:

$$C_2(t) = C_\infty\left(1 - e^{-qt/V_2}\right) + C_\infty\left(\frac{V_1}{V_1-V_2}\right)\left(e^{-qt/V_2} - e^{-qt/V_1}\right) \tag{8.82}$$

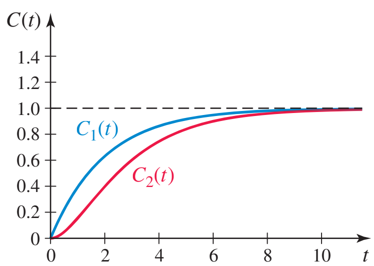


Figure 8.53 Solutions of (8.79) and (8.80) with $C_\infty = 1, q = 1, V_1 = 2, V_2 = 1, C_1(0) = C_2(0) = 0$.

Figure 8.53 shows one solution of the model.

Our solution for $C_2(t)$ in Equation (8.82) is quite complicated, and it is hard, even using the techniques from Section 5.6, to determine how the shape of the graph will depend on all of the constants in our model. From our solution for $C_1(t)$ we can see that $C_1(t) \to C_\infty$ as $t \to \infty$ because $C_1(t)$ can be decomposed as C_∞ plus an exponential term that decays to 0 as $t \to \infty$.

Since all of the exponential terms in (8.82) decay to 0 as $t \to \infty$, $C_2(t)$ must also converge to C_∞ as $t \to \infty$. That is, the concentration of solute in both tanks converges to C_∞, (i.e., to match the concentration in the water flowing into the first tank).

To understand how the solutions behave, it is often helpful to consider what happens if one of the constants in the equation is very small or very large. We will show how this kind of analysis can be used, focusing on the effect of the size of the second tank, V_2, on $C_2(t)$.

If $V_2 \ll V_1$ (the second tank is much smaller than the first), then $\frac{V_1}{V_1-V_2} \approx \frac{V_1}{V_1} = 1$, because we can neglect the V_2 term in the denominator. We can approximate $\frac{V_2}{V_1-V_2} \approx 0$, so:

$$C_2(t) \approx C_\infty\left(1 - e^{-qt/V_1}\right)$$

But this is the same expression as we found for $C_1(t)$. So the concentrations in the two tanks are almost identical. If V_2 is very small, then the concentration in the second tank quickly reaches equilibrium with the concentration in the inflow to the second tank (i.e., to $C_1(t)$). But $C_1(t)$ changes with time t. So the second tank matches the concentration of the first tank and eventually reaches the same equilibrium concentration C_∞. Because $C_2(t)$ is being fed with an inflow whose concentration is $C_1(t)$ rather than C_∞, it is not possible for the second tank to converge to the equilibrium concentration C_∞ faster than the first tank.

On the other hand, if $V_2 \gg V_1$ (the second tank is much larger than the first), then $\frac{V_1}{V_1 - V_2} \approx -\frac{V_1}{V_2} \approx 0$. In this case we may approximate (8.82) by

$$C_2(t) \approx C_\infty \left(1 - e^{-qt/V_2}\right)$$

Our expression for $C_2(t)$ is then the same as the expression for $C_1(t)$, only with the volume V_1 replaced by the (much larger) volume V_2. In fact $C_2(t)$ doesn't depend on V_1 at all in this limit. Our solution for $C_2(t)$ is the same as for a tank of volume V_2 that receives an inflow with constant concentration C_∞. We can understand this as follows: If $V_1 \ll V_2$, then the concentration in the first tank converges to C_∞ much quicker than the second tank. So quickly, in fact, that when considering the second tank we can assume that the first tank reaches C_∞ effectively instantly. So the second tank receives an almost constant inflow concentration C_∞, and can be analyzed as a single tank reaching equilibrium at this concentration.

As we noted in the introduction, one of the most important applications of two-compartment models is the study of how drugs move through the human body. Here we don't have physical tanks with flow between them, but matter moves from one compartment to another, and equations very similar to the above can be used to represent this movement.

EXAMPLE 5 **A Two-Compartment Model for Drug Metabolization** A patient takes a pill containing a drug. The pill enters her gut. From there it then passes into her bloodstream. A fraction f of the drug from the gut enters the patient's bloodstream in each unit of time. Once the drug enters the blood it is used by the patient's body or eliminated. A fraction k of the drug present in the patient's blood is removed in each unit of time. Calculate the amount of drug in the patient's blood as a function of time. Assume that f and k are constants.

Solution Here we model the flow of drug using two compartments: the gut and the blood. Let the amount of medication in the patient's gut be $g(t)$ and the amount in her blood be $b(t)$. We start with word equations for $g(t)$ and $b(t)$.

$$\begin{array}{c}\text{Rate of change of}\\ \text{drug in gut}\end{array} = -\left(\begin{array}{c}\text{Rate at which drug}\\ \text{passes from gut to blood}\end{array}\right)$$

$$\frac{dg}{dt} = -fg \qquad (8.83)$$

No extra drug enters gut unless the patient takes more pills.

At time $t = 0$, when the pill first enters the patient's gut, if no drug was previously present, then $g(t)$ will be equal to the total amount of drug contained in a pill. Let's call the amount of drug in one pill g_0. Then $g(0) = g_0$.

Since the drug leaving the gut enters the patient's blood,

$$\begin{array}{c}\text{Rate of change of}\\ \text{drug in blood}\end{array} = \begin{array}{c}\text{Rate at which}\\ \text{drug passes from}\\ \text{gut to blood}\end{array} - \begin{array}{c}\text{Rate at which}\\ \text{drug is removed}\\ \text{from the blood}\end{array}$$

$$\frac{db}{dt} = fg - kb \qquad (8.84)$$

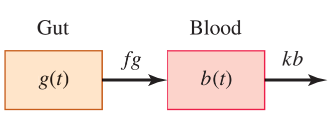


Figure 8.54 Flow of drug between gut and blood for Example 5.

If we assume that this is the first time the patient has taken this particular medication, then at time $t = 0$ there should be no medication present in the patient's blood (i.e., $b(0) = 0$).

We can represent the flow of drug between the two compartments by a diagram (Figure 8.54).

Just as in the two tank problem, the flow to the gut (there is no inflow) does not depend on the level of drug in the blood. That is, Equation (8.83) can be solved independently of Equation (8.84). In fact we can solve (8.83) by either of the methods (separation of variables or integrating factors) that we have studied in this chapter. We will use separation of variables:

$$\int \frac{dg}{g} = -\int f\,dt \quad \Rightarrow \quad g(t) = g_0 e^{-ft}$$

To solve (8.84) we substitute our expression for $g(t)$ into the equation:

$$\frac{db}{dt} = fg - kb = fg_0 e^{-ft} - kb \tag{8.85}$$

Equation (8.85) is not separable but can be solved using integrating factors:

$$\frac{db}{dt} + kb = fg_0 e^{-ft} \qquad \text{First, write in the form of (8.74).}$$

$$e^{kt}\frac{db}{dt} + ke^{kt}b = fg_0 e^{(k-f)t} \qquad \text{IF} = e^{kt}$$

$$\frac{d}{dt}(e^{kt}b) = \int fg_0 e^{(k-f)t}\,dt.$$

Let's assume that $k \neq f$ (the case where $k = f$ is left to you; see Problem 30). Then we can evaluate the integral on the right hand side, giving

$$e^{kt}b(t) = \frac{fg_0}{k-f}e^{(k-f)t} + C \qquad C \text{ is a constant of integration.}$$

so

$$b(t) = \frac{fg_0}{k-f}e^{-ft} + Ce^{-kt}.$$

To find the constant C, apply initial conditions:

$$0 = \frac{fg_0}{k-f} + C \quad \Rightarrow \quad C = -\frac{fg_0}{k-f}$$

Thus

$$b(t) = \frac{fg_0}{k-f}\left(e^{-ft} - e^{-kt}\right). \tag{8.86}$$

To understand our solution better it is useful to sketch a graph of how $b(t)$ varies over time. We can use the techniques from Section 5.6 to draw this graph. We will assume for definiteness that $k > f$. Modifications for $k < f$ will be discussed as we proceed.

Step 1 Find zeros. $b(t) = 0$ only when:

$$e^{-ft} - e^{-kt} = 0 \text{ (i.e., } e^{(k-f)t} = 1).$$

Since $k \neq f$, this occurs only when $t = 0$.

Step 2 Find where function is positive or negative. Since $b(t) = 0$ only when $t = 0$, $b(t)$ must be positive over the entire interval $(0, \infty)$ or negative over the entire interval). Since $b(t)$ represents an amount of drug, we expect only the first case to be possible. But to reason from Equation (8.86) rather than from the science of the model, consider the two factors in (8.86): $(e^{-ft} - e^{-kt})$ and $\frac{fg_0}{k-f}$. Since we are assuming $k > f$, $k - f > 0$, so $\frac{fg_0}{k-f} > 0$. Meanwhile $e^{(k-f)t} > 1$ since e raised to any power is positive, so $e^{-ft} > e^{-kt}$. Thus, both factors are positive, and so $b(t)$ is positive.

Step 3 and 4 Find $b'(t)$ and points where $b(t)$ or $b'(t)$ are undefined. Both $b(t)$ and db/dt are defined for all $t \geq 0$, and:

$$\frac{db}{dt} = \frac{fg_0}{k-f}\left(-fe^{-ft} + ke^{-kt}\right)$$

Steps 5 and 6 Locate local extrema and increasing and decreasing intervals of $b(t)$. From Step 3 we see $\frac{db}{dt} = 0$ only when $ke^{-kt} = fe^{-ft}$, that is, when $e^{(k-f)t} = k/f$ so the only candidate for a local extrema with $t > 0$ is at time $t_1 = \frac{\ln(k/f)}{k-f}$.

Notice that t_1 is a positive number whether $k > f$ or $k < f$. If $k > f$, then $k - f > 0$ and $\ln(k/f) > 0$ (since $\ln x > 0$ if $x > 1$). But if $k < f$, then $k - f < 0$ and $\ln(k/f) < 0$. Since our expression has both a negative numerator and a negative denominator, t_1 will be positive. The derivative $b'(t)$ can only change sign at points where $b'(t) = 0$ (i.e., at the local extremum t_1). We also observe that:

$$b'(0) = \frac{fg_0}{k-f}(k - f) = fg_0 > 0$$

so $b(t)$ starts off increasing for $0 \le t < t_1$. Moreover, since $b(t) \to 0$ as $t \to \infty$, and $b(t) > 0$, $b(t)$ must be decreasing as $t \to \infty$, so $b(t)$ is decreasing for $t > t_1$.

Step 7 Classify extrema. $b(t)$ switches from increasing to decreasing at $t = t_1$, and so $t = t_1$ is a local maximum.

Step 8 Behavior at end points of the interval. Since $b(t) > 0$ for $t > 0$ and $b(t) = 0$, $t = 0$ is a local (in fact a global) minimum point. The other end point of interest is the behavior of $b(t)$ as $t \to \infty$:

$$\lim_{t\to\infty} b(t) = \lim_{t\to\infty}\left(\frac{fg_0}{k-f}e^{-ft}\right) - \lim_{t\to\infty}\left(\frac{fg_0}{k-f}e^{-kt}\right)$$
$$= 0 - 0 = 0$$

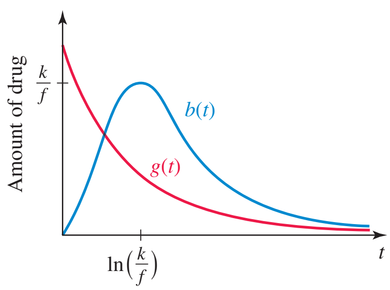


Figure 8.55 Sketch of the solution to the differential equations (8.83 and 8.84). Finding the maximum level of the drug in the patient's blood is left as an exercise (see Problem 30).

Figure 8.55 is a sketch of the solution, containing all of this information.

Two-compartment models are commonly used to predict how the body uses medications. For example, ***Albert and Gernaat (1984)*** collected data on the levels of ibuprofen (sold under the name Motrin®) in the blood of patients being treated for arthritis. We can fit the model to these data, that is, estimate the parameters k and f that make the mathematical model match experimental measurements: to determine how quickly the ibuprofen enters and leaves a patient's blood. The data and fit are shown in Figure 8.56.

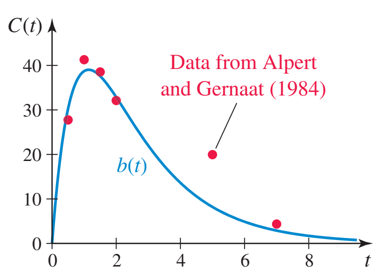


Figure 8.56 We fit our two-compartment model to the data of ***Albert and Gernaat (1984)*** (red crosses). From the fit we can estimate the parameters in our model: $k = 0.53\ \text{hr}^{-1}$ and $f = 1.35\ \text{hr}^{-1}$.

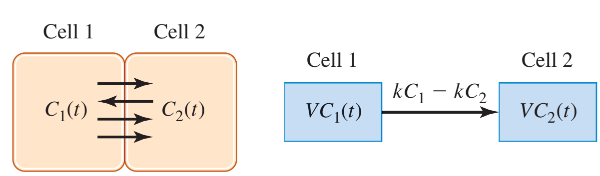


Figure 8.57 Two cells exchange a small molecule through the small region of membrane between them. In the second panel we represent the flow between the two cells.

EXAMPLE 6 **Equilibration Across a Membrane** Although the membrane of a cell may appear to be solid when the cell is viewed through a microscope, at the molecular scale the membrane is riddled with holes that allow small molecules to diffuse into and out of the cell, either into the cell's environment or into nearby cells. Controlling what enters and what leaves is one of the cell's most important tasks. Here we will consider a simplified model for the process by which two cells exchange small molecules.

Imagine the cells are in contact via a region of membrane. The two cells have the same volume V, and their respective concentrations of the molecule are $C_1(t)$ and $C_2(t)$ (see Figure 8.57).

The flow of molecules between the two cells will depend on the difference in concentration between the two cells. A commonly used model for the process of diffusion states that:

$$\begin{array}{c}\text{Rate of flow from}\\ \text{cell 1 to cell 2}\end{array} = k(C_1 - C_2). \tag{8.87}$$

The constant k is called the **permeability** of the membrane—it depends both on the properties of the membrane (the area of contact between the two cells, the size of the holes in the membrane) and upon the molecule (e.g., small molecules typically diffuse across the membrane more readily—that is, they have higher values of k than large molecules).

Now if $C_1 < C_2$, then the rate of flow predicted by (8.87) is negative. In that case molecules flow from cell 2 to cell 1. That is, molecules always tend to diffuse from the cell where the concentration is higher to the cell where the concentration is lower.

Hence, we can write down a word equation for the concentration in the two cells starting with the first cell:

$$\begin{array}{c}\text{Rate of change of}\\ \text{number of molecules}\\ \text{in cell 1}\end{array} = - \begin{array}{c}\text{Rate at which}\\ \text{molecules diffuse}\\ \text{from cell 1 to cell 2}\end{array}$$

which may then be written in mathematical form, i.e.:

$$\frac{d}{dt}(C_1 V) = -k(C_1 - C_2).$$

Similarly for cell 2:

$$\begin{array}{c}\text{Rate of change of}\\ \text{number of molecules}\\ \text{in cell 2}\end{array} = \begin{array}{c}\text{Rate at which}\\ \text{molecules diffuse}\\ \text{from cell 1 to cell 2}\end{array}$$

Molecules diffusing from cell 1 to cell 2 are subtracted from cell 1 and added to cell 2.

$$\frac{d}{dt}(C_2 V) = k(C_1 - C_2)$$

Or, on dividing all equations by the constant V:

$$\frac{dC_1}{dt} = -\frac{k}{V}(C_1 - C_2) \tag{8.88}$$

$$\frac{dC_2}{dt} = \frac{k}{V}(C_1 - C_2) \tag{8.89}$$

Previously we were able to solve one of our equations independently of the other. In this model, however, to calculate the concentration C_1, we need to solve an equation that involves C_2. But to solve the equation for C_2, we need to solve an equation that involves C_1. The two equations are *coupled*. In general, solving this kind of differential equation system requires techniques that will be introduced in Chapter 11. However, in this case a trick allows us to decouple the equations. Suppose that instead of solving (8.88) and (8.89) directly for C_1 and C_2 we try to solve for the quantity

$$C(t) = \frac{1}{2}(C_1(t) + C_2(t)),$$

which represents the average concentration of molecules in the two cells.

Why do we introduce this new dependent variable? If we look at how $C(t)$ changes with time we see:

$$\begin{aligned}\frac{dC}{dt} &= \frac{d}{dt}\left[\frac{1}{2}(C_1 + C_2)\right] = \frac{1}{2}\frac{dC_1}{dt} + \frac{1}{2}\frac{dC_2}{dt}\\ &= -\frac{k}{2V}(C_1 - C_2) + \frac{k}{2V}(C_1 - C_2) = 0\end{aligned} \tag{8.90}$$

So C is a constant. This makes sense, physically. When a molecule leaves cell 1, it enters cell 2, and conversely, so the total number of molecules in both cells must remain constant, so the average must also be constant.

Now if we know C, we can substitute for $C_2(t)$ in Equation (8.88) because $C_2(t) = 2C(t) - C_1(t)$. So:

$$\begin{aligned}\frac{dC_1}{dt} &= -\frac{k}{V}\Big(C_1 - (2C(t) - C_1(t)\Big)\\ &= -\frac{2k}{V}C_1 + \frac{2k}{V}C \qquad \text{$C(t)$ is a constant.}\end{aligned}$$

We can then solve this differential equation by the method of integrating factors.

$$\begin{aligned}\frac{dC_1}{dt} + \frac{2kC_1}{V} &= \frac{2kC}{V} \qquad \text{Write in the form of (8.74)}\\ \frac{dC_1}{dt}e^{2kt/V} + \frac{2kC_1}{V}e^{2kt/V} &= \frac{2kC}{V}e^{2kt/V} \qquad \textstyle\int a(t)\,dt = \frac{2kt}{V}\\ \frac{d}{dt}\left(C_1(t)e^{2kt/V}\right) &= \frac{2kC}{V}e^{2kt/V}\\ C_1(t)e^{2kt/V} &= \frac{2kC}{V}\int e^{2kt/V}\,dt\\ &= Ce^{2kt/V} + A\end{aligned}$$

that is,

$$C_1(t) = C + Ae^{-2kt/V} \tag{8.91}$$

where A is a constant that is determined by the initial conditions on $C_1(t)$. We then solve for $C_2(t)$ from:

$$\begin{aligned}C_2(t) &= 2C - C_1(t)\\ &= C - Ae^{-2kt/V}.\end{aligned} \tag{8.92}$$

Now let's consider a specific problem. Assume that at time $t = 0$, the concentration in cell 1 is A_1 and the concentration in cell 2 is A_2. Find $C_1(t)$ and $C_2(t)$.

Solution To solve the problem we need to use the initial conditions to solve for the coefficients C and A that appear in our expressions for $C_1(t)$ and $C_2(t)$.

First, note that:

$$C = \frac{1}{2}\left(C_1(t) + C_2(t)\right)$$

and since C is a constant:

$$C = \frac{1}{2}\Big(C_1(0) + C_2(0)\Big) = \frac{1}{2}(A_1 + A_2)$$

Now since $C_1(0) = A_1$, substituting into (8.91) we obtain:

$$C + A = A_1$$

$$\text{or} \quad A = A_1 - C = \frac{1}{2}(A_1 - A_2)$$

So by substituting for C and A in (8.91) and (8.92) we obtain:

$$C_1(t) = \frac{1}{2}(A_1 + A_2) + \frac{1}{2}(A_1 - A_2)e^{-2kt/V}$$

$$C_2(t) = \frac{1}{2}(A_1 + A_2) - \frac{1}{2}(A_1 - A_2)e^{-2kt/V}$$

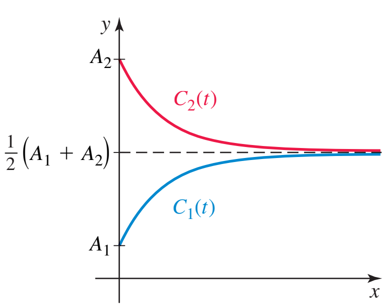


Figure 8.58 Plot of the solution of the two-compartment model for equilibration of the concentration of a small molecule between two cells.

That is, both C_1 and C_2 converge monotonically to $\frac{1}{2}(A_1 + A_2)$. The level of small molecule in cells equilibrates at the same value ($\frac{1}{2}(A_1 + A_2)$). A plot of the solution is given in Figure 8.58.

Section 8.4 Problems

8.4.1

Find the general solution of the differential equations in Problems 1–12 using the method of integrating factors:

1. $\frac{dy}{dt} + \frac{y}{t} = \frac{1}{t^2}$

2. $\frac{dy}{dt} + \frac{3y}{t} = t$

3. $\frac{dy}{dt} - y = t + 1$

4. $\frac{dy}{dt} + y = t^2$

5. $\frac{dy}{dx} + \frac{y}{x+2} = x - 1$

6. $\frac{dy}{dx} + \frac{y}{x+2} = x$

7. $\frac{dy}{dx} - \frac{y}{x(x+1)} = 1$

8. $\frac{dy}{dx} + \frac{2(x+1)y}{x(x+2)} = x$

9. $\frac{dx}{dt} + (t-1)x = t - 1$

10. $\frac{dx}{dt} + \frac{tx}{t^2+1} = t$

11. $\frac{dy}{dx} + \frac{y}{x} = y$

12. $\frac{dy}{dx} + x = y$

For each of the Problems 13–24 you should determine whether the problem needs to be solved using separation of variables or integrating factors (some of the problems may be solved using either method). Then solve the differential equation.

13. $\frac{dy}{dt} = \frac{y}{t} - t^2$

14. $\frac{dy}{dt} = \frac{y}{t+1}$

15. $\frac{dy}{dt} = y^2 t + y^2$

16. $\frac{dy}{dt} = y - yt$

17. $\frac{dy}{dt} = \cos t$

18. $\frac{dy}{dt} = 1 - y$

19. $\frac{dy}{dt} = t^3 + yt$

20. $\frac{dy}{dt} = t + yt$

21. $\frac{dy}{dx} = (x+1)y + (x+1)$

22. $\frac{dy}{dx} = (x+1)y + (x+1)y^2$

23. $\frac{dy}{dx} = \frac{xy}{x+1}$

24. $\frac{dy}{dx} = \frac{x}{y+1}$

8.4.2

In Problems 25–28 consider the two-compartment model for two tanks with respective volumes V_1 and V_2.

$$\frac{dC_1}{dt} = \frac{q}{V_1}(C_\infty - C_1) \tag{8.93}$$

$$\frac{dC_2}{dt} = \frac{q}{V_2}(C_1 - C_2) \tag{8.94}$$

where $C_1(t)$ is the concentration in the first tank and $C_2(t)$ is the concentration in the second tank, and q is the volume of water flowing between the two tanks in one unit of time.

25. When we analyzed (8.93) and (8.94) in the main text we assumed that $V_1 \neq V_2$. Now consider how the analysis must be modified if $V_1 = V_2$, and $C_1(0) = C_2(0) = 0$.

(a) Show that $C_1(t) = C_\infty(1 - e^{-qt/V_1})$ and $C_2(t) = C_\infty\left(1 - \left(1 + \frac{qt}{V_1}\right)e^{-qt/V_1}\right)$

(b) Show that $\lim_{t\to\infty} C_1(t) = C_\infty$ and $\lim_{t\to\infty} C_2(t) = C_\infty$.

26. Let $C_\infty = 0$, so that the fresh water is pumped into tank 1 and flushes solute from tank 1 into tank 2. Now assume that $C_1(0) = 1$ and $C_2(0) = 0$. If $q = 1$, $V_1 = 1$, and $V_2 = 2$, solve the pair of differential equations to find $C_1(t)$ and $C_2(t)$. Sketch both functions of time.

27. Let $C_\infty = 0$, so that the fresh water is pumped into tank 1 and flushes solute from tank 1 into tank 2. Now assume that $C_1(0) = 1$ and $C_2(0) = 0$. If $q = 1$, $V_1 = 3$, and $V_2 = 1$, solve the pair of differential equations to find $C_1(t)$ and $C_2(t)$, and sketch both functions of time.

28. Let $C_\infty = 0$, so that the fresh water is pumped into tank 1 and flushes solute from tank 1 into tank 2. Now assume that $C_1(0) = 1$ and $C_2(0) = 0$. If $q = 1$ and $V_1 = V_2 = 1$, solve the pair of differential equations to find $C_1(t)$ and $C_2(t)$, and sketch both functions of time.

29. Consider a two-compartment model where, instead of having a separate reservoir feeding into tank 1, the two tanks are separated by two pipes, one of which carries water from tank 1 to tank 2, at rate q, and the other carries water from tank 2 to tank 1, at the same rate q. A schematic and diagram of the flows is given in Figure 8.59.

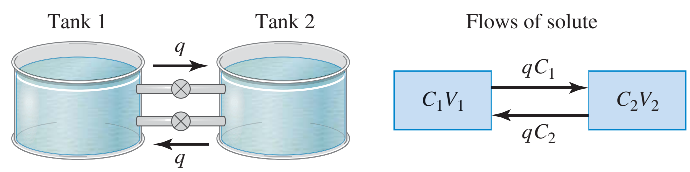


Figure 8.59 Schematic of two-compartment model for Problem 29.

(a) Explain why, although there is no net flow between the tanks, we would expect the concentrations in the tanks to change over time.

(b) Explain why the change in concentrations over time can be modeled using differential equations:

$$\frac{dC_1}{dt} = \frac{q}{V_1}(C_2 - C_1)$$
$$\frac{dC_2}{dt} = \frac{q}{V_2}(C_1 - C_2) \tag{8.95}$$

(c) To solve the differential equations in (8.95) start by assuming that $V_1 = V_2$. Then define $C(t) = \frac{1}{2}(C_1 + C_2)$ and by deriving a differential equations for dC/dt and explain why $C(t)$ is constant.

(d) Using the fact that $C(t)$ is a constant, eliminate $C_2(t)$ from the equation for $\frac{dC_1}{dt}$. Solve the equation you then obtain, and write down expressions for $C_1(t)$ and $C_2(t)$.

(e) Use the expression from part (d) to explain why, no matter what the starting values for $C_1(0)$ and $C_2(0)$ are, we expect $C_1(t)$ and $C_2(t)$ to converge to the same limit as $t \to \infty$.

(f) To solve the differential equations in (8.95) in the most general case ($V_1 \neq V_2$), let $C(t) = \frac{V_1C_1 + V_2C_2}{V_1 + V_2}$ (the weighted average of the concentrations in the two tanks). Explain why $C(t)$ is a constant.

(g) Using the fact that $C(t)$ is a constant, eliminate $C_2(t)$ from the equation for $\frac{dC_1}{dt}$. Solve the equation you then obtain, and write down expressions for $C_1(t)$ and $C_2(t)$.

(h) Use the expression from part (g) to explain why, no matter what the starting values for $C(t)$ and $C_2(t)$ are, we expect $C_1(t)$ and $C_2(t)$ to converge to the same limit as $t \to \infty$.

30. **Drug Modeling** In Example 5 we analyzed the flow of a medication from a patient's gut to their blood. Reanalyze this model assuming the rate of elimination of medication from the patient's blood is the same as the rate at which medication passes from the gut into blood. Then, if the amount of medication in the patient's blood is $b(t)$ and the amount in the patient's gut is $g(t)$:

$$\frac{dg}{dt} = -fg$$

and

$$\frac{db}{dt} = fg - fb$$

where f is a positive constant.

(a) Show that just as in Example 5, $g(t) = g_0 e^{-ft}$ where g_0 is the amount of drug in the pill the patient takes at time $t = 0$.

(b) Solve for $b(t)$, assuming $b(0) = 0$, and sketch the graph of $b(t)$ against t.

31. Find the maximum level of medication in a patient's blood, if the passage of medication through the patient's body is modeled using Equation (8.86).

32. Filling Box Models In Problem 10 of Section 8.3 we analyzed the concentration in a tank whose volume changes over time because the inflows and outflows are not matched. For such a tank it can be shown that if the concentration in the inflow is C_I, and the inflow and outflow rates are respectively q_{in} and q_{out}, then both concentration, $C(t)$, and volume of water in the tank, $V(t)$, vary with time and can be modeled by a pair of differential equations:

$$\frac{d}{dt}(CV) = q_{\text{in}} C_I - q_{\text{out}} C$$

and (8.96)

$$\frac{dV}{dt} = q_{\text{in}} - q_{\text{out}}.$$

(a) Show that the differential equations (8.96) imply that

$$((q_{\text{in}} - q_{\text{out}})t + V_0)\frac{dC}{dt} + q_{\text{in}} C = q_{\text{in}} C_I \tag{8.97}$$

where V_0 is the initial volume of water in the tank.

(b) Assuming that $q_{\text{in}} = 2$, $q_{\text{out}} = 1$, $V_0 = 1$, $C(0) = 0$, and $C_I = 1$, solve (8.97) using integrating factors to find $C(t)$.

(c) Assuming that $q_{\text{in}} = 1$, $q_{\text{out}} = 2$, $V_0 = 1$, $C(0) = 0$, and $C_I = 1$, solve (8.97) using integrating factors to find $C(t)$. What does your model predict will occur when $t = 1$? Explain whether this answer makes sense given that $V(1) = 0$.

Chapter 8 Review

Key Terms

Discuss the following definitions and concepts:

1. Differential equation
2. Separable differential equation
3. Solution of a differential equation
4. Pure-time differential equation
5. Autonomous differential equation
6. Exponential growth
7. Von Bertalanffy equation
8. Logistic equation
9. Allometric growth
10. Equilibrium
11. Stability
12. Eigenvalue
13. Single-compartment model
14. Patchy habitat model
15. Colonization rate
16. Mortality rate
17. Allee effect
18. Chemical reaction model
19. Reaction rate
20. Evolutionary game theory
21. Snow-drift game
22. Cooperator
23. Cheater
24. Pay-off matrix
25. Epidemic model
26. Susceptible
27. Infected
28. SI-model
29. Infection rate
30. Recovery rate
31. Endemic disease
32. Integrating factor
33. Two-compartment model
34. Coupled equations

Review Problems

1. ***For each of the following differential equations, find the general solution:***

(a) $dx/dt = 2 - x$,

(b) $dy/dx = \frac{1}{y} - \frac{1}{y^2}$,

(c) $dy/dx = yx - x$,

(d) $dy/dx = \frac{y}{x} + x^2$.

2. ***For each of the following differential equations, sketch the vector field plot, and identify any equilibria as well as determining their stability.***

(a) $\frac{dN}{dt} = N(N-1)(3-N)$,

(b) $dN/dt = N(1-N)$,

(c) $dx/dt = x^3 - 1$.

3. Newton's Law of Cooling and Time of Death Suppose that an object has temperature T and is brought into a room that is kept at a constant temperature T_a. Newton's law of cooling states that the rate of temperature change of the object is proportional to the difference between the temperature of the object and the surrounding medium.

(a) Denote the temperature at time t by $T(t)$, and explain why

$$\frac{dT}{dt} = k(T_a - T) \tag{8.98}$$

is the differential equation that expresses Newton's law of cooling.

(b) Derive the solution to the differential equation, assuming that at time $t = 0$, the temperature of the object is $T = T_0$.

Newton's law of cooling can be used to estimate the time of death of a person during criminal investigations. When we are alive, our bodies tend to maintain a constant temperature of around 37°C. On death, bodies start to cool. Assume that a cooling dead body obeys Equation (8.98). ***Molnar et al. (1969)*** found that cooling bodies have a cooling coefficient between $k = 0.04\ \text{hr}^{-1}$ and $k = 0.09\ \text{hr}^{-1}$.

(c) A body is found at 10 p.m. The body's temperature when it was found was 27.0°C, and the temperature of the room is 20.0°C. Estimate the range of possible times of death if k is between 0.04 hr and 0.09 hr.

(d) To make a more accurate estimate of time of death it can be helpful to measure k directly. To do this, the body's temperature is measured at 11 p.m. At this time, the body temperature is found to be 26.4°C (you may assume the room temperature remains constant at 20.0°C). Using the temperatures measured at 10 p.m. and 11 p.m., estimate k.

(e) Use your new measurement of k from part (d) to make a new estimate of the time of death.

4. Photosynthesis *(Adapted from* ***Horn, 1971****)* The following model is a simplified model of photosynthesis: Suppose that a leaf contains a number of traps that can capture light. If a trap captures light, the trap becomes energized. The energy in the trap can then be used to produce sugar, which causes the energized trap to become unenergized. The number of traps that can become energized is proportional to the number of unenergized traps and the intensity of the light. Denote by T the total number of traps (unenergized and energized) in a leaf, by I the light intensity, and by x the number of energized traps. Then the following differential equation describes how the number of energized traps changes over time:

$$\frac{dx}{dt} = k_1(T - x)I - k_2 x$$

Here, k_1, k_2, and I are positive constants. Find all equilibria and their stability.

5. Chemical Reactions A chemical reaction between chemical A and chemical B produces chemical C. Two molecules from chemical C can then combine to produce chemical D. The chain of reactions can be written as

$$\begin{array}{ll} A + B \to C & \text{with rate constant } k_{AB} \\ 2C \to D & \text{with rate constant } k_C. \end{array}$$

Chemicals A and B are continuously added to the system to keep A and B at constant concentrations a and b respectively. Then, the concentration $x(t)$ of C can be modeled using a differential equation:

$$\frac{dx}{dt} = k_{AB}ab - 2k_C x^2 \tag{8.99}$$

(a) Analyze (8.99); that is, find the equilibria of the equation and their stability.

(b) Assume that $k_{AB} = 1, k_C = 1, a = 1, b = 1$. Draw the vector field plot for (8.99).

(c) Now sketch the graph of the solution $x(t)$ against t if $x(0) = 0$ (it is not necessary to solve the differential equation to sketch the solution).

6. Island Biogeography ***Preston (1962)*** and ***MacArthur and Wilson (1963)*** investigated the effect of area on species diversity in oceanic islands. In their model, animals and plants of different species continuously travel to the island from the mainland. The fraction of mainland species that can be found on the island is p, which we call the **diversity of the island**. We want to derive a model for how $p(t)$ will change with time, t.

The number of different species present on the island will be affected by extinction and immigrations; species may die out, while other species from the mainland start new populations on the island. In general we expect:

$$\frac{dp}{dt} = \text{Rate of immigrations} - \text{Rate of extinctions}$$

Assume that a fraction m of species go extinct in one unit of time. Assume that, if a species is present on the mainland, then there is a probability c that it will emigrate to the island in one unit of time. But only species not present already on the island will add to the number of species there.

(a) Explain why we may model the diversity of the island using a model:

$$\frac{dp}{dt} = c(1 - p) - mp. \tag{8.100}$$

(b) Find the equilibria for p and determine their stability.

(c) Assume that, for a particular island, $c = 2$ and $m = 1$. If $p(0) = 1$ (that is, initially the island contains all species found on the mainland) solve Equation (8.100) to calculate $p(t)$.

(d) One question of interest in the field of biogeography—how diversity is affected by the physical environment—is how the area of the island affects its diversity. Assume that the extinction rate m decreases as island area, A, increases. But the rate of colonization, c, is unaffected. Will the equilibrium value of p increase or decrease if A is increased?

7. The Prisoner's Dilemma In this problem we will discuss a classic model for organism interactions based on ***Nowak (2006)***. This model is based on a game called **prisoner's dilemma**. Two individuals are charged with a crime. If neither confesses (i.e., they cooperate with each other) then it cannot be proven that they committed the crime, and they will be sentenced for a more minor crime. However, each prisoner has the option of confessing, and blaming their partner for the crime. In this case the person who confesses walks free, while the other person is sent to jail. We may regard this as a form of cheating—the cheater (person who confesses) receives a reward (no prison time), but at the expense of the cooperator (the partner who doesn't confess). But if both partners confess, then both will be sent to prison (though sharing the blame means that their sentences are reduced). We will imagine a community of organisms interacting via this game. We can summarize net benefit to each player using a payoff matrix.

We will put specific numbers in the payoff matrix to make the analysis clearer.

		Opponent	
		Cooperate	Cheat
Player	Cooperate	3	0
	Cheat	6	1

Assuming that each organism interacts with n others in unit time, and that reproductive rate is proportional to the total benefit from all of these interactions, then it can be shown that the proportion $x(t)$ of cooperators will obey a differential equation:

$$\frac{dx}{dt} = knx(x - 1)(2x + 1) \tag{8.101}$$

where $k > 0$ is a constant (as in Section 8.3.4). The proportion $y(t)$ of cheaters can be determined from $y = 1 - x$.

(a) Find the stable and unstable equilibria of Equation (8.101).

(b) What is the long time behavior of $x(t)$?

(c) Consider a population made up only of cooperators ($x = 1$). What is the total benefit that each individual in this population receives in one unit of time?

(d) Now consider a population made up only of cheaters ($x = 0$). What is the total benefit that each individual in this population receives in one unit of time?

Your answers from (c) and (d) illustrate an effect known as **tragedy of the commons**. Although in populations of cooperators all individuals get more benefits than in populations of cheaters, cheaters always prosper and eventually take over in a mixed population.

8. Insulin Infusion An insulin pump is used to treat diabetes by continuously infusing insulin into the fat in a patient's abdomen or thigh. From there the insulin enters the patient's bloodstream. We will model this process using a two-compartment model.

Assume that insulin is pumped into a small region of fat at a constant rate r. This region will be the first compartment in our model. A fraction f of the insulin in the fat then diffuses into the patient's blood in each unit of time. Within the patient's blood, a fraction k of the insulin present is eliminated in one unit of time r, and f and k are constants. So if we denote the amount of insulin in the patient's fat by x, and the amount in their blood by y then:

$$\frac{dx}{dt} = r - fx. \tag{8.102}$$

$$\frac{dy}{dt} = fx - ky. \tag{8.103}$$

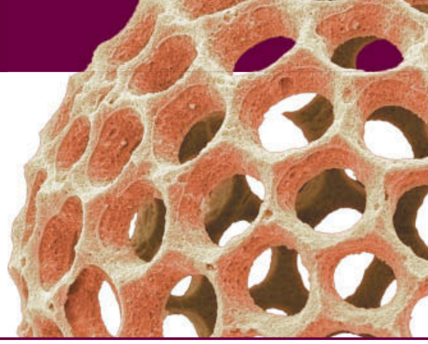
(a) By analyzing Equation (8.102) find the equilibrium level of insulin in the patient's fat (that is, the equilibrium value of x) and determine whether this equilibrium is stable or unstable.

(b) Assuming that $x(t)$ converges to its equilibrium value (found in part (a)), what is the equilibrium value for $y(t)$; that is, find the value of y that ensures that $dy/dt = 0$?

(c) Assuming the following parameter values: $r = 2$, $f = 1$, $k = 1$, and that at time $r = 0$, there is no insulin present either in the patient's fat or in their blood (that is, $x(0) = y(0) = 0$), solve the system of equations (8.102) and (8.103), and confirm that $x(t)$ and $y(t)$ converge to the equilibrium values that you calculated in parts (a) and (b).

CHAPTER

9



Linear Algebra and Analytic Geometry

In this chapter we will take the first steps toward extending the methods of calculus from functions of a single variable to functions of multiple variables (also called multivariate functions) and to systems of functions. Our first step will be to consider an important class of multivariate functions: functions that are linear in several variables. We will learn how to solve equations involving these functions, and how to represent the solutions graphically. Specifically, we will learn how to

- solve systems of linear equations;
- define matrices and perform algebraic operations on matrices;
- use Leslie matrices to model age-structured populations;
- define and analyze linear maps; and
- define lines and planes in more than two dimensions.

The material in this chapter focuses on what is needed for Chapters 10 and 11. We will not have space to discuss one of the most important applications of this material: vectors and matrices are also vital for the statistical models that are used in bioinformatics. Students planning to study bioinformatics further will find themselves using the material in this chapter extensively.

9.1 Linear Systems

An ecologist is studying how quickly different species of fungi spread when they are introduced into new habitats. The fungi spread using spores, which are blown about by the wind. The ecologist believes that fungi will spread faster in environments with higher wind speeds (for example, windspeeds are faster in grasslands than in forests). However, she also believes that smaller spores will disperse more easily than large spores, because they are lighter and more easily carried by the wind. So the distance that spores spread will depend both on the windspeed, and upon the size of the spores. Suppose that the distance traveled is f, the windspeed is x, and the spore size is y. Then we say that f is a function of both x and y. In Chapter 10 we will study functions of more than one variable (also called **multivariate functions**) at length. For now it is sufficient to know that, to predict f, we need to know both x and y.

The ecologist wants to know how f depends on x and y. A simple and commonly used model is that f may be a **linear function** of x and y. That is:

$$f(x, y) = ax + by$$

for some pair of coefficients a and b. Compare this with linear functions as we encountered them in Section 1.2. In Section 1.2 we said that $g(x)$ was a linear function of x if $g(x) = mx + c$ for a pair of constants m and c. Compare this expression with the one for $f(x, y)$. If we ignore the dependence of $f(x, y)$ upon y, and only look at how it depends upon x, we see that $f(x, y)$ is a linear function of x, with a playing the role of the slope parameter m. Similarly, if we ignore its dependence on x, then $f(x, y)$ is a linear function of y, with b now playing the role of the slope parameter m. That is, f is linear in both of its variables.

Once the form of the model has been decided, the ecologist can use her data to estimate the values of a and b. Using real data to determine unknown coefficients in a model is known as **data-fitting**.

How many measurements are necessary to estimate the parameters in the model? Suppose that x (the wind speed) is measured in meters per second (m/s) and y (the spore size) is measured in microns (usually abbreviated as μm). If the ecologist measures that 1 μm spores are blown 30 m by a 10 m/s wind, then:

$$f(10, 1) = 30$$
$$\text{so} \quad 10 \cdot a + 1 \cdot b = 30. \tag{9.1}$$

This is not enough information to uniquely determine a and b. For example, $f(10, 1) = 30$ is consistent with $a = 3$ and $b = 0$ (because then $f = 10 \cdot 3 + 1 \cdot 0 = 30$), or we could have $a = 2$ and $b = 10$, or $a = 1$ and $b = 20$, and so on. In fact we can rewrite (9.1) as:

$$10a = 30 - b$$
$$a = 3 - b/10. \tag{9.2}$$

No matter what the value of b is, (9.2) can be used to find a value for a that makes $f(10, 1)$ equal to 30. In general, with two unknown coefficients in our equation we need to know two values of $f(x, y)$ to solve for both a and b. Suppose that the ecologist measures the dispersal of a second species of fungus. This species grows in a habitat with wind speed $x = 20$ m/s, and has spore size $y = 3$ μm. The ecologist finds that its spores travel 55 m. So:

$$f(10, 1) = 10a + b = 30$$
$$f(20, 3) = 20a + 3b = 55.$$

Now we have **two linear equations in two unknowns** (a and b). Certainly one equation is not enough to solve for a and b. But if we have two equations, will we be able to find unique values of a and b that solve both of them? The answer is not always. In this section we will study the conditions under which systems of linear equations have unique solutions, as well as present a method for finding these solutions if they exist.

9.1.1 Graphical Solution

We want to solve a system of linear equations (that is, more than one linear equation). In this subsection we will restrict ourselves to systems of two linear equations in two variables. That is, we want to solve, in x and y, the pair of equations:

$$\begin{aligned} Ax + By &= C \\ Dx + Ey &= F \end{aligned} \tag{9.3}$$

where A, B, C, D, E, and F are constants and x and y are the two variables. (We require that A and B and that D and E are not both equal to 0.) When we say that we "solve" (9.3) for x and y, we mean that we find values for x and y that satisfy *each* equation of the system (9.3). Recall that the standard form of a linear equation in two variables is

$$Ax + By = C$$

where A, B, and C are constants, A and B are not both equal to 0, and x and y are the two variables; the graph of this equation is a straight line. (See Figure 9.1.) Any point (x, y) on this straight line satisfies (or solves) the equation $Ax + By = C$. Similarly the second equation in 9.3; $Dx + Ey = F$ also specifies a straight line. Because each equation in (9.3) describes a straight line, we are therefore asking for the point of intersection of these two lines. The following three cases are possible:

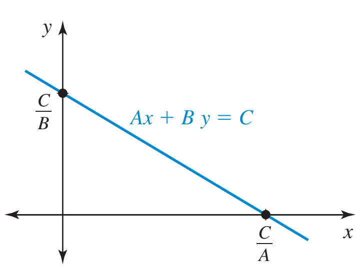


Figure 9.1 The graph of a linear equation in standard form.

1. The two lines have exactly one point of intersection. In this case, the system (9.3) has exactly one solution, as illustrated in Figure 9.2.
2. The two lines are parallel and do not intersect. In this case, the system (9.3) has no solution, as illustrated in Figure 9.3.
3. The two lines are parallel and intersect (i.e., they are identical). In this case, the system (9.3) has infinitely many solutions—namely, each point on the line, as illustrated in Figure 9.4.

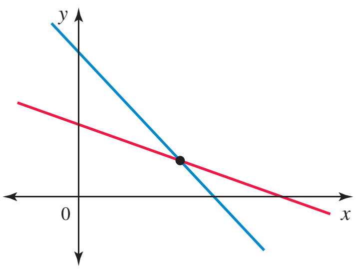


Figure 9.2 The two lines have exactly one point of intersection.

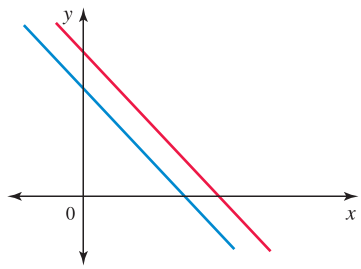


Figure 9.3 The two lines are parallel but do not intersect.

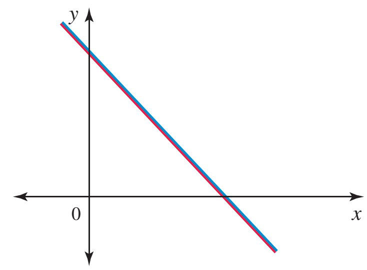


Figure 9.4 The two lines are identical.

Exactly One Solution

EXAMPLE 1 Find the solution of

$$\begin{aligned} 2x + 3y &= 6 \\ 2x + y &= 4. \end{aligned} \tag{9.4}$$

Solution The line corresponding to $2x + 3y = 6$ has y-intercept $(0, 2)$ and x-intercept $(3, 0)$; the line corresponding to $2x + y = 4$ has y-intercept $(0, 4)$ and x-intercept $(2, 0)$. (See Figure 9.5.) To find the solution of the linear system (9.4), we need to find the point of intersection of the two lines. Solving each equation for y in terms of x produces the new set of equations:

$$\begin{aligned} y &= 2 - \frac{2}{3}x \\ y &= 4 - 2x. \end{aligned}$$

Setting the right-hand sides equal to each other, we obtain

$$\begin{aligned} 2 - \frac{2}{3}x &= 4 - 2x \\ \frac{4}{3}x &= 2 \qquad \text{Solve for } x. \\ x &= \frac{3}{2}. \end{aligned}$$

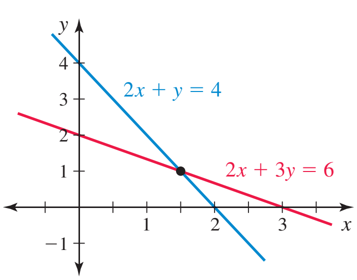


Figure 9.5 In Example 1, the two lines have exactly one point of intersection.

To find y, we substitute the value of x back into one of the two original equations (say, $2x + y = 4$; it does not matter which one you choose). Then

$$y = 4 - 2x = 4 - (2)\left(\frac{3}{2}\right) = 1$$

and the solution is the point $(x, y) = (3/2, 1)$.

Looking at Figure 9.5, we see that the two lines have exactly one point of intersection. This is so because the lines have different slopes: $-2/3$ and -2, respectively. The point of intersection is (3/2,1) and corresponds to the only solution of (9.4). •

No Solution

EXAMPLE 2 Solve

$$\begin{aligned} 2x + y &= 4 \\ 4x + 2y &= 6 \end{aligned} \tag{9.5}$$

Solution The line corresponding to $2x + y = 4$ has y-intercept $(0, 4)$ and x-intercept $(2, 0)$; the line corresponding to $4x + 2y = 6$ has y-intercept $(0, 3)$ and x-intercept $(3/2, 0)$. (See Figure 9.6.) Since

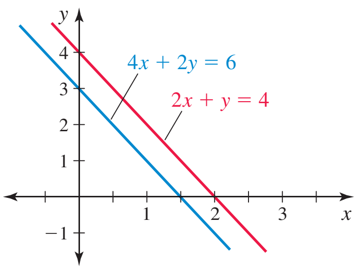


Figure 9.6 In Example 2, the two lines are parallel but do not intersect.

$$\begin{aligned} 2x + y = 4 \quad &\Longleftrightarrow \quad y = 4 - 2x \\ 4x + 2y = 6 \quad &\Longleftrightarrow \quad y = 3 - 2x \end{aligned}$$

both lines have the same slope, namely, -2, but different y-intercepts: 4 and 3, respectively. This implies that the two lines are parallel and do not intersect.

Let's see what happens when we solve the system (9.5). We equate the two equations $y = 4 - 2x$ and $y = 3 - 2x$:

$$4 - 2x = 3 - 2x$$

This implies that

$$4 = 3$$

The last expression is obviously wrong. We conclude that there is no point (x, y) that satisfies both equations in (9.5) simultaneously.

Looking at Figure 9.6, we see that the two lines are parallel. Since parallel lines that are not identical do not intersect, (9.5) has no solution. (Just look at the graph.) In this case, we write the solution as $\emptyset$, the symbol for the empty set. •

Infinitely Many Solutions

EXAMPLE 3 Solve

$$\begin{aligned} 2x + y &= 4 \\ 4x + 2y &= 8 \end{aligned} \tag{9.6}$$

Solution If we divide the second equation by 2, we find that both equations are identical, namely, $2x + y = 4$. That is, both equations describe the same line with x-intercept $(2, 0)$ and y-intercept $(0, 4)$, as shown in Figure 9.7. Every point (x, y) on this line is therefore a solution of (9.6). To find the solution algebraically, we use the same procedure as in Examples 1 and 2. We first solve each equation for y:

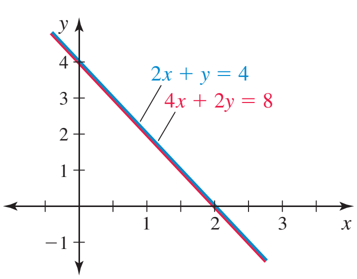


Figure 9.7 In Example 3, the two lines are identical.

$$\begin{aligned} 2x + y = 4 \quad &\Longrightarrow \quad y = 4 - 2x \\ 4x + 2y = 8 \quad &\Longrightarrow \quad y = 4 - 2x \end{aligned}$$

Equating the two equations $y = 4 - 2x$ and $y = 4 - 2x$ yields

$$4 - 2x = 4 - 2x$$

which simplifies to

$$0 = 0$$

This is a true statement, which implies that any value of x is a solution. A convenient way to write the solution is to introduce a new variable, say, t, to denote the x-coordinate. The new variable t can take on any real number. To find the corresponding y-coordinate, we substitute $t \in \mathbf{R}$ for x:

$$y = 4 - 2x = 4 - 2t$$

The solution can then be written as the set of points

$$\{(t, 4 - 2t) : t \in \mathbf{R}\}$$

reflecting the fact that the system (9.6) has infinitely many solutions, as expected from the graphical considerations. [Figure 9.7 shows that the two lines representing (9.6) are identical.] •

In Example 3, we introduced the variable t to describe the set of solutions. We call t a **dummy variable**; it stands for any real number. Introducing a dummy variable is a convenient way to describe the set of solutions when there are infinitely many.

9.1.2 Solving Equations Using Elimination

The graphical and algebraic way of solving systems of linear equations we have employed so far works only for systems in two variables. To solve systems of linear equations in more than two variables, we will need to develop a method that will work for a system of any size. The basic strategy will be to transform the system of linear equations into a new system of equations that has the same solutions as the original. The new system is called an **equivalent system**. It will be of a simpler form, so that we can solve for the unknown variables one by one and thus arrive at a solution. We illustrate this approach in the next example. We tag all equations with labels of the form (R_i); R_i stands for "ith row." The labels will allow us to keep track of our computations.

EXAMPLE 4 Solve

$$\begin{aligned} 3x + 2y &= 8 \qquad (R_1) \\ 2x + 4y &= 5 \qquad (R_2) \end{aligned}$$

Solution There are two basic operations that transform a system of linear equations into an equivalent system: (1) We can multiply an equation by a nonzero number, and (2) we can add any multiple of one equation to the other.

Our goal will be to eliminate x in the second equation. If we multiply the first equation by 2 and the second equation by -3, we obtain

$$\begin{array}{lrl} 2(R_1) & 6x + \;\;4y = & 16 & \\ -3(R_2) & -6x - 12y = & -15 & \quad x\text{-coefficients are now equal and opposite} \end{array}$$

If we add the two equations, we find that

$$2(R_1) - 3(R_2) \qquad -8y = 1$$

Through algebraic manipulations, we eliminated x from the second equation. We can replace the original system of equations by a new (equivalent) system by leaving the first equation unchanged and replacing the second equation by $-8y = 1$. We then obtain the equivalent system of equations [labeled (R_3) and (R_4)]

$$\begin{array}{rrl} (R_1) & 3x + 2y = 8 & (R_3) \\ 2(R_1) - 3(R_2) & -8y = 1 & (R_4) \end{array}$$

We can now successively solve the system. It follows from equation (R_4) that

$$y = -\frac{1}{8}$$

Substituting this value into equation (R_3) and solving for x, we find that

$$\begin{aligned} 3x + (2)\left(-\frac{1}{8}\right) &= 8 \\ 3x &= 8 + \frac{1}{4} \\ x &= \frac{11}{4} \end{aligned}$$

The solution is therefore $(x, y) = (11/4, -1/8)$. •

When we look at equations (R_3) and (R_4), we see that the left-hand side has the shape of a triangle:

$$\begin{aligned} Ax + By &= C \\ Dy &= E \end{aligned}$$

For some set of constants A, B, C, D, E. We therefore call a system that is written in this form **upper triangular**. Note that the first equation involves both x and y, but the second equation involves only y. We can therefore solve the second equation for y, and substitute into the first equation to solve for y. This method of solving equations, called **elimination** or **Gaussian elimination**, eliminates x from the second row of the system. As we will see in the next subsection, this method can be generalized to larger systems of equations.

9.1.3 Solving Systems of Linear Equations

In this subsection, we will extend the solution method of Example 4 to systems of m equations in n variables, which we write in the form

$$\begin{array}{ll} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 & (R_1) \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 & (R_2) \\ \dots\dots\dots\dots\dots\dots\dots\dots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m & (R_m) \end{array} \tag{9.7}$$

The variables are now $x_1, x_2, \dots, x_n$. The coefficients a_{ij} on the left-hand side have two subscripts. The first subscript indicates the equation, and the second subscript indicates to which variable (i.e., x_j) a_{ij} belongs. For instance, you would find a_{21} in the second equation as the coefficient of x_1; a_{43} would be in the fourth equation as the coefficient of x_3. Using double subscripts is a convenient way of labeling the coefficients. The subscripts on the numbers b_i on the right-hand side of (9.7) indicate the equation.

We will transform this system into an equivalent system in upper triangular form. (Recall that *equivalent* means that the new system has the same solutions as the old system.) That is, our goal is to reduce the rows of the system of equations so that the first row contains the variables $x_1, x_2, \dots, x_n$, while the second row contains only $x_2, x_3, \dots, x_n$ (i.e., x_1 has been eliminated), and the third row depends only on $x_3, x_4, \dots, x_n$ (i.e., both x_1 and x_2 have been eliminated), and so on. To do so, we will use the following three basic operations:

1. Multiplying an equation by a nonzero constant
2. Adding one equation to another
3. Rearranging the order of the equations

Elimination is an iterative process. To reduce a system like (9.7) into upper triangular form, first follow the steps below:

Gaussian Elimination Steps

1. Label the rows of the equation, e.g., (R_1), (R_2), (R_3), …, (R_m).
2. If (R_1) does not contain the first variable x_1, then reorder rows by swapping (R_1) with a row that does contain x_1.
3. Keep (R_1) for the first row.
4. To eliminate x_1 from (R_2), multiply (R_1) by a_{21} and multiply (R_2) by $-a_{11}$. Then add the two rows.
5. Repeat step 4 to eliminate x_1 from (R_3), (R_4), …, (R_m).

At the end of these steps you will have eliminated x_1, from rows (R_2), (R_3), …, (R_m). Don't worry about memorizing the precise manipulation needed for step 4; it is enough to remember that you need to use row (R_1) to eliminate x_1 from row (R_2). Compare the coefficients of x_1 in rows (R_1) and (R_2) and find what multiple of (R_1) and what multiple of (R_2) would have equal and opposite coefficients for x_1.

After following the Gaussian elimination steps, your new rows (R_2) through (R_m) will make up a system of $(m-1)$ equations in $(n-1)$ unknowns, namely $x_2, x_3, \dots, x_n$, because x_1 has been eliminated from all of them. Repeat the Gaussian elimination steps to eliminate x_2 from rows $(R_3), (R_4), \dots (R_m)$. Then use the same steps to eliminate x_3 from rows $(R_4), (R_5) \dots (R_m)$, and so on. With enough patience and care any system of linear equations can be reduced to upper triangular form by this method, no matter how many equations there are, or how many variables they contain.

As in the case of a linear system with two equations in two variables, the general system (9.7) may have

1. Exactly one solution
2. No solutions
3. Infinitely many solutions

When a system has no solutions, we say that the system is **inconsistent**.

Exactly One Solution

EXAMPLE 5 Solve

$$\begin{aligned} 2x - y + 3z &= 9 & (R_1)\\ 3x + 5y - z &= 10 & (R_2)\\ 4x + 2y - 3z &= -1 & (R_3) \end{aligned} \tag{9.8}$$

Solution Our goal is to reduce this system to upper triangular form. The first step is to eliminate x from the second and third equations.

We leave the first equation unchanged. Multiplying the first equation by 2 and the second equation by -3:

$$\begin{aligned} -3(R_1) \quad -6x + 3y - 9z &= -27 \\ 2(R_2) \quad 6x + 10y - 2z &= 20 \qquad x \text{ coefficients are now equal and opposite} \end{aligned}$$

Then add the equations to eliminate x:

$$-3(R_1) + 2(R_2) \qquad 13y - 11z = -7.$$

We transform the third equation by multiplying the first equation by 2 and the third equation by -1.

$$\begin{aligned} 2(R_1) \quad 4x - 2y + 6z &= 18 \\ -(R_3) \quad -4x - 2y + 3z &= 1 \qquad x \text{ coefficients are now equal and opposite} \end{aligned}$$

Adding the two equations eliminates x:

$$2(R_1) - (R_3) \qquad -4y + 9z = 19$$

These two steps transform the original set of equations into the following equivalent set of equations, labeled (R_4)–(R_6):

$$\begin{aligned} (R_1) \qquad 2x - y + 3z &= 9 & (R_4)\\ -3(R_1) + 2(R_2) \qquad 13y - 11z &= -7 & (R_5)\\ 2(R_1) - (R_3) \qquad -4y + 9z &= 19 & (R_6) \end{aligned}$$

(R_5) and (R_6) together make a system of two linear equations in two unknowns (y and z). We use Gaussian elimination on these rows, i.e., try to eliminate y from (R_6).

Multiply (R_5) by 4 and (R_6) by 13:

$$\begin{aligned} 4(R_5) \quad 52y - 44z &= -28 \\ 13(R_6) \quad -52y + 117z &= 247 \qquad y \text{ coefficients are now equal and opposite} \end{aligned}$$

Adding the two equations eliminates y:

$$4(R_5) + 13(R_6) \qquad 73z = 219.$$

We leave the first two equations unchanged. The new (equivalent) system of equations is thus

$$\begin{array}{rrcl} (R_4) & 2x - y + 3z = & 9 & (R_7) \\ (R_5) & 13y - 11z = & -7 & (R_8) \\ 4(R_5)+13(R_6) & 73z = & 219 & (R_9) \end{array} \tag{9.9}$$

The system of equations is now in upper triangular form. We use back-substitution to find the solution. Solving equation (R_9) for z yields

$$z = \frac{219}{73} = 3.$$

Solve (R_8) for y and substitute the value of z:

$$y = \frac{1}{13}(-7 + 11z) = \frac{1}{13}(-7 + (11)(3)) = \frac{26}{13} = 2.$$

Solve (R_7) for x and substitute the values of y and z:

$$\begin{aligned} x &= \frac{1}{2}(9 + y - 3z) \\ &= \frac{1}{2}(9 + 2 - 3(3)) = 1 \end{aligned}$$

Hence, the solution is $x = 1$, $y = 2$, and $z = 3$. •

In Example 5 we eliminated x from row (R_2) and x and y from row (R_3). However, we could have started with:

$$\begin{aligned} 3x + 5y - z &= 10 \\ 4x + 2y - 3z &= -1 \\ 2x - y + 3z &= 9 \end{aligned}$$

which is the same set of equations as (9.8), only in a different order. Accordingly, Gaussian elimination will produce a different equivalent set of upper triangular system of equations than (9.9). But the solutions will be the same.

With so much arithmetic being required to solve a system of linear equations, it is quite easy to make a mistake. It is, however, easy to check that your solution is correct: Substitute for x, y, z into the original equations and make sure that all are satisfied. In Example 5, if we substitute $x = 1$, $y = 2$, $z = 3$ into (R_1), (R_2), and (R_3), we obtain:

$$\begin{array}{rl} 2x - y + 3z = 2(1) - (2) + 3(3) = 2 - 2 + 9 = 9 & (R_1) \\ 3x + 5y - z = 3(1) + 5(2) - 3 = 3 + 10 - 3 = 10 & (R_2) \\ 4x + 2y - 3z = 4(1) + 2(2) - 3(3) = 4 + 4 - 9 = -1 & (R_3) \end{array}$$

So the system of equations is satisfied.

No Solution

EXAMPLE 6 Solve

$$\begin{array}{rl} 2x - y + z = 3 & (R_1) \\ 4x - 4y + 3z = 2 & (R_2) \\ 2x - 3y + 2z = 1 & (R_3) \end{array}$$

Solution Begin by eliminating terms involving x from the second and third equations. We leave the first equation unchanged. To eliminate x from (R_2) and (R_3) we form the following equations:

$$\begin{array}{rrl} (R_1) & 2x - y + z = 3 & (R_4) \\ 2(R_1) - (R_2) & 2y - z = 4 & (R_5) \\ (R_1) - (R_3) & 2y - z = 2 & (R_6) \end{array}$$

We calculated (R_5) by observing that, if we multiply (R_1) by 2, and (R_2) by (-1), then the two equations have equal and opposite coefficients for the term in x (i.e., $4x$ and $-4x$, respectively). (R_6) is calculated from observing that (R_1) and $-(R_3)$ have equal and opposite coefficients for the term in x (i.e., $2x$ and $-2x$, respectively).

You should notice that there is a problem with (R_5) and (R_6): The left-hand sides of the two equations are the same, but the right-hand sides are different. It is impossible that $2y - z$ can be equal to 4 and also to 2, so the system does not have a solution. Don't worry if you did not notice the inconsistency between (R_5) and (R_6); it will become apparent when we try to use (R_5) to eliminate y from (R_6).

$$\begin{array}{rrcl}
(R_4) & 2x - y + z &=& 3 \\
(R_5) & 2y - z &=& 4 \\
(R_5) - (R_6) & 0 &=& 2
\end{array}$$

The last equation, $0 = 2$, is a false statement, which means that this system does not have a solution. We therefore write the solution as $\emptyset$, the symbol denoting the empty set. •

Infinitely Many Solutions

EXAMPLE 7 Solve

$$\begin{array}{rcll}
x - 3y + z &=& 4 & (R_1) \\
x - 2y + 3z &=& 6 & (R_2) \\
2x - 6y + 2z &=& 8 & (R_3)
\end{array}$$

Solution We proceed as in Examples 5 and 6. We leave the first equation unchanged, and eliminate x from (R_2) and (R_3):

$$\begin{array}{rrcll}
(R_1) & x - 3y + z &=& 4 & (R_4) \\
(R_2) - (R_1) & y + 2z &=& 2 & (R_5) \\
(R_1) - \frac{1}{2}(R_3) & 0z &=& 0 & (R_6)
\end{array}$$

Notice that eliminating x from (R_3) also eliminates y from that row, putting the system in upper triangular form without additional steps being needed.

The third equation, $0z = 0$, is a correct statement. It means that we can substitute any number for z and still solve the system of equations. We introduce the dummy variable t and set

$$z = t \quad \text{for } t \in \mathbf{R}.$$

If we solve (R_5) for y and substitute $z = t$ into the resulting equation, we get

$$y = 2 - 2z = 2 - 2t.$$

If we solve (R_4) for x and substitute $y = 2 - 2t$ and $z = t$ into the resulting equation, we find that

$$\begin{aligned}
x &= 4 + 3y - z = 4 + 3(2 - 2t) - t \\
&= 4 + 6 - 6t - t = 10 - 7t
\end{aligned}$$

The solution is therefore the set

$$\{(x, y, z) : x = 10 - 7t, y = 2 - 2t, z = t, \text{for } t \in \mathbf{R}\}.$$ •

Just like in Example 3, the set of solutions for Example 7 forms a line of points, and our solution is one way of writing the equation for that line.

You should be prepared to solve systems of linear equations even when the variables are not named x, y, z, etc. In the next example we return to the problem from the introduction.

EXAMPLE 8 **Predicting Spore Dispersal** The distance f that mushroom spores travel is predicted to depend on the windspeed x and on the size of spores y. A simple model for this dependence is:

$$f(x, y) = ax + by$$

By fitting f to data the coefficients a and b can be estimated.

Suppose that when the windspeed is 10 m/s, 1 μm spores travel 30 m, while 3 μm spores in a wind of 20 m/s travel 55 m. Estimate the parameters a and b.

Solution We are given data:

$$\begin{aligned} & f(10, 1) = 10a + b = 30 \qquad (R_1) \\ \text{and} \qquad & f(20, 3) = 20a + 3b = 55 \qquad (R_2) \end{aligned} \tag{9.10}$$

with which to estimate a and b. (9.10) is a system of two equations in two unknowns (a and b). Although our unknowns are no longer x, y, or z, the equations can still be solved by Gaussian elimination.

$$\begin{aligned} (R_1) \qquad & 10a + b = 30 \qquad (R_3) \\ (R_2) - 2(R_1) \qquad & b = -5 \qquad (R_4), \quad \text{Eliminate } a \end{aligned}$$

So $b = -5$. Substitute for b into (R_3) to calculate a:

$$a = \frac{30 - b}{10} = 3 - \frac{1}{10}b = 3.5.$$

9.1.4 Representing Systems of Equations Using Matrices

When we transform a system of linear equations, we make changes only to the coefficients of the variables. It is convenient to introduce notation that will simply keep track of all the coefficients. This motivates the following definition:

Definition A **matrix** is a rectangular array of numbers:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

The elements a_{ij} of the matrix A are called **entries**. If the matrix has m rows and n columns, it is called an $\boldsymbol{m \times n}$ **matrix**.

We can use a matrix A to represent the coefficients of the linear system we introduced in (9.7):

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{9.11}$$

In this case A is called the **coefficient matrix** of the equation system. The entries a_{ij} of the matrix have two subscripts: The entry a_{ij} is located in the ith row and the jth column. You can think of this numbering system as the street address of the entries, using streets and avenues. Suppose that avenues go north–south and streets go east–west; then finding the corner of Second Street and Third Avenue would be like finding the element of A that is in the second row and third column, which is a_{23} in the matrix.

If a matrix has the same number of rows as columns, it is called a **square matrix**. An $m \times 1$ matrix is called a **column vector** and a $1 \times n$ matrix is called a **row vector**. Following are examples of a 3×3 square matrix, a 3×1 column vector, and a 1×4 row vector (the shape of each of the matrices explains their names):

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & -1 \\ 5 & 4 & 3 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 & 0 & 5 \end{bmatrix}.$$

If A is a square matrix, then the **diagonal line** of A consists of the elements $a_{11}, a_{22}, \ldots, a_{nn}$. In the preceding 3×3 square matrix, the diagonal line thus consists of the elements 1, 1, and 3.

To solve systems of linear equations with the use of matrices, we introduce the **augmented matrix**—the coefficient matrix of the linear system (9.11), augmented by an additional column representing the right-hand side of (9.11). The augmented matrix representing the linear system (9.11) is therefore

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right]$$

So how can we use augmented matrices to solve systems of linear equations?

EXAMPLE 9 Find the augmented matrix for the system given in Example 5, and solve the system by using the augmented matrix.

Solution The augmented matrix is given by

$$\begin{array}{c} \left[\begin{array}{ccc|c} 2 & -1 & 3 & 9 \\ 3 & 5 & -1 & 10 \\ 4 & 2 & -3 & -1 \end{array}\right] \end{array} \begin{array}{l} (R_1) \\ (R_2) \\ (R_3) \end{array}$$

Our goal is to transform the augmented matrix into the simpler, upper triangular form, in which all the entries below the diagonal line are zero. We use the same basic transformations as before, and you should compare the steps that follow with the steps in Example 5. We indicate on the left side of the matrix the transformation that we used, and we relabel the rows on the right side of the matrix:

$$\begin{array}{r} (R_1) \\ -3(R_1)+2(R_2) \\ 2(R_1)-(R_3) \end{array} \left[\begin{array}{ccc|c} 2 & -1 & 3 & 9 \\ 0 & 13 & -11 & -7 \\ 0 & -4 & 9 & 19 \end{array}\right] \begin{array}{l} (R_4) \\ (R_5) \\ (R_6) \end{array}$$

That is, we copied row 1, which we now call row 4; to get the second row, we formed $-3(R_1)+2(R_2)$ to eliminate the entry (a_{21}), calling the resulting row (R_5), and so on.

The next step is to eliminate the -4 in (R_6). We find that

$$\begin{array}{r} (R_4) \\ (R_5) \\ 4(R_5)+13(R_6) \end{array} \left[\begin{array}{ccc|c} 2 & -1 & 3 & 9 \\ 0 & 13 & -11 & -7 \\ 0 & 0 & 73 & 219 \end{array}\right] \begin{array}{l} (R_7) \\ (R_8) \\ (R_9) \end{array}$$

This is now the augmented matrix for the system of linear equations in (9.9), and we can proceed as in Example 5 to solve the system by back-substitution.

So far, all of the systems that we have considered have had the same number of equations as variables. This need not be the case, however. When a system has fewer equations than variables, we say that it is **underdetermined**. Although underdetermined systems can be inconsistent (have no solutions), they frequently have infinitely many solutions. When a system has more equations than variables, we say that it is **overdetermined**. Overdetermined systems are frequently inconsistent.

In Example 10 we will look at an underdetermined system, and in Example 11 an overdetermined system.

EXAMPLE 10 Solve the following underdetermined system:

$$\begin{aligned} 2x + 2y - z &= 1 \quad (R_1) \\ 2x - \ \ y + z &= 2 \quad (R_2) \end{aligned}$$

Solution The system has fewer equations than variables and is therefore underdetermined. We use the augmented matrix

$$\left[\begin{array}{rrr|r} 2 & 2 & -1 & 1 \\ 2 & -1 & 1 & 2 \end{array}\right] \begin{array}{l} (R_1) \\ (R_2) \end{array}$$

to solve the system. Transforming the augmented matrix into upper triangular form, we find that

$$\begin{array}{r} (R_1) \\ (R_1)-(R_2) \end{array} \left[\begin{array}{rrr|r} 2 & 2 & -1 & 1 \\ 0 & 3 & -2 & -1 \end{array}\right] \begin{array}{l} (R_3) \\ (R_4) \end{array}$$

Translating this matrix back into a system of equations, we obtain

$$\begin{aligned} 2x + 2y - \ \ z &= \ \ 1 \\ 3y - 2z &= -1 \end{aligned}$$

It then follows that

$$y = -\frac{1}{3} + \frac{2}{3}z$$

and

$$2x = 1 - 2y + z = 1 - 2\left(-\frac{1}{3} + \frac{2}{3}z\right) + z = \frac{5}{3} - \frac{1}{3}z$$

or

$$x = \frac{5}{6} - \frac{1}{6}z$$

We use a dummy variable again and set $z = t$, $t \in \mathbf{R}$; therefore, $x = \frac{5}{6} - \frac{1}{6}t$ and $y = -\frac{1}{3} + \frac{2}{3}t$. The solution can then be written as the equation for a straight line:

$$\left\{(x, y, z) : x = \frac{5}{6} - \frac{1}{6}t,\ y = -\frac{1}{3} + \frac{2}{3}t,\ z = t,\ t \in \mathbf{R}\right\}$$

•

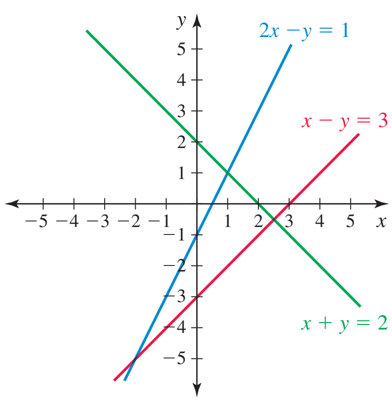


Figure 9.8 The straight lines in Example 11 do not intersect in one point.

EXAMPLE 11 Solve the following overdetermined system:

$$\begin{aligned} 2x - y &= 1 \quad (R_1) \\ x + y &= 2 \quad (R_2) \\ x - y &= 3 \quad (R_3) \end{aligned}$$

Solution The system has more equations than variables and is therefore overdetermined. To solve it, we write

$$\begin{array}{rrcl} (R_1) & 2x - \ \ y &=& \ \ 1 \quad (R_4) \\ 2(R_2)-(R_1) & 3y &=& \ \ 3 \quad (R_5) \\ (R_2)-(R_3) & 2y &=& -1 \quad (R_6) \end{array}$$

It follows from (R_6) that $y = -\frac{1}{2}$ and from (R_5) that $y = 1$. Since $1 \neq -\frac{1}{2}$, there cannot be a solution. The system is inconsistent, and the solution set is the empty set.

Since all three equations involve only the variables x and y, they can each be represented by straight lines. In order for the system to have solutions, all three lines would need to intersect in one point (or be identical). They don't (see Figure 9.8), so there is no solution.

•

EXAMPLE 12 **More Than One Dummy Variable** Solve

$$\begin{aligned} 3x - 3y + 6z &= 9 \qquad (R_1) \\ -x + y - 2z &= -3 \qquad (R_2) \end{aligned}$$

Solution We see that $(R_1) = -3(R_2)$. Thus, eliminating the variable x from (R_2) yields

$$\begin{array}{rr} (R_1) & 3x - 3y + 6z = 9 \\ (R_1)+3(R_2) & 0z = 0 \end{array}$$

We can write the solution by introducing the dummy variable t and setting $z = t, t \in \mathbf{R}$. To solve (R_1), we need a second dummy variable s, which satisfies $y = s, s \in \mathbf{R}$:

$$3x = 9 + 3s - 6t$$

or

$$x = 3 + s - 2t$$

The solution can be written as

$$\{(x, y, z) : x = 3 + s - 2t, y = s, z = t; s \in \mathbf{R}, t \in \mathbf{R}\}$$

We will discuss how to interpret this solution in Section 9.5.

Section 9.1 Problems

9.1.1, 9.1.2

In Problems 1–4, solve each linear system of equations. In addition, for each system, graph the two lines corresponding to the two equations in a single coordinate system and use your graph to explain your solution.

1. $\begin{aligned} x - y &= 1 \\ x - 2y &= -2 \end{aligned}$

2. $\begin{aligned} 2x + y &= 6 \\ x - 4y &= -4 \end{aligned}$

3. $\begin{aligned} x - 2y &= 2 \\ y &= 1 + \tfrac{1}{2}x \end{aligned}$

4. $\begin{aligned} 2x + y &= \tfrac{1}{3} \\ 6x + 3y &= 1 \end{aligned}$

5. Determine c such that

$$\begin{aligned} 2x - 3y &= 5 \\ 4x - 6y &= c \end{aligned}$$

has **(a)** infinitely many solutions and **(b)** no solutions. **(c)** Is it possible to choose a number for c so that the system has exactly one solution? Explain your answer.

6. **(a)** Determine the solution of

$$\begin{aligned} -2x + 3y &= 5 \\ ax - y &= 1 \end{aligned}$$

in terms of a.

(b) For which values of a are there no solutions, exactly one solution, and infinitely many solutions?

7. Show that the solution of

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

where $a_{11}, a_{12}, a_{21}, a_{22}, b_1$ and b_2 are all constants, is given by

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{21}a_{12}}$$

and

$$x_2 = \frac{-a_{21}b_1 + a_{11}b_2}{a_{11}a_{22} - a_{21}a_{12}}$$

8. Show that if

$$a_{11}a_{22} - a_{21}a_{12} \neq 0$$

then the system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= 0 \\ a_{21}x_1 + a_{22}x_2 &= 0 \end{aligned}$$

has exactly one solution, namely, $x_1 = 0$ and $x_2 = 0$.

In Problems 9–16, reduce the system of equations to upper triangular form and find all the solutions.

9. $\begin{aligned} 2x - y &= 3 \\ x - 3y &= 7 \end{aligned}$

10. $\begin{aligned} 5x - 3y &= 2 \\ 2x + 7y &= 3 \end{aligned}$

11. $\begin{aligned} 7x - y &= 4 \\ 3x + 2y &= 1 \end{aligned}$

12. $\begin{aligned} 5x + 2y &= 8 \\ -x + 3y &= 9 \end{aligned}$

13. $\begin{aligned} 3x - y &= 1 \\ -3x + y &= 4 \end{aligned}$

14. $\begin{aligned} 2x + 3y &= 5 \\ -y &= -2 + \tfrac{2}{3}x \end{aligned}$

15. $\begin{aligned} x + 2y &= 3 \\ 4y + 2x &= 6 \end{aligned}$

16. $\begin{aligned} x - 2y &= 2 \\ 4y - 2x &= -4 \end{aligned}$

17. **Predicting Heart Disease.** The risk of a person developing heart disease is correlated with their BMI (or body mass index), which is calculated by dividing their body mass by the square of their height (high BMIs typically mean the person is overweight). Heart disease is also related to a person's level of activity. Suppose a person's BMI is x and their physical activity is y (measured, for example, in minutes of exercise each day). Let h be a heart disease risk index (high values of h mean high risk of heart disease, low values of h mean a low risk of heart disease). h depends on x and y, that is, $h = h(x, y)$. A simple linear model is that if only the effects of BMI and activity are considered:

$$h(x, y) = ax + by$$

for some constants a and b.

(a) Would you expect $a > 0$ or $a < 0$? What about b?

(b) To fit the values of a and b, consider the following data.

A patient with BMI of 20, and who does 50 min of activity each day, has risk index $h = 25$.

A patient with BMI of 35, and who does 10 min of activity each day, has risk index $h = 315$.

Use these data to estimate the values of a and b.

(c) How accurate is this model likely to be? In other words name some other factors besides physical activity and BMI that could affect h, and should therefore be included in the model.

18. Plant Growth The rate of growth of a plant depends on the amount of light available to it, which depends on whether it is growing in shade or full sun. Let the light that the plant receives be x. x could, for example, be measured in lumens (lumen is a unit for total light absorption). Growth rate also depends on how many herbivorous insects graze on the plant. Denote the number of nearby herbivorous insects by y (y could, for example, represent the number of herbivorous insects per square meter of habitat). A simple model is that the rate of growth, r is a linear function of x and y; that is:

$$r(x, y) = ax + by$$

for some constants a and b.

(a) Would you expect $a > 0$ or $a < 0$? What sign do you expect b will have?

(b) Consider the following data:

A plant that receives an average of 3000 lumens of light, and that has 10 herbivorous insects per square meter, grows at a rate of 20 cm/yr.

A different plant, receiving an average of 5000 lumens of light, has 40 herbivorous insects per square meter, and grows at a rate of 10 cm/yr.

Use these data to fit a and b.

(c) How accurate is this model likely to be? In other words name some other factors besides light and number of herbivorous insects that could affect r and therefore should be in the model.

9.1.3

In Problems 19–24, solve each system of linear equations.

19. $\begin{aligned} 2x - 3y + z &= -1 \\ x + y - 2z &= -3 \\ 3x - 2y + z &= 2 \end{aligned}$

20. $\begin{aligned} 5x - y + 2z &= 6 \\ x + 2y - z &= -1 \\ 3x + 2y - 2z &= 1 \end{aligned}$

21. $\begin{aligned} 2x + y + z &= 7 \\ 3x + 2y + z &= 9 \\ x + y - z &= 0 \end{aligned}$

22. $\begin{aligned} -2x + 4y - z &= -1 \\ x + 7y + 2z &= -4 \\ 3x - 2y + 3z &= -3 \end{aligned}$

23. $\begin{aligned} 2x - y + 3z &= 3 \\ 2x + y + 4z &= 4 \\ 2x - 3y + 2z &= 2 \end{aligned}$

24. $\begin{aligned} 2x + y - 2z &= 3 \\ 2x - 3y - 4z &= 0 \\ x - 5y + 3z &= -6 \end{aligned}$

9.1.4

In Problems 25–28, find the augmented matrix and use it to solve the system of linear equations.

25. $\begin{aligned} 3x - 2y + z &= 4 \\ 4x + y - 2z &= -12 \\ 2x - 3y + z &= 7 \end{aligned}$

26. $\begin{aligned} -x - 2y + 3z &= -9 \\ 2x + y - z &= 5 \\ 4x - 3y + 5z &= -9 \end{aligned}$

27. $\begin{aligned} y + x &= 3 \\ z - y &= -1 \\ x + z &= 2 \end{aligned}$

28. $\begin{aligned} 2x + z &= 4y - 1 \\ x + 2y + 9 &= 3z \\ 3x + 2z &= 4 - 2y \end{aligned}$

In Problems 29–34, determine whether each system is overdetermined or underdetermined; then solve each system.

29. $\begin{aligned} x - 2y + z &= 3 \\ 2x - 3y + z &= 8 \end{aligned}$

30. $\begin{aligned} x - y &= 2 \\ x + y + z &= 3 \end{aligned}$

31. $\begin{aligned} 2x - y &= 3 \\ x - y &= 4 \\ x - 3y &= 1 \end{aligned}$

32. $\begin{aligned} 4y - 3z &= 6 \\ 2y + z &= 1 \\ y + z &= 0 \end{aligned}$

33. $\begin{aligned} 2x - 7y + z &= 2 \\ x + y - 2z &= 4 \end{aligned}$

34. $\begin{aligned} x + y &= -1 \\ 2x - y &= 7 \\ x - 2y &= 8 \end{aligned}$

35. SplendidLawn sells three types of lawn fertilizer: SL 24–4–8, SL 21–7–12 and SL 17–0–0. The three numbers refer to the percentages of nitrogen, phosphate, and potassium, in that order, of the contents. (For instance, 100 g of SL 24–4–8 contains 24 g of nitrogen.) Suppose that each year your lawn requires 500 g of nitrogen, 100 g of phosphate, and 180 g of potassium. How much of each of the three types of fertilizer should you apply?

36. Microbiological Diversity DNA sequencing allows the different bacteria and fungi present in a patch of soil to be identified. Many new species have been found by this method, and it also reveals the diversity of microorganism ecosystems.

You are an ecologist explaining differences in diversity between different soil habitats. (We have previously met the Gini-Simpson diversity index and Shannon diversity index as ways of quantifying diversity.) You believe that diversity is affected by the amount of rainfall (because rain introduces new microbes into the soil, and also leads to ground-water flows that redistribute microbes). Also, some microbes (e.g., *Streptomyces* bacteria) produce antibiotics that can suppress the growth of other microbes, reducing the overall diversity. Let x be the amount of rainfall (measured, e.g., in mm/day) and y be the number of antibiotic-producing species that are present. Then you hypothesize that diversity d is a linear function of x and y:

$$d(x, y) = ax + by + c$$

where a, b, c are all constants.

(a) Explain why including the constant c allows d to be non-zero even when $x = 0$ and $y = 0$. Does that make sense biologically?

(b) Do you expect $a > 0$ or $a < 0$? What sign do you expect b to have?

(c) Use the following data to fit the parameters a, b, and c.

A patch of soil with 3 mm/day average rainfall, and no antibiotic-producing species has diversity $d = 0.65$.

A patch of soil with 5 mm/day average rainfall, and 10 antibiotic-producing species has diversity $d = 0.65$.

A patch of soil with 1 mm/day average rainfall, and 5 antibiotic-producing species has diversity $d = 0.5$.

9.2 Matrices

We introduced matrices in the previous section; in this section, we will learn how to add and multiply matrices.

9.2.1 Matrix Operations

Recall that an $m \times n$ matrix A is a rectangular array of numbers with m rows and n columns. We write it as

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \left[\, a_{ij} \,\right]_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

We will also use the shorthand notation $A = [a_{ij}]$ if the size of the matrix is clear.

Definition Equality of Matrices Suppose that $A = [a_{ij}]$ and $B = [b_{ij}]$ are two $m \times n$ matrices. Then

$$A = B$$

if and only if, for all $1 \le i \le m$ and $1 \le j \le n$,

$$a_{ij} = b_{ij}$$

This definition says that we can compare matrices of the same size, and they are equal if all their corresponding entries are equal. The next definition shows how to add matrices.

Definition Addition of Matrices Suppose that $A = [a_{ij}]$ and $B = [b_{ij}]$ are two $m \times n$ matrices. Then

$$C = A + B$$

is an $m \times n$ matrix with entries

$$c_{ij} = a_{ij} + b_{ij} \quad \text{for } 1 \le i \le m, 1 \le j \le n$$

Note that addition is defined only for matrices of equal size. Matrix addition satisfies the following two properties:

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$

The matrix with all its entries equal to zero is called the **zero matrix** and is denoted by **0**. The following holds:

$$A + \mathbf{0} = A$$

We can multiply matrices by numbers. To be clear whether we are multiplying a matrix by a number, or multiplying two matrices (which we will learn to do later), we call in this section this process **multiplication by a scalar**. A **scalar** is another name for a number.

Definition Multiplication by a Scalar Suppose that $A = [a_{ij}]$ is an $m \times n$ matrix and c is a scalar. Then cA is an $m \times n$ matrix with entries ca_{ij} for $1 \le i \le m$ and $1 \le j \le n$.

EXAMPLE 1 Find $A + 2B - 3C$ if

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ -1 & -3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

Solution

$$A + 2B - 3C = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix} + 2\begin{bmatrix} 0 & 1 \\ -1 & -3 \end{bmatrix} - 3\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ -2 & -6 \end{bmatrix} + \begin{bmatrix} -3 & 0 \\ 0 & -9 \end{bmatrix}$$ Multiplication by a scalar

$$= \begin{bmatrix} 2+0-3 & 3+2+0 \\ 1-2+0 & 0-6-9 \end{bmatrix} = \begin{bmatrix} -1 & 5 \\ -1 & -15 \end{bmatrix}$$ Addition of matrices •

EXAMPLE 2 Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 4 \end{bmatrix}$. Find B so that

$$A + B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Solution If

$$A + B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

then

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 3 \\ 0 & 4 \end{bmatrix}$$

$$B = \begin{bmatrix} 1-1 & 1-3 \\ 1-0 & 1-4 \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 1 & -3 \end{bmatrix}$$ Addition of matrices •

It is often useful when working with matrices to interchange rows and columns. This operation is called **transposition**.

> **Definition Transposition of Matrices** Suppose that $A = [a_{ij}]$ is an $m \times n$ matrix. Then the **transpose** of A, denoted by A', is an $n \times m$ matrix with entries
>
> $$a'_{ij} = a_{ji}$$

The next example shows how this operation works.

EXAMPLE 3 Transpose the following matrices:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 4 \end{bmatrix}$$

Solution To find the transpose, we need to interchange rows and columns. Since A is a 2×3 matrix, its transpose is the 3×2 matrix

$$A' = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

To check this against our definition, note that

$$a'_{11} = a_{11} = 1, \quad a'_{12} = a_{21} = 4, \quad a'_{21} = a_{12} = 2, \ldots$$

The matrix B is a 2×1 matrix, which is a column vector. Its transpose is the 1×2 matrix

$$B' = \begin{bmatrix} 1 & 2 \end{bmatrix}.$$

which is a row vector. Similarly, C is a 1×2 row vector; its transpose is then the 2×1 column vector

$$C' = \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

•

Additional properties of the transpose of a matrix are discussed in Problems 17–20.

In the preceding example, we saw that the transpose of a row vector is a column vector and vice versa. When we need to write a column vector in text, such as $X = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, we can instead write $X = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}'$, which, because it is the transpose of a row vector, is really a column vector. A large column vector written as the transpose of a row vector takes up less space in text and is often easier to write.

9.2.2 Matrix Multiplication

The multiplication of two matrices is more complicated than any of the operations we learned in the previous subsection. We give the definition first.

Definition Matrix Multiplication Suppose that $A = [a_{ij}]$ is an $m \times l$ matrix and $B = [b_{ij}]$ is an $l \times n$ matrix. Then

$$C = AB$$

is an $m \times n$ matrix whose entries are given by

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{il}b_{lj} = \sum_{k=1}^{l} a_{ik}b_{kj}$$

for $1 \leq i \leq m$ and $1 \leq j \leq n$ (see Figure 9.9).

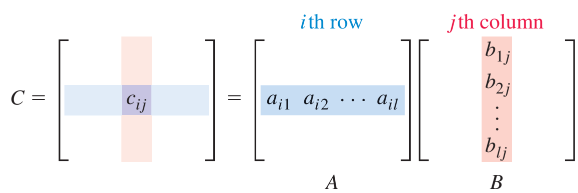


Figure 9.9 If $C = AB$, then to calculate C_{ij} multiply the ith row of A by the jth column of B.

To explain this mathematical definition in words, note that c_{ij} is the entry in C that is located in the ith row and the jth column. To obtain it, we multiply the entries of the ***i*th row** of A with the entries of the ***j*th column** of B, and add all of the products together (see Figure 9.9). For the product AB to be defined, the number of columns in A must equal the number of rows in B. The definition looks rather formidable; let's see how it actually works.

EXAMPLE 4 **Matrix Multiplication** Compute AB when

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 & 3 & -3 \\ 0 & -1 & 4 & 0 \\ -1 & 0 & -2 & 1 \end{bmatrix}$$

Solution First, note that A is a 2×3 matrix and B is a 3×4 matrix. That is, A has 3 columns and B has 3 rows. Therefore, the product AB is defined and AB is a 2×4 matrix. We write the product as

$$C = AB = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & -3 \\ 0 & -1 & 4 & 0 \\ -1 & 0 & -2 & 1 \end{bmatrix}$$

For instance, to find c_{11}, we multiply the first row of A with the first column of B as follows:

$$c_{11} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = (1)(1) + (2)(0) + (3)(-1) = -2$$

You can see from this calculation that the number of columns of A must be equal to the number of rows of B. Otherwise, we would run out of numbers when multiplying the corresponding entries.

To find c_{23}, we multiply the second row of A by the third column of B:

$$c_{23} = \begin{bmatrix} -1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ -2 \end{bmatrix} = (-1)(3) + (0)(4) + (4)(-2) = -11$$

The other entries are obtained in a similar way, and we find that

$$C = \begin{bmatrix} -2 & 0 & 5 & 0 \\ -5 & -2 & -11 & 7 \end{bmatrix}$$

•

When matrices are multiplied, the order is important. For instance, BA is not defined for the matrices in Example 4, since the number of columns of B (which is 4) is different from the number of rows of A (which is 2). Hence, typically, $AB \neq BA$; even if both AB and BA exist, the resulting matrices are usually not the same. They can differ both in the number of rows and columns and in the actual entries. The next two examples illustrate this important point.

EXAMPLE 5 **Order Is Important** Suppose that

$$A = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Show that both AB and BA are defined but $AB \neq BA$.

Solution Note that A is a 1×3 matrix and B is a 3×1 matrix. Thus, the product AB is defined, since the number of columns of A is the same as the number of rows of B (namely, 3); the product AB is a 1×1 matrix. When we carry out the matrix multiplication, we find that

$$AB = \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 - 1 + 0 \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix} = 1.$$

A 1×1 matrix is simply a number.

The product BA is also defined, since the number of columns of B is the same as the number of rows of A (namely, 1); the product BA is a 3×3 matrix. When we carry out the matrix multiplication, we obtain

$$BA = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 \\ -2 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Comparing AB and BA, we see immediately that $AB \neq BA$, since the two matrices are not even of the same size. •

EXAMPLE 6 **Order Is Important** Suppose that

$$A = \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix}.$$

Show that both AB and BA are defined but $AB \neq BA$.

Solution A and B are each 2×2 matrices, so both AB and BA are 2×2 matrices as well. We find that

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} -2 & 1 \\ -4 & 2 \end{bmatrix}$$

The entries in AB are not the same as in BA, so $AB \neq BA$. •

The product AB in Example 6 resulted in a matrix with all entries equal to 0. Earlier, we called this matrix the zero matrix and denoted it by $\mathbf{0}$. Example 6 shows that a product of two matrices can be the zero matrix without either matrix being the zero matrix. This is an important fact and is different from multiplying real numbers. When multiplying two real numbers, we know that if the product is equal to 0, at least one of the two factors must have been equal to 0. This rule, however, does not hold for matrix multiplication; if A and B are two matrices whose product AB is defined, then the product AB can be equal to the zero matrix $\mathbf{0}$ without either A or B being equal to $\mathbf{0}$.

The following properties of matrix multiplication assume that all matrices are of appropriate sizes so that all of the matrix multiplications are defined:

Matrix Multiplication Properties

1. $(A + B)C = AC + BC$
2. $A(B + C) = AB + AC$
3. $(AB)C = A(BC)$
4. $A\mathbf{0} = \mathbf{0}A = \mathbf{0}$

(We will practice applying these properties in the problems at the end of this section.)

Next, we note that if A is a square matrix (i.e., if A has the same number of rows as columns), we can define powers of A: If k is a positive integer, then

$$A^k = A^{k-1}A = AA^{k-1} = \underbrace{AA \ldots A}_{k \text{ factors}}$$

For instance, $A^2 = AA$, $A^3 = AAA$, and so on.

EXAMPLE 7 **Powers of Matrices** Find A^3 if

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

Solution To find A^3, we first need to compute A^2. We obtain

$$A^2 = AA = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}.$$

To compute A^3 now, we compute either A^2A or AA^2. Both will yield A^3. This is a case in which the order of multiplication does not matter, since $A^3 = (AA)A = A(AA) = AAA$. We do it both ways.

$$A^3 = A^2A = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 37 & 54 \\ 81 & 118 \end{bmatrix}$$

and

$$A^3 = AA^2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} = \begin{bmatrix} 37 & 54 \\ 81 & 118 \end{bmatrix}.$$ •

An important square matrix is the **identity matrix**, denoted by I_n. The identity matrix is an $n \times n$ matrix with 1's on its diagonal line and 0's elsewhere; that is,

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

For instance,

$$I_1 = [\,1\,], \qquad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(We frequently write I instead of I_n if the size of I_n is clear from the context.) The identity matrix serves the same role as the number 1 in the multiplication of real numbers: If A is an $m \times n$ matrix, then

$$AI_n = I_m A = A$$

It follows that $I^k = I$ for any positive integer k.

Why do we take the trouble of defining matrix multiplication? Matrix multiplication allows systems of linear equations to be written in matrix form. Consider, for example, the system of m equations in n unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Let A be the coefficient matrix of the linear system, that is, $A = [a_{ij}]$ is an $m \times n$ matrix. Let $B = [b_i]$ be the $m \times 1$ column vector made up of all of the constants on the right-hand side of the linear system. We also define a $n \times 1$ column vector $X = [x_i]$, which is made up of all of the unknowns in the system. Then the linear system of equations can be written as a single matrix equation:

$$AX = B$$

The matrix product can be formed because A has n columns, and X has n rows.

EXAMPLE 8 **Matrix Representation of Linear Systems** Write

$$\begin{aligned} 2x_1 + 3x_2 + 4x_3 &= 1 \\ -x_1 + 5x_2 - 6x_3 &= 7 \end{aligned}$$

in matrix form.

Solution In matrix form, we have

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$$

We see that A is a 2×3 matrix and X is a 3×1 matrix. We therefore expect (on the basis of the rules of multiplication) that B is a 2×1 matrix, which it is indeed. •

We will use matrix representation to solve systems of linear equations in the next subsection.

9.2.3 Inverse Matrices

In this subsection, we will learn how to solve systems of n linear equations in n unknowns when they are written in the matrix form

$$AX = B \tag{9.12}$$

where A is an $n \times n$ square matrix and X and B are $n \times 1$ column vectors. We start with a simple example in which $n = 1$. When we solve

$$5x = 10$$

for x, we simply divide both sides by 5 or, equivalently, multiply both sides by $\frac{1}{5} = 5^{-1}$, and obtain

$$5^{-1} \cdot 5x = 5^{-1} \cdot 10$$

Since $5^{-1} \cdot 5 = 1$ and $5^{-1} \cdot 10 = 2$, we find that $x = 2$. To solve (9.12), we therefore need an operation that is analogous to division, or multiplication by the "reciprocal" of A. We will define a matrix A^{-1} that will serve this function. It is called the **inverse matrix** of A. If this inverse matrix exists, then we can write the solution of (9.12) as

$$A^{-1}AX = A^{-1}B$$

In order for the inverse matrix to have the same effect as multiplying a number by its reciprocal, it should have the property $A^{-1}A = I$, where I is the identity matrix. The solution would then be of the form

$$X = A^{-1}B$$

Definition **Matrix Inverse** Suppose that $A = [a_{ij}]$ is an $n \times n$ square matrix. If there exists an $n \times n$ square matrix B such that

$$AB = BA = I_n$$

then B is called the inverse matrix of A and is denoted by A^{-1}.

If A has an inverse matrix, A is called **invertible** or **nonsingular**; if A does not have an inverse matrix, A is called **singular**. If A is invertible, its inverse matrix is unique; that is, if B and C are both inverse matrices of A, then $B = C$. (In other words, if you and your friend compute the inverse of A, you both should get the exact same answer.) To see why, assume that B and C are both inverse matrices of A. Using $BA = I_n$ and $AC = I_n$, we find that

$$B = BI_n = B(AC) = (BA)C = I_nC = C$$

Properties of Inverse Matrices

1. If A is an invertible $n \times n$ matrix, then
$$(A^{-1})^{-1} = A$$
2. If A and B are invertible $n \times n$ matrices, then
$$(AB)^{-1} = B^{-1}A^{-1}$$

The first property implies that if we apply the inverse operation twice, we get the original matrix back. This property is familiar from dealing with real numbers: Take the number 2; its inverse is $2^{-1} = 1/2$. If we take the inverse of $1/2$, we get 2 again. To see why the analogous property holds for matrices, assume that A is an $n \times n$ matrix with inverse matrix A^{-1}. Then, by definition,

$$AA^{-1} = A^{-1}A = I_n$$

But this equation also says that A is the inverse matrix of A^{-1}; that is, $A = (A^{-1})^{-1}$.

To derive the second property, we must show that

$$(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I_n$$

so $B^{-1}A^{-1}$ is the inverse of AB. Let's start with the first product:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

Also, for the second product:

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n.$$

In the next example, we will see how we can check whether two matrices are inverses of each other. The matrices will be of size 3×3. It is somewhat cumbersome to find the inverse of a matrix of size 3×3 or larger (unless you use a graphical calculator or computer software). Here, we start out with the two matrices and simply check whether they are inverse matrices of each other.

EXAMPLE 9 **Checking Whether Matrices Are Inverses of Each Other** Show that the inverse of

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 3 & 5 & -1 \\ 4 & 2 & -3 \end{bmatrix} \quad \text{is} \quad B = \begin{bmatrix} \frac{13}{73} & -\frac{3}{73} & \frac{14}{73} \\ -\frac{5}{73} & \frac{18}{73} & -\frac{11}{73} \\ \frac{14}{73} & \frac{8}{73} & -\frac{13}{73} \end{bmatrix}.$$

Solution Carrying out the matrix multiplications AB and BA, we see that

$$AB = I_3 \quad \text{and} \quad BA = I_3$$

Therefore, $B = A^{-1}$. •

In this book, we will need to invert only 2×2 matrices. In the next example, we will show one way to do this.

EXAMPLE 10 **Finding an Inverse** Find the inverse of

$$A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

Solution We need to find a matrix

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

such that $AB = I_2$. We will then check that $BA = I_2$ and, hence, that B is the inverse of A. We must solve

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

or on multiplying out the two matrices on the LHS:

$$\begin{aligned} 2b_{11} + 5b_{21} &= 1 \quad (R_1) \\ b_{11} + 3b_{21} &= 0 \quad (R_2) \end{aligned} \quad \text{and} \quad \begin{aligned} 2b_{12} + 5b_{22} &= 0 \quad (R_1) \\ b_{12} + 3b_{22} &= 1 \quad (R_2) \end{aligned} \tag{9.13}$$

Altogether we have four linear equations in four unknowns (b_{11}, b_{12}, b_{21}, and b_{22}), but the first pair of equations involves only b_{11} and b_{21}, while the second pair involves only b_{12} and b_{22}. We can therefore treat the equations as two systems of linear equations, each involving two unknowns, and solve the two systems separately.

We solve both systems by reducing them to upper triangular form.

$$\begin{aligned} (R_1) &\quad 2b_{11} + 5b_{21} = 1 \\ (R_1) - 2(R_2) &\quad -\ b_{21} = 1 \end{aligned} \quad \text{and} \quad \begin{aligned} (R_1) &\quad 2b_{12} + 5b_{22} = 0 \\ (R_1) - 2(R_2) &\quad -\ b_{22} = -2 \end{aligned} \tag{9.14}$$

The left set of equations has the solution

$$b_{21} = -1 \quad \text{and} \quad b_{11} = \frac{1}{2}(1 - 5b_{21}) = \frac{1}{2}(1 + 5) = 3$$

The right set of equations has the solution

$$b_{22} = 2 \quad \text{and} \quad b_{12} = \frac{1}{2}(-5b_{22}) = -5$$

Hence,

$$B = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$$

You can verify that both $AB = I_2$ and $BA = I_2$; thus, $B = A^{-1}$. •

In the optional Subsection 9.2.4, we will present a method that can be employed for larger matrices. It is essentially the same method as in Example 10, but uses augmented matrices. Since manually inverting larger matrices takes a long time, graphing calculators or computer software should be used to find inverse matrices.

At the end of Subsection 9.2.2, we mentioned the connection between systems of linear equations and matrix multiplication. A system of linear equations can be written as

$$AX = B \tag{9.15}$$

In a linear system of n equations in n unknowns, A is an $n \times n$ matrix. If A is invertible, then multiplying (9.15) by A^{-1} from the left, we find that

$$A^{-1}AX = A^{-1}B$$

Since $A^{-1}A = I_n$ and $I_nX = X$, it follows that

$$X = A^{-1}B$$

We will use this equation to redo Example 5 of Section 9.1.

EXAMPLE 11 **Using Inverse Matrices to Solve Linear Systems** Solve

$$\begin{aligned} 2x - y + 3z &= 9 \\ 3x + 5y - z &= 10 \\ 4x + 2y - 3z &= -1 \end{aligned}$$

Solution The coefficient matrix of A is

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 3 & 5 & -1 \\ 4 & 2 & -3 \end{bmatrix}$$

which we encountered in Example 9, where we saw that

$$A^{-1} = \begin{bmatrix} \frac{13}{73} & -\frac{3}{73} & \frac{14}{73} \\ -\frac{5}{73} & \frac{18}{73} & -\frac{11}{73} \\ \frac{14}{73} & \frac{8}{73} & -\frac{13}{73} \end{bmatrix}.$$

So $A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = B$ where $B = \begin{bmatrix} 9 \\ 10 \\ -1 \end{bmatrix}$. And it follows that:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1} \begin{bmatrix} 9 \\ 10 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{13}{73} & -\frac{3}{73} & \frac{14}{73} \\ -\frac{5}{73} & \frac{18}{73} & -\frac{11}{73} \\ \frac{14}{73} & \frac{8}{73} & -\frac{13}{73} \end{bmatrix} \begin{bmatrix} 9 \\ 10 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

That is, $x = 1$, $y = 2$, and $z = 3$, as we found in Example 5 of Section 9.1. •

We now have two methods for solving systems of linear equations when the number of equations is equal to the number of variables: We can reduce the system to upper triangular form and then use back-substitution, or we can write the system in the matrix form $AX = B$, find A^{-1}, and then compute $X = A^{-1}B$. Of course, the second method works only when A^{-1} exists. If A^{-1} does not exist, the system has either no solution or infinitely many solutions. But when A^{-1} exists, $AX = B$ has exactly one solution.

As we saw in Example 10, calculating the inverse of a matrix often entails more work than solving a system of equations using Gaussian elimination. However, for systems of two equations in two unknowns, we will find that there is a rule that you can memorize for writing down the matrix inverse.

The main advantage of the matrix inverse approach is that if you have to solve the same system of equations multiple times with different values for the vector of constants, B, then, once the matrix inverse is known, calculating the solution of the system from the product $A^{-1}B$ may require less work than using Gaussian elimination.

EXAMPLE 12 Solve the matrix equations:

(a) $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$,

(b) $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$,

(c) $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

Solution In all cases we could solve the matrix equation by writing it as a system of linear equations and solving these equations using Gaussian elimination. However, since the same coefficient matrix is used in each equation, it is quicker to find the inverse matrix A^{-1} (this was done in Example 10):

$$A^{-1} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}.$$

Then:

(a) $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -7 \\ 3 \end{bmatrix}$

(b) $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$

(c) $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$ •

There is a very useful criterion for quickly checking whether a 2×2 matrix is invertible. Deriving this criterion will also lead us to a formula for the inverse of an invertible 2×2 matrix. If we want to calculate the inverse of a matrix $A = [a_{ij}]$, we first set:

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

We wish to find B such that $AB = BA = I$ or $B = A^{-1}$. From

$$\overbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}^{A} \overbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}}^{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

we obtain the following set of equations:

$$\begin{aligned} a_{11}b_{11} + a_{12}b_{21} &= 1 \\ a_{21}b_{11} + a_{22}b_{21} &= 0 \end{aligned} \tag{9.16}$$

$$\begin{aligned} a_{11}b_{12} + a_{12}b_{22} &= 0 \\ a_{21}b_{12} + a_{22}b_{22} &= 1 \end{aligned} \tag{9.17}$$

In these equations $a_{11}, a_{21}, a_{12}, a_{22}$ are all known constants and b_{11}, b_{21}, b_{12}, and b_{22} are unknowns that need to be solved for.

We solve the system of linear equations (9.16) for b_{11} and b_{21}:

$$\begin{aligned} a_{11}b_{11} + a_{12}b_{21} &= 1 \qquad (R_1) \\ a_{21}b_{11} + a_{22}b_{21} &= 0 \qquad (R_2) \end{aligned}$$

Eliminating b_{11}:

$$\begin{aligned} a_{11}(R_2) - a_{21}(R_1) \qquad a_{11}a_{22}b_{21} - a_{12}a_{21}b_{21} &= -a_{21} \\ (a_{11}a_{22} - a_{12}a_{21})b_{21} &= -a_{21} \qquad \text{Factor } b_{21} \end{aligned}$$

If $a_{11}a_{22} - a_{12}a_{21} \neq 0$, then:

$$b_{21} = -\frac{a_{21}}{a_{11}a_{22} - a_{12}a_{21}} \qquad \text{Isolate } b_{21}$$

Substituting b_{21} into (R_1) and solving for b_{11}, we obtain, for $a_{11} \neq 0$,

$$\begin{aligned} b_{11} &= \frac{1}{a_{11}}[1 - a_{12}b_{21}] \\ &= \frac{1}{a_{11}}\left[1 - \frac{-a_{12}a_{21}}{a_{11}a_{22} - a_{12}a_{21}}\right] \qquad \text{Substitute for } b_{21} \\ &= \frac{1}{a_{11}}\frac{a_{11}a_{22}}{a_{11}a_{22} - a_{12}a_{21}} \qquad \text{Simplify} \\ &= \frac{a_{22}}{a_{11}a_{22} - a_{12}a_{21}}. \end{aligned}$$

When $a_{11} = 0$, the same final result holds. In that case we solve (R_2) for b_{11}.

When we solve (9.17), we find:

$$\begin{aligned} b_{12} &= -\frac{a_{12}}{a_{11}a_{22} - a_{12}a_{21}} \\ b_{22} &= \frac{a_{11}}{a_{11}a_{22} - a_{12}a_{21}} \end{aligned}$$

Our calculations yield two important results, which you should commit to memory.

Theorem Inverse of a 2 × 2 Matrix If

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

and $a_{11}a_{22} - a_{12}a_{21} \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

The expression $a_{11}a_{22} - a_{12}a_{21}$ is called the **determinant** of A and is denoted by $\det A$.

Definition Determinant of a 2 × 2 Matrix If

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

then the determinant of A is defined as

$$\det A = a_{11}a_{22} - a_{12}a_{21}$$

Looking back now at the formula for A^{-1}, where A is a 2×2 matrix whose determinant is nonzero, we see that, to find the inverse of A, we divide A by the determinant of A. Finally, switch the diagonal elements of A, and change the sign of the off-diagonal elements. If the determinant is equal to 0, then the inverse of A does not exist.

The determinant can be defined for any $n \times n$ matrix. We will not give the general formula, which is computationally complicated for $n \geq 3$. Graphing calculators or computer software can compute determinants. But we mention the following result, which allows us to determine whether or not an $n \times n$ matrix has an inverse:

Theorem Invertibility of Matrices Suppose that A is an $n \times n$ matrix. Then A is nonsingular if and only if $\det A \neq 0$.

EXAMPLE 13 **Using the Determinant to Find Inverse Matrices** Determine which of the following matrices is invertible, and in each case compute the inverse if it exists:

(a) $A = \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix}$ **(b)** $B = \begin{bmatrix} 2 & 6 \\ 1 & 3 \end{bmatrix}$.

Solution **(a)** To check whether A is invertible, we compute the determinant of A:

$$\det A = (3)(4) - (2)(5) = 2 \neq 0$$

Hence, A is invertible and we find that

$$A^{-1} = \frac{1}{2}\begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -\frac{5}{2} \\ -1 & \frac{3}{2} \end{bmatrix}$$

(b) Since

$$\det B = (2)(3) - (1)(6) = 0$$

B is not invertible. •

We mentioned that if A is invertible, then

$$AX = B$$

has exactly one solution, namely, $X = A^{-1}B$. Of particular importance is the case when $B = \mathbf{0}$. That is, assume that A is a 2×2 matrix and $B = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$; then $X = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a solution of

$$AX = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We call the solution $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ a **trivial** solution. It is the only solution of $AX = \mathbf{0}$ when A is invertible. If $AX = \mathbf{0}$ has a solution $X \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, then X is called a **nontrivial** solution. In order to get a nontrivial solution for

$$AX = \mathbf{0},$$

A must be singular. More generally:

Theorem Nontrivial Solution Matrix Equations Suppose that A is an $n \times n$ matrix, and X and $\mathbf{0}$ are $n \times 1$ matrices. Then the equation

$$AX = \mathbf{0}$$

has a nontrivial solution if and only if A is singular.

EXAMPLE 14 **Nontrivial Solutions** Let $A = \begin{bmatrix} a & 6 \\ 3 & 2 \end{bmatrix}$ Determine a so that $AX = \mathbf{0}$ has at least one nontrivial solution, and find the nontrivial solution(s).

Solution To determine when $AX = \mathbf{0}$ has a nontrivial solution, we must find conditions under which A is singular or, equivalently, $\det A = 0$.

For this matrix $\det A = 2a - 18$, so A is singular if $2a - 18 = 0$, or $a = 9$. Therefore, if $a = 9$, $AX = \mathbf{0}$ has a nontrivial solution. To compute that nontrivial solution, we must solve

$$\begin{bmatrix} 9 & 6 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which can be written as a system of two linear equations:

$$\begin{aligned} 9x_1 + 6x_2 &= 0 \\ 3x_1 + 2x_2 &= 0 \end{aligned}$$

We see that the first equation is three times the second equation, so if the second equation is satisfied, then so is the first automatically. Hence we only need to solve:

$$3x_1 + 2x_2 = 0$$

The system has infinitely many solutions of the form,

$$\left\{(x_1, x_2) : x_2 = t \text{ and } x_1 = -\frac{2}{3}t, \text{ for } t \in \mathbf{R}\right\}$$

In particular, the system has nontrivial solutions; for instance, choosing $t = 3$, we find that $x_1 = -2$ and $x_2 = 3$, and choosing $t = -1$, we find that $x_1 = 2/3$ and $x_2 = -1$. Any value for t that is different from 0 will yield a nontrivial solution. If $t = 0$, the trivial solution $(0, 0)$ results.

If we graphed the two lines corresponding to the two equations when $a = 9$, the lines would be identical. For all other values of a, the two lines would intersect only at the point $(0, 0)$, which corresponds to the trivial solution. (Try it!) •

9.2.4 Computing Inverse Matrices

In the preceding subsection, we saw how to invert 2×2 matrices by solving two systems of linear equations. We derived a formula for the inverse of a nonsingular 2×2 matrix. Manually inverting larger matrices takes a long time, and algorithms have been implemented into graphing calculators and computers to carry out these calculations. Here is one method: Inverting an $n \times n$ matrix results in n linear systems, each consisting of n equations in n unknowns. By solving these n systems simultaneously, we can speed up the process of finding the inverse matrix.

Recall Equation (9.13) of Subsection 9.2.3:

$$\begin{aligned} 2b_{11} + 5b_{21} &= 1 \\ b_{11} + 3b_{21} &= 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} 2b_{12} + 5b_{22} &= 0 \\ b_{12} + 3b_{22} &= 1 \end{aligned}$$

Let us rewrite both systems as augmented matrices:

$$\left[\begin{array}{cc|c} 2 & 5 & 1 \\ 1 & 3 & 0 \end{array}\right] \qquad \text{and} \qquad \left[\begin{array}{cc|c} 2 & 5 & 0 \\ 1 & 3 & 1 \end{array}\right]$$

We see that each augmented matrix has the same matrix A on its left side. To solve each system, we must perform row transformations until we can read off the solutions. Reading off the solutions is easiest when the matrix on the left is transformed into the identity matrix. This is slightly more involved than using Gaussian elimination to make the coefficient matrix upper triangular. We do this for the augmented matrix

$$\left[\begin{array}{cc|c} 2 & 5 & 1 \\ 1 & 3 & 0 \end{array}\right] \begin{array}{l} (R_1) \\ (R_2) \end{array}$$

We first use Gaussian elimination to make the coefficient matrix upper triangular.

$$\begin{array}{r} R_1 \\ (R_1)-2(R_2) \end{array} \left[\begin{array}{rr|r} 2 & 5 & 1 \\ 0 & -1 & 1 \end{array}\right] \begin{array}{l} (R_3) \\ (R_4) \end{array}$$

We now use (R_4) to eliminate everything in the coefficient matrix except for the terms on the diagonal. For this coefficient matrix, the only term that needs to be eliminated is the 5 in the top right corner.

$$\begin{array}{r} (R_3)+5(R_4) \\ (R_4) \end{array} \left[\begin{array}{rr|r} 2 & 0 & 6 \\ 0 & -1 & 1 \end{array}\right] \begin{array}{l} (R_5) \\ (R_6) \end{array}$$

Then we multiply each row by constants chosen so the diagonal terms all become 1s:

$$\begin{array}{r} \frac{1}{2}(R_5) \\ -(R_6) \end{array} \left[\begin{array}{rr|r} 1 & 0 & 3 \\ 0 & 1 & -1 \end{array}\right] \begin{array}{l} (R_7) \\ (R_8) \end{array}$$

Rewriting this matrix as a system of linear equations, we see that

$$b_{11} = 3 \quad (R_7)$$

$$b_{21} = -1 \quad (R_8)$$

To find b_{12} and b_{22}, we need to transform the augmented matrix

$$\left[\begin{array}{rr|r} 2 & 5 & 0 \\ 1 & 3 & 1 \end{array}\right]$$

so that the matrix on the left is the identity matrix. This involves the exact same transformations as before, since the coefficient matrix is the same for both systems of linear equations. But rather than do that we will instead show you a trick that speeds up the calculation: Instead of doing each augmented matrix separately, we do them simultaneously. That is, we write

$$\left[\begin{array}{rr|rr} 2 & 5 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right]$$

and then make row transformations until this matrix is of the form

$$\left[\begin{array}{rr|rr} 1 & 0 & b_{11} & b_{12} \\ 0 & 1 & b_{21} & b_{22} \end{array}\right]$$

We can then read off the inverse matrix:

$$A^{-1} = \left[\begin{array}{rr} b_{11} & b_{12} \\ b_{21} & b_{22} \end{array}\right]$$

Let's do the calculation:

$$\left[\begin{array}{rr|rr} 2 & 5 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right] \begin{array}{l} (R_1) \\ (R_2) \end{array}$$

$$\begin{array}{r} (R_1) \\ (R_1)-2(R_2) \end{array} \left[\begin{array}{rr|rr} 2 & 5 & 1 & 0 \\ 0 & -1 & 1 & -2 \end{array}\right] \begin{array}{l} (R_3) \\ (R_4) \end{array}$$

$$\begin{array}{r} (R_3)+5(R_4) \\ (R_4) \end{array} \left[\begin{array}{rr|rr} 2 & 0 & 6 & -10 \\ 0 & -1 & 1 & -2 \end{array}\right] \begin{array}{l} (R_5) \\ (R_6) \end{array}$$

$$\begin{array}{r} \frac{1}{2}(R_5) \\ -(R_6) \end{array} \left[\begin{array}{rr|rr} 1 & 0 & 3 & -5 \\ 0 & 1 & -1 & 2 \end{array}\right]$$

We recognize the first column on the right

$$\left[\begin{array}{r} 3 \\ -1 \end{array}\right]$$

which gives

$$b_{11} = 3 \quad \text{and} \quad b_{21} = -1.$$

The second column on the right,

$$\left[\begin{array}{r} -5 \\ 2 \end{array}\right]$$

yields

$$b_{12} = -5 \quad \text{and} \quad b_{22} = 2.$$

We conclude that

$$A^{-1} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$$

as we saw in Example 10.

This method works for larger matrices well as—it is quite efficient! We illustrate the technique in the next example on a 3×3 matrix.

EXAMPLE 15 Find the inverse (if it exists) of

$$\begin{bmatrix} 1 & -1 & -1 \\ 2 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}.$$

Solution We write

$$\left[\begin{array}{rrr|rrr} 1 & -1 & -1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & -1 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} (R_1) \\ (R_2) \\ (R_3) \end{array}$$

and perform appropriate row transformations in order to get the identity matrix on the left side. The matrix on the right side is then the inverse matrix. The first step is to make the matrix on the left side upper triangular:

$$\begin{array}{l} (R_1) \\ 2(R_1)-(R_2) \\ (R_1)+(R_3) \end{array} \left[\begin{array}{rrr|rrr} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & -1 & -3 & 2 & -1 & 0 \\ 0 & 0 & -2 & 1 & 0 & 1 \end{array}\right] \begin{array}{l} (R_4) \\ (R_5) \\ (R_6) \end{array}$$

Now we will eliminate all of the entries on the left-side matrix except for the entries that will be occupied by 1s when we have turned the matrix into the identity matrix. We start with the third column:

$$\begin{array}{c} (R_4)-\frac{1}{2}(R_6) \\ (R_5)-\frac{3}{2}(R_6) \\ (R_6) \end{array} \left[\begin{array}{rrr|rrr} 1 & -1 & \boxed{0} & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & -1 & \boxed{0} & \frac{1}{2} & -1 & -\frac{3}{2} \\ 0 & 0 & -2 & 1 & 0 & 1 \end{array}\right] \begin{array}{l} (R_7) \\ (R_8) \quad \text{Eliminate first two entries in third column.} \\ (R_9) \end{array}$$

We continue to the second column:

$$\begin{array}{c} (R_7)-(R_8) \\ (R_8) \\ (R_9) \end{array} \left[\begin{array}{rrr|rrr} 1 & \boxed{0} & 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & \frac{1}{2} & -1 & -\frac{3}{2} \\ 0 & 0 & -2 & 1 & 0 & 1 \end{array}\right] \begin{array}{l} (R_{10}) \\ (R_{11}) \quad \text{Eliminate first entry in second column} \\ (R_{12}) \end{array}$$

Finally we multiply each row by a constant to turn our diagonal entries into 1s:

$$\begin{array}{c} (R_{10}) \\ -(R_{11}) \\ -\frac{1}{2}(R_{12}) \end{array} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -\frac{1}{2} & 1 & \frac{3}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & 0 & -\frac{1}{2} \end{array}\right]$$

We succeeded in getting the identity matrix on the left side. Therefore,

$$A^{-1} = \begin{bmatrix} 0 & 1 & 1 \\ -\frac{1}{2} & 1 & \frac{3}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}$$

This method is reasonably quick as long as the matrices are not too big. If you cannot get the identity matrix on the left side, then the matrix does not have an inverse, as is illustrated in the next example.

EXAMPLE 16 Find the inverse (if it exists) of

$$A = \begin{bmatrix} 2 & 6 \\ 1 & 3 \end{bmatrix}$$

Solution Since $\det A = (2)(3) - (1)(6) = 0$, A is singular and therefore does not have an inverse. But let's see what happens if we try to find the inverse. First, we have

$$\left[\begin{array}{cc|cc} 2 & 6 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{array}\right] \quad \begin{array}{c} (R_1) \\ (R_2) \end{array}$$

As before, we start by making the left-side matrix upper triangular. We find that

$$\begin{array}{c} (R_1) \\ (R_1) - 2(R_2) \end{array} \quad \left[\begin{array}{cc|cc} 2 & 6 & 1 & 0 \\ 0 & 0 & 1 & -2 \end{array}\right] \quad \begin{array}{c} (R_3) \\ (R_4) \end{array}$$

Since the last row on the left side consists only of 0's, we cannot obtain the identity matrix on the left side. That is, our method fails to provide an inverse matrix, and we conclude that A does not have an inverse. •

Section 9.2 Problems

9.2.1, 9.2.2

In Problems 1–6, let

$$\mathbf{A} = \begin{bmatrix} -1 & 2 \\ 0 & -3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & -2 \\ -1 & -1 \end{bmatrix}$$

1. Find $A - B + 2C$.

2. Find $-2A + 3B$.

3. Determine D so that $A + B = 2A - B + D$.

4. Show that $A + B = B + A$.

5. Show that $(A + B) + C = A + (B + C)$.

6. Show that $2(A + B) = 2A + 2B$.

In Problems 7–12, let

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 3 & -1 \\ 0 & -2 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & -1 & 4 \\ -2 & 0 & -1 \\ 1 & 3 & 3 \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 1 \\ 2 & 0 & 2 \end{bmatrix}$$

7. Find $2A + 3B - C$.

8. Find $3C - B + \frac{1}{2}A$.

9. Determine D so that $A + B + C + D = \mathbf{0}$.

10. Determine D so that $A + 4B = 2(A + B) + D$.

11. Show that $A + B = B + A$.

12. Show that $(A + B) + C = A + (B + C)$.

13. Show that if $A + B = C$, then $A = C - B$.

14. Find the transpose of

$$A = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

15. Find the transpose of

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 3 & 1 & -4 \end{bmatrix}$$

16. Suppose A is a 2×2 matrix. Find conditions on the entries of A such that

$$A - A' = \mathbf{0}$$

17. Suppose that A and B are $m \times n$ matrices. Show that

$$(A + B)' = A' + B'$$

18. Suppose that A is an $m \times n$ matrix. Show that

$$(A')' = A$$

19. Suppose that A is an $m \times n$ matrix and k is a real number. Show that

$$(kA)' = kA'$$

20. Suppose that A is an $m \times k$ matrix and B is a $k \times n$ matrix. Show that

$$(AB)' = B'A'$$

In Problems 21–26, let

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 0 \\ -1 & -1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$$

21. Compute the following:

(a) AB **(b)** BA

22. Compute ABC.

23. Show that $AC \neq CA$.

24. Show that $(AB)C = A(BC)$.

25. Show that $(A + B)C = AC + BC$.

26. Show that $A(B + C) = AB + AC$.

27. Suppose that A is a 3×4 matrix and B is a 4×2 matrix. What is the size of the product AB?

28. Suppose A is a 3×4 matrix and B is an $m \times n$ matrix. What are values of m and n such that the following products are defined?

(a) AB **(b)** BA

29. Suppose that A is a 3×4 matrix, B is a 1×3 matrix, C is a 3×1 matrix, and D is a 4×3 matrix. Which of the matrix multiplications that follow are defined? Whenever it is defined, state the size of the resulting matrix.

(a) BD' **(b)** DA **(c)** ACB

30. Suppose that A is an $l \times p$ matrix, B is an $m \times q$, matrix, and C is an $n \times r$ matrix. What can you say about l, m, n, p, q, and r if the products that follow are defined? State the size of the resulting matrix.

(a) ABC **(b)** $AB'C$ **(c)** BAC' **(d)** $A'CB'$

31. Let

$$A = \begin{bmatrix} 1 & 3 \\ 0 & -2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 0 & -3 \\ 2 & 1 & -1 & 0 \end{bmatrix}$$

(a) Compute AB. **(b)** Compute $B'A$.

32. Let

$$A = \begin{bmatrix} 1 & 4 & -2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$$

(a) Compute AB. **(b)** Compute BA.

33. Let

$$A = \begin{bmatrix} 2 & 1 \\ 0 & -3 \end{bmatrix}$$

Find A^2, A^3, and A^4.

34. Suppose that

$$A = \begin{bmatrix} 1 & -1 \\ 3 & 0 \\ 5 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Show that $(AB)' = B'A'$.

35. Let

$$B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(a) Find B^2, B^3, B^4, and B^5.

(b) What can you say about B^k when **(i)** k is even and **(ii)** k is odd?

36. Let

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Show that $I_3 = I_3^2 = I_3^3$.

37. Let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \quad \text{and} \quad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Show that $AI_2 = I_2A = A$.

38. Let

$$A = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & -2 \\ -1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Show that $AI_3 = I_3A = A$.

In Problems 39–42, write each system in matrix form. (There is no need to solve the systems).

39.
$$\begin{aligned} 2x_1 + 3x_2 - x_3 &= 0 \\ 3x_2 + x_3 &= 1 \\ x_1 - x_3 &= 2 \end{aligned}$$

40.
$$\begin{aligned} 2x_2 - x_1 &= x_3 \\ 4x_1 + x_3 &= 7x_2 \\ x_2 - x_1 &= x_3 \end{aligned}$$

41.
$$\begin{aligned} 2x_1 - x_2 &= 4 \\ -x_1 + 2x_2 &= 3 \\ 3x_1 &= 4 \end{aligned}$$

42.
$$\begin{aligned} x_1 - 3x_2 + x_3 &= 1 \\ -2x_1 + x_2 - x_3 &= 0 \end{aligned}$$

9.2.3

43. Show that the inverse of

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$

is

$$B = \frac{1}{5}\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$$

44. Show that the inverse of

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

is

$$A^{-1} = \begin{bmatrix} \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

In Problems 45–48, let

$$\mathbf{A} = \begin{bmatrix} \mathbf{-1} & \mathbf{1} \\ \mathbf{2} & \mathbf{3} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{2} & \mathbf{0} \\ \mathbf{3} & \mathbf{2} \end{bmatrix}$$

45. Find the inverse (if it exists) of A.

46. Find the inverse (if it exists) of B.

47. Show that $(A^{-1})^{-1} = A$.

48. Show that $(AB)^{-1} = B^{-1}A^{-1}$.

49. Find the inverse (if it exists) of

$$C = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$$

50. Find the inverse (if it exists) of

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

51. Suppose that

$$A = \begin{bmatrix} -1 & 0 \\ 2 & -1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$$

Find X such that $AX = D$ by

(a) solving the associated system of linear equations and

(b) using the inverse of A.

52. **(a)** Show that if $X = AX + D$, then

$$X = (I - A)^{-1}D$$

provided that $I - A$ is invertible.

(b) Suppose that

$$A = \begin{bmatrix} 3 & 2 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$$

Compute $(I - A)^{-1}$, and use your result in (a) to compute X.

53. Use the determinant to determine whether the matrix

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$$

is invertible.

54. Use the determinant to determine whether the matrix

$$A = \begin{bmatrix} -1 & 3 \\ 1 & 1 \end{bmatrix}$$

is invertible.

55. Use the determinant to determine whether the matrix
$$A = \begin{bmatrix} 4 & -1 \\ 8 & -2 \end{bmatrix}$$
is invertible.

56. Use the determinant to determine whether the matrix
$$A = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix}$$
is invertible.

57. Suppose that
$$A = \begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}$$

(a) Compute $\det A$. Is A invertible?

(b) Suppose that
$$X = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$
Write $AX = B$ as a system of linear equations.

(c) Show that if
$$B = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$
then
$$AX = B$$
has infinitely many solutions. Graph the two straight lines associated with the corresponding system of linear equations, and explain why the system has infinitely many solutions.

(d) Find a column vector
$$B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$
so that
$$AX = B$$
has no solutions.

58. Suppose that
$$A = \begin{bmatrix} a & 8 \\ 2 & 4 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

(a) Show that when $a \neq 4$, $AX = B$ has exactly one solution.

(b) Suppose $a = 4$. Find conditions on b_1 and b_2 such that $AX = B$ has **(i)** infinitely many solutions and **(ii)** no solutions.

(c) Explain your results in (a) and (b) graphically.

In Problems 59–62, write down the inverse of A.

59. $A = \begin{bmatrix} 2 & 1 \\ -3 & -1 \end{bmatrix}$ **60.** $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$

61. $A = \begin{bmatrix} -1 & 4 \\ 5 & 0 \end{bmatrix}$ **62.** $A = \begin{bmatrix} -2 & -1 \\ 3 & 2 \end{bmatrix}$

63. Use the determinant to determine whether
$$A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$$
is invertible. If it is invertible, compute its inverse. In either case, solve $AX = \mathbf{0}$.

64. Use the determinant to determine whether
$$B = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$
is invertible. If it is invertible, compute its inverse. In either case, solve $BX = \mathbf{0}$.

65. Use the determinant to determine whether
$$C = \begin{bmatrix} 1 & 3 \\ 1 & 3 \end{bmatrix}$$
is invertible. If it is invertible, compute its inverse. In either case, solve $CX = \mathbf{0}$.

66. Use the determinant to determine whether
$$D = \begin{bmatrix} -3 & 6 \\ -4 & 8 \end{bmatrix}$$
is invertible. If it is invertible, compute its inverse. In either case, solve $DX = \mathbf{0}$.

9.2.4

In Problems 67–70, find the inverse matrix to each given matrix if the inverse matrix exists.

67. $A = \begin{bmatrix} 2 & -1 & -1 \\ 2 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix}$ **68.** $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ -1 & -1 & -1 \end{bmatrix}$

69. $A = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -2 & 0 \\ -1 & 1 & 2 \end{bmatrix}$ **70.** $A = \begin{bmatrix} -1 & 0 & 2 \\ -1 & -2 & 3 \\ 0 & 2 & -1 \end{bmatrix}$

9.3 Linear Maps, Eigenvectors, and Eigenvalues

In this section, we will use boldface lowercase letters (e.g., $\mathbf{x}$) to denote vectors. An alternative notation that is useful when you are writing equations out by hand is to underline the vectors: e.g., write $\underline{x}$. Consider a function of the form
$$\mathbf{x} \mapsto A\mathbf{x} \tag{9.18}$$
where A is a 2×2 matrix and $\mathbf{x}$ is a 2×1 column vector (or, simply, vector). Since $A\mathbf{x}$ is a 2×1 vector, this function takes a 2×1 vector and maps it into a 2×1 vector. That enables us to apply A repeatedly: We can compute $A(A\mathbf{x}) = A^2\mathbf{x}$, which is again a 2×1 vector, and so on. We will first look at vectors, then at maps $A\mathbf{x}$, and finally at iterates of the map A (i.e., $A^2\mathbf{x}$, $A^3\mathbf{x}$, and so on).

According to the properties of matrix multiplication, the function (9.18) satisfies the following conditions:

1. $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$, and
2. $A(\lambda\mathbf{x}) = \lambda(A\mathbf{x})$, where λ is a scalar.

Because of these two properties, we say that the function $\mathbf{x} \to A\mathbf{x}$ is **linear**. We also call these functions **linear maps** and as a shorthand call the matrix A a **map**.

Linear maps are important in other contexts as well. These topics are covered in courses on matrix or linear algebra. We will encounter them in Chapters 10 and 11. Here, we restrict our discussion to 2×2 matrices but point out that we can generalize the discussion that follows to arbitrary $n \times n$ matrices.

9.3.1 Graphical Representation

Vectors We begin with a graphical representation of vectors. We assume that $\mathbf{x}$ is a 2×1 matrix. We call $\mathbf{x}$ a *column vector* or simply a *vector*. Since a 2×1 matrix has just two components, we can represent a vector in the plane. For instance, to represent the vector

$$\mathbf{x} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

in the x_1–x_2 plane, we draw an arrow from the origin $(0, 0)$ to the point $(3, 4)$, as illustrated in Figure 9.10. We see from Figure 9.10 that a vector has a length and a direction. The length of the vector $\mathbf{x} = \left[\begin{smallmatrix} 3 \\ 4 \end{smallmatrix}\right]$, denoted by $|\mathbf{x}|$, is the distance from the origin $(0, 0)$ to the point $(3, 4)$; that is,

$$\text{length of } \mathbf{x} = |\mathbf{x}| = \sqrt{9 + 16} = 5 \qquad \text{Use Pythagoras' theorem}$$

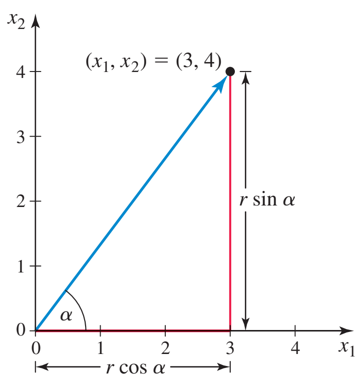


Figure 9.10 The vector $\left[\begin{smallmatrix} 3 \\ 4 \end{smallmatrix}\right]$ in the x_1–x_2 plane.

We define the direction of $\mathbf{x}$ as the angle α between the positive x_1-axis and the vector $\mathbf{x}$ (measured counterclockwise), as shown in Figure 9.10. The angle α is in the interval $[0, 2\pi)$. In this example, α satisfies $\tan \alpha = 4/3$.

More generally (see Figure 9.10 again), a vector $\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]$ has length

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2}$$

and direction α, where $\alpha \in [0, 2\pi)$ satisfies

$$\tan \alpha = \frac{x_2}{x_1}$$

The angle α is always measured counterclockwise from the positive x_1-axis.

If we denote the length of $\mathbf{x}$ by r (i.e., $r = |\mathbf{x}|$), then, as shown in Figure 9.10, since $x_1 = r \cos \alpha$ and $x_2 = r \sin \alpha$, the vector $\mathbf{x}$ can also be written as

$$\mathbf{x} = \begin{bmatrix} r \cos \alpha \\ r \sin \alpha \end{bmatrix}$$

We thus have two distinct ways of representing vectors in the plane: We can use either the endpoint (x_1, x_2) or the length and direction (r, α). The first representation leads to our familiar Cartesian coordinate system. The second representation, using the length and direction of the corresponding vector, leads to the **polar coordinate system**. We will use both representations in what follows.

EXAMPLE 1 If the length of the vector $\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]$ is 4 and its angle with the positive x_1-axis is 120° (measured clockwise), what is its representation in Cartesian coordinates?

Solution An angle of 120° measured clockwise from the positive x_1-axis corresponds to an angle of $360^\circ - 120^\circ = 240^\circ$ measured counterclockwise from the positive x_1-axis (Figure 9.11). Since the length of the vector is 4, we obtain

$$x_1 = 4\cos(240^\circ) = (4)\left(-\frac{1}{2}\right) = -2$$

$$x_2 = 4\sin(240^\circ) = (4)\left(-\frac{1}{2}\sqrt{3}\right) = -2\sqrt{3}$$

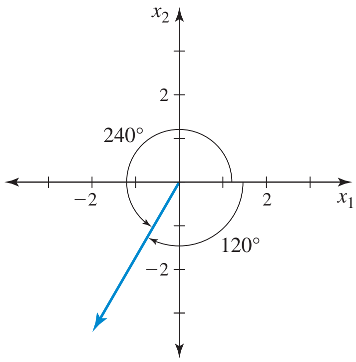


Figure 9.11 The vector in Example 1.

which yields the Cartesian coordinate representation

$$\mathbf{x} = \begin{bmatrix} -2 \\ -2\sqrt{3} \end{bmatrix}.$$

Because vectors are matrices, we can use matrix addition to add vectors. For instance,

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

This vector sum has a graphical representation. (See Figure 9.12.) To add $\left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$, we move the vector $\left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]$ to the tip of the vector $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ without changing the direction or the length of $\left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]$. The sum of the two vectors then starts at (0, 0) and ends at the tip of the vector that was moved. This series of operations can also be described in the following way: The sum is the diagonal in the parallelogram that is formed by the two vectors $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]$. The rules for vector addition are therefore referred to as the **parallelogram law**.

Multiplication of a vector by a scalar is carried out componentwise. For instance, if we multiply $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ by 2, we get

$$2\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \cdot 1 \\ 2 \cdot 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

This operation corresponds to changing the length of the vector. The vector $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ has length $\sqrt{1+4} = \sqrt{5}$; the vector $2\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ has length $\sqrt{4+16} = \sqrt{20} = 2\sqrt{5}$. That is, multiplying the vector $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ by 2 increases its length by the factor 2. Since 2 is positive, the resulting vector points in the same direction as the vector $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$. If we multiply $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ by -1, then the resulting vector is $\left[\begin{smallmatrix} -1 \\ -2 \end{smallmatrix}\right]$, which has the same length as the original vector, but points in the opposite direction, as illustrated in Figure 9.13.

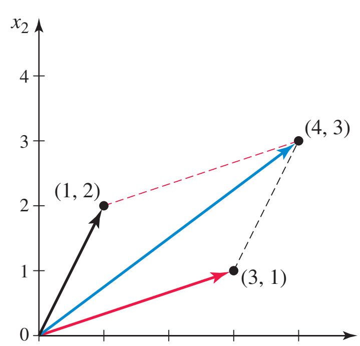


Figure 9.12 Addition of two vectors.

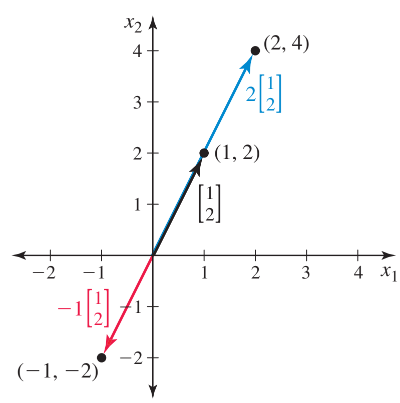


Figure 9.13 The vectors in Example 1.

EXAMPLE 2 Let

$$\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} -1 \\ -3 \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

Find $-\frac{1}{2}\mathbf{u}$, $\mathbf{u}+\mathbf{v}$, $\mathbf{v}+\mathbf{w}$, and $-\mathbf{v}$, and illustrate the results graphically.

Solution

$$-\frac{1}{2}\mathbf{u} = -\frac{1}{2}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -\frac{1}{2} \end{bmatrix}$$

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} -1 \\ -3 \end{bmatrix} + \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \end{bmatrix}$$

$$-\mathbf{v} = -\begin{bmatrix} -1 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

The results are illustrated in Figure 9.14. •

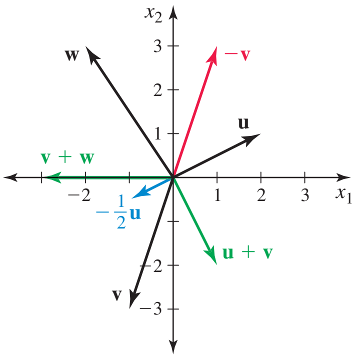


Figure 9.14 The vectors in Example 2.

Vector Addition and Multiplication of a Vector by a Scalar

Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Then

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}$$

If a is a scalar, then

$$a\mathbf{x} = \begin{bmatrix} ax_1 \\ ax_2 \end{bmatrix}$$

If $|\mathbf{x}|$ denotes the length of $\mathbf{x}$, then the length of $a\mathbf{x}$ is the absolute value of a times the length of $\mathbf{x}$—that is, $|a||\mathbf{x}|$.

Linear Maps We will first use a graphical approach to study functions of the form

$$\mathbf{x} \mapsto A\mathbf{x}$$

where A is a 2×2 matrix and $\mathbf{x}$ is a 2×1 vector. Since $A\mathbf{x}$ is a 2×1 vector as well, the map A takes the 2×1 vector $\mathbf{x}$ and maps it into a 2×1 vector.

The simplest such map is the **identity map**, represented by the identity matrix I_2:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Since

$$I_2\mathbf{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}$$

it follows that the identity matrix leaves the vector $\mathbf{x}$ unchanged.

Slightly more complicated is the map

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} ax_1 \\ bx_2 \end{bmatrix}$$

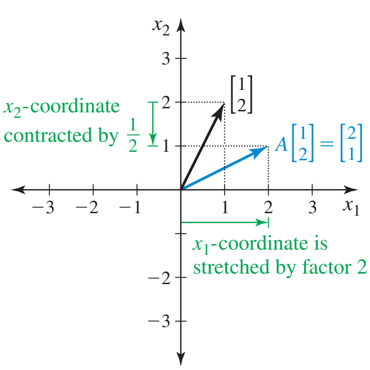


Figure 9.15 The action of the matrix $\begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$ on the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

This map stretches or contracts each coordinate separately. In Figure 9.15, we show the action of the map on $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ when $a = 2$ and $b = \frac{1}{2}$. We find that

$$\begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

This map stretches the first coordinate by a factor of 2 and contracts the second coordinate by a factor of 1/2. If a is negative, then in addition to being stretched (or contracted) by a factor $|a|$ the x_1-coordinate is also reflected around the x_2-axes. Similarly negative values of b lead to the x_2-coordinate being reflected around the x_1-axes. For example, if our map is $A = \begin{bmatrix} 2 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}$ (i.e., $a = 2$ and $b = 1/2$), then in addition to the transformations shown in Figure 9.15, the x_2-coordinate is also reflected (see Figure 9.16).

Another linear map is a **rotation**, which rotates a vector in the x_1–x_2 plane by a fixed angle.

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

rotates a vector by an angle θ.

If $\theta > 0$, then the rotation is counterclockwise; if $\theta < 0$, the rotation is clockwise by the angle $|\theta|$.

To check that this is indeed a rotation, we investigate the action of R_θ on a vector with coordinates (x_1, x_2), as illustrated in Figure 9.17. Using the polar coordinate system, we can write this vector as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix}$$

where r is the length of the vector and α is the angle it forms with the positive x_1-axis.

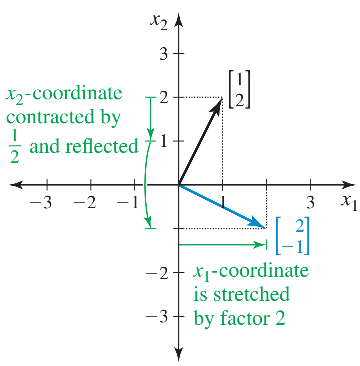


Figure 9.16 The action of the matrix $\begin{bmatrix} 2 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}$ on the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

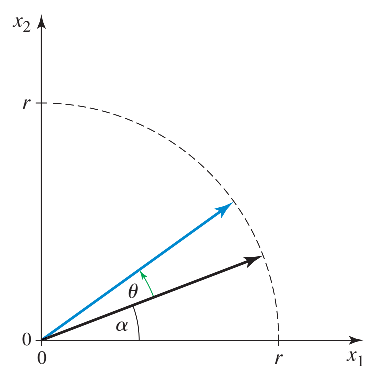


Figure 9.17 The rotation of a vector.

Applying R_θ, we find that

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix} = \begin{bmatrix} r(\cos\theta\cos\alpha - \sin\theta\sin\alpha) \\ r(\sin\theta\cos\alpha + \cos\theta\sin\alpha) \end{bmatrix}$$
$$= \begin{bmatrix} r\cos(\theta + \alpha) \\ r\sin(\theta + \alpha) \end{bmatrix}$$

where we used the trigonometric identities

$$\cos(\theta + \alpha) = \cos\theta\cos\alpha - \sin\theta\sin\alpha$$
$$\sin(\theta + \alpha) = \sin\theta\cos\alpha + \cos\theta\sin\alpha$$

We see that the resulting vector still has length r and that the angle with the x_1-axis is $\alpha + \theta$. If $\theta > 0$, as in Figure 9.17, then the vector is rotated counterclockwise by the angle θ.

EXAMPLE 3 Use a rotation matrix to rotate the vector $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ counterclockwise by the angle $\pi/3$.

Solution The rotation matrix we seek is

$$R_{\pi/3} = \begin{bmatrix} \cos(\pi/3) & -\sin(\pi/3) \\ \sin(\pi/3) & \cos(\pi/3) \end{bmatrix} = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}.$$

Hence, the rotated vector has coordinates

$$\begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} - \frac{3}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} + \frac{3}{2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 - 3\sqrt{3} \\ \sqrt{3} + 3 \end{bmatrix}$$ •

From this brief discussion of linear maps, we see that the map $\mathbf{x} \to A\mathbf{x}$ typically takes the vector $\mathbf{x}$ and rotates, stretches, or contracts it. For an arbitrary matrix A, vectors may be moved in a way that has no simple geometric interpretation.

EXAMPLE 4 Investigate the action of $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ on $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

Solution If $\mathbf{x} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$, then

$$A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$$

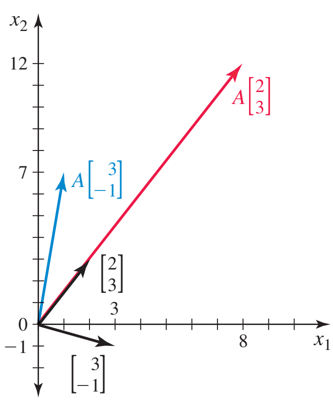


Figure 9.18 The action of the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ on two vectors.

Although we can compute the outcome of this map, there does not seem to be a straightforward geometric explanation of it. (See Figure 9.18.) On the other hand, if $\mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, then

$$A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 8 \\ 12 \end{bmatrix} = 4 \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

The result of $A\mathbf{x}$ is simply a multiple of $\mathbf{x}$. (See Figure 9.18.) Such a vector is called an **eigenvector**, and the stretching or contracting factor is called an **eigenvalue**. Eigenvectors and eigenvalues are the topic of the next subsection. •

9.3.2 Eigenvalues and Eigenvectors

In Subsection 9.3.1, we saw that there are matrices and vectors for which the map $\mathbf{x} \to A\mathbf{x}$ takes on a particularly simple form, namely,

$$A\mathbf{x} = \lambda\mathbf{x} \tag{9.19}$$

where λ is a scalar. Now we investigate this relationship. Again restricting our discussion to 2×2 matrices, we begin with the following definition:

> **Definition Eigenvectors and Eigenvalues** Assume that A is a square matrix. A nonzero vector $\mathbf{x}$ that satisfies the equation
>
> $$A\mathbf{x} = \lambda\mathbf{x}$$
>
> is an **eigenvector** of the matrix A, and the number λ is an **eigenvalue** of the matrix A.

Note that we assume that the vector $\mathbf{x}$ is different from the zero vector. (The zero vector $\mathbf{x} = \mathbf{0}$ always satisfies the equation $A\mathbf{x} = \lambda\mathbf{x}$ and thus would not be special.) The eigenvalue λ can be 0, however. We will see that λ can even be a complex number.

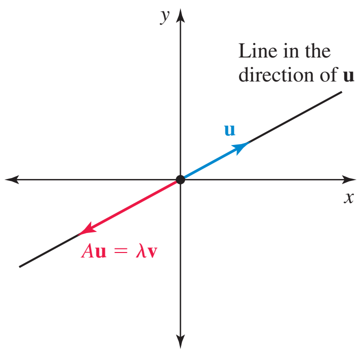


Figure 9.19 Any vector on the line in the direction of the eigenvector will remain on the line under the map A.

The action of A on eigenvectors produces a particularly simple form: If we apply A to an eigenvector $\mathbf{x}$ (i.e., if we compute $A\mathbf{x}$), the result is a constant multiple of $\mathbf{x}$. This property of an eigenvector has an important geometric interpretation when the eigenvalue λ is a real number: If we draw a straight line through the origin in the direction of an eigenvector, then any vector on this straight line will remain on the line after the map A is applied. (See Figure 9.19.)

How can we find eigenvalues and eigenvectors for 2×2 matrices? We will see that a 2×2 matrix has two eigenvalues, which are either distinct or identical. We will discuss only the case in which the two eigenvalues are distinct; the case in which the eigenvalues are identical is more complicated and is covered in courses on linear algebra.

We will show how to find eigenvalues and eigenvectors by way of example. We will use the same matrix as in Example 4.

EXAMPLE 5 Find all eigenvalues and eigenvectors of $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$.

Solution We are interested in finding $\mathbf{x} \neq \mathbf{0}$ and λ such that

$$A\mathbf{x} = \lambda\mathbf{x}$$

We can rewrite this equation as

$$A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$$

In order to factor $\mathbf{x}$, we must multiply $\lambda\mathbf{x}$ by the identity matrix $I = I_2$. (Because we will be dealing only with 2×2 matrices, we simply write I instead of I_2.) Multiplication by I yields

$$A\mathbf{x} - \lambda I\mathbf{x} = \mathbf{0}$$

We can now factor $\mathbf{x}$, resulting in

$$(A - \lambda I)\mathbf{x} = \mathbf{0}$$

In Section 9.2, we showed that in order to obtain a nontrivial solution ($\mathbf{x} \neq \mathbf{0}$), $A - \lambda I$ must be singular; that is,

$$\det(A - \lambda I) = 0$$

This is the key equation that will allow us to find eigenvalues. Now, since

$$A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{bmatrix}$$

$$\begin{aligned} \det(A - \lambda I) &= (1-\lambda)(2-\lambda) - (2)(3) \qquad \det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc \\ &= 2 - 3\lambda + \lambda^2 - 6 = \lambda^2 - 3\lambda - 4 \\ &= (\lambda + 1)(\lambda - 4) \end{aligned}$$

So:

$$\lambda_1 = -1 \qquad \text{and} \qquad \lambda_2 = 4. \qquad \text{Solve } (\lambda+1)(\lambda-4) = 0$$

These two numbers are the eigenvalues of the matrix A. Each eigenvalue will have its own eigenvector.

To compute the eigenvectors, we carry out the following calculations: If $\lambda_1 = -1$, we must find a nonzero vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ such that

$$A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (-1)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Writing this matrix equation as a system of equations gives

$$\begin{aligned} x_1 + 2x_2 &= -x_1 \\ 3x_1 + 2x_2 &= -x_2 \end{aligned}$$

Simplifying, we obtain

$$\begin{aligned} 2x_1 + 2x_2 &= 0 \\ 3x_1 + 3x_2 &= 0 \end{aligned}$$

The two equations both reduce to

$$x_1 + x_2 = 0$$

We are looking for a nonzero vector that satisfies this equation. For instance,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

would be a reasonable choice. To check that $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is indeed an eigenvector corresponding to the eigenvalue $\lambda_1 = -1$, we compute

$$\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = (-1)\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

That is,

$$A\begin{bmatrix} 1 \\ -1 \end{bmatrix} = (-1)\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

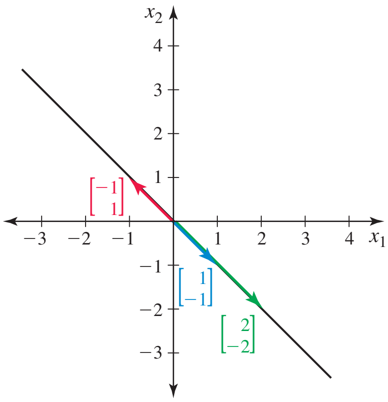


Figure 9.20 Three vectors that are all eigenvectors corresponding to the eigenvalue $\lambda_1 = -1$.

The vector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is not the only eigenvector associated with $\lambda_1 = -1$. In fact, any vector $a\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -a \\ a \end{bmatrix}$, $a \neq 0$, is an eigenvector associated with the eigenvalue -1. For instance, $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ or $\begin{bmatrix} 2 \\ -2 \end{bmatrix}$ are other choices. (See Figure 9.20.)

To find an eigenvector associated with $\lambda_2 = 4$, we must solve

$$A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

which yields

$$\begin{aligned} x_1 + 2x_2 &= 4x_1 \\ 3x_1 + 2x_2 &= 4x_2 \end{aligned}$$

Simplifying, we obtain

$$\begin{aligned} -3x_1 + 2x_2 &= 0 \\ 3x_1 - 2x_2 &= 0 \end{aligned} \qquad \text{The two equations are equivalent}$$

For instance,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

satisfies the preceding system. We see that A has two eigenvalues: -1 and 4. An eigenvector associated with -1 is $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, and an eigenvector associated with 4 is $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$. (See Figure 9.21.) As before, any vector $b\begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $b \neq 0$, is an eigenvector associated with the eigenvalue 4. •

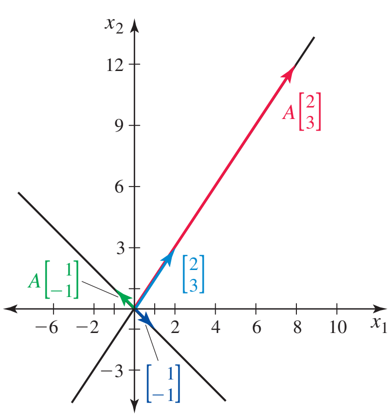


Figure 9.21 The two eigenvectors with their corresponding lines. The images of the eigenvectors under the map A remain on their respective lines.

Eigenvectors are not unique; they are determined only up to a multiplicative constant. When the eigenvalues are real, as in the previous example, all eigenvectors corresponding to a particular eigenvalue lie on the *same* straight line through the origin. For example, the line represented by the vector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is given by

$$l_1 = \{(x_1, x_2) : x_1 + x_2 = 0\}$$

and the line represented by the vector $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ is given by

$$l_2 = \{(x_1, x_2) : 3x_1 - 2x_2 = 0\}$$

The lines l_1 and l_2 are **invariant** under the map $\mathbf{x} \mapsto A\mathbf{x}$, in the sense that if we choose a point (x_1, x_2) on a line that is represented by an eigenvector, then since

$$A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

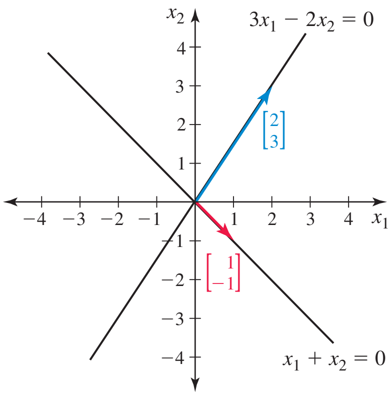


Figure 9.22 Eigenvectors and associated lines of the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$.

the result of the map is a point that is on the same line. We check this claim for the line l_1. Assume that $(x_1, x_2) \in l_1$; that is, $x_1 + x_2 = 0$. Then the point $(\lambda x_1, \lambda x_2) \in l_1$ as well, since $\lambda x_1 + \lambda x_2 = \lambda(x_1 + x_2) = 0$. This situation is illustrated in Figure 9.22, where we draw both eigenvectors and the corresponding lines. We will refer to a graph such as this as the **geometric interpretation** of eigenvalues and eigenvectors.

Since eigenvectors are often used in further calculations, you should choose values that are easy to work with. Small integers are good choices. How do we choose them? If we look at the straight line $3x_1 - 2x_2 = 0$ representing the eigenvectors corresponding to the eigenvalue $\lambda_2 = 4$ in Example 5, we see that the line goes through the origin and has slope $3/2$. The entries of the vector $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ are thus the smallest integers that produce a vector on this line.

In the next example, we use a different matrix, which will also have real and distinct eigenvalues, to show the procedure for finding eigenvalues and eigenvectors one more time.

EXAMPLE 6 **Finding Eigenvalues and Eigenvectors** Find the eigenvalues and eigenvectors of the matrix $A = \begin{bmatrix} 1 & 4 \\ 1 & -2 \end{bmatrix}$.

Solution To find the eigenvalues of A, we must solve $\det(A - \lambda I) = 0$. Now

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 1-\lambda & 4 \\ 1 & -2-\lambda \end{bmatrix} \\ &= (1-\lambda)(-2-\lambda) - 4 = \lambda^2 + \lambda - 6 \\ &= (\lambda - 2)(\lambda + 3) \end{aligned}$$

so $\det(A - \lambda I) = 0$ yields

$$\lambda_1 = 2 \qquad \text{and} \qquad \lambda_2 = -3.$$

To find an eigenvector associated with the eigenvalue $\lambda_1 = 2$, we must determine x_1 and x_2 (not both equal to 0) such that

$$\begin{bmatrix} 1 & 4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Writing this matrix equation as a system of linear equations, we get

$$\begin{aligned} x_1 + 4x_2 &= 2x_1 \\ x_1 - 2x_2 &= 2x_2 \end{aligned}$$

We see that both equations are equivalent to:

$$-x_1 + 4x_2 = 0.$$

Setting $x_2 = 1$, we find that $x_1 = 4$; that is, $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ is an eigenvector associated with the eigenvalue $\lambda_1 = 2$.

To find an eigenvector associated with the eigenvalue $\lambda_2 = -3$, we must determine x_1 and x_2 (not both equal to 0) such that

$$\begin{bmatrix} 1 & 4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -3 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Writing this matrix equation as a system of linear equations, we obtain

$$\begin{aligned} x_1 + 4x_2 &= -3x_1 \\ x_1 - 2x_2 &= -3x_2 \end{aligned}$$

Both equations are equivalent to $x_1 + x_2 = 0$. Setting $x_1 = 1$, we find that $x_2 = -1$; that is, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector associated with the eigenvalue $\lambda_2 = -3$.

Both eigenvectors and the corresponding lines are illustrated in Figure 9.23. •

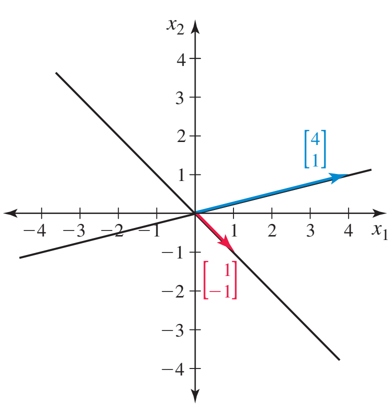


Figure 9.23 The eigenvectors and their corresponding lines in Example 6.

In the definition of eigenvectors and eigenvalues we argued that only non-zero vectors x are allowed for eigenvectors. However, as the next example shows, a matrix may still have 0 as an eigenvalue.

EXAMPLE 7 Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Find the eigenvalues and corresponding eigenvectors of A.

Solution To find the eigenvalues of A, we must calculate

$$\begin{aligned}\det(A - \lambda I) &= \det \begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \qquad \det(A-\lambda I)=0 \text{ for eigenvalues} \\ &= (1-\lambda)^2 - (1)(1) \\ &= 1 - 2\lambda + \lambda^2 - 1 = \lambda^2 - 2\lambda \\ &= \lambda(\lambda - 2)\end{aligned}$$

Solving $\det(A - \lambda I) = 0$ yields:

$$\lambda_1 = 0 \quad \text{and} \quad \lambda_2 = 2$$

The eigenvector corresponding to $\lambda_1 = 0$ satisfies

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Writing this matrix equation as a system of linear equations, we get

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_1 + x_2 &= 0 \end{aligned}$$

The two equations are the same, and we can choose $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ as an eigenvector corresponding to $\lambda_1 = 0$.

The eigenvector corresponding to $\lambda_2 = 2$ satisfies

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

which results in

$$\begin{aligned} x_1 + x_2 &= 2x_1 \\ x_1 + x_2 &= 2x_2 \end{aligned}$$

We can choose $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as an eigenvector corresponding to $\lambda_2 = 2$. •

Let's look at one last example of finding eigenvalues and eigenvectors. This time, we choose a matrix that illustrates a case when we can read off eigenvalues from the matrix directly.

EXAMPLE 8 **Reading Off Eigenvalues from a Matrix** Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix}.$$

Solution To find the eigenvalues of A, we must solve

$$\begin{aligned}\det(A - \lambda I) &= \det \begin{bmatrix} -2-\lambda & 1 \\ 0 & -1-\lambda \end{bmatrix} \\ &= (-2-\lambda)(-1-\lambda) - (0)(1) \\ &= (-2-\lambda)(-1-\lambda) = 0\end{aligned}$$

Solving gives

$$\lambda_1 = -2 \quad \text{and} \quad \lambda_2 = -1$$

Look back at the matrix A: The eigenvalues we found are identical to the diagonal elements of A! That is because one of the off-diagonal elements of A is equal to 0. Knowing the following facts simplifies finding eigenvalues:

Eigenvalues of a Triangular Matrix

If the matrix A is of the form:

$$A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \quad \text{or} \quad A = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \quad \text{or} \quad A = \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix}$$

then the eigenvalues of A are a and c.

We will prove these results in Problems 61–63.

Returning to our example, let's find the associated eigenvectors. To find an eigenvector associated with the eigenvalue $\lambda_1 = -2$, we must determine x_1 and x_2 (not both equal to 0) such that

$$\begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Writing this matrix equation as a system of linear equations, we obtain

$$\begin{aligned} -2x_1 + x_2 &= -2x_1 \\ -x_2 &= -2x_2. \end{aligned}$$

Simplifying either equation yields

$$x_2 = 0$$

This time, we cannot choose a value for x_2 (since $x_2 = 0$). But looking at the first equation, $-2x_1 + x_2 = -2x_1$, tells us that any value for x_1 will yield an identity, provided that $x_2 = 0$. Since $x_2 = 0$, we cannot choose $x_1 = 0$ (because $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is not an eigenvector); any value of $x_1 \neq 0$ will do, however, so let's choose $x_1 = 1$. Then $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector associated with $\lambda_1 = -2$.

To find an eigenvector associated with the eigenvalue $\lambda_2 = -1$, we must determine x_1 and x_2 (not both equal to 0) such that

$$\begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (-1) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Writing this matrix equation as a system of linear equations, we get

$$\begin{aligned} -2x_1 + x_2 &= -x_1 \\ -x_2 &= -x_2 \end{aligned}$$

which simplifies to

$$\begin{aligned} -x_1 + \quad x_2 &= 0 \\ 0x_2 &= 0 \end{aligned}$$

The first equation tells us that $x_1 = x_2$; the second equation tells us that we can choose any value for x_2. Choosing $x_2 = 1$, we need to set $x_1 = 1$. When we do, we find that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector associated with the eigenvalue $\lambda_2 = -1$. •

So far, we have seen only examples in which the eigenvalues were real and distinct. In the next example, we will see that eigenvalues can be complex. When they are, we will not compute the corresponding eigenvectors, because the eigenvectors will be complex as well and do not have a geometric interpretation.

EXAMPLE 9 **Complex Eigenvalues** Let

$$A = \begin{bmatrix} \cos 30^\circ & -\sin 30^\circ \\ \sin 30^\circ & \cos 30^\circ \end{bmatrix}.$$

Describe the action of A on the vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Compute the eigenvalues of A.

Solution We recognize that A is a matrix that describes a counterclockwise rotation by 30°. (The matrix A is the matrix R_θ for $\theta = 30^\circ$, as defined in Subsection 9.3.1.) The vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is thus rotated counterclockwise by 30°, and $A\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos 30^\circ \\ \sin 30^\circ \end{bmatrix} = \begin{bmatrix} \sqrt{3}/2 \\ 1/2 \end{bmatrix}$. (See Figure 9.24.)

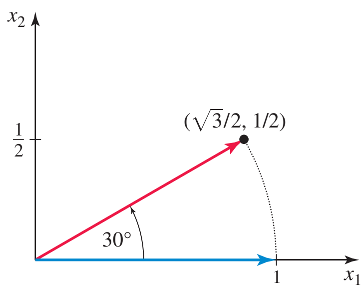


Figure 9.24 The vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is rotated counterclockwise by 30° in Example 9.

We will now compute the eigenvalues associated with this matrix. We set

$$\det(A - \lambda I) = 0$$

Using $\cos 30^\circ = \frac{1}{2}\sqrt{3}$ and $\sin 30^\circ = \frac{1}{2}$, we find that

$$\det \begin{bmatrix} \frac{1}{2}\sqrt{3} - \lambda & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2}\sqrt{3} - \lambda \end{bmatrix} = 0$$

or

$$\left(\frac{1}{2}\sqrt{3} - \lambda\right)^2 + \frac{1}{4} = 0$$

$$\frac{3}{4} - \lambda\sqrt{3} + \lambda^2 + \frac{1}{4} = 0$$

$$\lambda^2 - \sqrt{3}\lambda + 1 = 0$$

Solving this quadratic equation, we obtain

$$\lambda_{1,2} = \frac{\sqrt{3} \pm \sqrt{3-4}}{2} = \frac{1}{2}\left(\sqrt{3} \pm i\right)$$ Use quadratic formula from Section 1.2.6

where $i^2 = -1$. This solution shows that eigenvalues can be complex numbers. •

In case you are surprised that a real matrix like the one in Example 9 may have complex eigenvalues, notice that to calculate the eigenvalues we need to solve $\det(A - \lambda I) = 0$. For a 2×2 matrix this turns out to be a quadratic equation. Even with real coefficients a quadratic equation may have complex roots.

In Chapter 11, we will examine the stability of equilibria in systems of ordinary differential equations. This examination will lead us to investigate the eigenvalues of certain linear maps. It will be important to determine whether the real parts of the eigenvalues are positive or negative. In the case where the linear map is given by a 2×2 matrix, there is a useful criterion. Consider

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The determinant of A is

$$\det A = ad - bc$$

and the **trace** of A (denoted by $\operatorname{tr} A$) is defined as

$$\boxed{\operatorname{tr} A = a + d}$$

The trace is the sum of the diagonal elements of A. The trace and the determinant of a matrix are related to its eigenvalues.

The eigenvalues of A satisfy

$$\det(A - \lambda I) = 0$$

or

$$\det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = 0$$

which yields

$$(a - \lambda)(d - \lambda) - bc = 0$$
$$\lambda^2 - (a + d)\lambda + ad - bc = 0.$$

Since $\det A = ad - bc$ and $\operatorname{tr} A = a + d$, we can write the last equation for λ as the quadratic equation

$$\lambda^2 - (\operatorname{tr} A)\lambda + \det A = 0 \tag{9.20}$$

If λ_1 and λ_2 are the two solutions of (9.20), then λ_1 and λ_2 must satisfy

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = 0$$

Multiplying the left-hand side out, we find that λ_1 and λ_2 satisfy

$$\lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2 = 0$$

Comparing this equation with (9.20), we find the following important result:

Relationship of Eigenvalues to Determinant and Trace
If A is a 2×2 matrix with eigenvalues λ_1 and λ_2, then

$$\operatorname{tr} A = \lambda_1 + \lambda_2 \qquad \text{and} \qquad \det A = \lambda_1\lambda_2$$

To prepare for the next theorem, we make the following observations: Assume that λ_1 and λ_2 are both real and negative. Then the trace of A, which is the sum of the two eigenvalues, is negative, and the determinant of A, which is the product of the two eigenvalues, is positive.

In the case when the eigenvalues are complex conjugates, we have the same result. Assume that λ_1 and λ_2 are complex conjugates and that their real parts are negative. Then we can show that the trace of A is negative and the determinant is positive. (We will see an example of this fact later.)

That is, whenever the real parts of the two eigenvalues λ_1 and λ_2 are negative, it follows that $\operatorname{tr} A < 0$ and $\det A > 0$. The converse of this result is also true: If $\operatorname{tr} A < 0$ and $\det A > 0$, then both eigenvalues have negative real parts. This will be a useful criterion in Chapter 11, since it will enable us to determine whether or not both eigenvalues have negative real parts without computing the eigenvalues.

Theorem Let A be a 2×2 matrix with eigenvalues λ_1 and λ_2. Then the real parts of λ_1 and λ_2 are negative if and only if

$$\operatorname{tr} A < 0 \qquad \text{and} \qquad \det A > 0$$

EXAMPLE 10 **Trace and Determinant** Use the preceding theorem to show that both of the eigenvalues of

$$A = \begin{bmatrix} -1 & -3 \\ 1 & -2 \end{bmatrix}$$

have negative real parts. Then compute the eigenvalues, and use them to recompute the trace and the determinant of A.

Solution Since

$$\operatorname{tr} A = -1 - 2 = -3 < 0 \quad \text{and} \quad \det A = (-1)(-2) - (1)(-3) = 5 > 0$$

it follows from the theorem that both eigenvalues have negative real parts. To compute the eigenvalues, we solve

$$\det(A - \lambda I) = 0$$

or

$$\det \begin{bmatrix} -1-\lambda & -3 \\ 1 & -2-\lambda \end{bmatrix} = 0$$

Evaluating the preceding equation, we find that

$$\begin{aligned}(-1-\lambda)(-2-\lambda) - (1)(-3) &= 0 \\ \lambda^2 + 3\lambda + 5 &= 0\end{aligned}$$

That is,

$$\lambda_{1,2} = \frac{-3 \pm \sqrt{9-20}}{2} = -\frac{3}{2} \pm \frac{1}{2} i\sqrt{11}$$

So the real parts of both eigenvalues (which are equal to $-\frac{3}{2}$ for both λ_1 and λ_2) are negative.

We can now use the eigenvalues to recompute the trace and the determinant of A. For the trace, we have

$$\operatorname{tr} A = \lambda_1 + \lambda_2 = \left(-\frac{3}{2} + \frac{1}{2}i\sqrt{11}\right) + \left(-\frac{3}{2} - \frac{1}{2}i\sqrt{11}\right) = -3$$

This is the same result that we obtained previously. Note that since λ_1 and λ_2 are complex conjugates, the imaginary parts cancel when we add λ_1 and λ_2. The determinant of A is

$$\begin{aligned}\det A = \lambda_1\lambda_2 &= \left(-\frac{3}{2} + \frac{1}{2}i\sqrt{11}\right)\left(-\frac{3}{2} - \frac{1}{2}i\sqrt{11}\right) \\ &= \frac{9}{4} - \frac{1}{4}i^2(11) = \frac{9}{4} + \frac{11}{4} = \frac{20}{4} = 5\end{aligned}$$

which also is the same result that we obtained previously.

9.3.3 Iterated Maps (Needed for Sections 9.4 and 10.9)

We restrict ourselves to the case in which A is a 2×2 matrix with real eigenvalues. We saw that in this case the eigenvectors define lines through the origin that are invariant under the map A. If the two invariant lines are *not* identical, we say that the two eigenvectors are **linearly independent**. (See Figure 9.25.) This notion can be formulated as follows in terms of eigenvectors: If we denote the two eigenvectors by $\mathbf{u}_1$ and $\mathbf{u}_2$, then $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent if there does not exist a number a such that $\mathbf{u}_1 = a\mathbf{u}_2$. (Linear independence is defined not just for eigenvectors: *Any* two nonzero vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are linearly independent if there does not exist a number a such that $\mathbf{x}_1 = a\mathbf{x}_2$.)

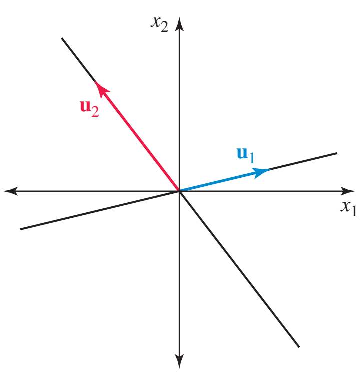


Figure 9.25 The two eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent.

The following criterion is useful:

> **Linear Independence of Eigenvectors** Let A be a 2×2 matrix with eigenvalues λ_1 and λ_2. Let $\mathbf{u}_1$ be the eigenvector associated with λ_1 and $\mathbf{u}_2$ the eigenvector associated with λ_2. If $\lambda_1 \neq \lambda_2$, then $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent.

There are also cases in which $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent even though $\lambda_1 = \lambda_2$. We will, however, be concerned primarily with cases in which $\lambda_1 \neq \lambda_2$. Hence, the preceding criterion will suffice for our purposes. (The other cases are covered in courses on linear algebra.)

A consequence of linear independence is that we can write any vector uniquely as a **linear combination** of two eigenvectors. Suppose that $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent eigenvectors of a 2×2 matrix; then any 2×1 vector $\mathbf{x}$ can be written as

$$\mathbf{x} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$$

where a_1 and a_2 are uniquely determined. We will not prove this statement but will examine what we can do with it.

If we apply A to $\mathbf{x}$ (written as a linear combination of the two eigenvectors of A), then, using the linearity of the map A, we find that

$$A\mathbf{x} = A(a_1\mathbf{u}_1 + a_2\mathbf{u}_2) = a_1 A\mathbf{u}_1 + a_2 A\mathbf{u}_2$$

Now, $\mathbf{u}_1$ and $\mathbf{u}_2$ are both eigenvectors corresponding to A. Hence, $A\mathbf{u}_1 = \lambda_1\mathbf{u}_1$ and $A\mathbf{u}_2 = \lambda_2\mathbf{u}_2$. We thus obtain

$$A\mathbf{x} = a_1\lambda_1\mathbf{u}_1 + a_2\lambda_2\mathbf{u}_2$$

This representation of $\mathbf{x}$ is particularly useful if we apply A repeatedly to $\mathbf{x}$. Applying A to $A\mathbf{x}$, we find that

$$\begin{aligned} A^2\mathbf{x} = A(A\mathbf{x}) &= A(a_1\lambda_1\mathbf{u}_1 + a_2\lambda_2\mathbf{u}_2) = a_1\lambda_1 A\mathbf{u}_1 + a_2\lambda_2 A\mathbf{u}_2 \\ &= a_1\lambda_1^2\mathbf{u}_1 + a_2\lambda_2^2\mathbf{u}_2 \end{aligned}$$

where we again used the fact that $\mathbf{u}_1$ and $\mathbf{u}_2$ are eigenvectors of the matrix A. Continuing in this way yields

$$A^n\mathbf{x} = a_1\lambda_1^n\mathbf{u}_1 + a_2\lambda_2^n\mathbf{u}_2 \tag{9.21}$$

Thus, instead of multiplying A with itself n times (which is rather time consuming), we can use the right-hand side of (9.21), which just amounts to adding two vectors (a much faster task) as the next example shows.

EXAMPLE 11 Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$. Find $A^{10}\mathbf{x}$ for $\mathbf{x} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$.

Solution We computed eigenvalues and eigenvectors for the matrix A earlier, and we found that

$$\lambda_1 = -1 \quad \text{with} \quad \mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and

$$\lambda_2 = 4 \qquad \text{with} \qquad \mathbf{u}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

We first represent $\mathbf{x} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ as a linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$. For this, we must find a_1 and a_2 so that

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + a_2 \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Writing this matrix equation as a system of linear equations yields

$$\begin{aligned} a_1 + 2a_2 &= 4 \qquad (R_1) \\ -a_1 + 3a_2 &= 1 \qquad (R_2) \end{aligned}$$

Using the method of elimination, we obtain

$$\begin{aligned} (R_1) \qquad a_1 + 2a_2 &= 4 \\ (R_1) + (R_2) \qquad 5a_2 &= 5 \end{aligned}$$

Hence, $a_2 = 1$ and $a_1 = 4 - 2a_2 = 4 - 2 = 2$. We now find that

$$\begin{bmatrix} 4 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

To compute $A^{10}\mathbf{x}$, we use the right-hand side of (9.21):

$$\begin{aligned} A^{10}\mathbf{x} = A^{10}\begin{bmatrix} 4 \\ 1 \end{bmatrix} &= A^{10}\left(2\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) \\ &= 2A^{10}\begin{bmatrix} 1 \\ -1 \end{bmatrix} + A^{10}\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= 2(-1)^{10}\begin{bmatrix} 1 \\ -1 \end{bmatrix} + 4^{10}\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= 2\begin{bmatrix} 1 \\ -1 \end{bmatrix} + 4^{10}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2{,}097{,}154 \\ 3{,}145{,}726 \end{bmatrix} \end{aligned}$$

To compute A^{10} directly would have taken a much longer time. •

Repeatedly multiplying a vector by the same matrix is known as an **iterated map**. Why would you ever need to do this? In the next section we will show how maps can be used to model the growth of populations, accounting for things like the age or sex structure of the population.

Section 9.3 Problems

9.3.1

1. Let

$$A = \begin{bmatrix} 2 & 2 \\ -1 & 4 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

(a) Show by direct calculation that $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$.

(b) Show by direct calculation that $A(\lambda\mathbf{x}) = \lambda(A\mathbf{x})$.

2. Let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

(a) Show by direct calculation that $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$.

(b) Show by direct calculation that $A(\lambda\mathbf{x}) = \lambda(A\mathbf{x})$.

In Problems 3–8, represent each given vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ in the x_1–x_2 plane, and determine its length and the angle that it forms with the positive x_1-axis (measured counterclockwise).

3. $\mathbf{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ **4.** $\mathbf{x} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$ **5.** $\mathbf{x} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

6. $\mathbf{x} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ **7.** $\mathbf{x} = \begin{bmatrix} -\sqrt{3} \\ -1 \end{bmatrix}$ **8.** $\mathbf{x} = \begin{bmatrix} 1 \\ -\sqrt{3} \end{bmatrix}$

In Problems 9–12, vectors are given in their polar coordinate representation (length r, and angle α measured counterclockwise from the positive x_1-axis). Find the representation of the vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ in Cartesian coordinates.

9. $r = 2, \alpha = 60^\circ$ **10.** $r = 3, \alpha = 120^\circ$

11. $r = 1, \alpha = 180^\circ$ **12.** $r = 5, \alpha = 270^\circ$

13. Suppose a vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ has length 3 and is 15° clockwise from the positive x_1-axis. Find x_1 and x_2.

14. Suppose a vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ has length 2 and is 140° clockwise from the positive x_1-axis. Find x_1 and x_2.

15. Suppose a vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ has length 3 and is 25° counterclockwise from the positive x_2-axis. Find x_1 and x_2.

16. Suppose a vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ has length 1 and is 90° counterclockwise from the negative x_2-axis. Find x_1 and x_2.

In Problems 17–22, find $\mathbf{x} + \mathbf{y}$ for each pair of vectors $\mathbf{x}$ and $\mathbf{y}$. Represent $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{x} + \mathbf{y}$ in the plane, and explain graphically how you add $\mathbf{x}$ and $\mathbf{y}$.

17. $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ **18.** $\mathbf{x} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

19. $\mathbf{x} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ **20.** $\mathbf{x} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

21. $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ **22.** $\mathbf{x} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

In Problems 23–28, compute $a\mathbf{x}$ for each vector $\mathbf{x}$ and scalar a. Represent $\mathbf{x}$ and $a\mathbf{x}$ in the plane, and explain graphically how you obtain $a\mathbf{x}$.

23. $\mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ and $a = 2$ **24.** $\mathbf{x} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ and $a = -1$

25. $\mathbf{x} = \begin{bmatrix} 0 \\ -4 \end{bmatrix}$ and $a = 0.5$ **26.** $\mathbf{x} = \begin{bmatrix} 3 \\ -9 \end{bmatrix}$ and $a = -1/3$

27. $\mathbf{x} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$ and $a = 1/4$ **28.** $\mathbf{x} = \begin{bmatrix} 0.5 \\ 0.25 \end{bmatrix}$ and $a = 4$

In Problems 29–34, let

$$\mathbf{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \textit{and} \quad \mathbf{w} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

29. Compute $\mathbf{u} + \mathbf{v}$ and illustrate the result graphically.

30. Compute $\mathbf{u} - \mathbf{v}$ and illustrate the result graphically.

31. Compute $\mathbf{w} - \mathbf{u}$ and illustrate the result graphically.

32. Compute $\mathbf{v} - \frac{1}{2}\mathbf{u}$ and illustrate the result graphically.

33. Compute $\mathbf{u} + \mathbf{v} + \mathbf{w}$ and illustrate the result graphically.

34. Compute $2\mathbf{v} - \mathbf{w}$ and illustrate the result graphically.

In Problems 35–40, give a geometric interpretation of the map $\mathbf{x} \mapsto A\mathbf{x}$ for each given map A.

35. $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

36. $A = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$

37. $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

38. $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

39. $A = \frac{1}{2}\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$

40. $A = \frac{1}{2}\begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & \sqrt{2} \end{bmatrix}$

41. Use a rotation matrix to rotate the vector $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ counterclockwise by the angle $\pi/3$.

42. Use a rotation matrix to rotate the vector $\begin{bmatrix} 4 \\ -1 \end{bmatrix}$ counterclockwise by the angle $\pi/6$.

43. Use a rotation matrix to rotate the vector $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ clockwise by the angle 45°.

44. Use a rotation matrix to rotate the vector $\begin{bmatrix} -2 \\ -3 \end{bmatrix}$ counterclockwise by the angle 45°.

45. Use a rotation matrix to rotate the vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ clockwise by the angle $\pi/3$.

46. Use a rotation matrix to rotate the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ counterclockwise by the angle $\pi/6$.

47. Use a rotation matrix to rotate the vector $\begin{bmatrix} 5 \\ -3 \end{bmatrix}$ clockwise by the angle $\pi/2$.

48. Use a rotation matrix to rotate the vector $\begin{bmatrix} -2 \\ -3 \end{bmatrix}$ counterclockwise by the angle $\pi/2$.

9.3.2

In Problems 49–56, find the eigenvalues λ_1 and λ_2 and corresponding eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ for each matrix A. Determine the equations of the lines through the origin in the direction of the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$, and graph the lines together with the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ and the vectors $A\mathbf{v}_1$ and $A\mathbf{v}_2$.

49. $A = \begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix}$

50. $A = \begin{bmatrix} 0 & 0 \\ 1 & -3 \end{bmatrix}$

51. $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

52. $A = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$

53. $A = \begin{bmatrix} -1 & 2 \\ 4 & 1 \end{bmatrix}$

54. $A = \begin{bmatrix} -1 & 0 \\ 4 & 3 \end{bmatrix}$

55. $A = \begin{bmatrix} 5 & 3 \\ -6 & -4 \end{bmatrix}$

56. $A = \begin{bmatrix} -2 & -1 \\ -2 & -1 \end{bmatrix}$

In Problems 57–60, find the eigenvalues λ_1 and λ_2 for each matrix A.

57. $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$

58. $A = \begin{bmatrix} -7 & 0 \\ 0 & 6 \end{bmatrix}$

59. $A = \begin{bmatrix} 1 & -3 \\ 0 & 2 \end{bmatrix}$

60. $A = \begin{bmatrix} -1 & 4 \\ 0 & -2 \end{bmatrix}$

61. Find the eigenvalues λ_1 and λ_2 for

$$A = \begin{bmatrix} a & 0 \\ c & b \end{bmatrix}$$

62. Find the eigenvalues λ_1 and λ_2 for

$$A = \begin{bmatrix} a & c \\ 0 & b \end{bmatrix}$$

63. **(a)** Show that the eigenvalues of the matrix $A = \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix}$ are $\lambda_1 = a$, and $\lambda_2 = c$.

(b) Show that the corresponding eigenvectors are $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

64. Let

$$A = \begin{bmatrix} -2 & -3 \\ -1 & 1 \end{bmatrix}$$

Without explicitly computing the eigenvalues of A, decide whether or not the real parts of both eigenvalues are negative.

65. Let

$$A = \begin{bmatrix} 1 & 4 \\ -4 & -3 \end{bmatrix}$$

Without explicitly computing the eigenvalues of A, decide whether or not the real parts of both eigenvalues are negative.

66. Let

$$A = \begin{bmatrix} 0 & -1 \\ 2 & -1 \end{bmatrix}$$

Without explicitly computing the eigenvalues of A, decide whether or not the real parts of both eigenvalues are negative.

67. Let

$$A = \begin{bmatrix} 2 & 2 \\ 2 & -3 \end{bmatrix}$$

Without explicitly computing the eigenvalues of A, decide whether or not the real parts of both eigenvalues are negative.

68. Let

$$A = \begin{bmatrix} -2 & 5 \\ 2 & -3 \end{bmatrix}$$

Without explicitly computing the eigenvalues of A, decide whether or not the real parts of both eigenvalues are negative.

9.3.3

69. Let

$$A = \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix}$$

(a) Show that

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

are eigenvectors of A and that $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent.

(b) Represent

$$\mathbf{x} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

as a linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$.

(c) Use your results in (a) and (b) to compute $A^{20}\mathbf{x}$.

70. Let

$$A = \begin{bmatrix} -1 & -2 \\ -4 & 1 \end{bmatrix}$$

(a) Show that

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

are eigenvectors of A and that $\mathbf{u}_1$ and $\mathbf{u}_2$ are linearly independent.

(b) Represent

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

as a linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$.

(c) Use your results in (a) and (b) to compute $A^{10}\mathbf{x}$.

71. Let

$$A = \begin{bmatrix} -1 & 0 \\ 2 & 1 \end{bmatrix}.$$

Find

$$A^{15} \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

without using a calculator.

72. Let

$$A = \begin{bmatrix} 4 & 3 \\ 2 & -1 \end{bmatrix}.$$

Find

$$A^{30} \begin{bmatrix} -4 \\ -2 \end{bmatrix}$$

without using a calculator.

73. Let

$$A = \begin{bmatrix} 5 & 7 \\ -2 & -4 \end{bmatrix}.$$

Find

$$A^{20} \begin{bmatrix} -3 \\ -2 \end{bmatrix}$$

without using a calculator.

74. Let

$$A = \begin{bmatrix} 1 & -1/4 \\ 1/2 & 1/4 \end{bmatrix}.$$

Find

$$A^{30} \begin{bmatrix} 1/2 \\ 3/2 \end{bmatrix}$$

without using a calculator.

9.4 Demographic Modeling

9.4.1 Modeling with Leslie Matrices

In Chapters 2 and 8, we learned how to use mathematical models to predict the growth of populations. However, in all of these mathematical models we assumed that all animals in the population could be treated in the same way. In fact, the processes that lead to population size changing over time (i.e., birth and death) depend on the **demographic structure** of the population, that is the distribution of ages and sexes within the population. For example, suppose we want to predict the number of births that will occur in a population of elephants within a Sri Lankan National Park (which is the problem that motivated us in Section 2.3). Our prediction for the number of births will be affected if we know that all of the elephants in the park are female, or if they are all elderly.

In this section we will discuss how to use linear maps to model the size of demographically structured populations over time. Our models will be discrete in the sense that they predict the population structure at discrete times $t = 0, 1, 2, \ldots$. Just as in Chapter 2, these times must be specified to develop the model, and will depend on how frequently we are able to observe the population; for example, for a growing population of microorganisms, the interval between measurements may be an hour, while for the elephants in the National Park, it could be a month or a year.

We begin with the simplest model. Suppose that at time t, there are $N(t)$ elephants in the population. We take measurements of the population size at discrete times $t = 0, 1, 2, \ldots$. For example, if the elephants are counted at yearly intervals, then $t = 0$ would be the start of the observation period, $t = 1$ would be one year later, and so on. Suppose that in one year a fraction b of the elephants in the population give birth to baby elephants, while a fraction m of the elephants die.

Then:

$$\begin{aligned} N(t+1) &= N(t) + \begin{matrix}\text{No.}\\ \text{births}\\ \text{in 1 year}\end{matrix} - \begin{matrix}\text{No.}\\ \text{deaths}\\ \text{in 1 year}\end{matrix} \\ &= N(t) + bN(t) - mN(t) \\ &= (1 + b - m)N(t) \end{aligned}$$

So if we define a new constant $R = 1 + b - m$, then

$$N(t+1) = RN(t) \quad \text{for } t = 0, 1, 2, \ldots. \tag{9.22}$$

To solve this model we need to know the initial size of the population; $N(0) = N_0$. The quantity R describes the relative change of the population size from year to year. We assume that $R \geq 0$. The solution of (9.22) can be found by first computing $N(t)$ for $t = 1, 2,$ and 3:

$$N(1) = RN(0)$$

$$N(2) = RN(1) = R[RN(0)] = R^2N(0)$$

$$N(3) = RN(2) = R[R^2N(0)] = R^3N(0)$$

Recognizing the pattern, we conclude that

$$N(t) = R^tN(0).$$

The population either grows exponentially (if $R > 1$), decays exponentially (if $R < 1$), or stays constant over time (if $R = 1$). However, knowing only the size of the population may not be enough to predict the number of births and deaths each year. Birth and death rates are affected by the ages and sexes of the animals in the population. For example, if the population consists mostly of elderly male elephants we would expect a high mortality rate and a low birth rate.

Reproduction is highly age dependent in many organisms, not just in mammals. For instance, periodical cicadas spend 13 to 17 years living underground, emerging to reproduce only once in their life. The purple coneflower, a prairie flower, does not reproduce until it is about three years old. To take the life history into account, it is not enough to know the total population size—our model needs to include how many animals or plants there are in each age group. These kinds of models were introduced by Patrick Leslie in 1945. Leslie used vectors to store all of the different demographic group sizes, and matrices to model how group sizes change over time. Leslie's approach is widely used not only in population biology, but also in the life insurance industry.

We begin with a specific example to illustrate the method. We will imagine that we are modeling the size of a colony of seabirds. The colony is made up of seabirds of different ages. Only one- and two-year-old birds are capable of breeding: Younger birds do not breed, nor do older birds. Since only females produce offspring, we will follow only female birds. We assume that birds breed no more than once per year and that we take a census of the population at the end of each breeding season. Birds born during a particular breeding season are of age 0 at the end of that season. If a zero-year-old survives until the end of the next breeding season, it will be age 1 when we take a census at the end of that season. If a one-year-old survives until the end of the next breeding season, it will be age 2 at the end of that breeding season, and so on. In our example, we will assume that no bird lives beyond age 3; that is, there are no birds age 4 or older in the colony. We define

$$N_x(t) = \text{number of female birds of age } x \text{ at time } t$$

where $t = 0, 1, 2, \ldots$. Then

$$\mathbf{N}(t) = \begin{bmatrix} N_0(t) \\ N_1(t) \\ N_2(t) \\ N_3(t) \end{bmatrix}$$

$N_x(t)$ are predicted average numbers. They do not need to be integers

is the vector that describes the number of females in each age class at time t.

We assume that 40% of the females age 0, 30% of the females age 1, and 10% of the females age 2 at time t are alive when we take the census at time $t + 1$. That is,

$$N_1(t + 1) = (0.4)N_0(t)$$

$$N_2(t + 1) = (0.3)N_1(t)$$

$$N_3(t + 1) = (0.1)N_2(t)$$

The number of zero-year-old females at time $t + 1$ is equal to the number of female offspring during the breeding season that survive until the end of the breeding season, when the census is taken. Let's assume that on average each 1-year-old female produces two female chicks that survive to the end of the year, while each 2-year-old female produces one and a half female chicks that survive to the end of year. (Remember that we are following only the number of female birds, so if a bird is producing equal numbers of male as female chicks, then a female bird will have to produce 3 chicks to have 1.5 female chicks.) Then our modeling assumptions imply that:

$$N_0(t+1) = \begin{array}{c}\text{No. female}\\ \text{chicks born to}\\ \text{1-year-old mothers}\end{array} + \begin{array}{c}\text{No. female}\\ \text{chicks born to}\\ \text{2-year-old mothers}\end{array}$$

$$N_0(t+1) = 2N_1(t) + 1.5N_2(t)$$

There is no contribution of three-year-olds to the newborn class: $N_0(t+1)$.

We can summarize the dynamics in matrix form:

$$\begin{bmatrix} N_0(t+1) \\ N_1(t+1) \\ N_2(t+1) \\ N_3(t+1) \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1.5 & 0 \\ 0.4 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \end{bmatrix} \begin{bmatrix} N_0(t) \\ N_1(t) \\ N_2(t) \\ N_3(t) \end{bmatrix} \tag{9.23}$$

The 4×4 matrix in this equation is called the **Leslie matrix**, which we denote by L. We can write the matrix equation (9.23) in short form as

$$\mathbf{N}(t+1) = L\mathbf{N}(t) \tag{9.24}$$

Compare this equation with (9.22); just as in (9.22) we get the population at time $t+1$ by multiplying the population at time t by a constant, but here it is a constant matrix (L), rather than a scalar (R).

To see how this equation works, let's assume that the population at time t has the following age distribution:

$$N_0(t) = 1000, \quad N_1(t) = 200, \quad N_2(t) = 100, \quad \text{and} \quad N_3(t) = 10$$

When we take a census after the next breeding season, we find that 40% of the females age 0 at time t are alive at time $t+1$; that is, $N_1(t+1) = (0.4)(1000) = 400$. Also, 30% of the females age 1 at time t are alive at time $t+1$; that is, $N_2(t+1) = (0.3)(200) = 60$. And 10% of the females age 2 at time t are alive at time $t+1$; that is, $N_3(t+1) = (0.1)(100) = 10$. The number of surviving female offspring at time $t+1$ is

$$N_0(t+1) = 2N_1(t) + 1.5N_2(t) = (2)(200) + (1.5)(100) = 550$$

We can compute these same population sizes using the Leslie matrix (9.23):

$$\begin{aligned} \begin{bmatrix} N_0(t+1) \\ N_1(t+1) \\ N_2(t+1) \\ N_3(t+1) \end{bmatrix} &= \begin{bmatrix} 0 & 2 & 1.5 & 0 \\ 0.4 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \end{bmatrix} \begin{bmatrix} 1000 \\ 200 \\ 100 \\ 10 \end{bmatrix} \\ &= \begin{bmatrix} (2)(200) + (1.5)(100) \\ (0.4)(1000) \\ (0.3)(200) \\ (0.1)(100) \end{bmatrix} = \begin{bmatrix} 550 \\ 400 \\ 60 \\ 10 \end{bmatrix} \end{aligned}$$

To obtain the age distribution at time $t+2$, we apply the Leslie matrix to the population vector at time $t+1$; that is,

$$\begin{bmatrix} N_0(t+2) \\ N_1(t+2) \\ N_2(t+2) \\ N_3(t+2) \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1.5 & 0 \\ 0.4 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \end{bmatrix} \begin{bmatrix} 550 \\ 400 \\ 60 \\ 10 \end{bmatrix} = \begin{bmatrix} 890 \\ 220 \\ 120 \\ 6 \end{bmatrix}$$

These three breeding seasons are illustrated in Figure 9.26.

Generalizing the preceding discussion, we can introduce the general form of the Leslie matrix for a population with age classes. We assume that the population

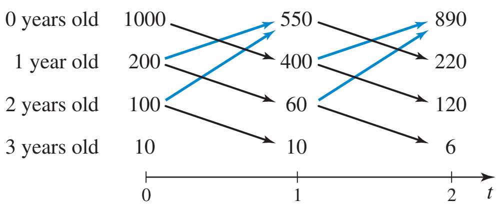


Figure 9.26 An illustration of seabird colony population changes.

is divided into $m + 1$ age classes. A census is taken at the end of each breeding season. Then, for the survival of each age class, we find that

$$N_1(t+1) = P_0 N_0(t)$$
$$N_2(t+1) = P_1 N_1(t)$$
$$\vdots$$
$$N_m(t+1) = P_{m-1} N_{m-1}(t)$$

where P_i denotes the fraction of females age i at time t that survive to time $t+1$. Since P_i denotes a fraction, it follows that $0 \le P_i \le 1$ for $i = 0, 1, \ldots, m-1$. The number of zero-year-old females is given by

$$N_0(t+1) = F_0 N_0(t) + F_1 N_1(t) + \cdots + F_m N_m(t)$$

where F_i is the average number of surviving female offspring per female individual with age i. Writing this equation in matrix form, we find that the Leslie matrix is given by

$$\begin{bmatrix} F_0 & F_1 & F_2 & \ldots & F_{m-1} & F_m \\ P_0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & P_1 & 0 & \cdots & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & 0 & P_{m-1} & 0 \end{bmatrix}$$

This is a matrix in which all elements are 0, except possibly those in the first row and in the first subdiagonal below the diagonal. The size of the matrix is $(m+1) \times (m+1)$, reflecting the $m + 1$ age classes.

EXAMPLE 1 Suppose that the Leslie matrix of a population is

$$\begin{bmatrix} 5 & 7 & 1.5 \\ 0.2 & 0 & 0 \\ 0 & 0.4 & 0 \end{bmatrix}$$

Interpret this matrix, and determine what happens if you follow a population for two seasons, starting with 1000 zero-year-old females.

Solution The population is divided into three age classes: zero-year-olds, one-year-olds, and two-year-olds. The elements below the diagonal represent the survival probabilities; that is, 20% of the zero-year-olds survive until the next census, and 40% of the one-year-olds survive until the next census. The maximum age is two years. Zero-year-olds produce an average of five surviving females per female; one-year-olds produce an average of seven surviving females per female; two-year-olds produce an average of 1.5 surviving females per female.

A population that starts with 1000 zero-year-old females at time 0 has the population vector

$$\mathbf{N}(0) = \begin{bmatrix} 1000 \\ 0 \\ 0 \end{bmatrix}$$

Using $\mathbf{N}(t+1) = L\mathbf{N}(t)$, we find that

$$\mathbf{N}(1) = \begin{bmatrix} 5 & 7 & 1.5 \\ 0.2 & 0 & 0 \\ 0 & 0.4 & 0 \end{bmatrix} \begin{bmatrix} 1000 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5000 \\ 200 \\ 0 \end{bmatrix}$$

and

$$\mathbf{N}(2) = \begin{bmatrix} 5 & 7 & 1.5 \\ 0.2 & 0 & 0 \\ 0 & 0.4 & 0 \end{bmatrix} \begin{bmatrix} 5000 \\ 200 \\ 0 \end{bmatrix} = \begin{bmatrix} 26{,}400 \\ 1000 \\ 80 \end{bmatrix}$$

The Leslie matrix method can be used to model population dynamics in organisms that have more complex life cycles, as the following example shows.

EXAMPLE 2 **Life Cycles in Nanoflagellates** Nanoflagellates are microscopic filter feeding organisms that live in both salt- and fresh-water environments and prey on bacteria. Typically nanoflagellates anchor themselves to objects in the water—for example, leaf debris, or even the skins of animals, like whales. They then use their tails (or flagella) to stir up the water and pull in nearby bacteria. This stage is known as the **sessile** life stage of the nanoflagellate. A problem with this strategy is that a nanoflagellate may find itself stuck in a place with few bacterial prey. In this case, the nanoflagellate can detach from the object and swim off in search of new places to feed. We call this the **freely swimming** life stage of the nanoflagellate. Finally, sessile nanoflagellates can reproduce: During reproduction a daughter cell detaches from the attached parent, and then swims off. Newly detached cells cannot attach straight away but must swim around while they grow feeding organs. We call this stage the **juvenile** life stage.

Derive a Leslie matrix–based model that describes the possible life stage transitions that can occur in one day in a population of nanoflagellates. Assume that:

- Of the sessile nanoflagellates a fraction m die in one day, and a fraction d detach and become freely swimming. Another fraction b reproduce (i.e., produce daughter cells during the day).
- Of the freely swimming cells, a fraction n die each day, while a fraction a attach and become sessile.
- Of the juvenile cells a fraction p die each day, and a fraction g mature into freely swimming cells.

Solution We first define a vector of population sizes for each of the three groups. Specifically, we define $N_s(t)$ to be the number of sessile nanoflagellates, $N_f(t)$ the number of freely swimming nanoflagellates, and $N_j(t)$ the number of juveniles, all at time t. To keep track of all the processes that change each of the population sizes it is helpful to draw a diagram (see Figure 9.27).

We then write down word equations for each process. For the sessile cells:

$$\begin{array}{c} \text{No. sessile cells} \\ \text{at time } t+1 \end{array} = \begin{array}{c} \text{No. sessile cells} \\ \text{at time } t \end{array} - \begin{array}{c} \text{No. sessile cells} \\ \text{that die between} \\ t \text{ and } t+1 \end{array}$$

$$- \begin{array}{c} \text{No. sessile cells} \\ \text{that detach between} \\ t \text{ and } t+1 \end{array} + \begin{array}{c} \text{No. freely swimming} \\ \text{cells that attach} \\ \text{between } t \text{ and } t+1 \end{array}$$

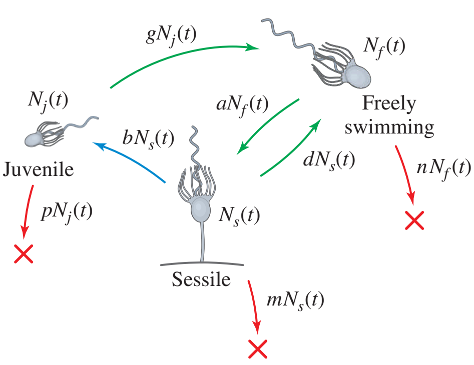


Figure 9.27 Graphical representation of the changes in population sizes among sessile, freely swimming, and juvenile nanoflagellates. The arrow between sessile and juvenile cells represents births (new cells added to the population).

So:

$$\begin{aligned} N_s(t+1) &= N_s(t) - mN_s(t) - dN_s(t) + aN_f(t) \\ &= (1-m-d)N_s(t) + aN_f(t). \end{aligned} \tag{9.25}$$

For the freely swimming cells:

$$\begin{array}{c} \text{No. freely swimming} \\ \text{cells at time } t+1 \end{array} = \begin{array}{c} \text{No. freely swimming} \\ \text{cells at time } t \end{array} - \begin{array}{c} \text{No. freely swimming cells} \\ \text{that die between} \\ t \text{ and } t+1 \end{array}$$

$$- \begin{array}{c} \text{No. freely swimming cells} \\ \text{that attach between} \\ t \text{ and } t+1 \end{array} + \begin{array}{c} \text{No. sessile cells} \\ \text{that detach between} \\ t \text{ and } t+1 \end{array}$$

$$+ \begin{array}{c} \text{No. juvenile cells that} \\ \text{mature into} \\ \text{freely swimming cells} \\ \text{between } t \text{ and } t+1 \end{array}$$

So

$$\begin{aligned} N_f(t+1) &= N_f(t) - nN_f(t) - aN_f(t) + dN_s(t) + gN_j(t) \\ &= (1-n-a)N_f(t) + dN_s(t) + gN_j(t) \end{aligned} \tag{9.26}$$

And finally for the juvenile cells:

$$\text{No. juvenile cells at time } t+1 = \begin{array}{c}\text{No. juvenile cells}\\ \text{at time } t\end{array} - \begin{array}{c}\text{No. juvenile cells}\\ \text{that die between}\\ t \text{ and } t+1\end{array}$$

$$- \begin{array}{c}\text{No. juvenile cells}\\ \text{that mature into}\\ \text{freely swimming cells}\\ \text{between } t \text{ and } t+1\end{array} + \begin{array}{c}\text{No. juvenile cells}\\ \text{born to}\\ \text{sessile cells}\\ \text{between } t \text{ and } t+1\end{array}$$

$$\begin{aligned} N_j(t+1) &= N_j(t) - pN_j(t) - gN_j(t) + bN_s(t) \\ &= (1-p-g)N_j(t) + bN_s(t) \end{aligned} \tag{9.27}$$

We define a vector $\mathbf{N}(t)$ that contains the sizes of all three populations,

$$\mathbf{N}(t) = \begin{bmatrix} N_s(t) \\ N_f(t) \\ N_j(t) \end{bmatrix}$$

then Equations (9.25) through (9.27) may be written in the form of a matrix equation.

$$\mathbf{N}(t+1) = \begin{bmatrix} 1-m-d & a & 0 \\ d & 1-n-a & g \\ b & 0 & 1-p-g \end{bmatrix} \mathbf{N}(t).$$

The constants m, d, n, a, g, b allow the model to be fit to different species of nanoflagellate and to different habitats. •

9.4.2 Stable Age Distributions in Demographic Models

We will now investigate what happens when we run a demographic model for a long time. Let's assume that the Leslie matrix is given by

$$L = \begin{bmatrix} 1.5 & 2 \\ 0.08 & 0 \end{bmatrix}$$

Using the equation

$$\mathbf{N}(t+1) = L\mathbf{N}(t)$$

we can compute successive population vectors; that is,

$$\begin{aligned} N_0(t+1) &= 1.5N_0(t) + 2N_1(t) \\ N_1(t+1) &= 0.08N_0(t) \end{aligned}$$

Suppose that we start with $N_0(0) = 100$ and $N_1(0) = 100$. Then

$$\begin{bmatrix} N_0(1) \\ N_1(1) \end{bmatrix} = \begin{bmatrix} 1.5 & 2 \\ 0.08 & 0 \end{bmatrix} \begin{bmatrix} 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 350 \\ 8 \end{bmatrix}$$

Continuing in this way, we find the population vectors at successive times, starting at time 0:

$$\begin{bmatrix} 100 \\ 100 \end{bmatrix}, \begin{bmatrix} 350 \\ 8 \end{bmatrix}, \begin{bmatrix} 541 \\ 28 \end{bmatrix}, \begin{bmatrix} 868 \\ 43 \end{bmatrix}, \begin{bmatrix} 1388 \\ 69 \end{bmatrix},$$

$$\begin{bmatrix} 2221 \\ 111 \end{bmatrix}, \begin{bmatrix} 3553 \\ 178 \end{bmatrix}, \begin{bmatrix} 5685 \\ 284 \end{bmatrix}, \ldots$$

(In these calculations, we rounded to the nearest integer.) The first thing you notice is that the total population is growing. [The population size at time t is $N_0(t) + N_1(t)$.] But we can say a lot more. If we look at the successive ratios

$$q_0(t) = \frac{N_0(t)}{N_0(t-1)}$$

we find the following:

t	1	2	3	4	5	6	7
$q_0(t)$	3.5	1.55	1.60	1.5991	1.6001	1.5997	1.6001

That is, $q_0(t)$ seems to approach a limiting value, namely, 1.6. The same happens when we look at the ratio

$$q_1(t) = \frac{N_1(t)}{N_1(t-1)}$$

We find

t	1	2	3	4	5	6	7
$q_1(t)$	0.08	3.5	1.536	1.605	1.609	1.604	1.596

Both ratios seem to approach 1.6. Moreover, if we look at the fraction of females in age class 0, namely,

$$p(t) = \frac{N_0(t)}{N_0(t) + N_1(t)}$$

we find

t	0	1	2	3	4	5	6	7
$p(t)$	0.5	0.9777	0.9508	0.9528	0.9526	0.9524	0.9523	0.9524

That is, this fraction also seems to converge. It looks as if about 95.2% of the population is in age class 0 when t is sufficiently large.

Although the population is increasing in size, the fraction of females in age class 0 (and hence also in age class 1) seems to converge. This constant fraction is referred to as the **stable age distribution**.

Not all matrix models have stable age distributions: We will give an example in Problem 11 in which the population does not reach a stable age distribution.

If we start the population in the stable age distribution, the fraction of females in age class 0 will remain the same, about 95.2%, and the population will increase by a constant by a factor of 1.6 each generation. Here is a numerical illustration of these two important properties: A stable age distribution for this population is

$$\mathbf{N}(0) = \begin{bmatrix} 2000 \\ 100 \end{bmatrix}$$

(We will learn in the next subsection how to find this vector.) If we start with the foregoing stable age distribution, then we have

$$\mathbf{N}(1) = L\mathbf{N}(0) = \begin{bmatrix} 1.5 & 2 \\ 0.08 & 0 \end{bmatrix} \begin{bmatrix} 2000 \\ 100 \end{bmatrix} = \begin{bmatrix} 3200 \\ 160 \end{bmatrix}$$

The fraction of females in age class 0 remains the same, namely, $3200/3360 = 2000/2100 \simeq 0.952$. We compare this with

$$(1.6) \begin{bmatrix} 2000 \\ 100 \end{bmatrix} = \begin{bmatrix} 3200 \\ 160 \end{bmatrix}$$

which yields the same result. That is, if $\mathbf{N}$ denotes a stable age distribution, then

$$L\mathbf{N} = \lambda \mathbf{N} \tag{9.28}$$

where $\lambda = 1.6$ and $N = [N_0, N_1]'$, with $N_0/(N_0 + N_1) \approx 95.2\%$.

Equation (9.28) is used to determine the stable age distribution. From Section 9.3 we recognize that the stable age distribution $\mathbf{N}$ is an eigenvector of the Leslie matrix, and the population growth rate λ is the corresponding eigenvalue.

We can therefore use rules for calculating iterative maps using eigenvalues and eigenvectors to determine how the sizes of different population classes will change over time.

EXAMPLE 3 In Subsection 9.4.1, we investigated an age-structured population with Leslie matrix

$$L = \begin{bmatrix} 1.5 & 2 \\ 0.08 & 0 \end{bmatrix}$$

Find both eigenvalues and eigenvectors, and derive a formula for the population size as a function of time if $\mathbf{N}(0) = \begin{bmatrix} 105 \\ 1 \end{bmatrix}$.

Solution To find the eigenvalues, we must solve $\det(L - \lambda I) = 0$. Since

$$L - \lambda I = \begin{bmatrix} 1.5 - \lambda & 2 \\ 0.08 & -\lambda \end{bmatrix}$$

it follows that

$$\begin{aligned} \det(L - \lambda I) &= (1.5 - \lambda)(-\lambda) - (2)(0.08) \\ &= -1.5\lambda + \lambda^2 - 0.16 \\ &= (\lambda - 1.6)(\lambda + 0.1). \qquad \text{Factorize} \end{aligned}$$

The eigenvalues are

$$\lambda_1 = 1.6 \quad \text{and} \quad \lambda_2 = -0.1$$

To compute the corresponding eigenvectors, we start with the larger eigenvalue, $\lambda_1 = 1.6$. We need to solve

$$\begin{bmatrix} 1.5 & 2 \\ 0.08 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (1.6) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

or

$$\begin{aligned} 1.5x_1 + 2x_2 &= 1.6x_1 \\ 0.08x_1 \qquad &= 1.6x_2 \end{aligned}$$

So:

$$\begin{aligned} -0.1x_1 + \quad 2x_2 &= 0 \qquad (R_1) \quad \text{Bring all terms to left hand side} \\ 0.08x_1 - 1.6x_2 &= 0 \qquad (R_2) \end{aligned}$$

or

$$\begin{aligned} -10(R_1) \qquad x_1 - 20x_2 &= 0 \\ 12.5(R_2) \qquad x_1 - 20x_2 &= 0 \end{aligned}$$

It follows that both equations are satisfied if:

$$x_1 = 20x_2$$

For instance, $x_2 = 1$ and $x_1 = 20$ satisfies the preceding equation. Therefore, an eigenvector corresponding to $\lambda_1 = 1.6$ is

$$\mathbf{u}_1 = \begin{bmatrix} 20 \\ 1 \end{bmatrix}.$$

The eigenvector corresponding to $\lambda_2 = -0.1$ satisfies

$$\begin{bmatrix} 1.5 & 2 \\ 0.08 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -0.1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

or

$$\begin{aligned} 1.5x_1 + 2x_2 &= -0.1x_1 \\ 0.08x_1 \qquad &= -0.1x_2 \end{aligned}$$

so:

$$\begin{aligned} 1.6x_1 + \quad 2x_2 &= 0 \qquad (R_1) \quad \text{Bring all terms to left hand side} \\ 0.08x_1 + 0.1x_2 &= 0 \qquad (R_2) \end{aligned}$$

or

$$\begin{aligned} 0.5\,(R_1) \qquad 0.8x_1 + x_2 &= 0 \\ 10\,(R_2) \qquad 0.8x_1 + x_2 &= 0 \end{aligned}$$

It follows that both equations are satisfied if

$$0.8x_1 = -x_2.$$

For instance, $x_1 = 5$ and $x_2 = -4$ satisfies the preceding equation. Therefore, an eigenvector corresponding to the eigenvalue $\lambda_2 = -0.1$ is

$$\mathbf{u}_2 = \begin{bmatrix} 5 \\ -4 \end{bmatrix}$$

Once we have the eigenvalues and eigenvectors we can use the theory of iterated maps from Section 9.3 to calculate the population sizes as functions of t.

Since $\mathbf{N}(0) = \begin{bmatrix} 105 \\ 1 \end{bmatrix}$ we can compute $\mathbf{N}(1)$ from:

$$\mathbf{N}(1) = L\mathbf{N}(0) = \begin{bmatrix} 1.5 & 2 \\ 0.08 & 0 \end{bmatrix} \begin{bmatrix} 105 \\ 1 \end{bmatrix} = \begin{bmatrix} 159.5 \\ 8.4 \end{bmatrix}$$

If we wanted to compute $\mathbf{N}(t)$ for some integer value t, we would need to find

$$\mathbf{N}(t) = L^t\mathbf{N}(0)$$

$\mathbf{N}(t)$ can be computed from (9.21). We need to write $\mathbf{N}(0)$ as a linear combination of the two eigenvectors and then apply L^t to that combination.

Looking at the eigenvectors, we see that $\mathbf{N}(0)$, as a linear combination of the two eigenvectors, is

$$\begin{bmatrix} 105 \\ 1 \end{bmatrix} = 5\begin{bmatrix} 20 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ -4 \end{bmatrix}$$

Now, if we want to compute $\mathbf{N}(1)$, we must find

$$\begin{aligned} \mathbf{N}(1) = L\mathbf{N}(0) = L\begin{bmatrix} 105 \\ 1 \end{bmatrix} &= L\left(5\begin{bmatrix} 20 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ -4 \end{bmatrix}\right) \\ &= 5L\begin{bmatrix} 20 \\ 1 \end{bmatrix} + L\begin{bmatrix} 5 \\ -4 \end{bmatrix} \end{aligned}$$

Since $\begin{bmatrix} 20 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_1 = 1.6$, it follows that $L\begin{bmatrix} 20 \\ 1 \end{bmatrix} = 1.6\begin{bmatrix} 20 \\ 1 \end{bmatrix}$. Likewise, since $\begin{bmatrix} 5 \\ -4 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_2 = -0.1$, it follows that $L\begin{bmatrix} 5 \\ -4 \end{bmatrix} = -0.1\begin{bmatrix} 5 \\ -4 \end{bmatrix}$. Hence,

$$\begin{aligned} \mathbf{N}(1) &= (5)(1.6)\begin{bmatrix} 20 \\ 1 \end{bmatrix} + (-0.1)\begin{bmatrix} 5 \\ -4 \end{bmatrix} \\ &= \begin{bmatrix} 159.5 \\ 8.4 \end{bmatrix} \end{aligned}$$

which, of course, is the same answer as before. Now, let's find $\mathbf{N}(t)$ for any t. For this, we need to compute

$$\begin{aligned} \mathbf{N}(t) = L^t\mathbf{N}(0) = L^t\begin{bmatrix} 105 \\ 1 \end{bmatrix} &= L^t\left(5\begin{bmatrix} 20 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ -4 \end{bmatrix}\right) \\ &= 5(1.6)^t\begin{bmatrix} 20 \\ 1 \end{bmatrix} + (-0.1)^t\begin{bmatrix} 5 \\ -4 \end{bmatrix}. \end{aligned} \tag{9.29}$$

Equation (9.29) from Example 3 can be used to prove our hypotheses about the stable age distribution and rate of growth of the population. In this equation the term $(1.6)^t$ grows much faster than $(-0.1)^t$. In fact, $(-0.1)^t$ tends to 0 as $t \to \infty$. Hence for large values of t we expect the first term in (9.29) to dominate over the second term. So for large t:

$$\mathbf{N}(t) \approx 5(1.6)^t\begin{bmatrix} 20 \\ 1 \end{bmatrix}.$$

From this equation we see that

$$q_0(t) = \frac{N_0(t)}{N_0(t-1)} \approx \frac{5(1.6)^t \cdot 20}{5(1.6)^{t-1} \cdot 20} = 1.6$$

Read off the first component of the solution vector

and

$$q_1(t) = \frac{N_1(t)}{N_1(t-1)} \approx \frac{5(1.6)^t \cdot 1}{5(16)^{t-1} \cdot 1} = 1.6.$$

So the populations in both age classes grow by a factor of 1.6 each year. Meanwhile the fraction of the population with age 0 is:

$$\frac{N_0(t)}{N_0(t) + N_1(t)} = \frac{5(1.6)^t \cdot 20}{5(1.6)^t \cdot 20 + 5(1.6)^t \cdot 1} = \frac{20}{21} \approx 0.952.$$

Both of these results are consistent with our earlier numerical experiments. Notice that neither prediction depends on the initial conditions. If we were to use a different initial condition then the coefficients in front of the terms in Equation (9.29) (i.e., 5 and 1) would be changed, but we would still expect that at large times:

$$\mathbf{N}(t) \approx c \cdot (1.6)^t \begin{bmatrix} 20 \\ 1 \end{bmatrix}$$

for some constant c.

We summarize the rules for deriving the stable age distribution before we prove them.

Rate of Growth and Stable Age Distribution for a Leslie Matrix Model.

Assume that the demographic groups within a population have sizes $N_0(t), N_1(t), \ldots, N_m(t)$, and obey a Leslie matrix model

$$\mathbf{N}(t+1) = L\mathbf{N}(t)$$

where

$$\mathbf{N}(t) = \begin{bmatrix} N_0(t) \\ N_1(t) \\ \vdots \\ N_m(t) \end{bmatrix},$$

and L is a $(m+1) \times (m+1)$ matrix. Then L will have $(m+1)$ eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{m+1}$. If λ_1 is the largest eigenvalue, in the sense that $|\lambda_1| > |\lambda_2|$, $|\lambda_1| > |\lambda_3|, \ldots, |\lambda_1| > |\lambda_{m+1}|$, and $\mathbf{u}_1$ is the associated eigenvector, then as t increases, the growth rate of the population will converge to λ_1. If

$$\mathbf{u} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_m \end{bmatrix}$$

then the proportion of organisms in class 0 will converge to $\frac{x_0}{x_0+x_1+\cdots+x_m}$, the proportion in class 1 will converge to $\frac{x_1}{x_0+\cdots+x_m}$, and so on.

If λ_1 is the largest eigenvalue and $0 < \lambda_1 < 1$, then the population size decreases over time. If $\lambda_1 > 1$, then the population size increases over time.

We will now prove these results only for the case where L is a 2×2 matrix. In this case L will have two eigenvalues: λ_1 and λ_2. We will assume that these eigenvalues are different, and that $|\lambda_1| > |\lambda_2|$; i.e., λ_1 is the larger eigenvalue. If the respective eigenvectors are $\mathbf{u}_1$ and $\mathbf{u}_2$, then $\mathbf{u}_1$ and $\mathbf{u}_2$ will be linearly independent. So for any initial condition $\mathbf{N}(0)$ we can find constants a_1 and a_2 such that $\mathbf{N}(0) = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$.

Then at time t the population sizes will be given by:

$$\begin{aligned}\mathbf{N}(t) &= L^t\mathbf{N}(0) = L^t(a_1\mathbf{u}_1 + a_2\mathbf{u}_2)\\ &= a_1L^t\mathbf{u}_1 + a_2L^t\mathbf{u}_2\\ &= a_1\lambda_1^t\mathbf{u} + a_2\lambda_2^t\mathbf{u}_2.\end{aligned}$$

Provided $a_1 \neq 0$, then as $t \to \infty$, the first term dominates over the second (because if $|\lambda_1| > |\lambda_2|$ then $\frac{\lambda_1^t}{\lambda_2^t} = \left(\frac{\lambda_1}{\lambda_2}\right)^t \to \infty$ as $t \to \infty$).

So for large t,

$$\mathbf{N}(t) \approx a_1\lambda_1^t\mathbf{u}_1.$$

Hence if $\mathbf{u}_1 = \begin{bmatrix} x \\ y \end{bmatrix}$ then $N_0(t) \approx a_1\lambda_1^t x$ and $N_1(t) \approx a_1\lambda_1^t y$. So both population classes grow or decay exponentially at a rate λ_1. Moreover, the proportions of organisms in the age 0 class converges to

$$\frac{N_0(t)}{N_0(t) + N_1(t)} \approx \frac{a_1\lambda_1^t x}{a_1\lambda_1^t x + a_1\lambda_1^t y} = \frac{x}{x+y},$$

while the proportion in the age 1 class converges to

$$\frac{N_1(t)}{N_0(t) + N_1(t)} \approx \frac{a_1\lambda_1^t y}{a_1\lambda_1^t x + a_1\lambda_1^t y} = \frac{y}{x+y},$$

and both limits are independent of a_1.

EXAMPLE 4 A seabird colony consists of two classes of birds—immature birds that do not breed and adult birds that do breed. Assume that the number of immature birds is denoted by $N_0(t)$, and the number of mature birds by $N_1(t)$. We model the changes in the sizes of two classes of birds from one year to the next using a Leslie matrix model:

$$\mathbf{N}(t+1) = L\mathbf{N}(t) \quad \text{where} \quad \mathbf{N}(t) = \begin{bmatrix} N_0(t) \\ N_1(t) \end{bmatrix}$$

and

$$L = \begin{bmatrix} 0.5 & 2 \\ 0.3 & 0.9 \end{bmatrix}.$$

Show that the bird population is predicted to grow without bound and find the stable distribution of the two classes as $t \to \infty$.

Solution This question about long-time behavior of the two classes requires us to find the eigenvalues of L. These eigenvalues solve $\det(L - \lambda I) = 0$ where

$$\begin{aligned}\det(L - \lambda I) &= \det\begin{bmatrix} 0.5-\lambda & 2 \\ 0.3 & 0.9-\lambda \end{bmatrix}\\ &= (0.5-\lambda)(0.9-\lambda) - 0.6\\ &= \lambda^2 - 1.4\lambda - 0.15\\ &= (\lambda + 0.1)(\lambda - 1.5).\end{aligned}$$

So the eigenvalues are $\lambda_1 = 1.5$ and $\lambda_2 = -0.1$.

For most initial conditions, the solution will be dominated by the eigenvector associated with λ_1 as $t \to \infty$. To calculate this eigenvector we solve:

$$\begin{bmatrix} 0.5 & 2 \\ 0.3 & 0.9 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1.5\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

or:

$$0.5x_1 + 2x_2 = 1.5x_1$$
$$0.3x_1 + 0.9x_2 = 1.5x_2.$$

so

$$-x_1 + 2x_2 = 0 \qquad (R_1)$$
$$0.3x_1 - 0.6x_2 = 0. \qquad (R_2)$$

Both (R_1) and (R_2) are satisfied if $x_1 = 2x_2$. One eigenvector is therefore $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Thus at large t, we expect:

$$\mathbf{N}(t) \approx c\,(1.5)^t \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

for some constant c. Note that the population in each class increases exponentially with rate of growth of 1.5.

The stable distribution of immature birds is $\frac{x_1}{x_1+x_2} = \frac{2}{3}$, while the stable distribution of mature birds is $\frac{x_2}{x_1+x_2} = \frac{1}{3}$. •

Section 9.4 Problems

9.4.1

In Problems 1–4, suppose that breeding occurs once a year and that a census is taken at the end of each breeding season.

1. Assume that a population is divided into three age classes and that 20% of the females age 0 and 70% of the females age 1 present at time t survive to time $t+1$. Assume further that females age 1 have an average of 2.4 female offspring and females age 2 have an average of 1.3 female offspring. If, at time 0, the population consists of 2000 females age 0, 800 females age 1, and 200 females age 2, find the Leslie matrix and the number of females in each age class at time 2.

2. Assume that a population is divided into three age classes and that 80% of the females age 0 and 10% of the females age 1 present at time t survive until time $t+1$. Assume further that females age 1 have an average of 1.6 female offspring and females age 2 have an average of 3.9 female offspring. If, at time 0, the population consists of 1000 females age 0, 100 females age 1, and 20 females age 2, find the Leslie matrix and the number of females in each age class at time 3.

3. A population is divided into four age classes. 70% of the females age 0, 50% of the females age 1, and 10% of the females age 2 present at time t survive until time $t+1$. Assume that females age 2 have an average of 4.6 female offspring and females age 3 have an average of 3.7 female offspring. If, at time 0, the population consists of 1500 females age 0, 500 females age 1, 250 females age 2, and 50 females age 3, find the Leslie matrix and the number of females in each age class at time 3.

4. A population is divided into four age classes. 65% of the females age 0, 40% of the females age 1, and 30% of the females age 2 present at time t survive until time $t+1$. Assume that females age 1 have an average of 2.8 female offspring, females age 2 have an average of 7.6 female offspring, and females age 3 have an average of 2.4 female offspring. If, at time 0, the population consists of 1000 females age 0, 500 females age 1, 200 females age 2, and 50 females age 3, find the Leslie matrix and the number of females in each age class at time 3.

In Problems 5–6, assume the given Leslie matrix L. Determine the number of age classes in the population, the fraction of one-year-olds present at time t that survive to time t + 1, and the average number of female offspring of a two-year-old female.

5. $L = \begin{bmatrix} 2 & 3 & 3 & 1 \\ 0.4 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.8 & 0 \end{bmatrix}$

6. $L = \begin{bmatrix} 0 & 5 & 0 \\ 0.8 & 0 & 0 \\ 0 & 0.4 & 0 \end{bmatrix}$

In Problems 7–8, assume the given Leslie matrix L. Determine the number of age classes in the population. What fraction of two-year-olds present at time t survive until time t + 1. Determine the average number of female offspring of a one-year-old female.

7. $L = \begin{bmatrix} 0 & 2.5 & 4 & 1.5 \\ 0.9 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.4 & 0 \end{bmatrix}$

8. $L = \begin{bmatrix} 0 & 4.2 & 3.7 \\ 0.7 & 0 & 0 \\ 0 & 0.1 & 0 \end{bmatrix}$

9. Assume that the Leslie matrix is

$$L = \begin{bmatrix} 0.5 & 1.5 \\ 1 & 0 \end{bmatrix}$$

Suppose that, at time $t = 0$, $N_0(0) = 100$ and $N_1(0) = 0$. Find the population vectors for $t = 0, 1, 2, \ldots, 10$. Compute the successive ratios

$$q_0(t) = \frac{N_0(t)}{N_0(t-1)} \quad \text{and} \quad q_1(t) = \frac{N_1(t)}{N_1(t-1)}$$

for $t = 1, 2, \ldots, 10$. What value do $q_0(t)$ and $q_1(t)$ approach as $t \to \infty$? (Take a guess.) Compute the fraction of females age 0 for $t = 0, 1, \ldots, 10$. Can you find a stable age distribution?

10. Assume that the Leslie matrix is

$$L = \begin{bmatrix} 0.2 & 3 \\ 0.33 & 0 \end{bmatrix}$$

Suppose that, at time $t = 0$, $N_0(0) = 10$ and $N_1(0) = 5$. Find the population vectors for $t = 0, 1, 2, \ldots, 10$. Compute the successive ratios

$$q_0(t) = \frac{N_0(t)}{N_0(t-1)} \quad \text{and} \quad q_1(t) = \frac{N_1(t)}{N_1(t-1)}$$

for $t = 1, 2, \ldots, 10$. What value do $q_0(t)$ and $q_1(t)$ approach as $t \to \infty$? (Take a guess.) Compute the fraction of females age 0 for $t = 0, 1, \ldots, 10$. Can you find a stable age distribution?

11. Assume that the Leslie matrix is

$$L = \begin{bmatrix} 0 & 2 \\ 0.6 & 0 \end{bmatrix}$$

Suppose that, at time $t = 0$, $N_0(0) = 5$ and $N_1(0) = 1$. Find the population vectors for $t = 0, 1, 2, \ldots, 10$. Compute the successive ratios

$$q_0(t) = \frac{N_0(t)}{N_0(t-1)} \quad \text{and} \quad q_1(t) = \frac{N_1(t)}{N_1(t-1)}$$

for $t = 1, 2, \ldots, 10$. Do $q_0(t)$ and $q_1(t)$ converge? Compute the fraction of females age 0 for $t = 0, 1, \ldots, 10$. Describe the long-term behavior of $q_0(t)$.

12. Assume that the Leslie matrix is

$$L = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}$$

Suppose that, at time $t = 0$, $N_0(0) = 1$ and $N_1(0) = 1$. Find the population vectors for $t = 0, 1, 2, \ldots, 10$. Compute the successive ratios

$$q_0(t) = \frac{N_0(t)}{N_0(t-1)} \quad \text{and} \quad q_1(t) = \frac{N_1(t)}{N_1(t-1)}$$

for $t = 1, 2, \ldots, 10$. Do $q_0(t)$ and $q_1(t)$ converge? Compute the fraction of females age 0 for $t = 0, 1, \ldots, 10$. Describe the long-term behavior of $q_0(t)$.

9.4.2

13. Suppose that

$$L = \begin{bmatrix} 2 & 4 \\ 0.3 & 0 \end{bmatrix}$$

is the Leslie matrix for a population with two age classes.

(a) Determine both eigenvalues.

(b) Give a biological interpretation of the larger eigenvalue.

(c) Find the stable age distribution.

14. Suppose that

$$L = \begin{bmatrix} 4 & 2 \\ 1 & 0.5 \end{bmatrix}$$

is the Leslie matrix for a population with two age classes.

(a) Determine both eigenvalues.

(b) Give a biological interpretation of the larger eigenvalue.

(c) Find the stable age distribution.

15. Suppose that

$$L = \begin{bmatrix} 3 & 2 \\ 1.5 & 1 \end{bmatrix}$$

is the Leslie matrix for a population with two age classes.

(a) Determine both eigenvalues.

(b) Give a biological interpretation of the larger eigenvalue.

(c) Find the stable age distribution.

16. Suppose that

$$L = \begin{bmatrix} 0 & 5 \\ 0.9 & 0 \end{bmatrix}$$

is the Leslie matrix for a population with two age classes.

(a) Determine both eigenvalues.

(b) Give a biological interpretation of the larger eigenvalue.

(c) Find the stable age distribution.

17. Suppose that

$$L = \begin{bmatrix} 0 & 5 \\ 0.09 & 0 \end{bmatrix}$$

is the Leslie matrix for a population with two age classes.

(a) Determine both eigenvalues.

(b) Give a biological interpretation of the larger eigenvalue.

(c) Find the stable age distribution.

9.5 Analytic Geometry

Analytic geometry combines techniques from algebra and geometry and provides important tools for multidimensional calculus, which will be introduced in Chapter 10. To get to this material we will therefore need a few results from analytic geometry. Our first task will be to generalize points and vectors in the plane to higher dimensions. We will then introduce a product between vectors that will allow us to determine the length of a vector and the angle between two vectors. Finally, we will give a vector representation of lines and planes in three-dimensional space.

9.5.1 Points and Vectors in Higher Dimensions

To represent points in a plane, we use a Cartesian coordinate system that consists of two axes—the x_1-axis and the x_2-axis—that are perpendicular to each other. Any point in the plane can be represented by an ordered pair (a_1, a_2) of real numbers, where a_1

is the x_1-coordinate and a_2 is the x_2-coordinate. Since we need two numbers to locate such a point, we call the plane "two dimensional."

The plane can thus be thought of as the set of all points (x_1, x_2) with $x_1 \in \mathbf{R}$ and $x_2 \in \mathbf{R}$. We introduce the notation $\mathbf{R}^2$ to denote the set of these points. The two-dimensional plane can be described as

$$\mathbf{R}^2 = \{(x_1, x_2) : x_1 \in \mathbf{R}, x_2 \in \mathbf{R}\}$$

To generalize this equation to n dimensions, we set

$$\mathbf{R}^n = \{(x_1, x_2, \ldots, x_n) : x_1 \in \mathbf{R}, x_2 \in \mathbf{R}, \ldots, x_n \in \mathbf{R}\}$$

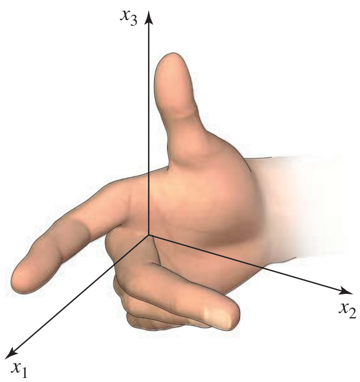


Figure 9.28 A right-handed three-dimensional coordinate system. The axes are perpendicular.

For instance, $\mathbf{R}^3$ is three-dimensional space; it consists of all points (x_1, x_2, x_3) with $x_i \in \mathbf{R}$ for $i = 1, 2,$ and 3. To represent points and vectors in $\mathbf{R}^3$, we use a coordinate system that consists of three mutually perpendicular axes. This coordinate system is shown in Figure 9.28. The axes need to be oriented in a "right-handed" manner. That is, the axes are perpendicular to each other and oriented so that the index finger of your right hand points along the positive x_1-axis, the middle finger of your right hand points along the positive x_2-axis, and the thumb of your right hand points along the positive x_3-axis.

In four and higher dimensions, we can no longer draw a coordinate system, although we can still represent such systems algebraically and work with them.

We introduced vectors in two dimensions in Section 9.3. A vector is a quantity that has a direction and a magnitude. A vector in two dimensions is an ordered pair that can be represented by a directed segment, as illustrated in Figure 9.29. In the figure, the vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is represented by a directed segment with initial point $(0, 0)$ and endpoint (x_1, x_2). The arrow at the tip indicates the direction of the vector. We will now generalize this representation to n dimensions.

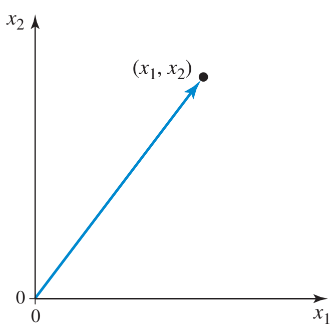


Figure 9.29 The vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ in the x_1–x_2 plane.

Definition A vector in n-dimensional space is an ordered n-tuple

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

of real numbers. The numbers $x_1, x_2, \ldots, x_n$ are called the **components** of $\mathbf{x}$.

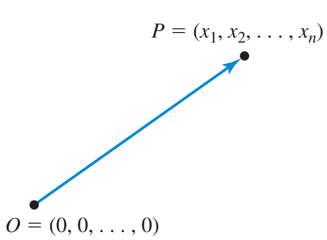


Figure 9.30 A vector in n-dimensional space.

Vectors in n-dimensional space also may be represented by directed segments with initial point $(0, 0, \ldots, 0)$ and endpoint $(x_1, x_2, \ldots, x_n)$. In three dimensions, we can use the three-dimensional right-handed Cartesian coordinate system to accurately visualize vectors. In four and higher dimensions, we can no longer draw the coordinate system. Instead, to represent a vector $\mathbf{x}$ whose endpoint has coordinates $(x_1, x_2, \ldots, x_n)$, we draw a directed arrow from the origin $(0, 0, \ldots, 0)$ to the point $P = (x_1, x_2, \ldots, x_n)$ (Figure 9.30).

Next, we generalize vector addition to n dimensions. In the previous section, we saw how to add vectors in two-dimensional space. In n dimensions, vector addition is defined in a similar way.

Vector Addition If

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

then

$$\mathbf{x}+\mathbf{y}=\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix}$$

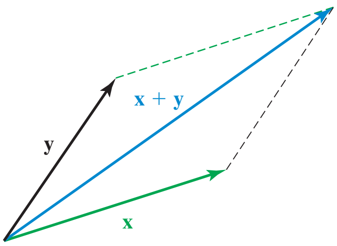


Figure 9.31 The parallelogram law for vector addition.

The geometric interpretation is the same as in two-dimensional space (although, again, visualization is impossible). This law of addition is similarly called the parallelogram law, because if $\mathbf{x}$ and $\mathbf{y}$ form two edges of a parallelogram then $\mathbf{x}+\mathbf{y}$ forms the diagonal of the parallelogram (see Figure 9.31).

Alternatively, vector addition can also be interpreted in the following way: To obtain $\mathbf{x}+\mathbf{y}$, we place the vector $\mathbf{y}$ at the tip of the vector $\mathbf{x}$. The sum $\mathbf{x}+\mathbf{y}$ is then the vector that starts at the same point as $\mathbf{x}$ and ends at the point where the moved vector $\mathbf{y}$ ends, as illustrated in Figure 9.32.

The multiplication of a vector by a scalar generalizes to higher dimensions as well and has the same geometric interpretation.

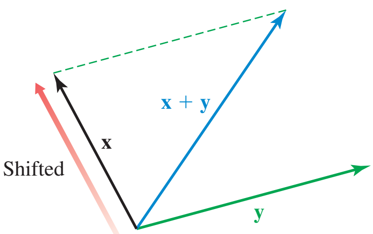


Figure 9.32 Adding two vectors by shifting y to lie at the tip of x.

Multiplication of a Vector by a Scalar If a is a scalar and $\mathbf{x}$ is a vector in n-dimensional space, then

$$a\mathbf{x}=a\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=\begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{bmatrix}$$

Vector Representation A vector $\mathbf{x}$ is a directed line segment $\overrightarrow{AB}$ from the initial point A to the terminal point B. A **vector representation** of this line segment has the origin at its initial point and the same direction and length as the directed segment $\overrightarrow{AB}$.

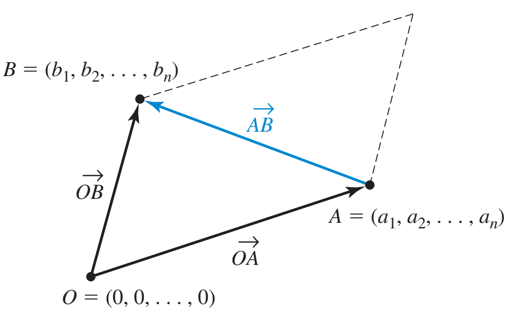


Figure 9.33 The vector representation of a directed segment.

We can use Figure 9.33 to find the vector representation of a directed segment $\overrightarrow{AB}$ from point A with coordinates $(a_1, a_2, \ldots, a_n)$ to point B with coordinates $(b_1, b_2, \ldots, b_n)$. Using the parallelogram law, we see that

$$\overrightarrow{OA}+\overrightarrow{AB}=\overrightarrow{OB}$$

Solving for $\overrightarrow{AB}$ yields

$$\overrightarrow{AB}=\overrightarrow{OB}-\overrightarrow{OA}$$

Now, $\overrightarrow{OB}$ has the vector representation $[b_1, b_2, \ldots, b_n]'$ and $\overrightarrow{OA}$ has the vector representation $[a_1, a_2, \ldots, a_n]'$. The difference $\overrightarrow{OB}-\overrightarrow{OA}$ is then the mathematical difference of the two vectors; that is,

$$\overrightarrow{OB}-\overrightarrow{OA}=\begin{bmatrix} b_1-a_1 \\ b_2-a_2 \\ \vdots \\ b_n-a_n \end{bmatrix}$$

The vector representing the line segment $\overrightarrow{AB}$ is thus given by

$$\overrightarrow{AB}=\begin{bmatrix} b_1-a_1 \\ b_2-a_2 \\ \vdots \\ b_n-a_n \end{bmatrix}$$

EXAMPLE 1 Find the vector representation of $\overrightarrow{AB}$ when $A = (2, -1)$ and $B = (1, 3)$.

Solution The vector representation of $\overrightarrow{AB}$ is given by

$$\overrightarrow{AB} = \begin{bmatrix} b_1 - a_1 \\ b_2 - a_2 \end{bmatrix} = \begin{bmatrix} 1 - 2 \\ 3 - (-1) \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$$

We illustrate this equation graphically in Figure 9.34.

We see from Figure 9.34 that if we shift the vector $\overrightarrow{AB}$ to the origin, its tip ends at $(-1, 4)$, which confirms that the vector representation of $\overrightarrow{AB}$ is $\left[\begin{smallmatrix} -1 \\ 4 \end{smallmatrix}\right]$.

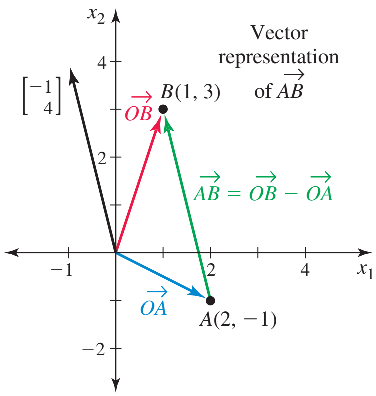


Figure 9.34 The vector representation of the vector $\overrightarrow{AB}$.

Length of a Vector The length of a vector in two dimensions is computed from the Pythagorean theorem. If $\mathbf{x} = \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]$, then the length of $\mathbf{x}$ is denoted by $|\mathbf{x}|$, and we find that

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2}$$

as illustrated in Figure 9.35. We can generalize this equation to n dimensions as follows: Recall that the transpose of a vector is denoted with the prime symbol; that is,

$$[x_1, x_2, \ldots, x_n]' = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

This representation allows us to write column vectors as the transposes of the row vectors—a convenient notation when one is writing large column vectors.

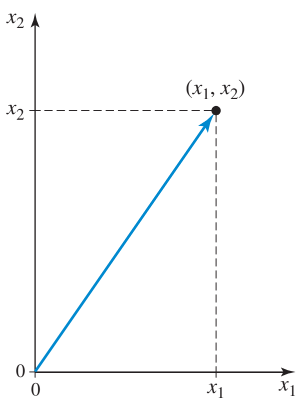


Figure 9.35 The length of the vector $\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]$.

> **Length of a Vector** The length of a vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]'$ is
>
> $$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

EXAMPLE 2 Find the length of

$$\mathbf{x} = \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}$$

Solution The length of $\mathbf{x}$ is given by

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{(1)^2 + (-3)^2 + (4)^2} = \sqrt{26}$$

If we know the length of a vector $\mathbf{x}$, we can **normalize** $\mathbf{x}$ to obtain a vector of length 1 in the same direction as $\mathbf{x}$. (See Figure 9.36.) We call such a vector a **unit vector** in the direction of $\mathbf{x}$ and denote it by $\hat{\mathbf{x}}$; that is,

$$\hat{\mathbf{x}} = \frac{\mathbf{x}}{|\mathbf{x}|}$$

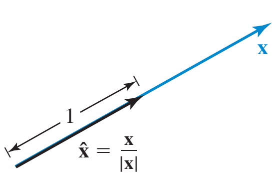


Figure 9.36 The normalized vector $\hat{\mathbf{x}}$ has the same direction as $\mathbf{x}$; its length is 1.

Of course, $|\hat{\mathbf{x}}| = 1$. We summarize all this as follows:

> $\dfrac{\mathbf{x}}{|\mathbf{x}|}$ is a vector of length 1 in the direction of $\mathbf{x}$.

EXAMPLE 3 Normalize the vector

$$\mathbf{x} = \begin{bmatrix} 3 \\ -6 \\ 6 \end{bmatrix}$$

Solution We must first find the length of $\mathbf{x}$:

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{(3)^2 + (-6)^2 + (6)^2} = \sqrt{81} = 9$$

The unit vector $\hat{\mathbf{x}}$ is then given by

$$\hat{\mathbf{x}} = \frac{\mathbf{x}}{|\mathbf{x}|} = \frac{1}{9}\begin{bmatrix} 3 \\ -6 \\ 6 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$$

We may check that $\hat{\mathbf{x}}$ is indeed a vector of length 1:

$$\hat{\mathbf{x}} = \sqrt{(1/3)^2 + (-2/3)^2 + (2/3)^2} = 1$$

9.5.2 The Dot Product

The dot product of two vectors allows us to determine the angle between the vectors.

Definition The **scalar product**, or **dot product**, of two vectors $\mathbf{x} = [x_1, x_2, \ldots, x_n]'$ and $\mathbf{y} = [y_1, y_2, \ldots, y_n]'$ is the number

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}'\mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i$$

Using Σ-notation.

Note that the dot product of two vectors is a scalar, i.e. a number, rather than a vector (hence the name "scalar product"). It is also called "dot product" because the notation uses a dot between $\mathbf{x}$ and $\mathbf{y}$.

EXAMPLE 4 Find the dot product of

$$\mathbf{x} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

Solution Using the definition of the dot product, we find that

$$\mathbf{x} \cdot \mathbf{y} = \begin{bmatrix} 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = -2 + 6 + 0 = 4$$

Properties of the Dot Product

1. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$
2. $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$

We can use the dot product to express the length of a vector. Recall that in n dimensions the length of a vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]'$ is defined as

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

If we compute the dot product of $\mathbf{x}$ with itself, we obtain

$$\mathbf{x} \cdot \mathbf{x} = \sum_{i=1}^{n} x_i^2$$

Comparing $|\mathbf{x}|$ and $\mathbf{x}\cdot\mathbf{x}$, we see that the following relationship holds:

$$\boxed{|\mathbf{x}|^2 = \mathbf{x}\cdot\mathbf{x}}$$

The length of a vector is therefore $|\mathbf{x}| = \sqrt{\mathbf{x}\cdot\mathbf{x}}$.

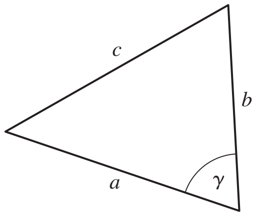


Figure 9.37 The law of cosines: $c^2 = a^2 + b^2 - 2ab\cos\gamma$.

The Angle between Two Vectors The other important application of the dot product is that it allows us to find the angle between two vectors. To derive this result, we need the trigonometric law of cosines, as illustrated in Figure 9.37.

Let $\mathbf{x}$ and $\mathbf{y}$ be two nonzero vectors whose initial points coincide. Then $\mathbf{x}-\mathbf{y}$ is a vector that connects the endpoint of $\mathbf{y}$ to the endpoint of $\mathbf{x}$, as illustrated in Figure 9.38.

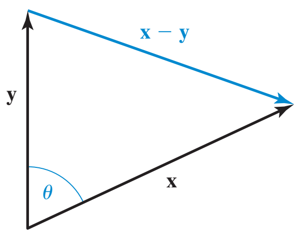


Figure 9.38 The angle between two vectors.

Using the law of cosines, we find that

$$|\mathbf{x}-\mathbf{y}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 - 2|\mathbf{x}||\mathbf{y}|\cos\theta$$

where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$. (See Figure 9.38.) Alternatively, the length of the vector $\mathbf{x}-\mathbf{y}$, denoted $|\mathbf{x}-\mathbf{y}|$, can be computed with the dot product:

$$\begin{aligned}|\mathbf{x}-\mathbf{y}|^2 &= (\mathbf{x}-\mathbf{y})\cdot(\mathbf{x}-\mathbf{y})\\ &= \mathbf{x}\cdot\mathbf{x} - \mathbf{x}\cdot\mathbf{y} - \mathbf{y}\cdot\mathbf{x} + \mathbf{y}\cdot\mathbf{y}\\ &= |\mathbf{x}|^2 - 2\mathbf{x}\cdot\mathbf{y} + |\mathbf{y}|^2\end{aligned}$$

Setting the two expressions for $|\mathbf{x}-\mathbf{y}|^2$ equal to each other gives

$$|\mathbf{x}|^2 + |\mathbf{y}|^2 - 2|\mathbf{x}||\mathbf{y}|\cos\theta = |\mathbf{x}|^2 - 2\mathbf{x}\cdot\mathbf{y} + |\mathbf{y}|^2$$

Solving this equation for $\mathbf{x}\cdot\mathbf{y}$, we find the following:

$$\boxed{\mathbf{x}\cdot\mathbf{y} = |\mathbf{x}||\mathbf{y}|\cos\theta \qquad (9.30)}$$

The significance of equation (9.30) is that it allows us to find the angle between the two nonzero vectors $\mathbf{x}$ and $\mathbf{y}$, because

$$\theta = \cos^{-1}\left(\frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}||\mathbf{y}|}\right)$$

Note that $\theta \in [0,\pi)$, since the interval $[0,\pi)$ is used to find the inverse of the cosine function.

EXAMPLE 5 Find the angle between

$$\mathbf{x} = \begin{bmatrix}2\\1\end{bmatrix} \quad\text{and}\quad \mathbf{y} = \begin{bmatrix}1\\1\end{bmatrix}.$$

Solution To determine the angle θ between $\mathbf{x}$ and $\mathbf{y}$, we use

$$\mathbf{x}\cdot\mathbf{y} = |\mathbf{x}||\mathbf{y}|\cos\theta$$

We find that

$$\begin{aligned}\mathbf{x}\cdot\mathbf{y} &= \begin{bmatrix}2\\1\end{bmatrix}\cdot\begin{bmatrix}1\\1\end{bmatrix} = 2+1 = 3\\ |\mathbf{x}| &= \sqrt{(2)^2+(1)^2} = \sqrt{5}\\ |\mathbf{y}| &= \sqrt{(1)^2+(1)^2} = \sqrt{2}\end{aligned}$$

Hence,

$$\cos\theta = \frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}||\mathbf{y}|} = \frac{3}{\sqrt{5}\sqrt{2}} = \frac{3}{\sqrt{10}}$$

and therefore,

$$\theta = \cos^{-1}\frac{3}{\sqrt{10}} \approx 18.4^\circ \quad \text{or} \quad 0.3218 \text{ radians}$$

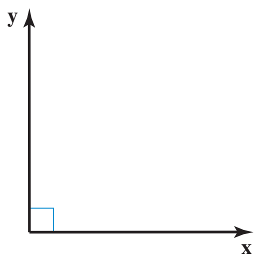


Figure 9.39 The vectors $\mathbf{x}$ and $\mathbf{y}$ are perpendicular.

We wish to single out the case in which $\theta = \pi/2$. We say that two vectors are **perpendicular** to each other if the angle between them is $\pi/2$; this situation is illustrated in Figure 9.39.

An important consequence of (9.30) is that it gives us a criterion with which to determine whether two vectors are perpendicular. Since $\cos(\pi/2) = 0$, we have the following theorem:

Theorem

$\mathbf{x}$ and $\mathbf{y}$ are perpendicular if $\mathbf{x} \cdot \mathbf{y} = 0$.

We will now give two examples that illustrate how to use this theorem.

EXAMPLE 6 Let $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Find $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ so that $\mathbf{x}$ and $\mathbf{y}$ are perpendicular.

Solution The vectors $\mathbf{x}$ and $\mathbf{y}$ are perpendicular if $\mathbf{x} \cdot \mathbf{y} = 0$. We find that

$$\mathbf{x} \cdot \mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = y_1 + 2y_2$$

We set the right-hand side equal to 0:

$$y_1 + 2y_2 = 0$$

Any choice of numbers (y_1, y_2) that satisfies this equation would thus give us a vector that is perpendicular to $\mathbf{x}$. For instance, if we choose $y_2 = 1$ and $y_1 = -2$, then

$$\mathbf{y} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

is perpendicular to $\mathbf{x}$.

EXAMPLE 7 Show that the coordinate axes in a two-dimensional Cartesian coordinate system are perpendicular.

Solution A two-dimensional Cartesian coordinate system is illustrated in Figure 9.40. We see that the x-axis can be represented by the vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and the y-axis by the vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Computing the dot product between these two vectors, we find that

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = (1)(0) + (0)(1) = 0$$

and we conclude that the two vectors are perpendicular. Therefore, the x-axis and the y-axis are perpendicular.

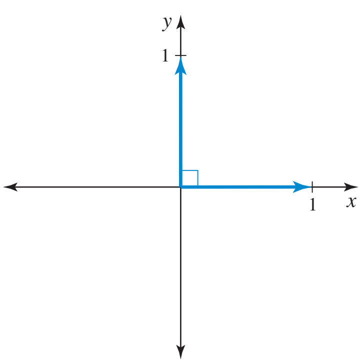


Figure 9.40 The coordinate axes in a two-dimensional Cartesian coordinate system are perpendicular.

We will now use the dot product and the result that the dot product between perpendicular vectors is zero to obtain the equation of a line in two-dimensional space and the equation of a plane in three-dimensional space.

Lines in the Plane We will use the dot product to write equations of lines in $\mathbf{R}^2$. Suppose that we wish to find the solution of the equation of a line through the point (x_0, y_0) and perpendicular to the vector $\mathbf{n} = \begin{bmatrix} a \\ b \end{bmatrix}$. If $r = (x, y)$ is any other point of the line, then the vector $\mathbf{r} - \mathbf{r}_0$ is perpendicular to $\mathbf{n}$, as illustrated in Figure 9.41.

Therefore,

$$\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$$

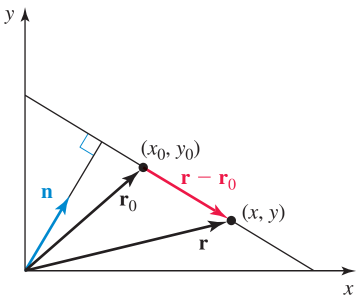


Figure 9.41 The vector $\mathbf{r} - \mathbf{r}_0$ is perpendicular to $\mathbf{n}$, for any point on the line.

This equation is called the **vector equation** of a line in the plane.

To obtain the **scalar equation** of this line, we set

$$\mathbf{n} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \text{and} \quad \mathbf{r}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

Then

$$\begin{aligned} \mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) &= \begin{bmatrix} a \\ b \end{bmatrix} \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\ &= a(x - x_0) + b(y - y_0) = 0 \end{aligned}$$

That is, we obtain the following result:

Scalar Equation of a Line in $\mathbf{R}^2$

The line through (x_0, y_0) and perpendicular to $\begin{bmatrix} a \\ b \end{bmatrix}$ has the equation

$$a(x - x_0) + b(y - y_0) = 0 \qquad (9.31)$$

Equation (9.31) should be thought of as an equation for (x, y). Only points (x, y) that solve (9.31) will lie on the line.

EXAMPLE 8 Find an equation of the line through (4, 3) and perpendicular to $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Solution Using (9.31), we find that

$$(1)(x - 4) + (2)(y - 3) = 0$$

Simplifying yields

$$x + 2y = 10.$$

Planes in $\mathbf{R}^3$ We can define a plane if we know a point P in the plane, and the normal vector to the plane, $\mathbf{n}$, as illustrated in Figure 9.42. If P is the endpoint of $\mathbf{r}_0$, then a point x lies in the plane if and only if $\mathbf{r} - \mathbf{r}_0$ is perpendicular to $\mathbf{n}$.

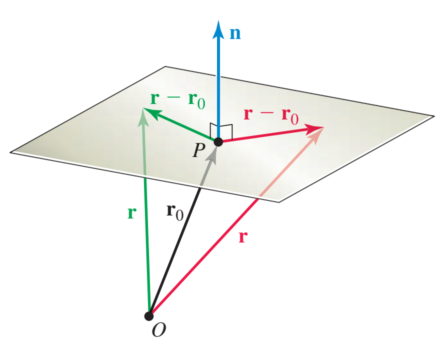


Figure 9.42 The vector $\mathbf{r} - \mathbf{r}_0$ is perpendicular to $\mathbf{n}$ for any point $\mathbf{r}$ located in the plane perpendicular to $\mathbf{n}$.

$$\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$$

This equation is called the **vector equation of a plane**. Note that, written in vector form, the equation looks identical to the vector equation of a line in $\mathbf{R}^2$. The difference is that the vectors are now in $\mathbf{R}^3$ and not in $\mathbf{R}^2$. Alternatively, we can rewrite the equation in a form that does not involve vectors. If we set

$$\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \text{and} \quad \mathbf{r}_0 = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$$

Then evaluating the dot product yields

> **Scalar Equation for a Plane in $\mathbf{R}^3$**
> The plane through (x_0, y_0, z_0) and perpendicular to $\left[a,\ b,\ c\right]'$ has the equation
> $$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \tag{9.32}$$

EXAMPLE 9 Find the equation of a plane in three-dimensional space through (2, 0, 3) and perpendicular to $\left[-1,\ 4,\ 1\right]'$.

Solution Using (9.32), we find that

$$(-1)(x - 2) + (4)(y - 0) + (1)(z - 3) = 0$$

Simplifying yields

$$-x + 4y + z = 1$$

9.5.3 Parametric Equations of Lines

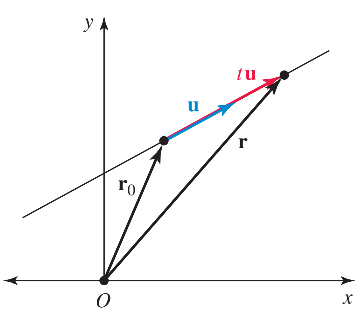


Figure 9.43 The parametric equation of a line.

There is another way to write lines that can be used in $\mathbf{R}^2$ and in $\mathbf{R}^3$. Looking at Figure 9.43, we can define the line shown there by a single point $\mathbf{r}_0$ on the line, and the vector $\mathbf{u}$ that points in the direction of the line: Any point $\mathbf{r}$ on the line can be thought of as the endpoint of the sum of the vector $\mathbf{r}_0$ and the vector $\mathbf{r} - \mathbf{r}_0$. Now, $\mathbf{r} - \mathbf{r}_0$ is a multiple of the vector $\mathbf{u}$. Thus, for the equation of the line in vector form, we have

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{u} \tag{9.33}$$

for some $t \in \mathbf{R}$. If $\mathbf{r}_0$ has coordinates (x_0, y_0), $\mathbf{r}$ has coordinates (x, y), and $\mathbf{u} = \left[\begin{smallmatrix} u_1 \\ u_2 \end{smallmatrix}\right]$, then

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + t \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{9.34}$$

for some $t \in \mathbf{R}$. By varying t, we can reach *any* point on the line. Equation (9.33) [or (9.34)] is called a **parametric equation**, in vector form, of a line and t is called a **parameter**. We can also write the equation of the line in parametric form for each coordinate separately:

$$x = x_0 + tu_1$$
$$y = y_0 + tu_2$$

for $t \in \mathbf{R}$.

EXAMPLE 10 Find the parametric equation of the line in the x–y plane that goes through the point (2, 1) in the direction of $\left[\begin{smallmatrix} -1 \\ -3 \end{smallmatrix}\right]$.

Solution We find that

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + t \begin{bmatrix} -1 \\ -3 \end{bmatrix}, \quad t \in \mathbf{R}$$

or $x = 2 - t$ and $y = 1 - 3t$, for $t \in \mathbf{R}$. By eliminating t, we can write the equation in the familiar standard form of a line in the x–y plane, namely, $t = 2 - x$. Therefore,

$$y = 1 - 3(2 - x) = 1 - 6 + 3x$$

and

$$3x - y - 5 = 0$$

is the standard form of the equation of this line.

EXAMPLE 11 Find the parametric equation of the line in the x–y plane that goes through the points $(-1, 2)$ and $(3, 5)$.

Solution We designate one of the two points, say, $(-1, 2)$, as the point $\mathbf{r}_0$ and let $\mathbf{u}$ be the vector that starts at $(-1, 2)$ and ends at $(3, 5)$. Then

$$\mathbf{u} = \begin{bmatrix} 3-(-1) \\ 5-2 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

and the parametric equation of this line is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} + t\begin{bmatrix} 4 \\ 3 \end{bmatrix}, \quad t \in \mathbf{R}$$

•

In Example 10, we saw that, by eliminating t, we can obtain the standard form of a linear equation. We can also go from the standard form to the parametric form by introducing a parameter t.

EXAMPLE 12 Find a parametric form of the line that is given in standard form by the equation:

$$2x - 3y + 1 = 0$$

Solution The easiest way to parameterize this equation is to set $x = t$, solve the equation for y, and substitute t for x:

$$3y = 2x + 1 \quad \text{is equivalent to} \quad y = \frac{2}{3}x + \frac{1}{3}$$

With $x = t$, we can write the parametric equation as

$$\begin{aligned} x &= t \\ y &= \frac{2}{3}t + \frac{1}{3} \end{aligned}$$

for $t \in \mathbf{R}$.

This is by no means the only way to parameterize the line. For example, if we had chosen $t = \frac{1}{3}(x-1)$, we would have found that

$$\begin{aligned} x &= 3t + 1 \\ y &= \frac{2}{3}(3t+1) + \frac{1}{3} = 2t + 1. \end{aligned}$$

Although these equations look very different, they define the same line. •

The vector representation of the line can be used in higher dimensions. For instance, a line in $\mathbf{R}^3$ that goes through a point $\mathbf{r}_0 = (x_0, y_0, z_0)$ in the direction of $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ would have the form

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + t\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \quad t \in \mathbf{R}$$

or, if we write out the coordinates separately,

$$\begin{aligned} x &= x_0 + tu_1 \\ y &= y_0 + tu_2 \\ z &= z_0 + tu_3 \end{aligned}$$

for $t \in \mathbf{R}$.

EXAMPLE 13 Find the parametric equation of the line in x–y–z space that goes through the points $(1, -1, 3)$ and $(2, 4, -1)$.

Solution We designate one point as $\mathbf{r}_0$, say $(2, 4, -1)$, and let $\mathbf{u}$ be the vector connecting the two points (it does not matter which of the two points we select as the starting point):

$$\mathbf{u} = \begin{bmatrix} 2-1 \\ 4-(-1) \\ -1-3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ -4 \end{bmatrix}$$

Then the parametric equation of this line is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ -1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 5 \\ -4 \end{bmatrix}, \quad t \in \mathbf{R}$$

or

$$x = 2 + t$$
$$y = 4 + 5t$$
$$z = -1 - 4t$$

for $t \in \mathbf{R}$.

Eliminating t in the parametric equation for a line in $\mathbf{R}^3$ is not very useful, since doing so does not yield just one equation as in the case of a line in $\mathbf{R}^2$. We will therefore forgo eliminating t in $\mathbf{R}^3$.

Section 9.5 Problems

9.5.1

1. Let $\mathbf{x} = [1, 0, -1]'$ and $\mathbf{y} = [-2, 1, 0]'$.

(a) Find $\mathbf{x} + \mathbf{y}$. **(b)** Find $2\mathbf{x}$. **(c)** Find $-3\mathbf{y}$.

2. Let $\mathbf{x} = [-4, 3, 2]'$ and $\mathbf{y} = [0, -2, 3]'$.

(a) Find $\mathbf{x} - \mathbf{y}$. **(b)** Find $2\mathbf{x} + 3\mathbf{y}$. **(c)** Find $-\mathbf{x} - 2\mathbf{y}$.

3. Let $A = (2, 3)$ and $B = (1, 1)$. Find the vector representation of $\overrightarrow{AB}$.

4. Let $A = (-1, 0)$ and $B = (2, -3)$. Find the vector representation of $\overrightarrow{AB}$.

5. Let $A = (0, 1, -3)$ and $B = (-1, 1, 2)$. Find the vector representation of $\overrightarrow{AB}$.

6. Let $A = (1, 3, -2)$ and $B = (0, -1, -1)$. Find the vector representation of $\overrightarrow{AB}$.

7. Find the length of $\mathbf{x} = [2, 2]'$.

8. Find the length of $\mathbf{x} = [-2, 7]'$.

9. Find the length of $\mathbf{x} = [0, 1, 5]'$.

10. Find the length of $\mathbf{x} = [2, 3, -1]'$.

11. Normalize $[1, 3, -1]'$. **12.** Normalize $[2, 0, -4]'$.

13. Normalize $[0, 6, 0]'$. **14.** Normalize $[-3, 0, 1, 3]'$.

9.5.2

15. Find the dot product of $\mathbf{x} = [-1, 2]'$ and $\mathbf{y} = [-3, -4]'$.

16. Find the dot product of $\mathbf{x} = [1, 2]'$ and $\mathbf{y} = [3, -1]'$.

17. Find the dot product of $\mathbf{x} = [0, -1, 3]'$ and $\mathbf{y} = [-3, 0, 1]'$.

18. Find the dot product of $\mathbf{x} = [2, -3, 1]'$ and $\mathbf{y} = [3, 1, -2]'$.

19. Use the dot product to compute the length of $[0, -1, 2]'$.

20. Use the dot product to compute the length of $[1, 1, 3]'$.

21. Use the dot product to compute the length of $[1, 2, 3, 4]'$.

22. Use the dot product to compute the length of $[1, 2, 3, 0]'$.

23. Find the angle between $\mathbf{x} = [3, 1]'$ and $\mathbf{y} = [3, -1]'$.

24. Find the angle between $\mathbf{x} = [-1, 2]'$ and $\mathbf{y} = [-2, 4]'$.

25. Find the angle between $\mathbf{x} = [0, -1, 3]'$ and $\mathbf{y} = [-3, 1, 1]'$.

26. Find the angle between $\mathbf{x} = [1, -3, 2]'$ and $\mathbf{y} = [3, 1, -4]'$.

27. Let $\mathbf{x} = [1, 1]'$. Find $\mathbf{y}$ so that $\mathbf{x}$ and $\mathbf{y}$ are perpendicular.

28. Let $\mathbf{x} = [-2, 1]'$. Find $\mathbf{y}$ so that $\mathbf{x}$ and $\mathbf{y}$ are perpendicular.

29. Let $\mathbf{x} = [3, -2, 1]'$. Find any vector $\mathbf{y}$ so that $\mathbf{x}$ and $\mathbf{y}$ are perpendicular. [Your solution will not be unique.]

30. Let $\mathbf{x} = [2, 0, 1]'$. Find $\mathbf{y}$ so that $\mathbf{x}$ and $\mathbf{y}$ are perpendicular.

31. A triangle has vertices at coordinates $P = (0, 0)$, $Q = (0, 3)$, and $R = (4, 3)$.

(a) Use basic trigonometry to compute the lengths of all three sides and the measures of all three angles.

(b) Use the results of this section to repeat (a).

32. A triangle has vertices at coordinates $P = (0, 0)$, $Q = (0, 3)$, and $R = (2, 0)$.

(a) Use basic trigonometry to compute the lengths of all three sides and the measures of all three angles.

(b) Use the results of this section to repeat (a).

33. A triangle has vertices at coordinates $P = (1, 2, 3)$, $Q = (1, 1, 2)$, and $R = (4, 2, 2)$.

(a) Compute the lengths of all three sides.

(b) Compute all three angles in both radians and degrees.

34. A triangle has vertices at coordinates $P = (2, 1, 5)$, $Q = (-1, 3, 7)$, and $R = (2, -4, 1)$.

(a) Compute the lengths of all three sides.

(b) Compute all three angles in both radians and degrees.

35. Find the equation of the line through $(2, 1)$ and perpendicular to $[1, 2]'$.

36. Find the equation of the line through $(3, 2)$ and perpendicular to $[-1, 1]'$.

37. Find the equation of the line through $(1, -2)$ and perpendicular to $[4, 1]'$.

38. Find the equation of the line through $(0, 1)$ and perpendicular to $[1, 0]'$.

39. Find the equation of the plane through $(0, 0, 0)$ and perpendicular to $[1, 1, 1]'$.

40. Find the equation of the plane through $(1, 0, -3)$ and perpendicular to $[1, -2, -1]'$.

41. Find the equation of the plane through $(0, 0, 0)$ and perpendicular to $[1, 0, 0]'$.

42. Find the equation of the plane through $(1, -1, 2)$ and perpendicular to $[-1, 1, 2]'$.

9.5.3

In Problems 43–46, find the parametric equation of the line in the x–y plane that goes through the indicated point in the direction of the indicated vector.

43. $(1, -1), \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

44. $(3, -4), \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

45. $(-1, -2), \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

46. $(-1, 4), \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

In Problems 47–50, find the parametric equation of the line in the x–y plane that goes through the given points. Then eliminate the parameter to find the equation of the line in standard form.

47. $(-1, 2)$ and $(3, 0)$

48. $(2, 1)$ and $(3, 5)$

49. $(1, -3)$ and $(4, -3)$

50. $(2, 3)$ and $(-1, -4)$

In Problems 51–54, parameterize the equation of the line given in standard form.

51. $3x + 2y - 1 = 0$

52. $x - 2y + 5 = 0$

53. $2x + y - 3 = 0$

54. $2x - y + 4 = 0$

In Problems 55–58, find the parametric equation of the line in x–y–z space that goes through the indicated point in the direction of the indicated vector.

55. $(1, -1, 2), \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

56. $(2, 0, 4), \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

57. $(-1, 3, -2), \begin{bmatrix} -1 \\ -2 \\ 4 \end{bmatrix}$

58. $(2, 1, -3), \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$

In Problems 59–62, find the parametric equation of the line in x–y–z space that goes through the given points.

59. $(5, 4, -1)$ and $(2, 0, 3)$

60. $(2, 0, -3)$ and $(4, 1, 1)$

61. $(2, -3, 1)$ and $(-5, 2, 1)$

62. $(1, 0, 4)$ and $(3, 2, 0)$

63. Where do a plane through $(1, -1, 2)$ and perpendicular to $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ and a line through the points $(0, -3, 2)$ and $(1, -2, 3)$ intersect?

64. Where do a plane through $(2, 0, -1)$ and perpendicular to $\begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}$ and a line through the points $(1, 0, -2)$ and $(1, -1, 1)$ intersect?

65. Given a plane through $(0, -2, 1)$ and perpendicular to $\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$, find a line through $(5, -1, 0)$ that is parallel to the plane.

66. Given the plane $x + 2y - z + 1 = 0$, find a line in parametric form that is perpendicular to the plane.

Chapter 9 Review

Key Terms

Discuss the following definitions and concepts:

1. Linear system of equations
2. Solving a linear system of equations
3. Upper triangular form
4. Gaussian elimination
5. Matrix
6. Basic matrix operations; addition, multiplication by a scalar
7. Transposition
8. Matrix multiplication
9. Identity matrix
10. Inverse matrix
11. Determinant
12. Vector
13. Linear map
14. Eigenvalues and eigenvectors
15. Leslie matrix
16. Stable age distribution
17. Interpretation of the largest eigenvalue of the Leslie matrix.
18. Vector addition
19. Parallelogram law
20. Multiplication of a vector by a scalar
21. Length of a vector
22. Dot product
23. Angle between two vectors
24. Perpendicular vectors
25. Line in the plane and in space
26. Equation of a plane
27. Parametric equation of a line

Review Problems

1. Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$$

(a) Find $A\mathbf{x}$ when $\mathbf{x} = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$. Graph both $\mathbf{x}$ and $A\mathbf{x}$ in the same coordinate system.

(b) Find the eigenvalues λ_1 and λ_2, and the corresponding eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$, of A.

(c) If $\mathbf{u}_i$ is the eigenvector corresponding to λ_i, find $A\mathbf{u}_i$ and explain graphically what happens when you apply A to $\mathbf{u}_i$.

(d) Write $\mathbf{x}$ from part (a) as a linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$; that is, find a_1 and a_2 so that

$$\mathbf{x} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$$

Show that

$$A\mathbf{x} = a_1\lambda_1\mathbf{u}_1 + a_2\lambda_2\mathbf{u}_2$$

and illustrate this equation graphically.

2. Let

$$A = \begin{bmatrix} 3 & 1/2 \\ -5 & -1/2 \end{bmatrix}$$

(a) Find $A\mathbf{x}$ when $\mathbf{x} = \left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$. Graph both $\mathbf{x}$ and $A\mathbf{x}$ in the same coordinate system.

(b) Find the eigenvalues λ_1 and λ_2, and the corresponding eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$, of A.

(c) If $\mathbf{u}_i$ is the eigenvector corresponding to λ_i, find $A\mathbf{u}_i$, and explain graphically what happens when you apply A to $\mathbf{u}_i$.

(d) Write $\mathbf{x}$ from part (a) as a linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$; that is, find a_1 and a_2 so that

$$\mathbf{x} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$$

Show that

$$A\mathbf{x} = a_1\lambda_1\mathbf{u}_1 + a_2\lambda_2\mathbf{u}_2$$

and illustrate this equation graphically.

3. Given the Leslie matrix

$$L = \begin{bmatrix} 0.75 & 0.5 \\ 0.5 & 0 \end{bmatrix}$$

find the growth rate of the population as $t \to \infty$ and determine its stable age distribution.

4. Given the Leslie matrix

$$L = \begin{bmatrix} 0.5 & 2.99 \\ 0.25 & 0 \end{bmatrix}$$

find the growth rate of the population and determine its stable age distribution.

5. Let

$$AB = \begin{bmatrix} 0 & 1 \\ 2 & -1 \end{bmatrix} \quad \text{and} \quad A^{-1} = \begin{bmatrix} 4 & -1 \\ 8 & -1 \end{bmatrix}$$

Find B.

6. Let

$$(AB)^{-1} = \begin{bmatrix} -1 & 3 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}$$

Find A.

7. Explain two different ways to solve a system of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

when $a_{11}a_{22} - a_{12}a_{21} \neq 0$.

8. Suppose that

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

has infinitely many solutions. If you wrote this system in matrix form $AX = B$, could you find X by computing $A^{-1}B$?

9. Let

$$\begin{aligned} ax + 3y &= 0 \\ x - y &= 0 \end{aligned}$$

How do you need to choose a so that this system has infinitely many solutions?

10. Let A be a 2×2 matrix and X and B be 2×1 matrices. Assume that $\det A = 0$. Explain how the choice of B affects the number of solutions of $AX = B$.

11. Suppose that

$$L = \begin{bmatrix} 0.75 & 0.5 \\ a & 0 \end{bmatrix}$$

is the Leslie matrix of a population with two age classes. For which values of a does this population grow?

12. Suppose that

$$L = \begin{bmatrix} 0.5 & 2.0 \\ 0.1 & 0 \end{bmatrix}$$

is the Leslie matrix of a population with two age classes.

(a) If you were to manage this population, would you need to be concerned about its long-term survival?

(b) Suppose that you can improve either the fecundity (i.e., rate of births) to zero-year-olds or the survival of the zero-year-olds, but due to physiological and environmental constraints, the rate of births of zero-year-olds will not exceed 1.5 and the survival of zero-year-olds will not exceed 0.4. Investigate how the growth rate of the population is affected by changing either the survival or the fecundity of zero-year-olds, or both. What would be the maximum achievable growth rate?

(c) In real situations, what other factors might you need to consider when you decide on management strategies?

13. Show that the eigenvalues of

$$\begin{bmatrix} a & c \\ 0 & b \end{bmatrix}$$

are equal to a and b.

14. Show that the eigenvalues of

$$\begin{bmatrix} a & 0 \\ c & b \end{bmatrix}$$

are equal to a and b.

15. **Island Colonies** Sea birds inhabit two separate colonies on different islands. Denote by $N_1(t)$ the number of seabirds on the first island and by $N_2(t)$ the number of seabirds on the second island. In one year a fraction 0.3 of the seabirds on the first island die and a fraction 0.4 of seabirds on the second island die. On average each seabird present at time t produces 3 offspring that survive to time $t + 1$.

(a) Given this information explain why:

$$N_1(t+1) = 2.7N_1(t) \quad \text{and} \quad N_2(t+1) = 2.6\,N_2(t)$$

(b) We now want to include the effect of migration between islands in our model. Assume that of the seabirds present on island 1 at time t, a fraction 0.2 will migrate to island 2 between time t and $t+1$. Conversely a fraction 0.1 of seabirds present on island 2 will migrate to island 1 between time t and $t+1$. Explain how to represent this additional data by a Leslie matrix model.

(c) Find the limiting growth rate of the total population of seabirds across both islands as $t \to \infty$.

(d) Find the fraction of seabirds that inhabit island 1, and the fraction that inhabit island 2 as $t \to \infty$. That is, find the distribution of seabirds as $t \to \infty$.

16. The Spore Bank Filamentous fungi typically produce billions of spores when they reproduce. It is thought that there is an adaptive advantage to making more spores than are needed to colonize the hosts available in the fungus' habitat. The remaining spores form a spore bank—that is, they persist in the soil until conditions trigger their germination and growth. We will build a model for the numbers of fungi, $N_f(t)$, and numbers of spores $N_s(t)$. $N_f(t)$ will be measured in the usual units (i.e., $N_f = 1$ represents 1 fungus) but spores are measured in billions (i.e., $N_s = 1$ represents 1 billion spores). The number of spores and number of fungi is surveyed on a yearly basis. Assume that of the spores present at the beginning at the year, 90% will die during the year and a fraction, 10^{-9} (i.e., one in a billion), will germinate and form fungi. Of the fungi present at the beginning of the year, 60% will die during the year. Another fraction (not given right now) will produce spores.

(a) If we define a vector $\mathbf{N}(t) = \begin{bmatrix} N_f(t) \\ N_s(t) \end{bmatrix}$ to represent the number of fungi and spores, explain why the data above leads to a Leslie matrix model:

$$\mathbf{N}(t+1) = L\mathbf{N}(t)$$

with $L = \begin{bmatrix} 0.4 & 1 \\ a & 0.1 \end{bmatrix}$. Interpret in words what the constant a represents.

(b) Show that if $a = 0.1$, then the number of fungi will decrease over time.

(c) What is the minimum value of a for the fungal population to not decline over time?

17. Cost of Flight The energy needs of hummingbirds in flight are known to depend both on the weight of the bird x (measured in grams) and on the elevation that they live at, y (measured in thousands of feet; flying is harder in the thinner air found at higher elevations). Assume that the cost of flight depends linearly on x and y, i.e., $f(x, y) = ax + by$ for some constants a and b.

(a) How many measurements are needed to estimate the constants a and b?

(b) Suppose that in some units $f = 5.5$ when a 20 g giant hummingbird flies at 3000 ft elevation. But when a rufous hummingbird weighing 4 g flies at 10,000 ft, $f = 5.6$. Estimate the constants a and b.

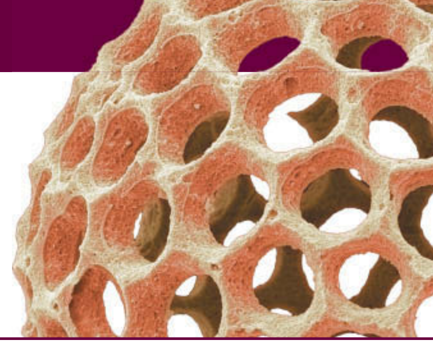
18. Zombie Ants *Cordyceps* is a species of fungus that infests insects. It has the remarkable ability to control the brains of the insects that it infects. When an unlucky ant becomes infected with *Cordyceps*, other ants in the nest will rush it out of the nest, and take it as far away as possible. Evicting infected ants is the only hope of preventing the infection from spreading. The likelihood f of a nest being overtaken by zombies therefore depends on two factors; the diversity of *Cordyceps* spores in the ant's habitat, x, and the amount of energy that the nest devotes to policing for infection and evicting infected ants, y. Assume a linear model, i.e., $f(x, y) = ax + by$ for some constants a and b.

(a) Do you expect a to be positive or negative? What about b?

(b) Use the following data to estimate a and b: For one species $x = 0.5$ and $y = 0.5$, and $f = 0.875$. Conversely, a second species, living in the same habitat, invests three times more energy into patrolling and evicting sick ants, and has a likelihood of becoming infected of 0.25.

CHAPTER

10



Multivariable Calculus

The primary focus of this chapter is differential calculus for functions that depend on more than one variable. Specifically, we will learn how to

- find limits;
- calculate partial derivatives;
- linearize functions of two variables; and
- calculate and graphically interpret the gradient and find extrema
- apply these methods to solve optimization methods, including fitting models to data.

Have you ever heard someone say "It's not the heat that's the problem—it's the humidity" when describing hot summer weather? How comfortable we feel on a hot day is affected both by the air temperature and by the humidity. We sweat to lower our body temperatures on hot days: As the sweat evaporates, its change in state from liquid to vapor removes heat from our bodies. But on humid days, our sweat evaporates more slowly so we feel hotter.

Overheating can lead to heat-stroke, so it is important from a public health point of view for weather forecasting services to keep track of, as well as predict, the apparent temperature that people feel, as well as the air temperature. The **heat index** combines the air temperature and the humidity into a single measurement that represents the apparent temperature. On hot days weather forecasting services often report both the air temperature and the heat index so that people can plan their outdoor activities appropriately.

Mathematically we may think of the heat index, H, as being a function of both the air temperature, T, and the relative humidity, R. (Relative humidity is usually given as a percentage from 0 to 100%, and it represents the amount of water vapor contained in the air. As the relative humidity increases, the evaporation rate of water deceases. At 100% relative humidity, sweat does not evaporate.)

When T is measured on the Fahrenheit scale, a widely used formula for calculating H is

$$\begin{aligned} H(T,R) = &-42.38 + 2.049T + 10.14R - 6.838 \times 10^{-3}T^2 \\ &- 0.2248TR - 5.482 \times 10^{-2}R^2 + 1.229 \times 10^{-3}T^2R \\ &+ 8.528 \times 10^{-4}TR^2 - 1.99 \times 10^{-6}T^2R^2 \end{aligned} \tag{10.1}$$

which gives the heat index, also on a Fahrenheit scale. The function H represents the apparent temperature under a particular combination of air temperature, T, and humidity, R. Equation (10.1) should only be used if $T \geq 80°$F and if $R \geq 40\%$; otherwise, the formula can produce nonsensical results. Since both T and R can be varied, we say that H is a function of two independent variables. For example, according to Equation (10.1), on a hot summer's day in Minneapolis ($T = 94.8°$F, $R = 64.8\%$), the apparent temperature will be:

$$\begin{aligned}H(94.8, 64.8) &= -42.38 + 2.049(94.8) + 10.14(64.8) - 6.838 \times 10^{-3}(94.8)^2 \\ &\quad - 0.2248(94.8)(64.8) - 5.482 \times 10^{-2}(64.8)^2 \\ &\quad + 1.229 \times 10^{-3}(94.8)^2(64.8) + 8.528 \times 10^{-4}(94.8)(64.8)^2 \\ &\quad - 1.99 \times 10^{-6}(94.8)^2(64.8)^2 \\ &= 116.4°\text{F}\end{aligned}$$

Conversely, on a hot summer day in Los Angeles ($T = 80.0°$F, $R = 76.6\%$) the apparent temperature will be:

$$\begin{aligned}H(80, 76.6) &= -42.38 + 2.049(80) + 10.14(76.6) - 6.838 \times 10^{-3}(80.0)^2 \\ &\quad - 0.2248(80)(76.6) - 5.482 \times 10^{-2}(76.6)^2 \\ &\quad + 1.229 \times 10^{-3}(80)^2(76.6) + 8.528 \times 10^{-4}(80)(76.6)^2 \\ &\quad - 1.99 \times 10^{-6}(80)^2(76.6)^2 \\ &= 83.4°\text{F}\end{aligned}$$

Notice that, although Minneapolis and Los Angeles are around 15°F apart in air temperature, the apparent temperature difference is much higher (around 30°F).

Equation (10.1) shows some of the additional complications that are associated with studying functions of more than one variable. In Chapter 5 we learned how to use calculus tools to interpret the behavior of functions of single variables; for example, to determine whether a function increases or decreases as the independent variable increases. It is not clear how similar inferences can be made for the heat index; e.g., we may wish to know: Does increasing T always increase H? How does increasing R affect H?

One way we can study the function H is to imagine holding one of the independent variables constant, and to study how H changes if the other independent variable is changed. To do this we could, for example, fix R (at, say, 50%) and vary T: This would correspond to comparing the apparent temperatures on different days when the relative humidity is always 50%, but on which there are different air temperatures. Since the relative humidity is the same, and only the air temperature is changed, we can treat H as if it were a function of a single independent variable (T). We can then make a graph showing the heat index as a function of the air temperature (see Figure 10.1).

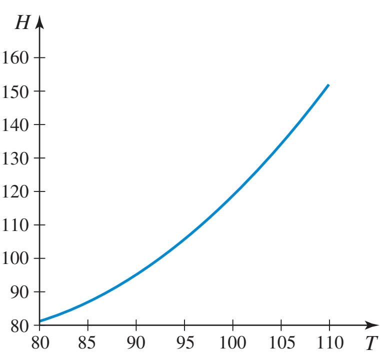


Figure 10.1 The graph of heat index as a function of T when $R = 50\%$.

What about the effect of relative humidity on the heat index? We can explore this effect by holding T constant (say $T = 90°$F) and treating R as an independent variable, which we assign values between 40% and 100%. This corresponds to comparing apparent temperatures from different days, all of which have the same air temperature $T = 90°$F but have different relative humidities. Again, since there is only one independent variable (R), we can make a graph of H as a function of R (see Figure 10.2).

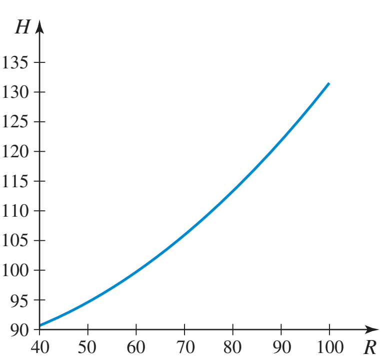


Figure 10.2 The graph of heat index as a function of R when $T = 90°$F.

Figures 10.1 and 10.2 suggest that the apparent temperature increases either if T is increased or if R is increased. The dependence of H on both T and R makes sense because of the physics; increasing air temperature or increasing humidity both increase the apparent temperature. But the figures only show that H increases if T increases for one particular value of R ($R = 50\%$), and that H increases if R increases for one particular value of $T (T = 90°\text{F})$.

In the next section we will describe other methods for graphing functions that produce a more complete picture of the dependence of H on T and R. In the following sections we show how methods of differential calculus can be used to determine how the value of a function changes as the independent variables change, without requiring a graph of the function.

10.1 Functions of Two or More Independent Variables

10.1.1 Defining a Function of Two or More Variables

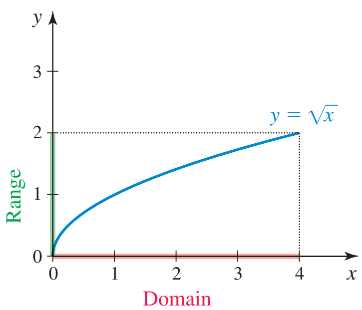


Figure 10.3 The domain and range of a function.

To recall the notation and terminology that we used when we considered functions of one variable, let

$$\begin{aligned} f : [0, 4] &\longrightarrow \mathbf{R} \\ x &\mapsto \sqrt{x} \end{aligned}$$

The function $y = f(x)$ depends on one variable: x. Its domain is the set of possible values of x, namely, the interval $[0, 4]$. Its range is the set of all possible values $y = f(x)$ for x in the domain of f. We see from Figure 10.3 that the range of $f(x)$ is the interval $[0, 2]$.

We now consider functions for which the domain consists of pairs of real numbers (x, y) with $x, y \in \mathbf{R}$ or, more generally, of n-tuples of real numbers $(x_1, x_2, \ldots, x_n)$ with $x_1, x_2, \ldots, x_n \in \mathbf{R}$. We also call n-tuples *points*. We use the notation $\mathbf{R}^n$ to denote the set of all n-tuples $(x_1, x_2, \ldots, x_n)$ with $x_1, x_2, \ldots, x_n \in \mathbf{R}$; that is,

$$\mathbf{R}^n = \{(x_1, x_2, \ldots, x_n) : x_1 \in \mathbf{R}, x_2 \in \mathbf{R}, \ldots, x_n \in \mathbf{R}\}$$

For $n = 1$, $\mathbf{R}^1 = \mathbf{R}$, which is the set of all real numbers. For $n = 2$, $\mathbf{R}^2$ is the set of all points in the plane, and so on. Note that n-tuples are *ordered*; for instance, $(2, 3)$ is not the same as $(3, 2)$. We consider functions whose ranges are subsets of the real numbers; such functions are called *real valued*.

Definition Suppose $D \subset \mathbf{R}^n$. Then a **real-valued function** f on D assigns a real number to each element in D, and we write

$$\begin{aligned} f : D &\longrightarrow \mathbf{R} \\ (x_1, x_2, \ldots, x_n) &\mapsto f(x_1, x_2, \ldots x_n) \end{aligned}$$

The set D is the **domain** of the function f, and the set

$$\{w \in \mathbf{R} : f(x_1, x_2, \ldots, x_n) = w \text{ for some } (x_1, x_2, \ldots, x_n) \in D\}$$

is the **range** of the function f.

If a function f depends on just two independent variables, we will often denote the independent variables by x and y, and write $f(x, y)$. In the case of three variables, we will often write $f(x, y, z)$. If f is a function of more than three independent variables, it is more convenient to use subscripts to label the variables—for example, $f(x_1, x_2, x_3, x_4)$.

The first example of this chapter shows how a function of more than one independent variable is evaluated at given points.

EXAMPLE 1 Evaluate the function

$$f(x, y, z) = \frac{xy}{z^2}$$

at the points $(2, 3, -1)$ and $(-1, 2, 3)$.

Solution Since $f(x, y, z)$ lists the independent variables in the order x, y, and z, to evaluate the function at $(2, 3, -1)$, we need to substitute 2 for x, 3 for y, and -1 for z:

$$f(2, 3, -1) = \frac{(2)(3)}{(-1)^2} = 6$$

Similarly,

$$f(-1, 2, 3) = \frac{(-1)(2)}{(3)^2} = -\frac{2}{9}$$

Notice that although the same three numbers appear in $(2, 3, -1)$ as in $(-1, 2, 3)$, these points are different, and f takes different values at the two points. •

Just as with functions of one variable, to completely specify a function of more than one variable we need to specify the domain. For example, the heat index function can only be evaluated when $T \geq 80°F$ and $R \geq 40\%$, so the domain of the function H is

$$\{(T, R) : T \geq 80 \text{ and } R \geq 40\}.$$

In the next two examples we will explore the domains and ranges of other functions of two variables.

EXAMPLE 2 Let $D = \{(x, y) : 0 \leq x \leq 1,\ 0 \leq y \leq 1\}$ and

$$\begin{aligned} f : D &\longrightarrow \mathbf{R} \\ (x, y) &\longrightarrow x + y \end{aligned}$$

Draw the domain of f in the x–y plane and determine the range of f.

Solution The domain of f is the set D, which consists of all points (x, y) whose x- and y-coordinates are between 0 and 1. This is the square shown in Figure 10.4.

To find the range of f, we need to determine what values f can take when we plug in points (x, y) from the domain D. The function $z = f(x, y)$ is smallest when x and y take the smallest values allowed within the set D; i.e., when $x = 0$ and $y = 0$ or at the point $(x, y) = (0, 0)$; where, $f(0, 0) = 0$. The function $z = f(x, y)$ is largest when x and y take the largest values allowed within the set D, i.e., when $x = 1$ and $y = 1$, at the point $(1, 1)$; where, $f(1, 1) = 2$. The function takes on all values in between. Hence, the range of f is the set $\{z : 0 \leq z \leq 2\}$. •

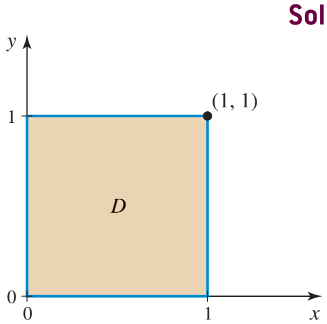


Figure 10.4 The domain of the function in Example 2.

As in the case of functions of a single variable, the domain sometimes needs to be restricted. The next example illustrates how to find the largest possible domain.

EXAMPLE 3 Let $D = \{(x, y) : -1 \leq x \leq 1, 0 \leq y \leq 1, \text{ and } y^2 \geq x\}$ and $f : D \to \mathbb{R}$ with $f(x, y) = \sqrt{y^2 - x}$, Draw the domain D of f, and find the function's range.

Solution The domain here has a complicated form, so it is worth breaking down the conditions necessary for (x, y) to be within D. The first two conditions are that $-1 \leq x \leq 1$ and $0 \leq y \leq 1$; these conditions describe a rectangle, shown in Figure 10.5a. The second condition is that $y^2 \geq x$. $y^2 = x$ is the equation for a parabola (it is easiest to draw this parabola by plotting x as a function of y). If $y^2 \geq x$, then (x, y) must be to the left of this parabola since $x \leq y^2$ implies that (x, y) lies to the left of (y^2, y); see Figure 10.5b. •

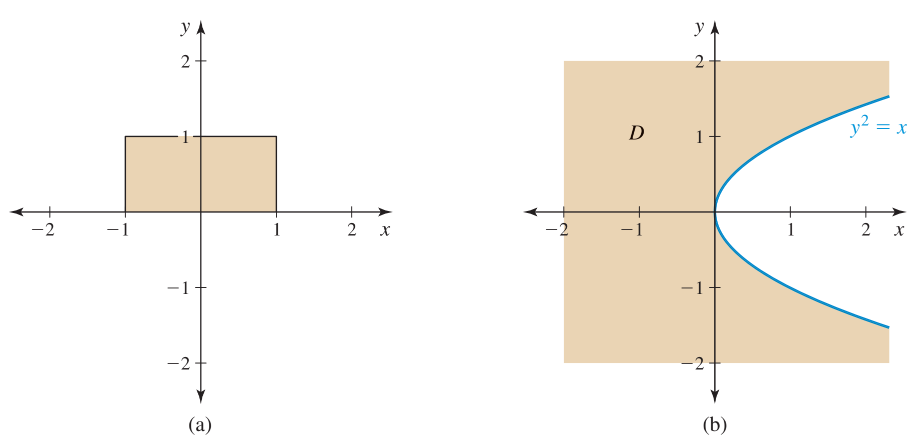


Figure 10.5 In Example 3 a point (x, y) lies in D if $-1 \leq x \leq 1$ and $0 \leq y \leq 1$ (shaded region in panel a) *and* if $y^2 \geq x$ (shaded region in panel b).

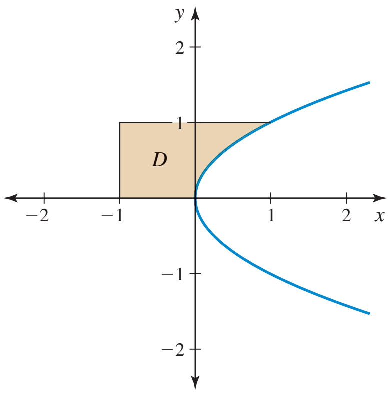


Figure 10.6 The domain D of the function f defined in Example 3.

Why is the condition $y^2 \geq x$ necessary? In order to calculate $f(x, y)$ we must take the square root of $y^2 - x$. This is only possible if $y^2 - x \geq 0$; i.e., if $y^2 \geq x$, which gives rise to the third condition used to define D.

In order for (x, y) to lie in D, it must lie both in the rectangular region shown in Figure 10.5a and on the left-hand side of the parabola shown in Figure 10.5b. D is therefore the **intersection** of these two regions and is shown in Figure 10.6.

To find the range of f we must find the smallest and largest values that $f(x, y)$ can take for any point $(x, y) \in D$. First we find the smallest value; because f is defined by taking a square root, it must be non-negative, i.e., the smallest value it can take is 0. This value is indeed attained at any point (x, y) at which $y^2 = x$ (i.e., along the parabola that makes up the right-hand border of D). To find the largest value of f note that the largest value of f coincides with the largest value of $y^2 - x$ in D. This quantity is largest when y^2 takes its *largest* value and x takes its *smallest* value. In D, this occurs when $y = 1$ or -1, and when $x = -1$. Then $f(-1, -1) = f(-1, 1) = \sqrt{1+1} = \sqrt{2}$. The function takes on all values between its smallest and largest values. Hence the range of f is the interval $[0, \sqrt{2}]$.

10.1.2 The Graph of a Function of Two Independent Variables—Surface Plot

When we were studying functions of a single independent variable, graphing the function was a useful way to visualize how the value of the function changes when the independent variable is changed. We would like to do the same thing for functions of more than one variable. We showed one way to do this in the introduction (where we held R fixed, and made a plot of how the function H changed as T was varied), but we would like to find a way of visualizing how the function depends on both of its variables. All of the following plotting methods can only be used for functions of two variables. There are specialized methods for plotting functions of three or more variables, but these usually require computers and are beyond the scope of this book.

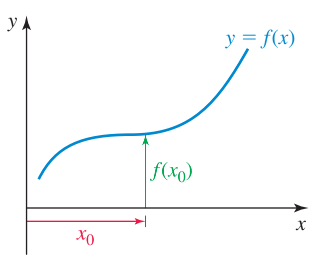


Figure 10.7

When we graph a function of one variable such as $f(x)$ we write $y = f(x)$ and then draw a curve made up of points (x, y) in two-dimensional space. For any point x in the domain of f, the height of the curve at that point is given by $y = f(x)$ (see Figure 10.7).

We can analogously draw graphs of functions of two variables, such as $f(x, y)$: To do so, we write $z = f(x, y)$ and then locate the point (x, y, z) in three-dimensional space.

> **Definition** If f is a function of two independent variables with domain D, then the graph of f is the set of all points (x, y, z) such that $z = f(x, y)$ for $(x, y) \in D$. That is, the graph of f is the set
>
> $$S = \{(x, y, z) : z = f(x, y),\ \ (x, y) \in D\}$$

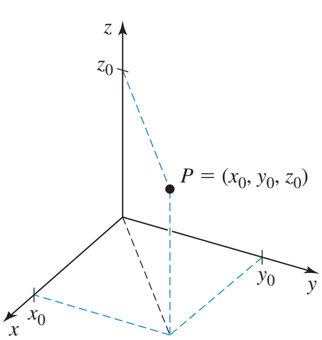


Figure 10.8 The right-handed coordinate system with a point.

To locate a point (x, y, z) in three-dimensional space, we use the Cartesian coordinate system, which we introduced in Chapter 9. This system consists of three mutually perpendicular axes that emanate from a common point, called the *origin*, which has coordinates $(0, 0, 0)$. The axes are oriented in a right-handed coordinate frame, as explained in Chapter 9. This is shown in Figure 10.8, together with a point P that has coordinates (x_0, y_0, z_0).

Figure 10.8 suggests that we can visualize the graph of $f(x, y)$, $(x, y) \in D$, as lying directly above (or below) the domain D in the x–y plane. The graph of $f(x, y)$ is therefore a **surface** in three-dimensional space, as illustrated in Figure 10.9 for $f(x, y) = 2x^2 - y^2$.

Graphing a surface in three-dimensional space is difficult, and it practically requires use of computer software. We show a few such surfaces, generated by computer software, in Figures 10.9, 10.11, and 10.12. You do not need to learn how to graph such

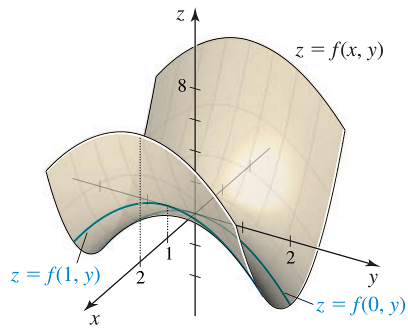


Figure 10.9 The graph of $z = f(x, y)$ is a surface in three-dimensional space [here, $f(x, y) = 2x^2 - y^2$].

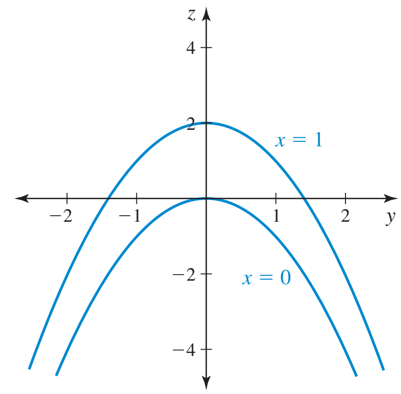


Figure 10.10 Curves produced by walking along the surface $z = 2x^2 - y^2$, along a line $x =$ constant.

functions, but it is helpful to be able to recognize how the surface shapes can be produced by a function.

To understand the shape of the graph of $f(x, y) = 2x^2 - y^2$ in Figure 10.9, we can fix a value for x and then "walk" on the surface in the y-direction. That is, we treat x as if it were a constant, and see how f changes as the variable y changes. For example, we could fix $x = 1$, in which case $f(1, y) = 2 - y^2$. This is the equation of a parabola that bends downward, and intercepts the line $y = 0$ at $f(1, 0) = 2$. Similarly, if we fix $x = 0$, $f(0, y) = -y^2$; this is another downward bending parabola. In Figure 10.10 we show the curves produced when x is held constant and y is varied. Looking at the surface in Figure 10.9, we see that we do indeed trace out downward bending parabolas if we walk along the surface in the y-direction. If, now, we instead fix y and walk in the direction of x, then if, say, we fix $y = 0$, we obtain $f(x, 0) = 2x^2$, which is the equation of an upward bending parabola. Similarly if we fix $y = 1$, then $f(x, 1) = 2x^2 - 1$, which is again an upward bending parabola. In general, no matter what value is fixed y at, if y is fixed and x is varied, then $f(x, y) = 2x^2-$ constant; so this is generally an upward bending parabola. We do indeed trace out upward bending parabolas if we walk along the surface in Figure 10.9 in the x-direction.

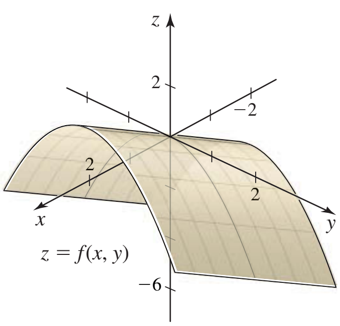


Figure 10.11 The graph of $f(x, y) = x - y^2$.

Figure 10.11 can be analyzed in a similar way. The function depicted is $f(x, y) = x - y^2$. If we fix $x = 0$ and vary y, then $f(0, y) = -y^2$, which is again a downward bending parabola. Similarly fixing $x = 1$, and varying y, we have $f(1, y) = 1-y^2$, which is also a downward bending parabola. In general, fixing any value of x and varying y gives a downward bending parabola $f(x, y) =$ constant $-y^2$, and this is indeed the curve traced out if we walk along the surface in Figure 10.11 in the y-direction. If on the other hand we fix y and vary x, then $f(x, y) = x-$constant, which is the equation for a straight line. Walking along the surface in Figure 10.11 in the x-direction traces out a straight line.

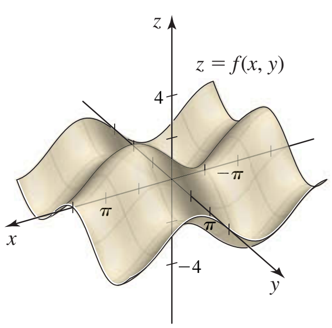


Figure 10.12 The graph of $f(x, y) = \sin x + \cos y$.

The surface in Figure 10.12 represents the function $f(x, y) = \sin x + \cos y$. If x is held constant and y is varied, then $f(x, y) =$ constant $+ \cos y$, i.e., walking along the surface in the y-direction, we trace out a cosine curve. If y is held constant and x is varied, then $f(x, y) = \sin x+$ constant, so by walking along the surface in the x-direction, we trace out a sine curve.

10.1.3 Heat Maps

Surface plots are very difficult to draw by hand, and can also be difficult to interpret for more complicated functions $f(x, y)$. An alternate method for visualizing a function $f(x, y)$ is to use a **heat map** in which each point (x, y) is assigned a color corresponding to the value of $f(x, y)$. This type of plot is called a heat map because it is often used to represent temperature data (for example, the forecasted temperatures on a map of the country), but any function of two variables can be represented in this way. As an example, in Figure 10.13 we use a heat map to represent the heat index H as a function of temperature (T) and relative humidity (R).

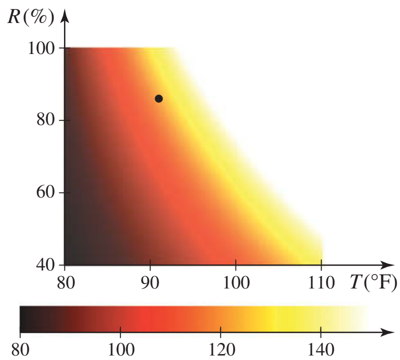


Figure 10.13 The heat index represented using a heat map. Note that we have colored all points with $H \geq 150°$F white. Heat index is not reliable when $H \geq 150°$F

An essential part of the heat map is a color bar, which allows us to interpret the colors used in the heat map. In Figure 10.13 the colors range from black for the lowest heat index (80°F) through red, to yellow, to white as the heat index increases. In our heat map we use a single color (white) to represent all heat indices above 150°F because that is the largest value for which heat index is usually reported. Equation (10.1) becomes unreliable when it predicts a heat index above 150°F; if the heat index exceeds 150°F, people are advised to remain indoors.

The heat map in Figure 10.13 shows us more than our previous plots of H (in Figures 10.1 and 10.2), namely the heat index always increases if T is increased or if R is increased. To use the heat map to read off a heat index value, pick a specific ordered pair of coordinates, e.g., $(T, R) = (91, 86)$ (these are the temperature and humidity measurements on a hot summer morning in Tallahassee, Florida). Locate this point on the heat map, and determine the color at this point—for this specific example, the heat map in Figure 10.13 is a bright orange. Using the color bar, we determine that $H \approx 120°$F (the actual value, according to Equation 10.1, is 123°F).

Creating a heat map for yourself is a little easier than making a surface plot. We will work through an example to show how it can be done by hand, but in practice this work is best done using a computer (spreadsheets and mathematical software like Mathematica, Maple, or Matlab are all capable of making heat maps).

EXAMPLE 4 Create a heat map using five colors and nine points to represent the function $f(x, y) = x + y$ on the domain:

$$D = \{(x, y), 0 \leq x \leq 1, \text{ and } 0 \leq y \leq 1\}$$

Solution We are told to use five colors in our heat map. We will use the *jet* color scale in which the smallest values of f are represented by blue, and the largest values by red or orange. The color scale is shown in Figure 10.14.

To represent f using this color scale we need to match colors to values of f, which means we need to know the range of f. In Example 1 we found that $f(D) = [0, 2]$, so our color scale will need to represent all values between 0 and 2. Since we have five colors, each color will need to represent an interval of size $\frac{2-0}{5} = 0.4$. The first color (dark blue) will represent values between 0 and 0.4, the second color (sky blue) will represent values between 0.4 and 0.8, and so on (see Figure 10.14)

We now divide the domain into 9 even tiles (see Figure 10.15) (i.e., three across and three down). Our heat map will consist of coloring each tile to represent the value of f in that tile. To do this we pick the point in the center of each tile (see Figure 10.15), and evaluate f at each point. We then color the tile according to the value of f at the center point. For this example, the bottom left tile has boundaries $x = 0, x = 1/3$, and $y = 0, y = 1/3$, so the center point is $x = \frac{1}{2}(0 + \frac{1}{3}) = \frac{1}{6}$ and $y = \frac{1}{2}(0 + \frac{1}{3}) = \frac{1}{6}$; i.e., $(x, y) = (1/6, 1/6)$. At this point $f(1/6, 1/6) = 1/6 + 1/6 = 1/3$. Since f takes a value between 0 and 0.4, we color this tile dark blue.

Just to the right of this tile, there is a tile whose boundaries are $x = 1/3, x = 2/3$, and $y = 0, y = 1/3$, so the center point of this tile is $(x, y) = (1/2, 1/6)$. At this center

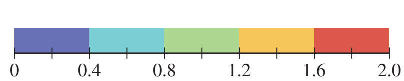


Figure 10.14 Color scale for Example 4

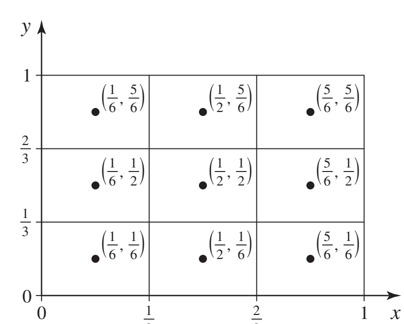


Figure 10.15 Tiling the domain in Example 4 with 9 evenly sized tiles.

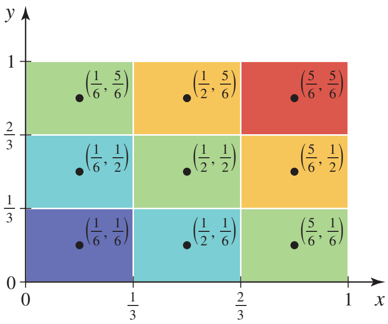


Figure 10.16 Heat map of the function $f(x, y) = x + y$ using 5 colors and 9 tiles.

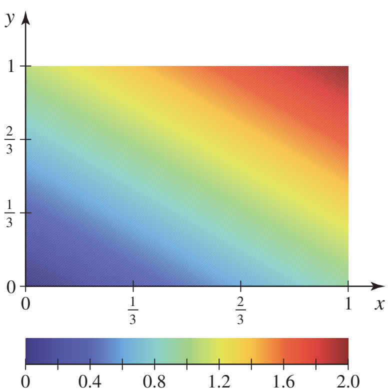


Figure 10.17 Heat map of the function $f(x, y) = x + y$ using 10,000 tiles.

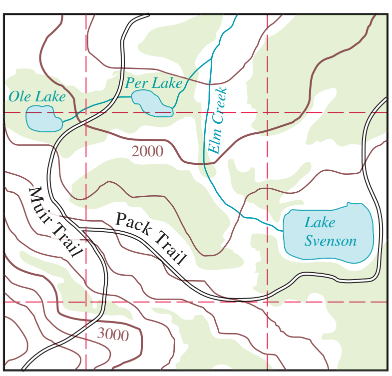


Figure 10.18 A topographical map with level curves.

point $f(1/2, 1/6) = 1/2 + 1/6 = 2/3$. Since f takes a value between 0.4 and 0.8, we color this tile sky blue. We continue in the same manner to color in all of the remaining tiles; see Figure 10.16. •

Drawing heat maps by hand is quite laborious, but it provides good practice for their use. Computer packages can accelerate the calculations, allowing more precise heat maps with more tiles and colors. As an example of this, we regenerated the heat map from Example 4 using 10,000 tiles (i.e., 100×100 tiles) and 100 colors. (With this many tiles, the colors blend smoothly; see Figure 10.17.)

10.1.4 Contour Plots

Although heat maps are a powerful tool for visualizing functions of two variables, drawing a heat map takes a lot of time. Also, using colors to represent the values of f has the disadvantage that it needs careful judgment of where points lie on a color scale to read values of f off the heat map.

Another way to visualize functions that does not use colors is with **level curves** or **contour lines**. This approach is used, for instance, in topographical maps. (See Figure 10.18.) There is a subtle distinction between level curves and contour lines, in that level curves are drawn in the function domain whereas contour lines are drawn on the surface. This distinction is not always made, and often the two terms are used interchangeably. In this text, we will almost exclusively use level curves, for which we now give the precise definition:

> **Definition** Suppose that $f : D \to \mathbf{R}$, $D \subset \mathbf{R}^2$. Then the level curves of f comprise the set of points (x, y) in the x–y plane where the function f has a constant value; that is, $f(x, y) = c$.

To show how this works in Figure 10.19 we redraw the heat index plot using level curves, rather than a heat map. On the plot the curve labeled 100 represents all combinations of T and R that give a heat index of 100°F. When we plot the T and R values for Tallahassee, Florida, on this plot $[(T, R) = (91, 86)]$, we see that the point lies close to the level curve $H = 130$°F, so the heat index $H(91, 86)$ is close to 120°F, (in fact slightly above 120°F).

To explore how functions may be represented by level curves, we will examine some simpler functions.

In Figure 10.20 we graph the surface of $f(x, y) = (x^2 + y^2)e^{-(x+y)^2}$; in Figure 10.21 we graph its level curves. To get an informative picture from the graph of the level curves, you should choose equidistant values for c—for instance, $c = 0, 1, 2, \ldots$ or $c = 0, -0.1, -0.2, \ldots$—so that you can infer the steepness of the curve from how

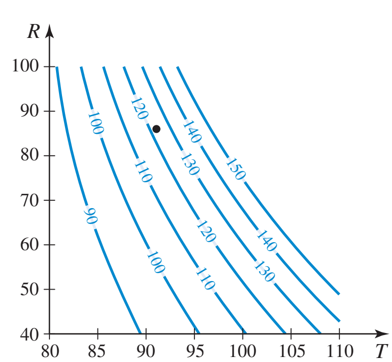


Figure 10.19 Level curve plot for the heat index $H(T, R)$.

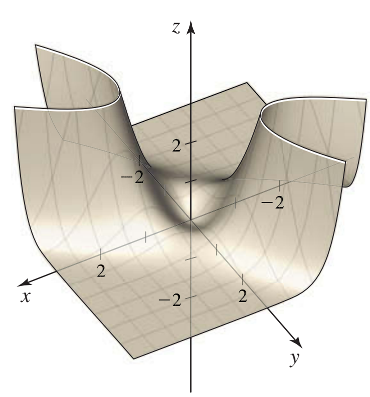


Figure 10.20 The graph of $f(x, y) = (x^2 + y^2)e^{-(x+y)^2}$.

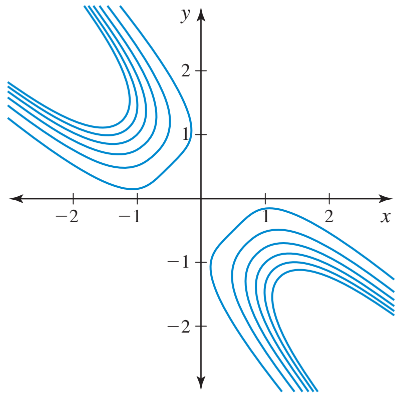


Figure 10.21 Level curves for $f(x, y) = (x^2 + y^2)e^{-(x+y)^2}$.

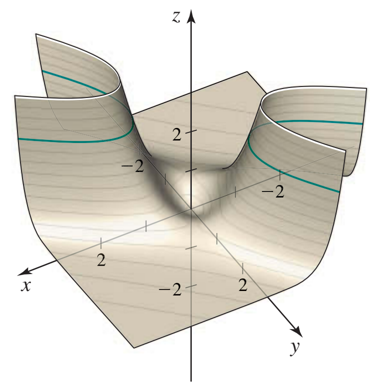


Figure 10.22 A contour line for $f(x, y) = (x^2 + y^2)e^{-(x+y)^2}$.

close together the level curves are. In Figure 10.21, the level curves are equidistant, with $c = 0.5, 1, 1.5, \ldots$.

If we project a level curve $f(x, y) = c$ up to its height c, the curve is called a **contour line**. The curve lies on the surface of $f(x, y)$ and traces the graph of f in a horizontal plane at height c (Figure 10.22).

EXAMPLE 5 Set $D = \{(x, y) : x^2 + y^2 \leq 4\}$. Draw the level curves of

$$f(x, y) = 4 - x^2 - y^2 \quad \text{for } (x, y) \in D$$

and

$$g(x, y) = \sqrt{4 - x^2 - y^2} \quad \text{for } (x, y) \in D$$

and compare your graphs.

Solution The level curve $f(x, y) = c$ is the set of all points (x, y) that satisfy

$$4 - x^2 - y^2 = c \quad \text{or} \quad x^2 + y^2 = 4 - c$$

We recognize the latter equation as that of a circle with center at the origin and radius $\sqrt{4 - c}$, illustrated in Figure 10.23 for $c = 0, 0.5, 1, 1.5$, and 2.

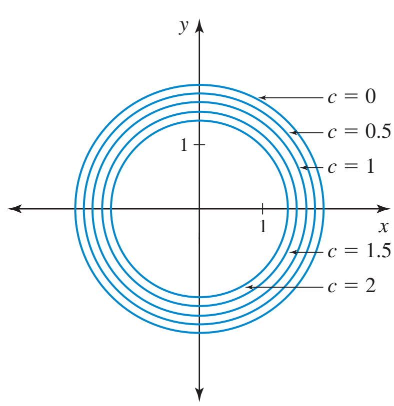


Figure 10.23 Level curves for $f(x, y) = 4 - x^2 - y^2$.

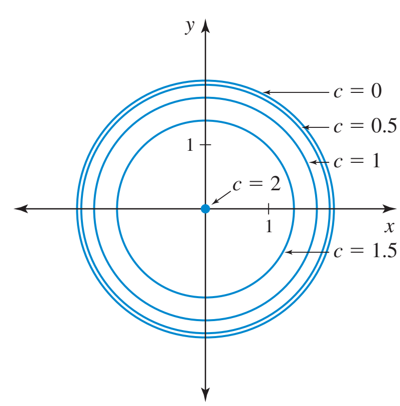


Figure 10.24 Level curves for $g(x, y) = \sqrt{4 - x^2 - y^2}$.

The level curve $g(x, y) = c$ satisfies

$$\sqrt{4 - x^2 - y^2} = c, \qquad \text{or} \qquad x^2 + y^2 = 4 - c^2$$

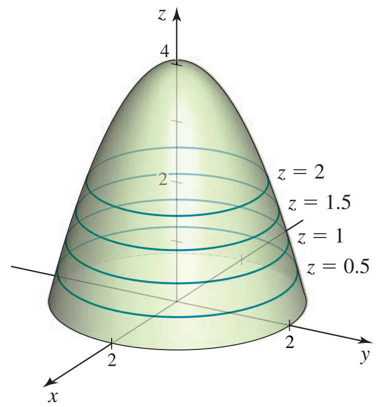


Figure 10.25 The surface of $f(x, y) = 4 - x^2 - y^2$ is a paraboloid.

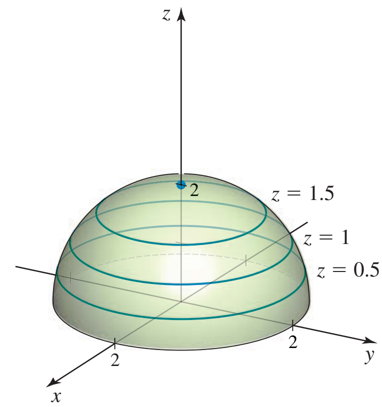


Figure 10.26 The surface of $g(x, y) = \sqrt{4 - x^2 - y^2}$ is the top half of a sphere.

So level curves are also circles with center at the origin, but the radius of a level curve is $\sqrt{4 - c^2}$; the circle is illustrated in Figure 10.24 for $c = 0, 0.5, 1, 1.5$, and 2.

To compare the level curves of two functions, we choose $c = 0, 0.5, 1, 1.5$, and 2 for both $f(x, y)$ and $g(x, y)$. The level curves for $c = 0$ and 1 are the same for the two functions, but the contour lines of $f(x, y)$ for $c = 1.5$ and 2 are a lot closer to the contour line for $c = 1$ than are those of $g(x, y)$. We can understand why when we look at the surfaces. The graph of $f(x, y)$ is obtained by rotating the parabola $z = 4 - x^2$ in the x–z plane about the z-axis; the surface thereby generated is called a *paraboloid*. (See Figure 10.25.)

The graph of $g(x, y)$ is obtained by rotating the half circle $z = \sqrt{4 - x^2}$ in the x–z plane about the z-axis; the surface generated is the top half of a sphere. (See Figure 10.26.)

The paraboloid is steeper than the sphere for c between 1 and 2, but shallower for c between 0 and 1, the contour lines of the paraboloid are closer together for c between 1 and 2, whereas the contour lines of the sphere are closer together for c between 0 and 1. •

EXAMPLE 6 In Figure 10.27, we show level curves for oxygen concentration (in mg/l) in Long Lake, Clear Water County (Minnesota), as a function of date and depth. For instance, on day 140 (May 20, 1998) at a 10-m depth, the oxygen concentration was 12 mg/l.

The water flea *Daphnia* needs a minimum of 3 mg/l oxygen to survive. Suppose that you went out to Long Lake on day 200 (July 19, 1998) and wanted to look for *Daphnia* in the lake. What is the maximum depth at which *Daphnia* can survive?

Solution Because *Daphnia* needs a minimum of 3 mg/l of oxygen, we need to find the depth on July 19, 1998 below which the oxygen concentration is always less than 3 mg/l. We see that the 3-mg/l oxygen-level curve goes through the point $(200, -17.5)$. Thus, on July 19, 1998, *Daphnia* could only survive above 17.5 m. •

EXAMPLE 7 In Figure 10.28, we show temperature **isoclines** (also called **isotherms**) for Long Lake, Clear Water County (Minnesota). Temperature isoclines are lines of equal temperature—that is, level curves. The isoclines shown are functions of date and depth. For instance, on day 130 (May 10, 1998), the temperature at 5 m was approximately 12°C. If we are interested in how temperature changes over time at a fixed depth, we would draw a horizontal line through the appropriate depth and then read off the days and the temperature where the horizontal line intersects the isoclines: For Example, at a depth of 5 meters, we find that the temperature climbs from roughly 8°C in the spring to more than 20°C in the summer and then falls back to below 12°C later in the fall. •

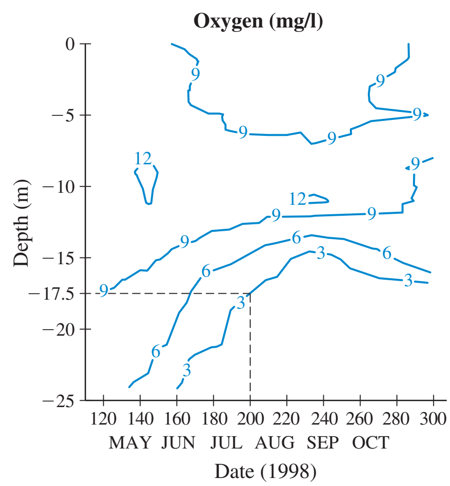


Figure 10.27 Level curves for oxygen concentration on Long Lake, Clear Water County (Minnesota). Courtesy of Leif Hembre.

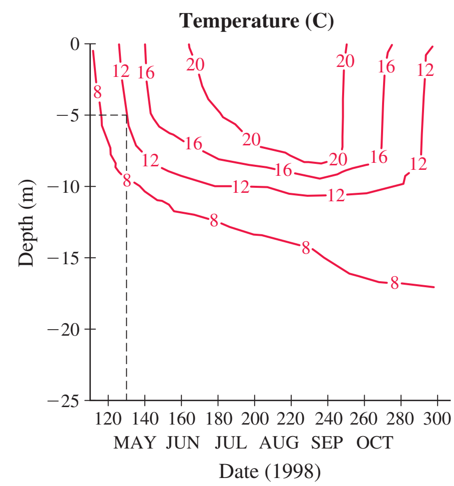


Figure 10.28 Isotherms for Long Lake, Clear Water County (Minnesota). Courtesy of Leif Hembre.

Section 10.1 Problems

10.1.1

1. Body Mass Index Body mass index (or BMI) is often used as an indicator for whether a person is over- or underweight. A person's BMI is calculated from their mass (in kg) and their height (in m). To calculate a person's BMI, divide their mass by the square of their height.

(a) If a person's mass is m, and their height is h, write down the formula that would be used to calculate their BMI.

(b) Jesse is 1.75 m tall, and he weighs 82 kg. What is his BMI?

(c) In a particular population, heights range from 1.50 m to 1.90 m, and masses range from 45 kg to 160 kg. Calculate the maximum possible range of BMI's for this population.

2. Heat Index

(a) Using Equation (10.1), calculate the heat index in the following locations:

(i) Lafayette, Louisiana in August ($T = 91°$F, $R = 95\%$)

(ii) Mojave, California in August ($T = 97°$F, $R = 40\%$)

(iii) Davis, California in August ($T = 91°$F, $R = 85\%$)

Which location will feel hottest?

(b) This part will show the danger of trying to evaluate a function outside of its correct domain. We are told that the heat index should only be used if $40\% \leq R \leq 100\%$ and $T \geq 80°$F. Calculate the heat index for Minneapolis in January ($T = 14°$F, $R = 85\%$). Does your answer make sense?

3. Wind-Chill In cold weather the temperature that we feel depends both on the actual air temperature and on whether the wind is blowing or not. Strong winds blow heat away from our bodies, increasing their rate of heat loss and making us feel colder, a phenomenon known as *wind chill*. In cold winters, weather forecasters report both the real temperature and the apparent temperature including wind chill.

One widely used formula for the apparent temperature W, when the air temperature (measured in °F) is T, and the wind speed (measured in mph) is V, is:

$$W = 35.74 + 0.6215T - 35.75V^{0.16} + 0.4275TV^{0.16} \tag{10.2}$$

This formula is only accurate when $T \leq 50°$F and $V \geq 3$ mph.

(a) Use (10.2) to calculate the apparent temperatures felt by a person in each of the three following locations:

(i) Boston in January ($T = 14°$F, $V = 11$ mph)

(ii) Minneapolis in January ($T = 23°$F, $V = 13$ mph)

(iii) Chicago in January ($T = 24°$F, $V = 18$ mph)

In which location will a person feel coldest?

(b) In this part we will see the danger of trying to evaluate a function outside of its correct domain. Calculate the wind chill for Los Angeles in January ($T = 65°$F, $V = 0.8$ mph). [Note that these (T, V) values are outside the domain for $W(T, V)$.] Does your answer make sense?

4. The Ideal Gas Law The ideal gas law is an equation that can be used to predict how the pressure of a container filled with gas will change if either the temperature of the container is changed (the container is heated or cooled) or if its volume is changed (the container is compressed or expanded). The ideal gas law states that, if the pressure of the gas is P the volume is V, the temperature is T, then

$$P = NkT/V \tag{10.3}$$

where N and k are both constants.

(a) Use Equation (10.3) to show that a container of gas obeys *Boyle's law*, which states that, if the temperature of the container is held constant, then the pressure of the gas (P) is inversely proportional to its volume (V).

(b) Use Equation (10.3) to show that a container of gas obeys *Gay-Lussac's law*, which states that, if the volume of the container is held constant, then the pressure of the gas (P) is proportional to its temperature (T).

(c) If the pressure of the gas is held constant, according to Equation (10.3) how is the temperature of the gas related to its volume? Draw a sketch showing how V varies as a function of T if P is held constant.

5. Cardiac Output Cardiac output (CO) is a way of measuring the amount of blood pushed out by a patient's heart. It is calculated as the product of heart rate (HR) and stroke volume (SV). Write cardiac output as a function of heart rate and stroke volume. If heart rate is measured in beats per minute and stroke volume in liters per beat, what is the unit for cardiac output?

6. Blood Pressure Mean arterial blood pressure (MAP) is a function of systolic blood pressure (SP) (that is, the pressure during the heart beat) and diastolic blood pressure (DP) (that is, the pressure between heart beats). At a resting heart rate,

$$\text{MAP} \approx \text{DP} + \frac{1}{3}(\text{SP} - \text{DP})$$

If systolic pressure is greater than diastolic pressure and both are nonnegative, what is the range of the function describing mean arterial pressure?

7. Locate the following points in a three-dimensional Cartesian coordinate system:

(a) $(1, 3, 2)$ **(b)** $(-1, -2, 1)$

(c) $(0, 1, -1)$ **(d)** $(2, 0, 2)$

8. Describe in words the set of all points in $\mathbf{R}^3$ that satisfy the following expressions:

(a) $x = 0$ **(b)** $y = 0$ **(c)** $z = 0$

(d) $z \geq 0$ **(e)** $y \leq 0$

In Problems 9–16, evaluate each function at the given point.

9. $f(x, y) = x^2 + y$ at $(2, 3)$

10. $f(x, y, z) = x^2 - 3y + z$ at $(3, -1, 1)$

11. $f(x, y) = \sqrt{2x + 3y^2}$ at $(-1, 2)$

12. $f(x, y) = \dfrac{x^2}{y}$ at $(3, 2)$

13. $h(x, t) = \exp\left[-\dfrac{(x-2)^2}{2t}\right]$ at $(1, 5)$

14. $g(n, p) = np(1-p)^{n-1}$ at $(5, 0.1)$

15. $h(x_1, x_2) = x_2 e^{-x_1}$ at $(2, -1)$

16. $g(x_1, x_2, x_3) = x_1\sqrt{x_2 x_3}$ at $(1, 2, 1)$

10.1.2

In Problems 17–22 find the range of each function $f(x, y)$, when defined on the specified domain D.

17. $f(x, y) = x^2 + y^2$; $D = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}$

18. $f(x, y) = \sqrt{9 - x^2 - y^2}$; $D = \{(x, y) : x^2 + y^2 \leq 9\}$

19. $f(x, y) = \ln(y - x)$; $D = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1, y > x\}$

20. $f(x, y) = \exp[-(x^2 + y^2)]$; $D = \{(x, y) : -1 \leq x \leq 1, 0 \leq y \leq 2\}$

21. $f(x, y) = x - y^2$; $D = \{(x, y) : -1 \leq x \leq 1, 1 \leq y \leq 2\}$

22. $f(x, y) = \dfrac{x}{y}$; $D = \{(x, y) : 0 \leq x \leq 1, 1 \leq y \leq 2\}$

23. $f(x, y) = \frac{x}{y}$; $D = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1, y > x\}$

24. $f(x, y) = x^2 y$; $D = (x, y) : -2 \leq x \leq 1, 0 \leq y \leq 1, y < x^2\}$

In Problems 25–30, explain in words why the surface plot in the indicated figure matches the function. Do this by describing what happens if you walk along the surface in the x-direction and in the y-direction.

25. $f(x, y) = 2x^2 + y^2 + 1$ (Fig. 10.29)

26. $f(x, y) = xy$ (Fig. 10.30)

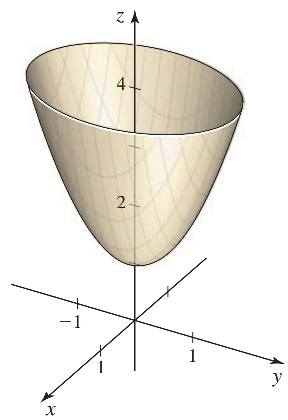


Figure 10.29

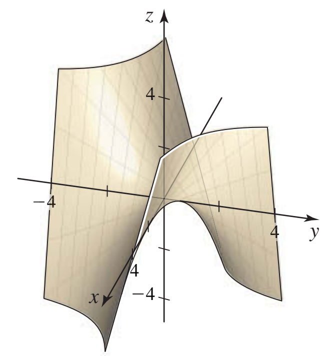


Figure 10.30

27. $f(x, y) = 2x^2 - y$ (Fig. 10.31)

28. $f(x, y) = xy^3$ (Fig. 10.32)

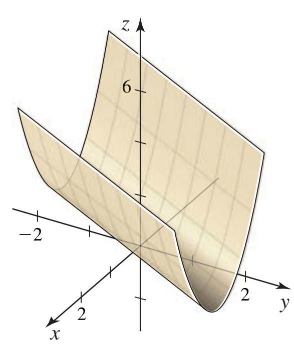


Figure 10.31

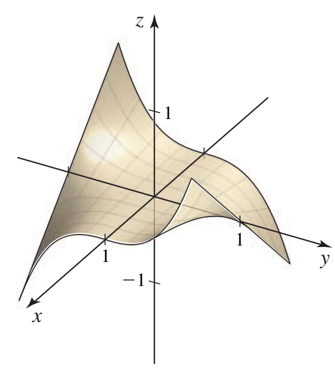


Figure 10.32

29. $f(x, y) = xe^{-y}$ (Fig. 10.33)

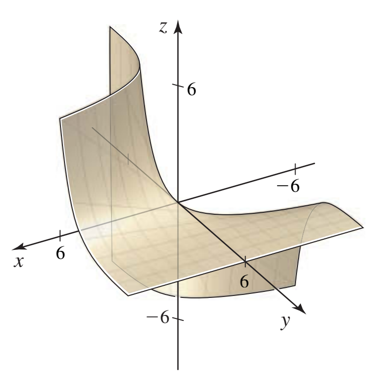


Figure 10.33

30. $f(x, y) = x - y$ (Fig. 10.34)

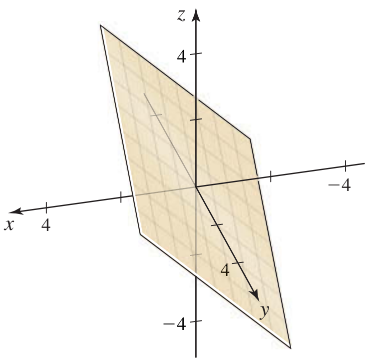


Figure 10.34

10.1.3

In Problems 31–34 use nine evenly spaced points and five colors to draw heat maps of the following functions, defined on their specified domains.

31. $f(x, y) = x^2 + y^2$ on $D = \{(x, y); 0 \leq x \leq 1, 0 \leq y \leq 1\}$

32. $f(x, y) = x - y$ on $D = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}$

33. $f(x, y) = x^2 - y^2$ on $D = \{(x, y) : -1 \leq x \leq 1, -1 \leq y \leq 1\}$

34. $f(x, y) = xy$ on $D = \{(x, y) : -1 \leq x \leq 1, -1 \leq y \leq 1\}$

35. **Heat-Index** At the beginning of this chapter we introduced the heat index as a way of calculating how temperature and humidity affect the apparent temperature. The equation for the heat index is:

$$H(T, R) = -42.38 + 2.049T + 10.14R - 6.838 \times 10^{-3}T^2 - 0.2248TR - 5.482 \times 10^{-2}R^2 + 1.229 \times 10^{-3}T^2R + 8.528 \times 10^{-4}TR^2 - 1.99 \times 10^{-6}T^2R^2$$

where T is the actual air temperature (in °F) and R is the relative humidity (in %). Using nine evenly spaced points and five colors, make a heat map for the heat index for the domain $D = \{(T, R) : 80 \leq T \leq 100, 40 \leq R \leq 60\}$. (You will find it easiest to calculate the heat index, H, if you program the formula for the heat index into a graphing calculator.)

36. **Heat-Index** Using the heat-map of the heat index from Figure 10.13, estimate the heat index for the following (T, R) values.

(a) $(T, R) = (90, 60)$

(b) $(T, R) = (100, 60)$

(c) $(T, R) = (90, 100)$

37. **Ecosystem Diversity** In Chapter 3 we introduced the Shannon diversity as a way of measuring the diversity of an ecosystem. Suppose that an ecosystem contains exactly three different species, and that a fraction p_1 of all organisms belong to species 1, a fraction p_2 belong to species 2, and a fraction p_3 belong to species 3:

(a) Explain why $p_1 + p_2 + p_3 = 1$.

(b) If we know p_1 and p_2, explain how we can calculate p_3.

(c) Explain why the domain of possible values for p_1 and p_2 is:

$$D = \{(p_1, p_2); p_1 \geq 0, p_2 \geq 0, \text{ and } p_1 + p_2 \leq 1\}$$

(d) For the ecosystem described in this problem the Shannon diversity index is defined to be:

$$S(p_1, p_2) = -p_1 \ln p_1 - p_2 \ln p_2 - (1 - p_1 - p_2) \ln(1 - p_1 - p_2) \tag{10.4}$$

on the domain D defined in part (c)

Figure 10.35 shows the diversity index as a function of p_1 and p_2. Use the heat map to answer the following questions:

(i) Estimate $S(0.1, 0.1)$ from the heat map, and compare with the numerical value that you calculate using Equation (10.4).

(ii) What is the largest possible value of S, and for what values of p_1 and p_2 is this value attained?

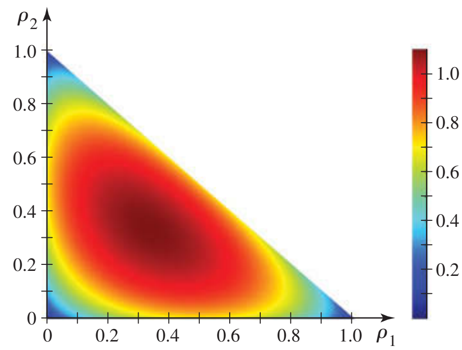


Figure 10.35 Heat map of the Shannon diversity index for Problem 37.

38. **Fitness Landscapes** The evolution of traits can be understood using the concept of a fitness landscape, which is a plot showing how fitness (the likelihood of the organism surviving to reproductive age) depends on the traits that are evolving over time. ***Martin and Wainwright (2013)*** directly measured the fitness landscape of a genus of pupfish by studying a population of hybrid pupfish that were bred to have different jaw lengths and head sizes. (Jaw length and head size are measured using units that allow both quantities to be negative.) Their data showing how fitness depends on these two traits is shown in Figure 10.36.

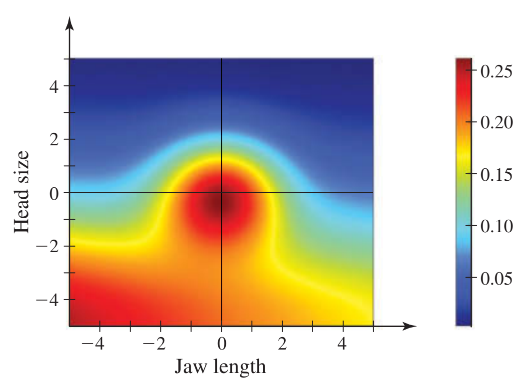


Figure 10.36 A fitness landscape shows how the probability of a pupfish surviving to adulthood depends on two traits (Problem 38). Adapted from ***Martin and Wainwright (2013)***.

(a) From the heat map estimate the survival probability for a fish whose jaw length is 0 and head size is 0.

(b) From the heat map estimate the survival probability for a fish whose jaw length is −5 and head size is −2.

(c) What is the maximum survival probability of a pupfish, and what combination of jaw length and head length produces this maximum survival probability?

10.1.4

In Problems 39–44 determine the equation of the level curves $f(x, y) = c$ and sketch the level curves for the specified values of c.

39. $f(x, y) = x + y;\ c = 0, -1, 1$

40. $f(x, y) = y - x^2;\ c = 0, 1, 2$

41. $f(x, y) = xy;\ c = 0, 1, 2$

42. $f(x, y) = x^2 - y^2;\ c = 0, 1, -1$

43. $f(x, y) = y/2;\ c = 0, 1, 2$

44. $f(x, y) = \frac{x-y}{x+y};\ c = 0, 1, 2$

45. Let

$$f_a(x, y) = ax^2 + y^2$$

for $(x, y) \in \mathbf{R}$, where a is a positive constant.

(a) Assume that $a = 1$ and describe the level curves of f_1. The graph of $f_1(x, y)$ intersects both the x–z and the y–z planes; show that these two curves of intersection are parabolas.

(b) Assume that $a = 4$. Then

$$f_4(x, y) = 4x^2 + y^2$$

and the level curves satisfy

$$4x^2 + y^2 = c$$

Use a graphing calculator to sketch the level curves for $c = 0, 1, 2, 3$, and 4. These curves are ellipses. Find the curves of intersection of $f_4(x, y)$ with the x–z and the y–z planes.

(c) Repeat (b) for $a = 1/4$.

(d) Explain in words how the surfaces of $f_a(x, y)$ change when a changes.

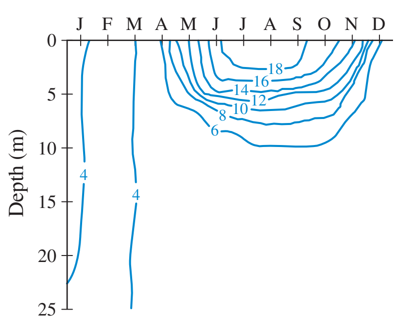


Figure 10.37 Isotherms for a typical lake in the Northern Hemisphere.

46. The graph in Figure 10.37 shows isotherms of a lake, as a function of depth and of the time of year.

(a) Use this plot to sketch the temperature profiles in March and June. That is, plot the temperature as a function of depth for a day in March and for a day in June.

(b) Explain how it follows from your temperature plots that the lake is **homeothermic**—that is, has the same temperature from the surface to the bottom—in March.

(c) Explain how it follows from your temperature plots that the lake is **stratified**—that is, has a warm layer on top, followed by a region where the temperature changes quickly, followed by a cold layer deeper down—in June.

47. Figure 10.38 shows the oxygen concentration for Long Lake, Clear Water County (Minnesota). The water flea *Daphnia* can survive only if the oxygen concentration is higher than 3 mg/l. Suppose that you wanted to sample the *Daphnia* population in 1997 on days 180, 200, and 220. What is the maximum depth that *Daphnia* can survive at on each of these days?

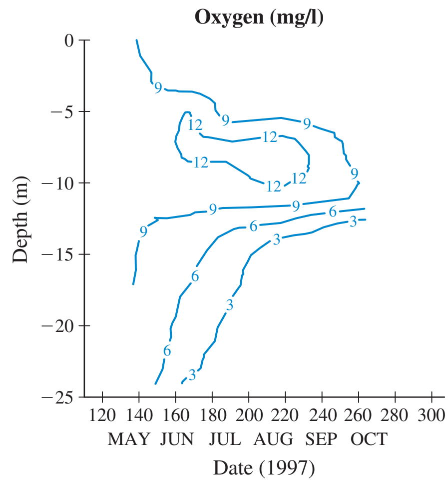


Figure 10.38 Level curves for oxygen concentration on Long Lake, Clear Water County (Minnesota). Courtesy of Leif Hembre.

48. **Diversity of the Deep Ocean** ***Woolley et al. (2016)*** studied the diversity of brittle stars (a close relative of starfish) living in the deep ocean to try to understand how many undiscovered species may be hidden in the ocean. Their data, showing how brittle star diversity depends upon depth and latitude, is shown in Figure 10.39.

(a) What is the diversity of brittle stars at a depth of 5500 m and at a latitude of 10°N?

(b) What is the diversity of brittle stars at a depth of 3000 m and at a latitude of 10°N?

(c) What is the maximum diversity of brittle stars, and at what depth and latitude is this maximum diversity attained?

(d) What is the minimum diversity of brittle stars, and in what interval of latitudes is this minimum diversity seen?

(e) Make a sketch of a graph showing how brittle star diversity varies with ocean depth at a latitude of −62.5°N.

(f) Make a sketch of a graph showing how brittle star diversity varies with latitude at an ocean depth of 3500 m.

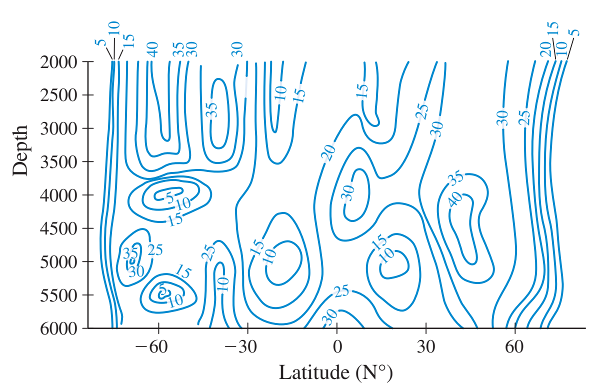


Figure 10.39 Diversity of brittle stars in the deep ocean (Problem 48). Adapted from ***Woolley et al. (2016).***

49. **Ultrasound Surgery** A promising way to destroy deep soft tissue tumors is to use ultrasound (high-frequency sound). An ultrasound beam is focused into the tumor, heating the tissue and destroying the cells in the tumor. The effectiveness of this approach depends on how tightly the ultrasound beam (and thus the heating effect) can be focused, and how much surrounding non-tumor tissue will be destroyed. ***Cline et al. (1994)*** directly measured the heating effect in cow muscle that was subjected to ultrasound surgery. Some of their data is shown by a level curve plot in Figure 10.40.

Use the level curves shown in Figure 10.40 to answer the following questions:

(a) What is the temperature increase at the point $(-3, 3)$?

(b) What is the temperature increase at the point $(1, 0)$?

(c) What is the maximum heating effect produced by the ultrasound beam?

(d) Assume that heating cells by 50°C is enough to kill them. What are the dimensions (height and width) of the region of cells that would be killed in Figure 10.40?

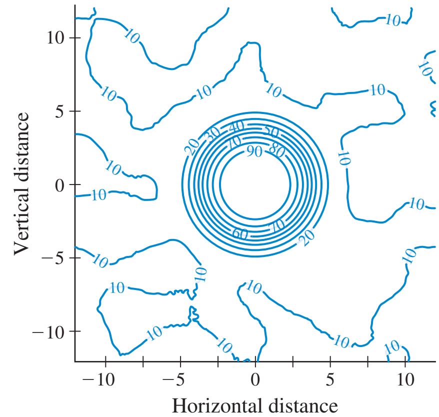


Figure 10.40 Thermal heating produced in ultrasound surgery; level curves show increase in temperature across an ultrasound beam. Adapted from ***Cline et al. (1994)***.

10.2 Limits and Continuity

Our first steps toward defining the derivative of a function of a single independent variable in Chapter 4 were taken when we defined limits in Chapter 3. To extend our definition of a derivative to multivariate functions, we must first extend the concepts of limits and continuity to the multivariable setting. However, the mechanical steps that are required to calculate these derivatives can actually be carried out without going deeply into the definitions of limits and continuity. Thus, it is possible to learn partial derivatives (which are covered in Section 10.3) without covering the material in this section first. We therefore regard this section as optional, except for students who plan to study multivariable calculus beyond the material covered in this textbook. Even these students may wish to skip Section 10.2.3, which covers the formal definition of limits, analogously to Section 3.6.

We will discuss only the two-dimensional case, but note that everything in this section can be generalized to higher dimensions.

10.2.1 Informal Definition of Limits

Let's start with an informal definition of limits. We say that the "limit of $f(x, y)$ as (x, y) approaches (x_0, y_0) is equal to L" if $f(x, y)$ can be made arbitrarily close to L whenever the point (x, y) is sufficiently close (but not equal) to the point (x_0, y_0). We denote this concept by

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L$$

As in the one-dimensional case in Chapter 3, there is a formal definition of limits, which is difficult to use. Fortunately, laws similar to those in the one-dimensional case allow us to compute limits in the two-dimensional case. We thus extend the limit laws from Chapter 3 to the two-dimensional case.

Limit Laws for the Two-Dimensional Case If a is a constant and if

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L_1 \qquad \text{and} \qquad \lim_{(x,y)\to(x_0,y_0)} g(x, y) = L_2$$

where L_1 and L_2 are real numbers, then the following hold:

1. Addition Rule

$$\lim_{(x,y)\to(x_0,y_0)} [f(x, y) + g(x, y)] = \lim_{(x,y)\to(x_0,y_0)} f(x, y) + \lim_{(x,y)\to(x_0,y_0)} g(x, y) = L_1 + L_2$$

2. **Constant-Factor Rule**

$$\lim_{(x,y)\to(x_0,y_0)} af(x,y) = a \lim_{(x,y)\to(x_0,y_0)} f(x,y) = aL_1$$

3. **Multiplication Rule**

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)g(x,y) = \left[\lim_{(x,y)\to(x_0,y_0)} f(x,y)\right]\left[\lim_{(x,y)\to(x_0,y_0)} g(x,y)\right] = L_1L_2$$

4. **Quotient Rule**

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y)}{g(x,y)} = \frac{\lim_{(x,y)\to(x_0,y_0)} f(x,y)}{\lim_{(x,y)\to(x_0,y_0)} g(x,y)} = \frac{L_1}{L_2}$$

provided that $L_2 \neq 0$.

The next two examples show how to compute limits by using the limit laws. In each example, we use the limit laws to break the function whose limit we wish to determine into functions of only x or only y. The limits of these functions can be calculated using the techniques that we learned in Chapter 3.

Limits of Polynomials When the Limits Exist.

EXAMPLE 1

(a) $$\lim_{(x,y)\to(0,0)}(x^2+y^2) = \lim_{(x,y)\to(0,0)} x^2 + \lim_{(x,y)\to(0,0)} y^2$$ Addition Rule

$$= 0^2 + 0^2 = 0.$$ First limit is a function of x only. Second limit is a function of y only.

(b) $$\lim_{(x,y)\to(4,-3)}(x^2+y^2) = \lim_{(x,y)\to(4,-3)} x^2 + \lim_{(x,y)\to(4,-3)} y^2$$ First limit is a function of x only. Second limit is a function of y only.

$$= 4^2 + (-3)^2 = 25.$$

(c) $$\lim_{(x,y)\to(-1,2)} x^2y = \left(\lim_{(x,y)\to(-1,2)} x^2\right)\left(\lim_{(x,y)\to(-1,2)} y\right)$$ Multiplication rule

$$= (-1)^2(2) = 2.$$ First limit is a function of x only. Second limit is a function of y only.

(d) $$\lim_{(x,y)\to(1,2)}(x^2y+3x) = \left(\lim_{(x,y)\to(1,2)} x^2y\right) + \left(3\lim_{(x,y)\to(1,2)} x\right)$$ Constant-factor rule on second term

$$= \left(\lim_{(x,y)\to(1,2)} x^2\right)\left(\lim_{(x,y)\to(1,2)} y\right) + \left(3\lim_{(x,y)\to(1,2)} x\right)$$

$$= (1)^2(2) + (3)(1) = 5.$$ •

We see from Example 1 that limits of polynomials are calculated by evaluating the functions at the respective points, provided that the functions are defined at those points. This is the case for rational functions as well.

Limits of Rational Functions When the Limits Exist.

EXAMPLE 2

(a) $$\lim_{(x,y)\to(-1,3)} \frac{3x}{y} = \frac{(3)(-1)}{3} = -1$$ Quotient rule

(b) $$\lim_{(x,y)\to(2,0)} \frac{4y+2x}{x^2+2xy-3} = \frac{(4)(0)+(2)(2)}{(2)^2+(2)(2)(0)-3} = \frac{4}{4-3} = 4$$ •

Limits That Do Not Exist. In the one-dimensional case, there were only two ways in which we could approach a number: from the left or from the right. If the two limits were different, we said that the limit did not exist. In two dimensions, there are many more ways that we can approach the point (x_0, y_0), namely, by any curve in the x–y plane that ends up at the point (x_0, y_0). We call such curves **paths**.

> If $f(x, y)$ approaches L_1 as $(x, y) \to (x_0, y_0)$ along path C_1 and $f(x, y)$ approaches L_2 as $(x, y) \to (x_0, y_0)$ along path C_2, and if $L_1 \neq L_2$, then $\lim_{(x,y)\to(x_0,y_0)} f(x, y)$ does not exist.

EXAMPLE 3 Show that $\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2}$ does not exist. Since $\lim_{(x,y)\to(0,0)} x^2 + y^2 = 0$, we cannot use the Quotient Rule to calculate a limit. In fact, as $(x, y) \to (0, 0)$ both $x^2 - y^2 \to 0$ and $x^2 + y^2 \to 0$, so the limit is of the form $\frac{0}{0}$.

Solution We first let $(x, y) \to (0, 0)$ along the positive x-axis; this is the curve C_1 in Figure 10.41. On C_1, $y = 0$ and $x > 0$. Then

$$\lim_{x\to 0+} \frac{x^2}{x^2} = 1$$

Next, we let $(x, y) \to (0, 0)$ along the positive y-axis; this is the curve C_2 in Figure 10.41. On C_2, $x = 0$ and $y > 0$. Then

$$\lim_{y\to 0+} \frac{-y^2}{y^2} = -1$$

Since $1 \neq -1$, we conclude that the limit does not exist. •

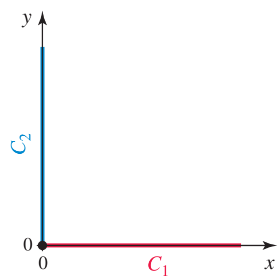


Figure 10.41 The paths C_1 and C_2 for Example 3.

Unless we have a lot of experience, it is not easy to find paths for which limits differ. Therefore, in the Problems section, in the cases where the limit does not exist, we will always provide the paths along which you should check the limits.

EXAMPLE 4 Show that

$$\lim_{(x,y)\to(0,0)} \frac{4x}{x + y^2}$$

does not exist.

Solution A natural choice for paths to $(0, 0)$ are straight lines of the form $y = mx$. If we substitute $y = mx$ in the preceding limit, then $(x, y) \to (0, 0)$ reduces to $x \to 0$ and we find that

$$\lim_{x\to 0} \frac{4x}{x + (mx)^2} = \lim_{x\to 0} \frac{4}{1 + m^2 x} = 4$$

That is, as long as we approach $(0, 0)$ along the straight line $y = mx$, the limit is always 4, irrespective of m. Such a path, labeled C_1, is shown in Figure 10.42.

You might be tempted to say that the limit exists. But not all straight line paths can be written in the form $y = mx$. In particular, if $x = 0$ (that is, we approach $(0, 0)$ along the y-axis shown as path C_2 in Figure 10.42), then

$$\frac{4x}{x + y^2} = \frac{0}{0 + y^2} = 0$$

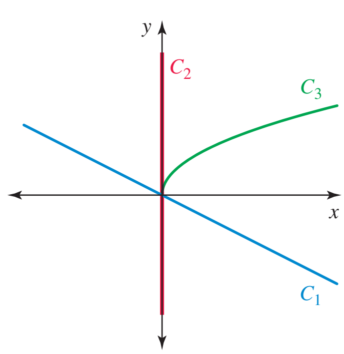


Figure 10.42 The paths C_1 and C_2 for Example 4.

So $\lim_{(x,y)\to(0,0)} \frac{4x}{x+y^2} = 0$ along this path. Since $0 \neq 4$ we have found paths along which the limits differ. Therefore, the limit does not exist.

In fact, the situation is even more complicated: Let's approach $(0, 0)$ along a curved path. We will use the parabola; $x = y^2$. This is the curve C_3 in Figure 10.42. Substituting y^2 for x then yields

$$\lim_{y\to 0} \frac{4y^2}{y^2 + y^2} = 2$$

So other limits are possible besides 0 and 4. •

To show that a limit does not exist, we must identify two paths along which the limits differ. To show that a limit exists, we cannot use paths, since the limits along *all* possible paths must be the same and there is no way to check all possible paths. Accordingly, to show that a limit exists, we proceed as in the one-dimensional case: We combine the formal definition of limits and the limit laws to rigorously compute limits.

10.2.2 Continuity

The definition of continuity is also analogous to that in the one-dimensional case:

> A function $f(x, y)$ is **continuous** at (x_0, y_0) if the following hold:
>
> **1.** $f(x, y)$ is defined at (x_0, y_0).
>
> **2.** $\lim\limits_{(x,y)\to(x_0,y_0)} f(x, y)$ exists.
>
> **3.** $\lim\limits_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0)$.

EXAMPLE 5 Show that

$$f(x, y) = 2 + x^2 + y^2$$

is continuous at (0, 0).

1. $f(x, y)$ is defined at (0, 0); specifically,

$$f(0, 0) = 2$$

2. To show that the limit exists, we refer to Example 1(a), where we showed that

$$\lim_{(x,y)\to(0,0)} (x^2 + y^2) = 0$$

which indicates that the limit exists. By the Addition Rule:

$$\lim_{(x,y)\to(0,0)} f(x, y) = 2 + \lim_{(x,y)\to(0,0)} (x^2 + y^2)$$

and the limit exists.

3. Using the fact that

$$\lim_{(x,y)\to(0,0)} (x^2 + y^2) = 0$$

we find that

$$\lim_{(x,y)\to(0,0)} (2 + x^2 + y^2) = 2 + \lim_{(x,y)\to(0,0)} (x^2 + y^2) = 2 + 0 = 2$$

Since $f(0, 0) = 2$, $f(x, y)$ is continuous at (0, 0). •

EXAMPLE 6 Show that

$$f(x, y) = \begin{cases} \frac{x^2 - y^2}{x^2 + y^2} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{for } (x, y) = (0, 0) \end{cases}$$

is discontinuous at (0, 0).

Solution The function $f(x, y)$ is defined at (0,0). Hence, condition 1 of the definition of continuity holds. But in Example 3 we showed that

$$\lim_{(x,y)\to(0,0)} f(x, y)$$

does not exist. Therefore, condition 2 of the definition is violated, and we conclude that $f(x, y)$ is discontinuous at $(x, y) = (0, 0)$. •

Composition of Functions. Using the definition of continuity and the rules for finding limits, we can show that polynomial functions of two variables (i.e., functions that are sums of terms of the form $ax^n y^m$, where a is a constant and n and m are nonnegative integers) are continuous.

We can obtain a much larger class of continuous functions when we allow the composition of functions. For instance, the function $h(x, y) = e^{x^2+y^2}$ can be written as a composition of two functions. To see this, set $z = f(x, y) = x^2 + y^2$ and $g(z) = e^z$. Then

$$h(x, y) = (g \circ f)(x, y) = g[f(x, y)] = e^{x^2+y^2}$$

Another example is $h(x, y) = \sqrt{y^2 - x}$. Here, $z = f(x, y) = y^2 - x$ and $g(z) = \sqrt{z}$. Then $h(x, y) = (g \circ f)(x, y)$. More generally,

> **Continuity of Composed Functions:**
>
> If:
> $$f : D \to \mathbf{R}, \quad D \subset \mathbf{R}^2$$
> and
> $$g : I \to \mathbf{R}, \quad I \subset \mathbf{R}$$
> then the composition $(g \circ f)(x, y)$ is defined as
> $$h(x, y) = (g \circ f)(x, y) = g[f(x, y)]$$
> If f is continuous at (x_0, y_0) and g is continuous at $z = f(x_0, y_0)$, then
> $$h(x, y) = (g \circ f)(x, y) = g[f(x, y)]$$
> is continuous at (x_0, y_0).

As an example, let's return to the function

$$h(x, y) = e^{x^2+y^2}$$

With $z = f(x, y) = x^2 + y^2$ and $g(z) = e^z$, $h(x, y) = g[f(x, y)]$ is continuous, since $f(x, y)$ is continuous for all $(x, y) \in \mathbf{R}^2$ and $g(z)$ is continuous for all z in the range of $f(x, y)$. Likewise,

$$h(x, y) = \sqrt{y^2 - x}$$

is continuous for all $(x, y) \in \{(x, y) : y^2 - x \geq 0\}$, since $z = f(x, y) = y^2 - x$ is continuous for all $(x, y) \in \mathbf{R}^2$ and $g(z) = \sqrt{z}$ is continuous for all $z \geq 0$.

10.2.3 Formal Definition of Limits

We recall the formal definition of limits in the one-dimensional case. To show that

$$\lim_{x \to 2} x^2 = 4$$

we must show that whenever x is close (but not equal) to 2, then x^2 is close to 4. Formally, this means that we must show that, for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|x^2 - 4| < \epsilon \qquad \text{whenever} \qquad 0 < |x - 2| < \delta$$

To generalize this idea to higher dimensions, we need to generalize what it means for x to be close to 2. In the one-dimensional case, we chose an interval centered at $x = 2$ that contained all points within a distance δ of the center of the interval, except for the center of the interval. In two dimensions, we replace the interval by a disk. A disk of radius δ centered at the point (x_0, y_0) is the set of all points that are within a distance δ of (x_0, y_0).

As with open and closed intervals, there are open and closed disks. We have the following definition:

Definition An **open disk** with radius r centered at $(x_0, y_0) \in \mathbf{R}^2$ is the set

$$B_r(x_0, y_0) = \left\{(x, y) \in \mathbf{R}^2 : \sqrt{(x - x_0)^2 + (y - y_0)^2} < r\right\}$$

A **closed disk** with radius r centered at $(x_0, y_0) \in \mathbf{R}^2$ is the set

$$\overline{B}_r(x_0, y_0) = \left\{(x, y) \in \mathbf{R}^2 : \sqrt{(x - x_0)^2 + (y - y_0)^2} \le r\right\}$$

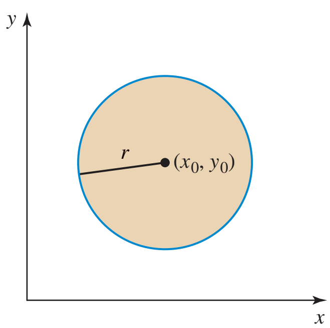


Figure 10.43 A closed disk with radius r centered at (x_0, y_0).

An open disk is thus a disk in which the boundary line is not part of the disk, whereas a closed disk contains the boundary line. This concept is analogous to that of intervals: An open interval does not contain the two endpoints (which are the boundary of the interval), whereas a closed interval does contain them. A closed disk is shown in Figure 10.43.

When we defined $\lim_{x \to x_0} f(x)$, we emphasized that the value of f at x_0 is not important. This idea is expressed in the formal ϵ–δ definition by excluding x_0 from the interval $(x_0 - \delta, x_0 + \delta)$, stated as $0 < |x - x_0| < \delta$. We will need to do the same when we generalize the limit to two dimensions. We write $B_\delta(x_0, y_0) - \{(x_0, y_0)\}$ to denote the open disk with radius δ centered at (x_0, y_0), where the center (x_0, y_0) is removed. We can now generalize the notion of limits to two dimensions:

Definition The limit of $f(x, y)$ as (x, y) approaches (x_0, y_0), denoted by

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y)$$

is the number L such that, for every $\epsilon > 0$, there exists a $\delta > 0$ so that

$$|f(x, y) - L| < \epsilon \qquad \text{whenever} \qquad (x, y) \in B_\delta(x_0, y_0) - \{(x_0, y_0)\}$$

This definition is similar to the one in one dimension: We require that whenever (x, y) is close (but not equal) to (x_0, y_0), it follows that $f(x, y)$ is close to L.

We provide one example in which we use the formal definition of limits. Using the formal definition is difficult, and we will not need to do so subsequently. In practice it is sufficient to be familiar with the limit laws that were given in Section 10.2.1. But seeing an example might help you to understand the definition.

EXAMPLE 7 Let $f(x, y) = x^2 + y^2$. Show that

$$\lim_{(x,y)\to(0,0)} f(x, y) = 0$$

Solution We must show that, for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|x^2 + y^2 - 0| < \epsilon \qquad \text{whenever} \qquad (x, y) \in B_\delta(0, 0) - \{(0, 0)\}$$

Now, $(x, y) \in B_\delta(0, 0)$ if $\sqrt{x^2 + y^2} < \delta$, or $x^2 + y^2 < \delta^2$. We need to show that for some value of δ we can define $B_\delta(0, 0)$ so that, if $(x, y) \in B_\delta(0, 0) - \{(0, 0)\}$, then $|x^2 + y^2| < \epsilon$. This suggests that we should choose δ so that $\delta^2 = \epsilon$. Let's try this. We set $\delta = \sqrt{\epsilon}$ for $\epsilon > 0$. Then

$$\sqrt{x^2 + y^2} < \delta = \sqrt{\epsilon}$$

implies that

$$|x^2 + y^2| < \epsilon$$

But this is exactly what we need to show. In other words, we have shown that, for every $\epsilon > 0$, we can find a $\delta > 0$ (namely $\delta = \sqrt{\epsilon}$) such that whenever (x, y) is close to $(0, 0)$, it follows that $x^2 + y^2$ is close to 0. •

Section 10.2 Problems

10.2.1

In Problems 1–14, use the properties of limits to calculate the following limits:

1. $\lim_{(x,y)\to(1,0)}(x^2 - 3y^2)$

2. $\lim_{(x,y)\to(-1,1)}(2xy + y^2)$

3. $\lim_{(x,y)\to(2,-1)}(x^2y^3 - xy)$

4. $\lim_{(x,y)\to(1,-1)}(2x^3 - 3y)(xy + 1)$

5. $\lim_{(x,y)\to(-1,3)} x^2(y^2 - 3xy)$

6. $\lim_{(x,y)\to(-5,1)} y(xy + x^2y^2)$

7. $\lim_{(x,y)\to(0,2)} \left(4xy^2 - \frac{x+1}{y^2}\right)$

8. $\lim_{(x,y)\to(1,1)} \frac{xy}{x^2+y^2}$

9. $\lim_{(x,y)\to(2,1)} \frac{x^2+y^2}{x^2-y^2}$

10. $\lim_{(x,y)\to(-1,1)} \frac{x^2+y}{2x+y}$

11. $\lim_{(x,y)\to(0,1)} \frac{2xy-3}{x^2+y^2+1}$

12. $\lim_{(x,y)\to(-1,-2)} \frac{x^2-y^2}{2xy+2}$

13. $\lim_{(x,y)\to(2,0)} \frac{2x+4y^2}{y^2+3x}$

14. $\lim_{(x,y)\to(1,-2)} \frac{2x^2+y}{2xy+3}$

15. Show that

$$\lim_{(x,y)\to(0,0)} \frac{2x^2 - y^2}{x^2 + y^2}$$

does not exist by computing the limit along the positive x-axis and the positive y-axis.

16. Show that

$$\lim_{(x,y)\to(0,0)} \frac{(x - y)(x + 2y)}{x^2 + y^2}$$

does not exist by computing the limit along the positive x-axis and the positive y-axis.

17. Compute

$$\lim_{(x,y)\to(0,0)} \frac{4xy}{x^2 + y^2}$$

along the x-axis, the y-axis, and the line $y = x$. What can you conclude?

18. Compute

$$\lim_{(x,y)\to(0,0)} \frac{3x(y + x)}{x^2 + y^3}$$

along lines of the form $y = mx$, for $m \neq 0$. What can you conclude?

19. Compute

$$\lim_{(x,y)\to(0,0)} \frac{2xy}{x^3 + yx}$$

along lines of the form $y = mx$, for $m \neq 0$, and along the parabola $y = x^2$. What can you conclude?

20. Compute

$$\lim_{(x,y)\to(0,0)} \frac{3x^2y}{(2x^4 + y^2)}$$

along lines of the form $y = mx$, for $m \neq 0$, and along the parabola $y = x^2$. What can you conclude?

10.2.2

21. Use the definition of continuity to show that

$$f(x, y) = 2x^2 + y^2 + 1$$

is continuous at $(0, 0)$.

22. Use the definition of continuity to show that

$$f(x, y) = \sqrt{9 + x^2 + y^2}$$

is continuous at $(0, 0)$.

23. Show that

$$f(x, y) = \begin{cases} \frac{4xy}{x^2+y^2} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{for } (x, y) = (0, 0) \end{cases}$$

is discontinuous at $(0, 0)$. (*Hint*: Use Problem 17.)

24. Show that

$$f(x, y) = \begin{cases} \frac{3x(y+x)}{x^2+y^3} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{for } (x, y) = (0, 0) \end{cases}$$

is discontinuous at $(0, 0)$. (*Hint*: Use Problem 18.)

25. Show that

$$f(x, y) = \begin{cases} \frac{2xy}{x^3+yx} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{for } (x, y) = (0, 0) \end{cases}$$

is discontinuous at $(0, 0)$. (*Hint*: Use Problem 19.)

26. Show that

$$f(x, y) = \begin{cases} \frac{3x^2y}{(2x^4+y^2)} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{for } (x, y) = (0, 0) \end{cases}$$

is discontinuous at $(0, 0)$. (*Hint*: Use Problem 20.)

27. (a) Write

$$h(x, y) = \sin(x^2 + y^2)$$

as a composition of two functions.

(b) For which values of (x, y) is $h(x, y)$ continuous?

28. (a) Write

$$h(x, y) = \sqrt{x + y}$$

as a composition of two functions.

(b) For which values of (x, y) is $h(x, y)$ continuous?

29. (a) Write

$$h(x, y) = e^{xy}$$

as a composition of two functions.

(b) For which values of (x, y) is $h(x, y)$ continuous?

30. (a) Write

$$h(x, y) = \cos(y - x)$$

as a composition of two functions.

(b) For which values of (x, y) is $h(x, y)$ continuous?

10.2.3

31. Draw an open disk with radius 2 centered at $(1, -1)$ in the x–y plane, and give a mathematical description of this set.

32. Draw a closed disk with radius 3 centered at $(2, 0)$ in the x–y plane, and give a mathematical description of this set.

33. Give a geometric interpretation of the set

$$A = \left\{(x, y) \in \mathbf{R}^2 : \sqrt{x^2 + y^2 - 4y + 4} < 3\right\}$$

34. Give a geometric interpretation of the set

$$A = \left\{(x, y) \in \mathbf{R}^2 : \sqrt{x^2 + 6x + y^2 - 2y + 10} < 2\right\}$$

35. Let

$$f(x, y) = 2x^2 + y^2$$

Use the ϵ–δ definition of limits to show that

$$\lim_{(x,y)\to(0,0)} f(x, y) = 0$$

36. Let

$$f(x, y) = 2x^2 + 3y^2$$

Use the ϵ–δ definition of limits to show that

$$\lim_{(x,y)\to(0,0)} f(x, y) = 0$$

10.3 Partial Derivatives

10.3.1 Functions of Two Variables

Suppose that the behavior of an organism depends on a number of independent variables. To investigate this dependency, a common experimental design is to measure the response when one variable is changed while all other variables are kept fixed. As an example, ***Pisek et al. (1969)*** measured the net assimilation of CO_2 by buttercups (*Ranunculus glacialis*) as a function of temperature and light intensity. In their experiments, they varied the temperature while keeping the light intensity constant. Repeating this experiment at different light intensities, they were able to determine how the net assimilation of CO_2 changes as a function of both temperature and light intensity.

This experimental design illustrates the idea behind partial derivatives. Suppose that we want to know how the function $f(x, y)$ changes when x and y change. Instead of changing both variables simultaneously, we might get an idea of how $f(x, y)$ depends on x and y when we change one variable while keeping the other variable fixed. In Section 10.1 we explored this idea by making graphs of the function $f(x, y)$ as a function of x only, or as a function of y. Now we will use calculus to study the same ideas.

To illustrate, we look at

$$f(x, y) = x^2 y$$

We want to know how $f(x, y)$ changes if we change, say, x and keep y fixed. So we fix $y = y_0$. Then the change in f with respect to x is simply the derivative of f with respect to x when $y = y_0$. That is,

$$\frac{d}{dx} f(x, y_0) = \frac{d}{dx} x^2 y_0 = 2xy_0$$

Such a derivative is called a *partial derivative*.

Definition Suppose that f is a function of two independent variables x and y. The **partial derivative** of f with respect to x is defined by

$$\frac{\partial f(x, y)}{\partial x} = \lim_{h\to 0} \frac{f(x + h, y) - f(x, y)}{h}$$

The partial derivative of f with respect to y is defined by

$$\frac{\partial f(x, y)}{\partial y} = \lim_{h\to 0} \frac{f(x, y + h) - f(x, y)}{h}$$

To denote partial derivatives, we use "∂" instead of "d." We will also use the notation

$$f_x(x, y) = \frac{\partial f(x, y)}{\partial x} \qquad \text{and} \qquad f_y(x, y) = \frac{\partial f(x, y)}{\partial y}$$

You should see the similarity between the definition of partial derivatives, and the formal definition of derivatives of Chapter 4. That is, to compute $\partial f/\partial x$, we look at the quotient given when we divide the difference in the f-values, $f(x+h, y) - f(x, y)$, by the difference in the x-values, $(x+h) - x$. The other variable, y, is not changed. We then let h tend to 0.

To compute $\partial f(x, y)/\partial x$, we differentiate f with respect to x while treating y as a constant. When we read $\partial f(x, y)/\partial x$, we can say "the partial derivative of f of x and y with respect to x." To read $f_x(x, y)$, we say "f sub x of x and y."

Finding partial derivatives is no different from finding derivatives of functions of one variable, since, by keeping all but one variable fixed, computing a partial derivative is equivalent to computing a derivative of a function of one variable. We treat the fixed variable as a coefficient, carrying out the derivative similarly to the functions we encountered in Section 4.4 that contained constants. You just need to keep straight which of the variables you have fixed and which one you will vary.

EXAMPLE 1 Find $\partial f/\partial x$ and $\partial f/\partial y$ when

$$f(x, y) = ye^{xy}.$$

Solution To compute $\partial f/\partial x$, we treat y as a constant and use the chain rule to differentiate f with respect to x:

$$\frac{\partial f(x, y)}{\partial x} = \frac{\partial}{\partial x}(ye^{xy}) = ye^{xy}y = y^2e^{xy} \qquad \text{Treat } y \text{ as a constant.}$$

To compute $\partial f/\partial y$, we treat x as a constant and use the product rule combined with the chain rule to differentiate f with respect to y:

$$\begin{aligned}\frac{\partial f(x, y)}{\partial y} &= \frac{\partial}{\partial y}(ye^{xy}) \qquad \text{Treat } x \text{ as a constant.}\\ &= 1 \cdot e^{xy} + ye^{xy}x\\ &= e^{xy}(1 + xy).\end{aligned}$$

•

EXAMPLE 2 Find $\partial f/\partial x$ when

$$f(x, y) = \frac{\sin(xy)}{x^2 + \cos y}$$

Solution We treat y as a constant and use the quotient rule:

$$u = \sin(xy) \qquad v = x^2 + \cos y$$

$$\frac{\partial u}{\partial x} = y\cos(xy) \qquad \frac{\partial v}{\partial x} = 2x$$

Hence,

$$\frac{\partial f(x, y)}{\partial x} = \frac{y(x^2 + \cos y)\cos(xy) - 2x\sin(xy)}{(x^2 + \cos y)^2}.$$

•

Geometric Interpretation. As with ordinary derivatives, partial derivatives are slopes of lines that are tangent to certain curves. We can find these curves on the surface $z = f(x, y)$.

Let's start with the interpretation of $\partial f/\partial x$. We fix $y = y_0$; then $f(x, y_0)$ as a function of x is given geometrically by intersecting the surface $z = f(x, y)$ with a vertical plane that is parallel to the x–z plane and goes through $y = y_0$. The curve of intersection where the vertical plane meets the surface is the graph of $z = f(x, y_0)$, as illustrated in Figure 10.44.

We can now project this curve onto the x–z plane, as illustrated in Figure 10.45. The curve is the graph of a function that depends only on x. Consequently, we can find the slope of the tangent line at any point (x_0, z_0), where $z_0 = f(x_0, y_0)$. The slope of

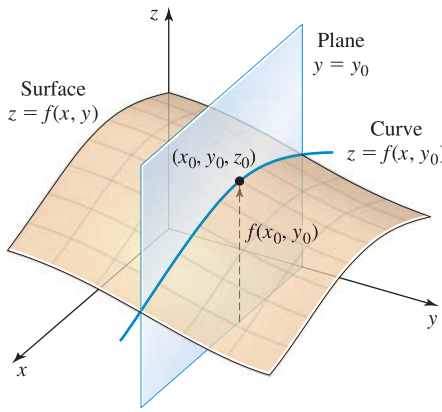


Figure 10.44 The surface of $f(x, y)$ intersected with the plane $y = y_0$.

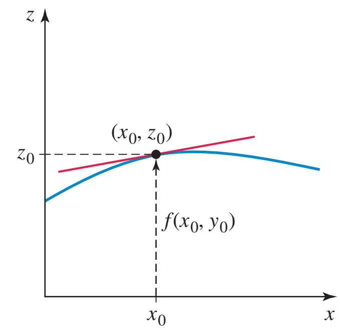


Figure 10.45 The projection of the curve $z = f(x, y_0)$ onto the x–z plane.

the tangent line is given by the derivative of the function f in the x-direction—that is, by $\partial f/\partial x$. Hence:

Geometric Meaning of $\frac{\partial f}{dx}$:

The partial derivative $\partial f/\partial x$ at (x_0, y_0) is the slope of the tangent line to the curve $z = f(x, y_0)$ at the point $x = x_0$.

The derivative $\partial f/\partial y$ has a similar meaning. This time, we intersect the surface $z = f(x, y)$ with the vertical plane $x = x_0$, which is parallel to the y–z plane and goes through $x = x_0$. The curve of intersection is the graph of $z = f(x_0, y)$, as illustrated in Figure 10.46.

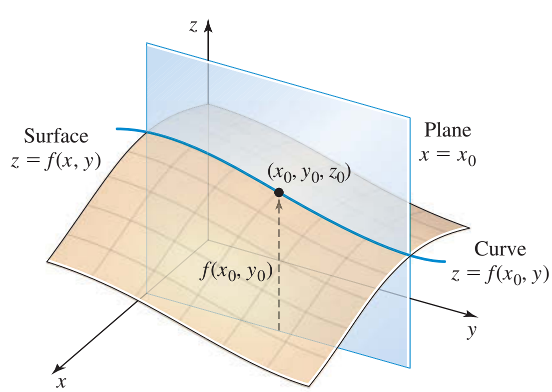


Figure 10.46 The surface of $f(x, y)$ intersected with the plane $x = x_0$.

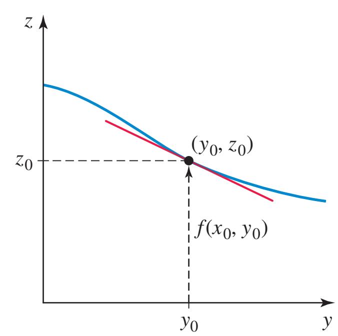


Figure 10.47 The projection of the curve $z = f(x_0, y)$ onto the y–z plane.

The projection of this curve onto the y–z plane, together with the tangent line at (y_0, z_0), is illustrated in Figure 10.47.

Geometric Meaning of $\frac{\partial f}{\partial y}$

The partial derivative $\partial f/\partial y$ evaluated at (x_0, y_0) is the slope of the tangent line to the curve $z = f(x_0, y)$ at the point $y = y_0$.

EXAMPLE 3 Let

$$f(x, y) = 3 - x^3 - y^2$$

Find $f_x(1, 1)$ and $f_y(1, 1)$, and interpret your results geometrically.

Solution We have

$$f_x(x, y) = -3x^2 \qquad \text{and} \qquad f_y(x, y) = -2y$$

Hence,

$$f_x(1, 1) = -3 \qquad \text{and} \qquad f_y(1, 1) = -2$$

To interpret $f_x(1, 1)$ geometrically, we fix $y = 1$. The vertical plane $y = 1$ intersects the graph of $f(x, y)$. The curve of intersection has slope -3 when $x = 1$. The projection of the curve of intersection is shown in Figure 10.48.

Similarly, to interpret $f_y(1, 1)$, we intersect the graph of $f(x, y)$ with the vertical plane $x = 1$. The tangent line at the curve of intersection has slope -2 when $y = 1$. The projection of the curve of intersection is shown in Figure 10.49. •

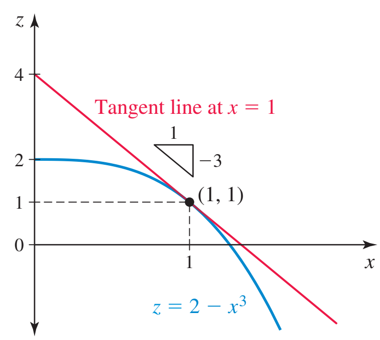


Figure 10.48 The curve of intersection of the graph $z = 3 - x^3 - y^2$ with the plane $y = 1$. The curve in the plane of intersection is shown, together with its tangent line at $(x, z) = (1, 1)$.

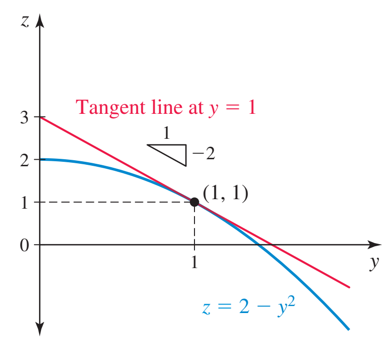


Figure 10.49 The curve of intersection of the graph $z = 3 - x^3 - y^2$ with the plane $x = 1$. The curve in the plane of intersection is shown, together with its tangent line at $(y, z) = (1, 1)$.

A Biological Application – Prey Capture

EXAMPLE 4 ***Holling (1959)*** derived an expression for the number of prey items P eaten by a predator during one unit of time as a function of the number of nearby prey N and the time that the predator needs to eat each prey (called T_h, or the handling time:)

$$P = \frac{aN}{1 + aT_hN} \tag{10.5}$$

Here, a is a positive constant called the *predator attack rate*. Equation (10.5) is called *Holling's disk equation*. [Holling came up with the equation when he measured how many sandpaper disks (representing prey) a blindfolded assistant (representing the predator) could pick up during a certain interval.] We can consider P as a function of N and T_h.

To determine how handling time influences the number of prey eaten, we compute

$$\begin{aligned}\frac{\partial P(N, T_h)}{\partial T_h} &= aN(-1)(1 + aT_hN)^{-2}aN \\ &= -\frac{a^2N^2}{(1 + aT_hN)^2} < 0\end{aligned}$$

since $a^2N^2 > 0$ and $(1 + aT_hN)^2 > 0$. Because $\partial P/\partial T_h$ is negative, the number of prey items eaten decreases with increasing handling time. That is, increasing the time the predator spends handling each prey item (i.e., increasing T_h) decreases the number of prey items it eats in unit time, P.

To determine how P changes with N, we compute

$$\begin{aligned}\frac{\partial P(N, T_h)}{\partial N} &= \frac{a(1+aT_hN) - aNaT_h}{(1+aT_hN)^2} &&\text{Quotient law}\\ &= \frac{a}{(1+aT_hN)^2} > 0\end{aligned}$$

since $aT > 0$ and $(1+aT_hN)^2 > 0$. Because $\partial P/\partial N$ is positive, the number of prey items eaten increases if the number of nearby prey is increased, again as expected. •

10.3.2 Functions of More Than Two Variables

The definition of partial derivatives extends in a straightforward way to functions of more than two variables. Partial derivatives are ordinary derivatives with respect to one variable while all other variables are treated as constants.

EXAMPLE 5 Let f be a function of three independent variables x, y, and z:

$$f(x, y, z) = e^{yz}(x^2 + z^3)$$

Find $\partial f/\partial x$, $\partial f/\partial y$, and $\partial f/\partial z$.

Solution To compute $\partial f/\partial x$, we fix y and z and differentiate f with respect to x.

$$\frac{\partial f}{\partial x} = e^{yz}2x. \qquad \text{Treat } y \text{ and } z \text{ as constants.}$$

Likewise,

$$\frac{\partial f}{\partial y} = ze^{yz}(x^2 + z^3). \qquad \text{Treat } x \text{ and } z \text{ as constants.}$$

$$\frac{\partial f}{\partial z} = ye^{yz}(x^2 + z^3) + e^{yz}3z^2 = e^{yz}(yx^2 + yz^3 + 3z^2). \qquad \text{Treat } x \text{ and } y \text{ as constants.}$$

•

10.3.3 Higher-Order Partial Derivatives

As in the case of functions of one variable, we can define higher-order partial derivatives for functions of more than one variable. For instance, to find the second partial derivative of $f(x, y)$ with respect to x, denoted by $\partial^2 f/\partial x^2$, we compute

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$$

We can write $\partial^2 f/\partial x^2$ as f_{xx}.

We can also compute **mixed derivatives**. For instance,

$$f_{yx} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

Note the order of yx in the subscript of f and the order of $\partial x\,\partial y$ in the denominator: Either notation means that we differentiate with respect to y first, and then differentiate with respect to x.

EXAMPLE 6 Set $f(x, y) = \sin x + xe^y$. Find f_{xx}, f_{yx}, and f_{xy}.

Solution

$$f_{xx} = \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial x}(\sin x + xe^y)\right] = \frac{\partial}{\partial x}[\cos x + e^y] = -\sin x.$$

$$f_{yx} = \frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}(\sin x + xe^y)\right] = \frac{\partial}{\partial x}[0 + xe^y] = e^y.$$

$$f_{xy} = \frac{\partial^2 f}{\partial y\,\partial x} = \frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial}{\partial y}\left[\frac{\partial}{\partial x}(\sin x + xe^y)\right] = \frac{\partial}{\partial y}[\cos x + e^y] = e^y.$$

•

In the preceding example, $f_{xy} = f_{yx}$, implying that the order of differentiation did not matter. Although this is not always the case, there are conditions under which the order of differentiation in mixed partial derivatives does not matter. To state this theorem, we need the notion of an open disk which was introduced in Section 10.2.3. We have the following definition:

Definition An **open disk** with radius r centered at $(x_0, y_0) \in \mathbf{R}^2$ is the set

$$B_r(x_0, y_0) = \left\{(x, y) \in \mathbf{R}^2 : \sqrt{(x - x_0)^2 + (y - y_0)^2} < r\right\}$$

An open disk is therefore made up of all points whose distance from (x_0, y_0) is less than r; i.e., it is the inside of a circle of radius r centered at (x_0, y_0).

We can now state the mixed-derivative theorem:

Theorem The Mixed-Derivative Theorem If $f(x, y)$ and its partial derivatives f_x, f_y, f_{xy}, and f_{yx} are continuous on an open disk centered at the point (x_0, y_0), then

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$$

It is even possible to define partial derivatives of even higher order. For instance, if f is a function of two independent variables x and y, then

$$\frac{\partial^3 f}{\partial x^2\, \partial y} = \frac{\partial}{\partial x} \frac{\partial^2 f}{\partial x\, \partial y} = \frac{\partial}{\partial x} \frac{\partial}{\partial x} \frac{\partial f}{\partial y}.$$

We illustrate this definition with

$$f(x, y) = y^2 \sin x$$

for which

$$\begin{aligned}\frac{\partial^3 f}{\partial x^2\, \partial y} &= \frac{\partial}{\partial x} \frac{\partial}{\partial x} \frac{\partial}{\partial y} (y^2 \sin x) \\ &= \frac{\partial}{\partial x} \frac{\partial}{\partial x} (2y \sin x) \\ &= \frac{\partial}{\partial x} (2y \cos x) = -2y \sin x\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^3 f}{\partial x\, \partial y^2} &= \frac{\partial}{\partial x} \frac{\partial}{\partial y} \frac{\partial}{\partial y} (y^2 \sin x) \\ &= \frac{\partial}{\partial x} \frac{\partial}{\partial y} (2y \sin x) = \frac{\partial}{\partial x} (2 \sin x) = 2 \cos x\end{aligned}$$

and so on.

In the applications of partial derivatives that we will meet in this book only the first order derivatives (f_x and f_y) and second order derivatives (f_{xx}, f_{xy}, f_{yx}, and f_{yy}) will be needed.

Section 10.3 Problems

10.3.1

In Problems 1–16, find $\partial f/\partial x$ and $\partial f/\partial y$ for the given functions.

1. $f(x, y) = x^2y + xy^2$

2. $f(x, y) = \frac{2x}{y} - \frac{3}{xy^2}$

3. $f(x, y) = (xy)^{3/2} - (xy)^{2/3}$

4. $f(x, y) = \frac{y^4}{x^3} - \frac{x^3}{y^4}$

5. $f(x, y) = \sin(x + y)$

6. $f(x, y) = \tan(x - 2y)$

7. $f(x, y) = \cos^2(x^2 - 2y)$

8. $f(x, y) = \sin^2(y^2x - x^3)$

9. $f(x, y) = e^{\sqrt{x+y}}$

10. $f(x, y) = x^2e^{x+2xy}$

11. $f(x, y) = e^x \sin(xy)$

12. $f(x, y) = e^{-y^2} \sin(x^2 + y^2)$

13. $f(x, y) = \ln(2x + y)$

14. $f(x, y) = \ln(3x^2 - xy)$

15. $f(x, y) = \ln\left(\frac{x^2+y}{y}\right)$

16. $f(x, y) = \ln\left(\frac{xy}{x^2+xy}\right)$

In Problems 17–24, find the indicated partial derivatives.

17. $f(x, y) = 3x^2 - y + 2y^2$; $f_x(1, 0)$

18. $f(x, y) = x^{1/3}y - xy^{-1/3}$; $f_y(1, 1)$

19. $g(x, y) = e^{x-y}$; $g_y(2, 1)$

20. $h(u, v) = e^u \sin(u + v)$; $h_u(1, -1)$

21. $f(x, z) = \ln(xz)$; $f_z(e, 1)$

22. $g(v, w) = \frac{w^2}{v+w}$; $g_v(1, 1)$

23. $f(x, y) = \frac{xy}{x^2+2}$; $f_x(-1, 2)$

24. $f(u, v) = e^{u+3v^2}$; $f_u(2, 1)$

25. Let

$$f(x, y) = 4 - x^2 - y^2$$

Compute $f_x(1, 1)$ and $f_y(1, 1)$, and interpret these partial derivatives geometrically.

26. Let

$$f(x, y) = \sqrt{4 - x^2 - y^2}$$

Compute $f_x(1, 1)$ and $f_y(1, 1)$, and interpret these partial derivatives geometrically.

27. Let

$$f(x, y) = 1 - x^2y + y^2$$

Compute $f_x(-2, 1)$ and $f_y(-2, 1)$, and interpret these partial derivatives geometrically.

28. Let

$$f(x, y) = 2x^3 - 3yx$$

Compute $f_x(1, 2)$ and $f_y(1, 2)$, and interpret these partial derivatives geometrically.

29. In Example 4, we investigated Holling's disk equation

$$P = \frac{aN}{1 + aT_hN}$$

(See Example 4 for the meaning of this equation.) We will now consider P as a function of the predator attack rate a. Determine how the predator attack rate a influences the number of prey eaten per predator.

30. Blood Oxygenation The oxygen content of blood depends on the partial pressure of oxygen in surrounding tissues (P) and on a reaction rate constant (K). Blood oxygenation is often modeled using Hill's equation, which predicts that the fraction of hemoglobin molecules in blood that are bound to oxygen will be given by a function of P and K:

$$f(P, K) = \frac{P^3}{K^3 + P^3}.$$

(a) Explain why, if $K > 0$ and $P \geq 0$, $f(P, K) < 1$ and $f(P, K) \geq 0$

(b) Use partial differentiation to determine the effect of increasing P on f.

(c) Use partial differentiation to determine the effect of increasing K on f.

10.3.2

In Problems 31–38, find $\partial f/\partial x$, $\partial f/\partial y$, and $\partial f/\partial z$ for the given functions.

31. $f(x, y, z) = x^2z + yz^2 - xy$

32. $f(x, y, z) = xy(z + x)$

33. $f(x, y, z) = x^3y^2z + \frac{x^2}{yz}$

34. $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$

35. $f(x, y, z) = e^{x+y+z}$

36. $f(x, y, z) = \sin(x + y - z)$

37. $f(x, y, z) = \ln(x + y + z)$

38. $f(x, y, z) = \exp(\sqrt{x^2 + y^2 + z^2})$

10.3.3

In Problems 39–48, find the indicated partial derivatives.

39. $f(x, y) = x^2y^2 + 2xy^2$; $\frac{\partial^2 f}{\partial x^2}$

40. $f(x, y) = y^2(x - 3y)$; $\frac{\partial^2 f}{\partial y^2}$

41. $f(x, y) = xe^y$; $\frac{\partial^2 f}{\partial x \partial y}$

42. $f(x, y) = \sin(x - y)$; $\frac{\partial^2 f}{\partial y \partial x}$

43. $f(u, w) = \tan(u + w)$; $\frac{\partial^2 f}{\partial u^2}$

44. $g(s, t) = \ln(s + t^2)$; $\frac{\partial^2 g}{\partial s^2}$

45. $f(x, y) = x \cos y$; $\frac{\partial^2 f}{\partial x \partial y}$

46. $f(x, y) = e^{x^2-y}$; $\frac{\partial^2 f}{\partial x \partial y}$

47. $f(x, y) = \ln(x + y)$; $\frac{\partial^2 f}{\partial x^2}$

48. $f(x, y) = \sin(3xy)$; $\frac{\partial^2 f}{\partial y^2}$

49. Predation Rate The prey-density responses of some predators are sigmoidal: the number of prey attacked has a sigmoidal shape when plotted as a function of prey density. If we denote the number of nearby prey by N, and the handling time of each prey item by T_h, then the rate of prey attacks per predator as a function of N and T_h can be expressed as

$$P(N, T_h) = \frac{b^2N^2}{1 + cN + bT_hN^2}$$

where b and c are positive constants.

(a) Investigate how an increase in the prey density N affects the function $P(N, T_h)$.

(b) Investigate how an increase in the handling time T_h affects the function $P(N, T_h)$.

(c) Graph $P(N, T_h)$ as a function of N when $T_h = 0.2$ hours, $b = 0.8$, and $c = 0.5$.

50. Parasite Interference Parasites live by stealing resources from hosts. When parasites reproduce their offspring must find new hosts. However, if a potential host is already infected by parasites, then new parasites will not be able to infect it. This leads to interference between parasites, and we will build a model for these effects in this Problem. We assume that N is the number of hosts in a given area, and P is the number of parasites. A frequently used model for host–parasite interactions is the **Nicholson–Bailey model** (see ***Nicholson and Bailey, 1935***), in which it is assumed that the number of parasitized hosts, denoted by N_a, is given by

$$N_a = N[1 - e^{-bP}] \tag{10.6}$$

where b is the searching efficiency.

(a) Let's treat N and P as independent variables and N_a as a function of N and P. By calculating the appropriate partial derivatives investigate how:

(i) an increase in the number of hosts N affects the number of parasitized hosts $N_a(N, P)$

(ii) an increase in the number of parasites affects the number of parasitized hosts $N_a(N, P)$

(b) Show that

$$b = \frac{1}{P} \ln \frac{N}{N - N_a}$$

by solving (10.6) for b.

(c) Consider

$$b = f(P, N, N_a) = \frac{1}{P} \ln \frac{N}{N - N_a}.$$

That is, we regard searching efficiency as a function of P, N, and N_a. How is the searching efficiency b affected when the, number of parasites increases?

51. Wind-Chill In Section 10.1, Problem 3, we introduced wind chill as a way of calculating the apparent temperature a person would feel as a function of the real air temperature, T, and V in mph. Then the wind chill (i.e., the apparent temperature) is:

$$W(T, V) = 35.74 + 0.6215T - 35.75V^{0.16}$$

$$0.4275TV^{0.16}$$

(a) By calculating the appropriate partial derivative, show that increasing T always increases W.

(b) Under what conditions does increasing V decrease W? Your answer will take the form of an inequality involving T.

(c) Assuming that W should always decrease when V is increased, use your answer from (b) to determine the largest domain in which this formula for W can be used.

10.4 Tangent Planes, Differentiability, and Linearization

When we first encountered differentiation of functions of a single variable we were able to show that the derivative had a geometric interpretation for the graph $y = f(x)$. In this section we will extend this idea to multivariate functions, and to a new category of functions; a **vector-valued function**, which takes on values that are represented by vectors. Vector-valued functions will be covered in more depth in Chapter 11, in which we discuss systems of differential equations, which can be used to model many different biological phenomena.

10.4.1 Functions of Two Variables

Tangent Planes. Suppose that $z = f(x)$ is differentiable at $x = x_0$. Then the equation of the tangent line of $z = f(x)$ at (x_0, z_0) with $z_0 = f(x_0)$ is given by

$$z - z_0 = f'(x_0)(x - x_0) \tag{10.7}$$

The curve z and the tangent line are illustrated in Figure 10.50.

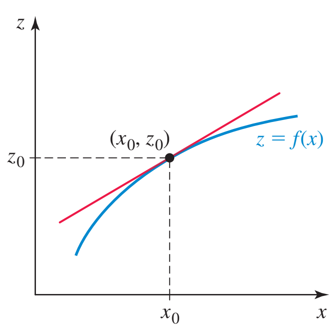


Figure 10.50 The curve $z = f(x)$ and its tangent line at the point (x_0, z_0).

We now generalize this situation to functions of two variables. The analogue of a tangent line is called a **tangent plane**, an example of which is shown in Figure 10.51. Let $z = f(x, y)$ be a function of two variables. We saw in the previous section that the partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$, evaluated at (x_0, y_0), are the slopes of tangent lines at the point (x_0, y_0, z_0), with $z_0 = f(x_0, y_0)$, to certain curves through (x_0, y_0, z_0) on the surface $z = f(x, y)$. These two tangent lines, one in the x-direction, the other in the y-direction, define a unique plane. If, in addition, $f(x, y)$ has partial derivatives that are continuous on an open disk containing (x_0, y_0), then we can show that at (x_0, y_0, z_0) the tangent line for any smooth curve on the surface $z = f(x, y)$ through (x_0, y_0, z_0) is contained in this plane. This plane is called the *tangent plane*.

We will use the two original tangent lines to find the equation of the tangent plane at a point (x_0, y_0, z_0) on the surface $z = f(x, y)$. We consider the curve that is obtained as the intersection of the surface $z = f(x, y)$ with the plane that is parallel to the y–z plane and contains the point (x_0, y_0, z_0)—that is, the plane $x = x_0$—and we denote this curve by C_1. Its tangent line at (x_0, y_0, z_0) is contained in the tangent plane, (see Figure 10.52.) Likewise, we consider the curve of intersection between $z = f(x, y)$ and the plane that is parallel to the x–z plane and contains the point (x_0, y_0, z_0)—that is, the plane $y = y_0$—and we denote this curve by C_2. Its tangent line, too, is contained in the tangent plane. (See Figure 10.52.)

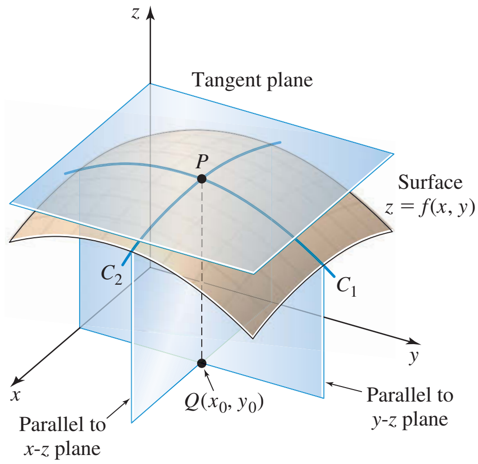


Figure 10.51 The curves C_1 and C_2 are formed by the intersection of the surface $z = f(x, y)$ and planes parallel to the $y - z$ plane and the $x - z$ plane.

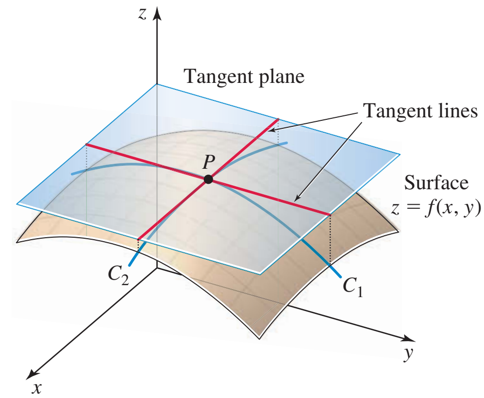


Figure 10.52 The tangent plane at P contains the tangent lines to C_1 and C_2 at P.

In Section 9.5, we gave the general equation of a plane:

$$z - z_0 = A(x - x_0) + B(y - y_0) \tag{10.8}$$

We will use curves C_1 and C_2 to determine the constants A and B. C_1 satisfies the equation

$$z = f(x_0, y)$$

The tangent line to C_1 at (x_0, y_0, z_0) therefore satisfies

$$z - z_0 = \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0) \tag{10.9}$$

This tangent line is contained in the tangent plane. The equation of the tangent line at (x_0, y_0, z_0) can therefore also be obtained by setting $x = x_0$ in (10.8). Doing so yields

$$z - z_0 = (A)(0) + B(y - y_0) = B(y - y_0)$$

Comparing this with (10.9), we find that

$$B = \frac{\partial f(x_0, y_0)}{\partial y}$$

Similarly, using C_2, we find that

$$A = \frac{\partial f(x_0, y_0)}{\partial x}$$

We thus arrive at the following result:

Equation of a Tangent Plane

If the tangent plane to the surface $z = f(x, y)$ at the point (x_0, y_0, z_0) exists, then that tangent plane has the equation

$$z - z_0 = \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0)$$

You should observe the similarity of this equation to (10.7), the equation of the tangent line. We will see that the mere existence of the partial derivatives $\frac{\partial f(x_0,y_0)}{\partial x}$ and $\frac{\partial f(x_0,y_0)}{\partial y}$ is *not* enough to guarantee the existence of a tangent plane at (x_0, y_0); something stronger is needed. But before we discuss the conditions under which tangent planes exist, let's look at an example.

EXAMPLE 1 It can be shown that the tangent plane to the surface

$$z = f(x, y) = 4x^2 + y^2$$

at the point $(1, 2, 8)$ exists. Find its equation.

Solution First note that the point $(1, 2, 8)$ is contained in the surface $z = f(x, y)$, since $8 = f(1, 2)$. To find the tangent plane, we need to compute the partial derivatives of $f(x, y)$, namely,

$$\frac{\partial f}{\partial x} = 8x \qquad \text{and} \qquad \frac{\partial f}{\partial y} = 2y$$

and evaluate them at $(x_0, y_0) = (1, 2)$:

$$\frac{\partial f(1, 2)}{\partial x} = 8 \qquad \text{and} \qquad \frac{\partial f(1, 2)}{\partial y} = 4$$

Hence, the equation of the tangent plane is

$$z - 8 = 8(x - 1) + 4(y - 2)$$

We can rewrite this equation by expanding terms on the right-hand side

$$z - 8 = 8x - 8 + 4y - 8$$

or

$$8x + 4y - z = 8. \qquad \text{Isolate } x, y, \text{ and } z$$

•

Differentiability. In discussing the conditions under which tangent planes exist, we need to define what it means for a function of two variables to be *differentiable*. To make the connection with functions of one variable clear, recall that the tangent line is used to linearly approximate $f(x)$ at $x = x_0$. The linear approximation of $f(x)$ at $x = x_0$ is given by

$$L(x) = f(x_0) + f'(x_0)(x - x_0) \tag{10.10}$$

The distance between $f(x)$ and its linear approximation at $x = x_0$ is then

$$|f(x) - L(x)| = |f(x) - f(x_0) - f'(x_0)(x - x_0)| \tag{10.11}$$

From the definition of the derivative, we have

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{10.12}$$

If we divide (10.11) by the distance between x and x_0, $|x - x_0|$, we find that

$$\begin{aligned} \frac{|f(x) - L(x)|}{|x - x_0|} &= \left| \frac{f(x) - L(x)}{x - x_0} \right| = \left| \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} \right| \\ &= \left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| \end{aligned}$$

Taking the limit $x \to x_0$ and using (10.12), we obtain

$$\lim_{x \to x_0} \left| \frac{f(x) - L(x)}{x - x_0} \right| = 0 \tag{10.13}$$

We say that $f(x)$ is differentiable at $x = x_0$ if (10.13) holds.

For functions of two variables, the definition of differentiability is based on the same idea. Notice, however, in the preceding discussion, we divided by the distance between x and x_0. In two dimensions, the distance between two points (x, y) and (x_0, y_0) is $\sqrt{(x - x_0)^2 + (y - y_0)^2}$.

Definition Differentiability of $f(x, y)$ Suppose that $f(x, y)$ is a function of two independent variables and that both $\partial f/\partial x$ and $\partial f/\partial y$ are defined throughout an open disk containing (x_0, y_0). Set

$$L(x, y) = f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0)$$

Then $f(x, y)$ is differentiable at (x_0, y_0) if

$$\lim_{(x,y)\to(x_0,y_0)} \left| \frac{f(x, y) - L(x, y)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \right| = 0$$

Furthermore, if $f(x, y)$ is differentiable at (x_0, y_0), then $z = L(x, y)$ defines the tangent plane to the graph of f at $(x_0, y_0, f(x_0, y_0))$. We say that $f(x, y)$ is differentiable if it is differentiable at every point of its domain.

The key idea in both the one- and the two-dimensional case is to approximate functions by linear functions, so that the error in the approximation vanishes as we approach the point at which we approximated the function [x_0 in the one-dimensional case, (x_0, y_0) in the two-dimensional case].

As in the one-dimensional case, the following theorem holds:

Theorem If $f(x, y)$ is differentiable at (x_0, y_0), then f is continuous at (x_0, y_0).

If $f(x, y)$ is differentiable at (x_0, y_0) then the function $f(x, y)$ is close to the tangent plane at (x_0, y_0) for all (x, y) close to (x_0, y_0). The mere existence of the partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ at (x_0, y_0), however, is *not* enough to guarantee differentiability (and, consequently, the existence of a tangent plane at a certain point). The next, very simple, example will show what can go wrong with that assumption.

EXAMPLE 2 Assume that

$$f(x, y) = \begin{cases} 0 & \text{if } xy \neq 0 \\ 1 & \text{if } xy = 0 \end{cases}$$

Show that $\frac{\partial f(0,0)}{\partial x}$ and $\frac{\partial f(0,0)}{\partial y}$ exist, but $f(x, y)$ is not continuous and, consequently, not differentiable, at $(0, 0)$. •

Solution The graph of $f(x, y)$ is shown in Figure 10.53. To compute $\partial f/\partial x$ at $(0, 0)$, we set $y = 0$. Then $f(x, 0) = 1$, and therefore,

$$\frac{\partial f(0, 0)}{\partial x} = 0$$

Likewise, setting $x = 0$, we find that $f(0, y) = 1$ and

$$\frac{\partial f(0, 0)}{\partial y} = 0$$

That is, both partial derivatives exist at $(0, 0)$. However, $f(x, y)$ is not continuous at $(0, 0)$. To see this, it is enough to show that $f(x, y)$ has different limits along two different paths as (x, y) approaches $(0, 0)$. For the first path, denoted by C_1, we choose $y = 0$. Then

$$\lim_{\substack{(x,y)\to(0,0) \\ \text{along } C_1}} f(x, y) = 1$$

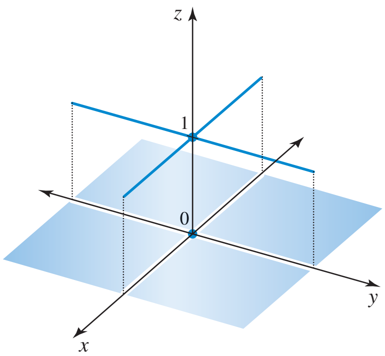


Figure 10.53 The graph of the function

$$f(x, y) = \begin{cases} 0 & \text{if } xy \neq 0 \\ 1 & xy = 0 \end{cases}$$

Even though $\frac{\partial f(0,0)}{\partial x}$ and $\frac{\partial f(0,0)}{\partial y}$ exist, the function is not continuous at (0, 0).

For the second path, denoted by C_2, we choose $y = x$. Then

$$\lim_{\substack{(x,y)\to(0,0) \\ \text{along } C_2}} f(x, y) = 0$$

Since $1 \neq 0$, $f(x, y)$ is not continuous at $(0, 0)$, and because differentiability implies continuity, a function that is not continuous cannot be differentiable. •

The definition of differentiability is not easy to use if we actually want to check whether a function is differentiable at a certain point. Fortunately, there is another criterion, which suffices for all practical purposes, that can be used to check whether $f(x, y)$ is differentiable. We saw in Example 2 that the mere existence of the partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ at a point (x_0, y_0) is not enough to guarantee that $f(x, y)$ is differentiable. However, if the partial derivatives are continuous on a disk centered at (x_0, y_0), that is enough to guarantee differentiability.

Sufficient Condition for Differentiability Suppose $f(x, y)$ is defined on an open disk centered at (x_0, y_0) and the partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ are continuous on an open disk centered at (x_0, y_0). Then $f(x, y)$ is differentiable at (x_0, y_0).

EXAMPLE 3 Show that $f(x, y) = 2x^2y - y^2$ is differentiable for all $(x, y) \in \mathbf{R}^2$.

Solution We use the sufficient condition for differentiability. First, observe that $f(x, y)$ is defined for all $(x, y) \in \mathbf{R}^2$. The partial derivatives are given by

$$\frac{\partial f}{\partial x} = 4xy \qquad \text{and} \qquad \frac{\partial f}{\partial y} = 2x^2 - 2y$$

Since both $\partial f/\partial x$ and $\partial f/\partial y$ are polynomials, both are continuous for all $(x, y) \in \mathbf{R}^2$ and, hence, $f(x, y)$ is differentiable for all $(x, y) \in \mathbf{R}^2$. •

Linearization. Just as we use tangent lines to locally approximate functions of a single variable, we may use tangent planes to locally approximate a function of two variables.

Definition Linearization of $f(x, y)$ Suppose that $f(x, y)$ is differentiable at (x_0, y_0). The linearization of $f(x, y)$ at (x_0, y_0) is the function

$$L(x, y) = f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0)$$

The approximation

$$f(x, y) \approx L(x, y)$$

is the **standard linear approximation**, or the **tangent plane approximation**, of $f(x, y)$ at (x_0, y_0).

EXAMPLE 4 Find the linear approximation of

$$f(x, y) = x^2y + 2xe^y$$

at the point (2, 0).

Solution The linearization of $f(x, y)$ at $(2, 0)$ is given by

$$L(x, y) = f(2, 0) + \frac{\partial f(2, 0)}{\partial x}(x - 2) + \frac{\partial f(2, 0)}{\partial y}(y - 0)$$

Now, $f(2, 0) = 4$,

$$\frac{\partial f}{\partial x} = 2xy + 2e^y, \quad \text{and} \quad \frac{\partial f}{\partial y} = x^2 + 2xe^y.$$

Hence,

$$\frac{\partial f(2,0)}{\partial x} = 2 \quad \text{and} \quad \frac{\partial f(2,0)}{\partial y} = 4 + 4 = 8$$

and we find that

$$\begin{aligned} L(x, y) &= 4 + 2(x - 2) + 8(y - 0) = 4 + 2x - 4 + 8y \\ &= 2x + 8y. \end{aligned}$$

•

EXAMPLE 5 Find the linear approximation of

$$f(x, y) = \ln(x - 2y^2)$$

at the point (3, 1), and use it to find an approximation for $f(3.05, 0.95)$. Use a calculator to compute the value of $f(3.05, 0.95)$ and compare it with the approximation.

Solution The linearization of $f(x, y)$ at (3, 1) is given by

$$L(x, y) = f(3, 1) + \frac{\partial f(3,1)}{\partial x}(x - 3) + \frac{\partial f(3,1)}{\partial y}(y - 1)$$

Now, $f(3, 1) = \ln(3 - 2) = \ln 1 = 0$,

$$\frac{\partial f(x,y)}{\partial x} = \frac{1}{x - 2y^2}, \quad \text{and} \quad \frac{\partial f(x,y)}{\partial y} = \frac{-4y}{x - 2y^2}$$

Hence,

$$\frac{\partial f(3,1)}{\partial x} = \frac{1}{3 - 2} = 1 \quad \text{and} \quad \frac{\partial f(3,1)}{\partial y} = \frac{-4}{3 - 2} = -4$$

and we find that

$$\begin{aligned} L(x, y) &= 0 + (1)(x - 3) + (-4)(y - 1) \\ &= x - 3 - 4y + 4 = x - 4y + 1 \end{aligned}$$

Using (3.05, 0.95), we get

$$L(3.05, 0.95) = 3.05 - (4)(0.95) + 1 = 0.25$$

Comparing this with $f(3.05, 0.95) \approx 0.2191$, we see that the error of approximation is about $|0.25 - 0.2191| = 0.031$.

•

10.4.2 Vector-Valued Functions

So far, we have considered only real-valued functions

$$f : \mathbf{R}^n \to \mathbf{R}$$

that is, functions that take n real numbers as their arguments and return a single real number value. We will now extend our discussion to functions whose range is a subset of $\mathbf{R}^m$—that is,

$$f : \mathbf{R}^n \to \mathbf{R}^m$$

Such functions are **vector-valued functions**, since they take on values that are represented by vectors:

$$\begin{aligned} f &: \mathbf{R}^n \to \mathbf{R}^m \\ (x_1, x_2, \ldots, x_n) &\mapsto \begin{bmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{bmatrix} \end{aligned}$$

Here, each function $f_i(x_1, \ldots, x_n)$ is a real-valued function:

$$\begin{aligned} f_i &: \mathbf{R}^n \to \mathbf{R} \\ (x_1, x_2, \ldots, x_n) &\mapsto f_i(x_1, x_2, \ldots, x_n) \end{aligned}$$

We will encounter vector-valued functions where $n = m = 2$ extensively in Chapter 11. For example, let u represent the number of sharks in an ocean, and v the number of sardines (sharks feed on sardines). Let f and g represent the growth rates of the shark and sardine populations, respectively. In Chapter 8 we considered how the growth rate of a population may depend on the size of that population (for example, if a large number of sharks is present offspring are likely to be born at a high rate). Hence, we expect the shark population growth rate to be a function of u. At the same time, the amount of resources that sharks can devote to their offspring will depend on how much prey is available. If the sardine population is small, then sharks may start to die off. Thus, the shark population growth rate will be a function of v also. Hence, the growth rate is a multivariable function $f(u, v)$. Conversely, if the shark population is large, then high rates of predation will affect the growth rate of the sardine population; so the sardine population growth rate will also depend on both u and v, i.e., will be a multivariable function $g(u, v)$. The growth rates of the two populations therefore form a map:

$$(u, v) \mapsto \begin{bmatrix} f(u, v) \\ g(u, v) \end{bmatrix}$$

from $\mathbf{R}^2$ to $\mathbf{R}^2$.

Our main task in this subsection will be to define the linearization of vector-valued functions whose domain and range are $\mathbf{R}^2$. We will motivate this definition by analogy with the cases of functions $f : \mathbf{R} \to \mathbf{R}$ and $f : \mathbf{R}^2 \to \mathbf{R}$. At the end of the section, we will mention how to generalize our results to arbitrary vector-valued functions.

We begin with differentiable functions $f : \mathbf{R} \to \mathbf{R}$. The linearization of $f : \mathbf{R} \to \mathbf{R}$ about x_0 is given by

$$L(x) = f(x_0) + f'(x_0)(x - x_0) \tag{10.14}$$

EXAMPLE 6 Find the linearization of

$$f(x) = 2 \ln x$$

at $x_0 = 1$.

Solution The linearization of $f(x)$ at $x = 1$ is given by

$$\begin{aligned} L(x) &= f(1) + f'(1)(x - 1) \\ &= 0 + (2)(x - 1) = 2x - 2 \end{aligned}$$

since $f'(x) = 2/x$. •

We found the linearization of a real-valued function $f : \mathbf{R}^2 \to \mathbf{R}$ in the previous subsection, namely,

$$L(x, y) = f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0)$$

We can write this equation in matrix notation as

$$L(x, y) = f(x_0, y_0) + \begin{bmatrix} \frac{\partial f(x_0, y_0)}{\partial x} & \frac{\partial f(x_0, y_0)}{\partial y} \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}$$

EXAMPLE 7 Using matrix notation, find the linearization of $f(x, y) = \ln x + \ln y$ at $(1, 1)$.

Solution $f(1, 1) = 0$. Since $f_x(x, y) = \frac{1}{x}$ and $f_y(x, y) = \frac{1}{y}$, it follows that

$$\frac{\partial f(1, 1)}{\partial x} = 1 \qquad \text{and} \qquad \frac{\partial f(1, 1)}{\partial y} = 1 \qquad \text{so:} \begin{bmatrix} \frac{\partial f(1, 1)}{\partial x} & \frac{\partial f(1, 1)}{\partial y} \end{bmatrix} = [1, 1].$$

Hence,

$$\begin{aligned} L(x, y) &= 0 + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x - 1 \\ y - 1 \end{bmatrix} \\ &= x - 1 + y - 1 = x + y - 2 \end{aligned}$$ •

We will denote vector-valued functions by boldface letters. Suppose that

$$\mathbf{h} : \mathbf{R}^2 \to \mathbf{R}^2$$
$$(x, y) \mapsto \begin{bmatrix} f(x, y) \\ g(x, y) \end{bmatrix}$$

and assume that all first partial derivatives are continuous on a disk centered at (x_0, y_0). We can linearize each component of the function $\mathbf{h}$. We denote the linearization of f by L_f, it is given by:

$$L_f(x, y) = f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0).$$

Similarly the linearization of g is

$$L_g(x, y) = g(x_0, y_0) + \frac{\partial g(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial g(x_0, y_0)}{\partial y}(y - y_0).$$

We define the vector-valued function $\mathbf{L}(x, y) = \begin{bmatrix} L_f(x, y) \\ L_g(x, y) \end{bmatrix}$. The linearization of $\mathbf{h}$ can thus be written in matrix form as

$$\mathbf{L}(x, y) = \begin{bmatrix} L_f(x, y) \\ L_g(x, y) \end{bmatrix} = \begin{bmatrix} f(x_0, y_0) \\ g(x_0, y_0) \end{bmatrix} + \begin{bmatrix} \frac{\partial f(x_0,y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0,y_0)}{\partial y}(y - y_0) \\ \frac{\partial g(x_0,y_0)}{\partial x}(x - x_0) + \frac{\partial g(x_0,y_0)}{\partial y}(y - y_0) \end{bmatrix}$$

or

$$\mathbf{L}(x, y) = \begin{bmatrix} f(x_0, y_0) \\ g(x_0, y_0) \end{bmatrix} + \begin{bmatrix} \frac{\partial f(x_0,y_0)}{\partial x} & \frac{\partial f(x_0,y_0)}{\partial y} \\ \frac{\partial g(x_0,y_0)}{\partial x} & \frac{\partial g(x_0,y_0)}{\partial y} \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}$$

where the matrix

$$\begin{bmatrix} \frac{\partial f(x_0,y_0)}{\partial x} & \frac{\partial f(x_0,y_0)}{\partial y} \\ \frac{\partial g(x_0,y_0)}{\partial x} & \frac{\partial g(x_0,y_0)}{\partial y} \end{bmatrix}$$

is a 2×2 matrix that is called the **Jacobi matrix** (often abbreviated as **Jacobian**) or the **derivative matrix**. We denote the Jacobi matrix by $(D\mathbf{h})$; that is,

$$(D\mathbf{h})(x_0, y_0) = \begin{bmatrix} \frac{\partial f(x_0,y_0)}{\partial x} & \frac{\partial f(x_0,y_0)}{\partial y} \\ \frac{\partial g(x_0,y_0)}{\partial x} & \frac{\partial g(x_0,y_0)}{\partial y} \end{bmatrix}$$

EXAMPLE 8 Assume that

$$\mathbf{f} : \mathbf{R}^2 \to \mathbf{R}^2$$
$$(x, y) \to \begin{bmatrix} u \\ v \end{bmatrix}$$

with

$$u(x, y) = x^2 y - y^3 \qquad \text{and} \qquad v(x, y) = 2x^3 y^2 + y$$

Compute the Jacobi matrix and evaluate it at $(1, 2)$.

Solution The Jacobi matrix is

$$\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} = \begin{bmatrix} 2xy & x^2 - 3y^2 \\ 6x^2 y^2 & 4x^3 y + 1 \end{bmatrix}.$$

At $(1, 2)$, we find that

$$D\mathbf{f}(1, 2) = \begin{bmatrix} 4 & -11 \\ 24 & 9 \end{bmatrix}$$

•

EXAMPLE 9 Let $\mathbf{f}: \mathbf{R}^2 \to \mathbf{R}^2$

$$\mathbf{f}(x, y) = \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix}$$

with

$$u(x, y) = ye^{-x} \qquad \text{and} \qquad v(x, y) = \sin x + \cos y$$

Find the linear approximation to $\mathbf{f}(x, y)$ at $(0, 0)$. Compare $\mathbf{f}(0.1, -0.1)$ with its linear approximation.

Solution We compute the Jacobi matrix first:

$$(D\mathbf{f})(x, y) = \begin{bmatrix} -ye^{-x} & e^{-x} \\ \cos x & -\sin y \end{bmatrix}$$

The linear approximation of $\mathbf{f}(x, y)$ at $(0, 0)$ is

$$\begin{aligned} \mathbf{L}(x, y) = \begin{bmatrix} L_u(x, y) \\ L_v(x, y) \end{bmatrix} &= \begin{bmatrix} u(0, 0) \\ v(0, 0) \end{bmatrix} + (D\mathbf{f})(0, 0) \begin{bmatrix} x - 0 \\ y - 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} y \\ 1 + x \end{bmatrix}. \end{aligned}$$

Using $(x, y) = (0.1, -0.1)$, we find that

$$\mathbf{L}(0.1, -0.1) = \begin{bmatrix} -0.1 \\ 1 + 0.1 \end{bmatrix} = \begin{bmatrix} -0.1 \\ 1.1 \end{bmatrix}.$$

We can also calculate $f(0.1, -0.1)$ directly:

$$\mathbf{f}(0.1, -0.1) = \begin{bmatrix} -0.1e^{-0.1} \\ \sin 0.1 + \cos(-0.1) \end{bmatrix} = \begin{bmatrix} -0.09 \\ 1.09 \end{bmatrix}.$$

We see that the linear approximation is close to the actual value. •

In this book we will only consider vector valued functions $f: \mathbf{R}^2 \to \mathbf{R}^2$, but we can generalize the Jacobi matrix to functions $\mathbf{f}: \mathbf{R}^n \to \mathbf{R}^m$. If

$$\mathbf{f}(x_1, x_2, \ldots, x_n) = \begin{bmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{bmatrix}$$

where $f_i: \mathbf{R}^n \to \mathbf{R}$, $i = 1, 2, \ldots, m$, are real-valued functions of n independent variables, then the Jacobi matrix is an $m \times n$ matrix of the form

$$D\mathbf{f} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

The linearization of $\mathbf{f}$ about the point $(x_1^*, x_2^*, \ldots, x_n^*)$ is then

$$L(x_1, \ldots, x_n) = \begin{bmatrix} f_1(x_1^*, \ldots, x_n^*) \\ f_2(x_1^*, \ldots, x_n^*) \\ \vdots \\ f_m(x_1^*, \ldots, x_n^*) \end{bmatrix} + D\mathbf{f}(x_1^*, \ldots, x_n^*) \begin{bmatrix} x_1 - x_1^* \\ \vdots \\ x_n - x_n^* \end{bmatrix}.$$

Section 10.4 Problems

10.4.1

In Problems 1–10, the tangent plane at the indicated point (x_0, y_0, z_0) exists. Find its equation.

1. $f(x, y) = 2x^3 + y^2$; $(1, 2, 6)$

2. $f(x, y) = x^2 - 3y^2$; $(1, 2, -11)$

3. $f(x, y) = xy$; $(-1, 2, -2)$

4. $f(x, y) = \sin x \cdot \cos y$; $(0, 0, 0)$

5. $f(x, y) = \sin(xy)$; $(1, 0, 0)$

6. $f(x, y) = e^{x-y}$; $(1, 1, 1)$

7. $f(x, y) = e^{2x^2+y^2}$; $(1, 0, e^2)$

8. $f(x, y) = \sqrt{x^2 + y^2}$; $(1, 1, \sqrt{2})$

9. $f(x, y) = \ln(x + y)$; $(2, -1, 0)$

10. $f(x, y) = x^2 e^{-y}$; $(1, 0, 1)$

In Problems 11–16, show that $f(x, y)$ is differentiable at the indicated point.

11. $f(x, y) = x + y^2$; $(1, 1)$

12. $f(x, y) = xy + 3x^2$; $(1, 0)$

13. $f(x, y) = \cos(x + y)$; $(0, 0)$

14. $f(x, y) = \sin(x - y)$; $(1, 0)$

15. $f(x, y) = xe^{-y}$; $(0, 1)$

16. $f(x, y) = (x^2 + y^2)e^{-x^2-y^2}$; $(1, 1)$

In Problems 17–24, find the linearization of $f(x, y)$ at the indicated point (x_0, y_0).

17. $f(x, y) = x - 3y$; $(3, 1)$

18. $f(x, y) = 2xy$; $(1, -1)$

19. $f(x, y) = \sqrt{x} + 2y$; $(1, 0)$

20. $f(x, y) = \cos(x^2 y)$; $\left(\frac{\pi}{2}, 0\right)$

21. $f(x, y) = \tan(x + y)$; $(0, 0)$

22. $f(x, y) = e^{3x+2y}$; $(1, 2)$

23. $f(x, y) = \ln(x^2 + y)$; $(1, 1)$

24. $f(x, y) = x^2 e^y$; $(1, 0)$

25. Find the linear approximation of

$$f(x, y) = e^{x+y}$$

at $(0, 0)$, and use it to approximate $f(0.1, 0.05)$. Using a calculator, compare the approximation with the exact value of $f(0.1, 0.05)$.

26. Find the linear approximation of

$$f(x, y) = \sin(x + 2y)$$

at $(0, 0)$, and use it to approximate $f(-0.1, 0.2)$. Using a calculator, compare the approximation with the exact value of $f(-0.1, 0.2)$.

27. Find the linear approximation of

$$f(x, y) = \sqrt{x + y^2}$$

at $(1, 0)$, and use it to approximate $f(1.1, 0.1)$. Using a calculator, compare the approximation with the exact value of $f(1.1, 0.1)$.

28. Find the linear approximation of

$$f(x, y) = (x^2 + y^2)e^{-(x^2+y^2)}$$

at $(0, 0)$, and use it to approximate $f(0.01, 0.05)$. Using a calculator, compare the approximation with the exact value of $f(0.01, 0.05)$.

10.4.2

In Problems 29–36, find the Jacobi matrix for each given function.

29. $\mathbf{f}(x, y) = \begin{bmatrix} x + y \\ x^2 - y^2 \end{bmatrix}$

30. $\mathbf{f}(x, y) = \begin{bmatrix} 2x - 3y \\ 4x^2 \end{bmatrix}$

31. $\mathbf{f}(x, y) = \begin{bmatrix} e^{x-y} \\ e^{x+y} \end{bmatrix}$

32. $\mathbf{f}(x, y) = \begin{bmatrix} (x - y)^2 \\ \sin(x - y) \end{bmatrix}$

33. $\mathbf{f}(x, y) = \begin{bmatrix} \cos(x - y) \\ \cos(x + y) \end{bmatrix}$

34. $\mathbf{f}(x, y) = \begin{bmatrix} \ln(x + y) \\ e^{x+y} \end{bmatrix}$

35. $\mathbf{f}(x, y) = \begin{bmatrix} 2x^2 y - 3y + x \\ e^x \sin y \end{bmatrix}$

36. $\mathbf{f}(x, y) = \begin{bmatrix} \sqrt{x^2 + y^2} \\ e^{-x^2} \end{bmatrix}$

In Problems 37–42, find a linear approximation to each function $\mathbf{f}(x, y)$ at the indicated point.

37. $\mathbf{f}(x, y) = \begin{bmatrix} 2x^2 y \\ \frac{1}{xy} \end{bmatrix}$ at $(1, 1)$

38. $\mathbf{f}(x, y) = \begin{bmatrix} 3x - y^2 \\ 4y \end{bmatrix}$ at $(-1, -2)$

39. $\mathbf{f}(x, y) = \begin{bmatrix} e^{2x-y} \\ \ln(x - y) \end{bmatrix}$ at $(2, 1)$

40. $\mathbf{f}(x, y) = \begin{bmatrix} e^x \sin y \\ e^{-y} \cos x \end{bmatrix}$ at $(0, 0)$

41. $\mathbf{f}(x, y) = \begin{bmatrix} \frac{x}{y} \\ \frac{y}{x} \end{bmatrix}$ at $(1, 2)$

42. $\mathbf{f}(x, y) = \begin{bmatrix} (x + y)^2 \\ xy \end{bmatrix}$ at $(-1, 1)$

43. Find a linear approximation to

$$\mathbf{f}(x, y) = \begin{bmatrix} x^2 - xy \\ 3y^2 - 1 \end{bmatrix}$$

at $(1, 2)$. Use your result to find an approximation for $\mathbf{f}(1.1, 1.9)$, and compare the approximation with the value of $\mathbf{f}(1.1, 1.9)$ that you get when you use a calculator.

44. Find a linear approximation to

$$\mathbf{f}(x, y) = \begin{bmatrix} x/y \\ 2xy \end{bmatrix}$$

at $(-1, 1)$. Use your result to find an approximation for $\mathbf{f}(-0.9, 1.05)$, and compare the approximation with the value of $\mathbf{f}(-0.9, 1.05)$ that you get when you use a calculator.

45. Find a linear approximation to

$$\mathbf{f}(x, y) = \begin{bmatrix} (x - y)^2 \\ 2x^2 y \end{bmatrix}$$

at $(2, -3)$. Use your result to find an approximation for $f(1.9, -3.1)$, and compare the approximation with the value of $f(1.9, -3.1)$ that you get when you use a calculator.

46. Find a linear approximation to

$$\mathbf{f}(x, y) = \begin{bmatrix} \sqrt{2x+y} \\ x - y^2 \end{bmatrix}$$

at $(1, 2)$. Use your result to find an approximation for $f(1.05, 2.05)$, and compare the approximation with the value of $f(1.05, 2.05)$ that you get when you use a calculator.

47. Predator-Prey Populations The Lotka-Volterra equations are often used to model the links between a particular of prey organisms (e.g., sardines) and a population of predatory organisms (e.g., sharks), (see Chapter 11.) In a particular ecosystem we will use u to represent the number of sharks and v to represent the number of sardines. Suppose the growth rate of the shark population is

$$f(u, v) = -0.5u + \frac{uv}{100}$$

and of the sardine population is

$$g(u, v) = 3v - 10uv$$

(a) Show that if $u = 0.3$ and $v = 50$, then $f(u, v) = 0$, and $g(u, v) = 0$. (The populations are said to be in equilibrium.)

(b) Find the linear approximation of the vector valued function

$$\mathbf{h} : (u, v) \mapsto \begin{bmatrix} f(u, v) \\ g(u, v) \end{bmatrix}$$

if u is close to 0.3 and v is close to 50.

48. Competing Populations Two different species of organisms may compete for the same limited resource, for example, hyenas and lions may compete for territory and prey. A large lion population will reduce the growth rate of the hyena population, and vice versa. Let the number of lions be x and the number of hyenas be y. The growth rates of the lion and hyena populations may be modeled using Lotka-Volterra equations (see Chapter 11). Suppose the growth rates are given by functions:

$$f(x, y) = 0.5x - \frac{x^2}{10} - \frac{xy}{20}$$

$$g(x, y) = 2y - \frac{y^2}{20} - \frac{xy}{5}$$

(a) Show that $f(x, y) = 0$ and $g(x, y) = 0$ when $x = 0$ and $y = 40$. The two populations are said to be in equilibrium when at this size.

(b) Linearize the vector valued function

$$\mathbf{h} : (x, y) \mapsto \begin{bmatrix} f(x, y) \\ g(x, y) \end{bmatrix}$$

when x is approximately 0, and y is approximately 40.

10.5 The Chain Rule and Implicit Differentiation

10.5.1 The Chain Rule for Functions of Two Variables

In Section 10.3, we described an experiment to determine how the amount of CO_2 stored by plants can change as a function of both temperature and light intensity. Suppose we follow a patch of these plants; studying how the amount of CO_2 they store changes over time. Over time the temperature of the plant's surroundings may change and so may the light intensity, I. If we denote the temperature at time t by $T(t)$, the light intensity at time t by $I(t)$, and the amount of stored CO_2 at time t by $N(t)$, then $N(t)$ is a function of both $T(t)$ and $I(t)$, and we can write

$$N(t) = f(T(t), I(t))$$

That is, $N(t)$ changes over time because both the temperature and the light intensity change.

To differentiate composite functions of one variable, we use the chain rule. Suppose that $w = f(x)$ is a function of one variable and that x depends on t. Then, by the chain rule, to differentiate w with respect to t, we have

$$\frac{dw}{dt} = \frac{dw}{dx}\frac{dx}{dt} \tag{10.15}$$

The chain rule can be extended to functions of more than one variable:

Chain Rule for Functions of Two Independent Variables If $w = f(x, y)$ is differentiable and x and y are differentiable functions of t, then w is a differentiable function of t and

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}$$

We will not rigorously prove this formula, but we will give a non-rigorous interpretation. We approximate $w = f(x, y)$ at (x_0, y_0) by its linear approximation

$$L(x, y) = f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}(y - y_0)$$

If we set $\Delta x = x - x_0$, $\Delta y = y - y_0$, and $\Delta w = f(x, y) - f(x_0, y_0)$, we can approximate $\Delta w = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$ by its linear approximation. We find that

$$\Delta w \approx \frac{\partial f(x_0, y_0)}{\partial x}\Delta x + \frac{\partial f(x_0, y_0)}{\partial y}\Delta y$$

Dividing both sides by Δt, we obtain

$$\frac{\Delta w}{\Delta t} \approx \frac{\partial f(x_0, y_0)}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial f(x_0, y_0)}{\partial y}\frac{\Delta y}{\Delta t}$$

If we let $\Delta t \to 0$, then

$$\frac{\Delta w}{\Delta t} \to \frac{dw}{dt}, \qquad \frac{\Delta x}{\Delta t} \to \frac{dx}{dt}, \qquad \frac{\Delta y}{\Delta t} \to \frac{dy}{dt}$$

and we get

$$\frac{dw}{dt} = \frac{\partial f(x_0, y_0)}{\partial x}\frac{dx}{dt} + \frac{\partial f(x_0, y_0)}{\partial y}\frac{dy}{dt}$$

or, in short,

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}.$$

EXAMPLE 1 Let

$$w = f(x, y) = x^2 y^3$$

with $x(t) = \sin t$ and $y(t) = e^{-t}$. Find the derivative of $w = f(x, y)$ with respect to t when $t = \pi/2$.

Solution

$$\begin{aligned}\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} && \text{Chain rule}\\ &= 2xy^3\frac{dx}{dt} + 3x^2y^2\frac{dy}{dt} && \tfrac{\partial w}{\partial x} = 2xy^3,\ \tfrac{\partial w}{\partial y} = 3x^2y^2\\ &= 2xy^3\cos t + 3x^2y^2(-1)e^{-t}\end{aligned}$$

We then substitute in: $t = \frac{\pi}{2}$ and

$$x\left(\frac{\pi}{2}\right) = \sin\frac{\pi}{2} = 1, \qquad \text{and} \qquad y\left(\frac{\pi}{2}\right) = e^{-\pi/2}.$$

We obtain:

$$\left.\frac{dw}{dt}\right|_{t=\pi/2} = (2)(1)(e^{-3\pi/2})\cos\left(\frac{\pi}{2}\right) - (3)(1)^2 e^{-\pi}e^{-\pi/2} = -3e^{-3\pi/2}$$

We can check the answer directly; by writing w as a function of t.

$$w(t) = (\sin^2 t)(e^{-3t})$$

Therefore,

$$\left.\frac{dw}{dt}\right|_{t=\pi/2} = 2(\sin t)(\cos t)e^{-3t} - 3(\sin^2 t)e^{-3t}\Big|_{t=\pi/2} = -3e^{-3\pi/2}.$$

EXAMPLE 2 Suppose that we wish to predict the abundance of a particular plant species. We suspect that the two major factors influencing the abundance of the plant are nitrogen levels and the level of disturbance due to grazing. Previous studies have shown that an increase in nitrogen in the soil has a negative effect on the abundance of this species; also, an increase in disturbance due to grazing seems to have a negative effect on abundance. If both nitrogen and disturbance due to grazing increase over the next few years, how would the abundance of the species be affected?

Solution We denote the abundance of the plant at time t by $B(t)$. We are interested in how $B(t)$ will change over time; that is, we want to find out whether $B(t)$ will increase or decrease over time. For this, we need to compute the derivative of $B(t)$. If we assume that the abundance of the plant is affected primarily by nitrogen and disturbance levels, we can consider B as a function of both N, the nitrogen level, and D, the disturbance level. N and D change over time and are thus functions of time. The function $B(t)$ is a function of both $N(t)$ and $D(t)$. Using the chain rule for functions of two independent variables, we find that

$$\frac{dB}{dt} = \underbrace{\frac{\partial B}{\partial N}}_{<0} \underbrace{\frac{dN}{dt}}_{>0} + \underbrace{\frac{\partial B}{\partial D}}_{<0} \underbrace{\frac{dD}{dt}}_{>0} < 0$$

since abundance B is a decreasing function of both N and D and since both nitrogen and disturbance levels are assumed to increase over the next few years. We thus find that the abundance of the plant will decrease over the next few years. •

EXAMPLE 3 **Rate of Change According to a Moving Observer** Temperature in the atmosphere varies both with time t, and with latitude, which we will denote by x. That is, the temperature is a function $T(t, x)$. The temperature experienced by a migrating bird may change over time, due both to the changes of air temperature with time, and because the bird's latitude changes as it migrates. Suppose that the bird migrates at velocity v, meaning its latitude increases at a rate v. Derive an expression for the rate of change of temperature that it observes.

Solution We are given that $dx/dt = v$. By the chain rule:

$$\begin{aligned}\frac{d}{dt}T(t, x(t)) &= \frac{dt}{dt}\frac{\partial T}{\partial t} + \frac{dx}{dt}\frac{\partial T}{\partial x} \\ &= \frac{\partial T}{\partial t} + v\frac{\partial T}{\partial x}. \qquad \frac{dt}{dt} = 1\end{aligned}$$

Notice that if $v = 0$, meaning the bird is not migrating, then the rate of change of the temperature is $\partial T/\partial t$. However, if v is non-zero, then dT/dt can be non-zero even if $\partial T/\partial t = 0$. If the temperature of the atmosphere is constant in time, but different latitudes have different temperatures, then the temperature felt by the bird will change as its latitude changes. •

10.5.2 Implicit Differentiation

We discussed implicit differentiation in Section 4.6. This was a useful technique for differentiating a function $y = f(x)$ when y was given implicitly, for example, a function like:

$$x^2 y - e^{-y} = 0 \tag{10.16}$$

To find dy/dx, we differentiate both sides with respect to x, keeping in mind that y is a function of x. We obtain

$$2xy + x^2\frac{dy}{dx} - e^{-y}\left(-\frac{dy}{dx}\right) = 0$$

$$2xy + \frac{dy}{dx}(x^2 + e^{-y}) = 0$$

Solving for dy/dx, we find that

$$\frac{dy}{dx} = -\frac{2xy}{x^2 + e^{-y}}. \tag{10.17}$$

The chain rule gives an alternative way of deriving equations like (10.17). We can define a function $F(x, y)$ by

$$F(x, y) = x^2 y - e^{-y}$$

$F(x, y)$ is a function of two variables. We say that the equation $F(x, y) = 0$ defines y implicitly as a function of x.

We will turn to the general case to see why rewriting an implicitly defined function as a function of two variables is a useful way to calculate derivatives. We think of y as a function of x and define a function $F(x, y) = 0$. This defines y implicitly as a function of x. To find the derivative of y with respect to x, we set

$$w = F(u, v) \quad \text{with } u(x) = x \text{ and } v(x) = y$$

This makes w a function of x; that is, $w = w(x)$. We can now use the chain rule to differentiate w with respect to x:

$$\frac{dw}{dx} = \frac{\partial F}{\partial u}\frac{du}{dx} + \frac{\partial F}{\partial v}\frac{dv}{dx}$$

Since $\dfrac{du}{dx} = 1$ and $\dfrac{dv}{dx} = \dfrac{dy}{dx}$, we obtain, with $u = x$ and $v = y$,

$$\frac{dw}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} \tag{10.18}$$

Because $F(x, y) = 0$, it follows that $w(x) = 0$ for all values of x and, therefore,

$$\frac{dw}{dx} = 0 \tag{10.19}$$

Equating (10.18) and (10.19), we obtain

$$0 = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx}$$

We can isolate dy/dx in this equation:

$$\frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} = -\frac{F_x}{F_y}$$

provided that $\partial F/\partial y \neq 0$. We summarize this result as follows:

Implicit Differentiation

Suppose that $w = F(x, y)$ is differentiable and $F(x, y) = 0$ defines y implicitly as a differentiable function of x. Then, at any point where $F_y \neq 0$,

$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

EXAMPLE 4 Find dy/dx if $x^2 y - e^{-y} = 0$.

Solution We set $F(x, y) = x^2 y - e^{-y}$. We need to find F_x and F_y:

$$F_x = \frac{\partial F}{\partial x} = 2xy$$

$$F_y = \frac{\partial F}{\partial y} = x^2 + e^{-y}$$

Then, since $F(x, y) = 0$,

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{2xy}{x^2 + e^{-y}}$$

as in (10.17). $x^2 + e^{-y} \neq 0$ for all $(x, y) \in \mathbf{R}^2$ because $x^2 \geq 0$ and $e^{-y} > 0$. Since $x^2 + e^{-y} > 0$, dy/dx is defined for all $(x, y) \in \mathbf{R}^2$ with $F(x, y) = 0$. •

The next example shows that we can use this rule to find the derivatives of inverse trigonometric functions.

EXAMPLE 5 Find dy/dx for

$$y = \arcsin x$$

Solution We already know the answer from Chapter 4. But let's see how we can use the results of this section to find the answer.

$$y = \arcsin x \quad \text{for } -1 \leq x \leq 1$$

is equivalent to

$$x = \sin y \quad \text{for } -\frac{\pi}{2} \leq y \leq \frac{\pi}{2}$$

we can define a function $F(x, y)$ that satisfies $F(x, y) = 0$ and defines y implicitly as a function of x, namely,

$$F(x, y) = x - \sin y$$

Then

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{1}{-\cos y} = \frac{1}{\cos y}$$

Since $\sin^2 y + \cos^2 y = 1$, we have

$$\cos y = \sqrt{1 - \sin^2 y} = \sqrt{1 - x^2}$$

Here, we must use the fact that $x = \sin y$ is defined for $-\pi/2 \leq y \leq \pi/2$ and $\cos y \geq 0$ for y in this interval. Using this representation for $\cos y$, we find that

$$\frac{dy}{dx} = \frac{1}{\sqrt{1 - x^2}}$$

which is defined for $-1 < x < 1$. •

Section 10.5 Problems

10.5.1

1. Let $f(x, y) = x^2 + y^2$ with $x(t) = 3t$ and $y(t) = t^2$. Find the derivative of $w = f(x, y)$ with respect to t when $t = 1$.

2. Let $f(x, y) = e^x$ with $x(t) = t$ and $y(t) = t^3$. Find the derivative of $w = f(x, y)$ with respect to t when $t = 0$.

3. Let $f(x, y) = \sqrt{x^2 + y^2}$ with $x(t) = t$ and $y(t) = \sin t$. Find the derivative of $w = f(x, y)$ with respect to t when $t = \pi/3$.

4. Let $f(x, y) = \ln(xy - x^2)$ with $x(t) = t^2$ and $y(t) = t$. Find the derivative of $w = f(x, y)$ with respect to t when $t = 5$.

5. Let $f(x, y) = \frac{1}{x} + \frac{1}{y}$ with $x(t) = t$ and $y(t) = 1 - t$. Find the derivative of $w = f(x, y)$ with respect to t when $t = 1/2$.

6. Let $f(x, y) = xe^y$ with $x(t) = e^t$ and $y(t) = t^2$. Find the derivative of $w = f(x, y)$ with respect to t when $t = 0$.

7. Write down an expression for $\frac{dz}{dt}$ where $z = f(x, y)$ with $x = u(t)$ and $y = v(t)$.

8. Write down an expression for $\frac{dw}{dt}$ where $w = e^{f(x,y)}$ with $x = u(t)$ and $y = v(t)$.

10.5.2

9. Find $\frac{dy}{dx}$ if $\sqrt{x^2 + y^2} = 1$.

10. Find $\frac{dy}{dx}$ if $\frac{xy}{x + y} = 1$.

11. Find $\frac{dy}{dx}$ if $x^2 + y^2 = \ln(xy)$.

12. Find $\frac{dy}{dx}$ if $\cos(x^2 + y^2) = \sin(x^2 - y^2)$.

13. Find $\frac{dy}{dx}$ if $y = \arccos x$.

14. Find $\frac{dy}{dx}$ if $y = \arctan x$.

15. The growth rate r of a particular organism is affected by both the availability of food and the number of competitors for the food source. Denote the amount of food available at time t by $F(t)$ and the number of competitors at time t by $N(t)$. The growth rate r can then be thought of as a function of the two time-dependent variables $F(t)$ and $N(t)$. Assume that the growth rate is an increasing function of the availability of food and a decreasing function of the number of competitors. How is the growth rate r affected if the availability of food decreases over time while the number of competitors increases?

16. Suppose that you travel along an environmental gradient, along which both temperature and precipitation increase. If the abundance of a particular plant species increases with both temperature and precipitation, would you expect to encounter this species more often or less often during your journey? (Use calculus to answer this question.)

17. Air Density The density of air changes with height. Under some conditions density ρ, depends on height z, and temperature T according to the equation:

$$\rho(z, T) = \rho_0 e^{-\lambda z/T}$$

where ρ_0 and λ are both constants.

A meteorological balloon ascends (i.e., starts at $z = 1$ and gains height) over the course of several hours.

(a) Assuming that the balloon ascends at a speed v (i.e., $dz/dt = v$) and that the temperature changes over time (i.e., that T is given by a function $T(t)$), derive, using the chain rule, an expression for the rate of change of air density, as measured by the weather balloon.

(b) Assume that $v = 1, \rho_0 = 1$, and $\lambda = 1$ and that when $t = 0, T = 1$. Are there any conditions under which the density, as measured by the balloon will not change in time? That is, find a differential equation that T must satisfy, if $d\rho/dt = 0$, and solve this differential equation.

18. A bacterium swims in a environment containing a chemical whose concentration is $c(x, t) = 1 - \frac{x}{1+t}$. The bacterium swims at a speed $\frac{dx}{dt} = -1$, starting at $x(0) = 1$. Calculate the rate of change of the chemical concentration, as measured by the bacterium when $t = 0$.

19. Maturity Time for Fish In Chapter 8 we met the von Bertalanffy equation as a model for the growth of a fish. The length $L(t)$ of the fish is modeled by a function of age t by the function:

$$L(t) = L_\infty + (L_0 - L_\infty)e^{-kt}$$

where L_0, L_∞, and k are all positive coefficients. We define the maturity age of the fish m to be the age x, at which the fish reaches 90% of its maximum length L_∞; that is

$$0.9L_\infty = L_\infty + (L_0 - L_\infty)e^{-km}$$

Show that, if L_∞ and L_0 are constants, then the maturity age decreases if k increases. [*Hint*: show that $dm/dk < 0$.]

20. Earthquake Frequency The time between two earthquakes is sometimes modeled using a Poisson process model. According to this model the probability that a second earthquake follows within a time t of the first earthquake is:

$$P(t) = 1 - e^{-\lambda t}$$

where λ is a positive constant. We define the average time between earthquakes to be a time T as follows: The probability that a second earthquake follows the first earthquake within time T is exactly 1/2, i.e.:

$$1/2 = 1 - e^{-\lambda T}$$

T is a function of λ. Show that, if the coefficient λ is increased in the model, then the average time between earthquakes will decrease.

10.6 Directional Derivatives and Gradient Vectors

Suppose that you are on a sloped surface, such as a hillside. Depending on which direction you walk, you must either go uphill, stay at the same level, or go downhill. That is, by choosing a particular direction, you have some control over the steepness of your path. How steep your path is can be described by the slope of the tangent line at your starting point in the direction of your path. This slope is given by the **directional derivative**. We will derive the directional derivative first using the chain rule (from Section 10.5.1). We will then re-derive the same result without using the chain rule, which means that the material in this section, which will be needed to solve optimization problems in Section 10.7, can be studied without needing to study Section 10.5.

10.6.1 Deriving the Directional Derivative

We assume that $z = f(x, y)$ is a differentiable function of two independent variables. We choose a point (x_0, y_0, z_0), with $z_0 = f(x_0, y_0)$ on the surface defined by $z = f(x, y)$. To define the slope of a tangent line at (x_0, y_0, z_0), we must specify a direction in which we wish to go. We know how to deal with this problem when we go in the x- or y-direction. In these cases, the partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ tell us how $f(x, y)$ changes. We will now explain how we can express the slope when we choose an arbitrary direction.

The first step is to find a way to express what we mean by "going in a certain direction." We start at a point (x_0, y_0) and wish to go in the direction of a unit vector $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$. (Recall that a unit vector has length 1.) This is illustrated in Figure 10.54, from which we see that

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{u} \tag{10.20}$$

where $\mathbf{r}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$, $\mathbf{r} = \begin{bmatrix} x \\ y \end{bmatrix}$, and t is a real number; different values of t get us to different points on the straight line through (x_0, y_0) that points in the direction of $\mathbf{u}$. We can also write (10.20) as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} tu_1 \\ tu_2 \end{bmatrix} \quad \text{for } t \in \mathbf{R} \tag{10.21}$$

Equation (10.21) is called the **parametric equation of a line**, which we introduced in Section 9.5; t is called the *parameter*. Since $x = x_0 + tu_1$ and $y = y_0 + tu_2$, it follows that

$$\frac{dx}{dt} = u_1 \qquad \text{and} \qquad \frac{dy}{dt} = u_2 \tag{10.22}$$

We can now find out what happens to the value of $f(x, y)$ when we start at (x_0, y_0) and go in the direction of the unit vector $\mathbf{u}$, as illustrated in Figure 10.55. To begin, we use the parametric equation (10.21) of the line that passes through (x_0, y_0) and that is oriented in the direction of $\mathbf{u}$:

$$x = x_0 + tu_1 \qquad \text{and} \qquad y = y_0 + tu_2$$

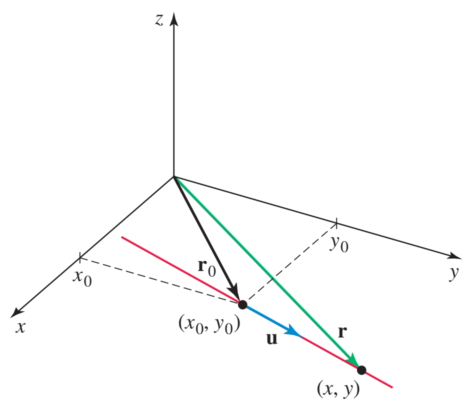


Figure 10.54 Going in the direction of $\mathbf{u}$ from (x_0, y_0).

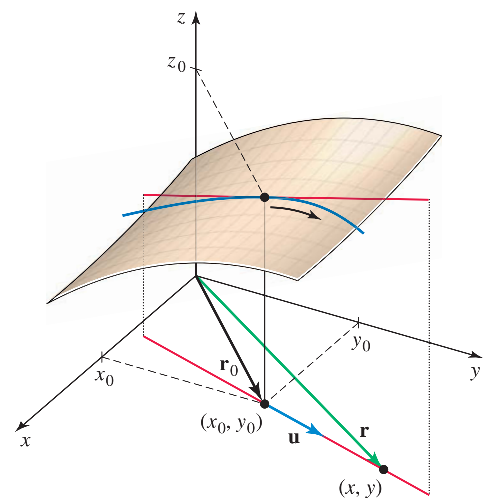


Figure 10.55 An illustration of the directional derivative.

Deriving the Directional Derivative Using the Chain Rule. Since $z(t) = f(x(t), y(t))$, we can use the chain rule to find out how f changes when we vary t (i.e., as we move along the straight line):

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = \frac{\partial f}{\partial x}u_1 + \frac{\partial f}{\partial y}u_2 \qquad \text{Substitute for } \tfrac{dx}{dt} \text{ and } \tfrac{dy}{dt} \text{ from (10.22)} \tag{10.23}$$

Deriving the Directional Derivative Without Using the Chain Rule. To understand directional derivatives (which are needed for Section 10.7) without first learning the chain rule for multivariable functions, we will show how Equation (10.21) can be derived directly from the standard linearization of a function, which was introduced in Section 10.4. Note that if

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} tu_1 \\ tu_2 \end{bmatrix}$$

then if x is close to x_0 and y is close to y_0 (i.e., t is small) then the standard linearization of $f(x, y)$ gives:

$$\begin{aligned} f(x, y) &= f(x_0, y_0) + (x - x_0)\frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0)\frac{\partial f}{\partial y}(x_0, y_0) \\ &= f(x_0, y_0) + tu_1\frac{\partial f}{\partial x}(x_0, y_0) + tu_2\frac{\partial f}{\partial y}(x_0, y_0) \end{aligned}$$

which may be rewritten as:

$$\frac{f(x,y)-f(x_0,y_0)}{t} \approx u_1\frac{\partial f}{\partial x}(x_0,y_0)+u_2\frac{\partial f}{\partial y}(x_0,y_0)$$

Now if t is made smaller and smaller (i.e., (x, y) gets closer and closer to (x_0, y_0)), the right-hand side of this equation agrees more and more accurately with the left-hand side. Moreover, as $t \to 0$, the left-hand side converges to the rate of change of f, as measured by the observer traveling along the straight line. So again we arrive at Equation (10.23).

Equation (10.23) can be written as a dot product:

$$\frac{dz}{dt} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

The first vector in the dot product is called the **gradient** of f.

Definition Assume that $z = f(x, y)$ is a function of two independent variables and that $\partial f/\partial x$ and $\partial f/\partial y$ exist. Then the vector

$$\nabla f(x,y) = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x} \\ \frac{\partial f(x,y)}{\partial y} \end{bmatrix}$$

is called the gradient of f at (x, y).

The notation ∇f is read "grad f" or "gradient of f." The symbol ∇ is called "del," so you can also say "del f." An alternative notation is grad f.

We can now define the derivative of f in a particular direction:

Definition The directional derivative of $f(x, y)$ at (x_0, y_0) in the direction of the unit vector $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is

$$D_{\mathbf{u}} f(x_0,y_0) = (\nabla f(x_0,y_0)) \cdot \mathbf{u} \qquad \text{Do not confuse } D_{\mathbf{u}} \text{ with a Jacobi matrix} \qquad (10.24)$$

Note that in the definition of the directional derivative, we assume that $\mathbf{u}$ is a *unit vector*. Choosing a unit vector (as opposed to a vector of some other length) ensures that the directional derivative of $f(x, y)$ agrees with the partial derivatives when we go along the positive x- or y-axis. That is, if $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, then

$$D_{\mathbf{u}} f(x_0,y_0) = \begin{bmatrix} \frac{\partial f(x_0,y_0)}{\partial x} \\ \frac{\partial f(x_0,y_0)}{\partial y} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\partial f(x_0,y_0)}{\partial x}$$

and if $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then

$$D_{\mathbf{u}} f(x_0,y_0) = \frac{\partial f(x_0,y_0)}{\partial y}$$

EXAMPLE 1 Compute the directional derivative of

$$f(x,y) = \sqrt{x^2+2y^2}$$

at the point $(-1, 2)$ in the direction $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$.

Solution We first compute the gradient vector

$$\nabla f(x, y) = \begin{bmatrix} \frac{x}{\sqrt{x^2+2y^2}} \\ \frac{2y}{\sqrt{x^2+2y^2}} \end{bmatrix}$$

Evaluating this vector at $(-1, 2)$, we find that

$$\nabla f(-1, 2) = \begin{bmatrix} \frac{-1}{\sqrt{1+8}} \\ \frac{4}{\sqrt{1+8}} \end{bmatrix} = \begin{bmatrix} -1/3 \\ 4/3 \end{bmatrix}$$

Since $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ is not a unit vector, we normalize it first. The vector $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ has length $\sqrt{1+9} = \sqrt{10}$; hence,

$$\mathbf{u} = \frac{1}{\sqrt{10}} \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

and

$$\begin{aligned} D_{\mathbf{u}} f(-1, 2) &= (\nabla f(-1, 2)) \cdot \mathbf{u} \\ &= \begin{bmatrix} -\frac{1}{3} \\ \frac{4}{3} \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{bmatrix} \\ &= \frac{1}{3\sqrt{10}} + \frac{12}{3\sqrt{10}} = \frac{13}{3\sqrt{10}}. \end{aligned}$$

•

EXAMPLE 2 Compute the directional derivative of

$$f(x, y) = x^2 y - 2y^2$$

at the point $(-3, 2)$ in the direction of $(-1, 1)$.

Solution We first compute the gradient vector

$$\nabla f(x, y) = \begin{bmatrix} 2xy \\ x^2 - 4y \end{bmatrix}$$

Evaluating this vector at $(-3, 2)$, we find that

$$\nabla f(-3, 2) = \begin{bmatrix} (2)(-3)(2) \\ (-3)^2 - (4)(2) \end{bmatrix} = \begin{bmatrix} -12 \\ 1 \end{bmatrix}$$

The vector that goes from $(-3, 2)$ to $(-1, 1)$ has the form

$$\begin{bmatrix} -1 - (-3) \\ 1 - 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

This vector has length $\sqrt{4+1} = \sqrt{5}$. Normalizing the vector yields

$$\mathbf{u} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

and the directional derivative is

$$\begin{aligned} D_{\mathbf{u}} f(-3, 2) &= (\nabla f(-3, 2)) \cdot \mathbf{u} \\ &= \frac{1}{\sqrt{5}} \begin{bmatrix} -12 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{5}}(-24 - 1) = -\frac{25}{\sqrt{5}} = -5\sqrt{5} \end{aligned}$$

•

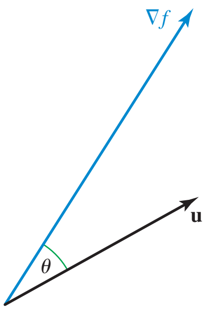


Figure 10.56 The angle θ between ∇f and the unit vector $\mathbf{u}$.

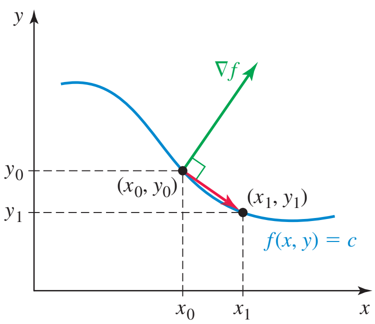


Figure 10.57 The gradient is perpendicular to the level curve.

10.6.2 Properties of the Gradient Vector

The directional derivative is a dot product. We can therefore write

$$\begin{aligned} D_{\mathbf{u}} f(x, y) &= (\nabla f(x, y)) \cdot \mathbf{u} \\ &= |\nabla f(x, y)||\mathbf{u}| \cos\theta \qquad \text{Use geometric interpretation of dot product.} \end{aligned}$$

where θ is the angle between ∇f and $\mathbf{u}$. (See Figure 10.56.) Since $|\mathbf{u}| = 1$ ($\mathbf{u}$ is a unit vector), we have

$$D_{\mathbf{u}} f(x, y) = |\nabla f(x, y)| \cos\theta$$

The angle θ is in the interval $[0, 2\pi)$, and $\cos\theta$ is maximal when $\theta = 0$. We therefore find that $D_{\mathbf{u}} f(x, y)$ is maximal when $\mathbf{u}$ is in the direction of ∇f. That is, the most rapid rate of increase of f is measured by an observer who travels in the direction of ∇f.

We will now show that, geometrically, the gradient vector at a point (x_0, y_0) is perpendicular to the level curve $f(x, y) = c$ that passes through this point. The level curve $f(x, y) = c$ is a curve in the x–y plane. Let's start at a point (x_0, y_0) on the level curve $f(x, y) = c$. If (x_1, y_1) is a nearby point on the same level curve (see Figure 10.57), then using the standard linearization we see that:

$$\begin{aligned} f(x_1, y_1) \approx f(x_0, y_0) &+ (x_1 - x_0)\frac{\partial f}{\partial x} f(x_0, y_0) \\ &+ (y_1 - y_0)\frac{\partial f}{\partial y} f(x_0, y_0) \end{aligned}$$

But since (x_0, y_0) and (x_1, y_1) lie on the same level curve, $f(x_0, y_0) = f(x_1, y_1) = c$ and so:

$$\begin{aligned} 0 &\approx (x_1 - x_0)\frac{\partial f}{\partial x}(x_0, y_0) + (y_1 - y_0)\frac{\partial f}{\partial y}(x_0, y_0) \\ &= \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) \\ \frac{\partial f}{\partial y}(x_0, y_0) \end{bmatrix} \cdot \begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \end{bmatrix} = \nabla f(x_0, y_0) \cdot \begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \end{bmatrix} \end{aligned}$$

And this equation is more and more exactly satisfied, the closer the point (x_1, y_1) is chosen to (x_0, y_0). Now $\begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \end{bmatrix}$ is a vector that points from the point (x_0, y_0) to the point (x_1, y_1). Since the dot product of this vector with $\nabla f(x_0, y_0)$ is equal to 0, the two vectors are orthogonal. Now, we know that as the point (x_1, y_1) approaches (x_0, y_0), the direction of the vector $\begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \end{bmatrix}$ will approach the tangent to the curve on which both points lie. So the above result shows us that ∇f is perpendicular to this tangent vector, i.e., perpendicular to the level curve.

Properties of the Gradient Suppose that $f(x, y)$ is a differentiable function. The gradient vector $\nabla f(x, y)$ has the following properties:

1. At each point (x_0, y_0), $f(x, y)$ increases most rapidly in the direction of the gradient vector $\nabla f(x_0, y_0)$.
2. The gradient vector of f at a point (x_0, y_0) is perpendicular to the level curve through (x_0, y_0).

EXAMPLE 3 Let $f(x, y) = x^2 y + y^2$. In what direction does $f(x, y)$ increase most rapidly at $(1, 1)$?

Solution The function $f(x, y)$ increases most rapidly at $(1, 1)$ in the direction of $\nabla f(1, 1)$. Since

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2xy \\ x^2 + 2y \end{bmatrix}$$

it follows that

$$\nabla f(1,1) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

That is, $f(x, y)$ increases most rapidly in the direction $\left[\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right]$ at the point $(1, 1)$.

EXAMPLE 4 Find a unit vector that is perpendicular to the level curve of the function $f(x, y) = x^2 - y^2$, at $(1, 2)$.

Solution The gradient of f at $(1, 2)$ is perpendicular to the level curve at $(1, 2)$. The gradient of f is given as

$$\nabla f(x, y) = \begin{bmatrix} 2x \\ -2y \end{bmatrix}.$$

Hence,

$$\nabla f(1, 2) = \begin{bmatrix} 2 \\ -4 \end{bmatrix}.$$

To normalize this vector, we divide $\nabla f(1, 2)$ by its length. Since

$$|\nabla f(1, 2)| = \sqrt{(2)^2 + (-4)^2} = \sqrt{4 + 16} = 2\sqrt{5},$$

the unit vector that is perpendicular to the level curve of $f(x, y)$ at $(1, 2)$ is

$$\mathbf{u} = \frac{1}{2\sqrt{5}} \begin{bmatrix} 2 \\ -4 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} \frac{1}{5}\sqrt{5} \\ -\frac{2}{5}\sqrt{5} \end{bmatrix}.$$

Section 10.6 Problems

10.6

In Problems 1–8, find the gradient of each function.

1. $f(x, y) = x^3 y^2$

2. $f(x, y) = \frac{xy}{x^2+y^2}$

3. $f(x, y) = \sqrt{x^3 - 3xy}$

4. $f(x, y) = x(x^2 - y^2)^{2/3}$

5. $f(x, y) = \exp\left[\sqrt{x^2 + y^2}\right]$

6. $f(x, y) = \tan \frac{x-y}{x+y}$

7. $f(x, y) = \ln\left(\frac{x}{y} + \frac{y}{x}\right)$

8. $f(x, y) = \cos(3x^2 - 2y^2)$

In Problems 9–14, compute the directional derivative of $f(x, y)$ at the given point in the indicated direction.

9. $f(x, y) = \sqrt{x^2 + 2y^2}$ at $(-1, 2)$ in the direction $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$

10. $f(x, y) = \sin(x + y)$ at $(-1, 0)$ in the direction $\left[\begin{smallmatrix} 2 \\ -1 \end{smallmatrix}\right]$

11. $f(x, y) = \exp(x + y^2)$ at $(0, 0)$ in the direction $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$

12. $f(x, y) = x^3 y^2$ at $(2, 3)$ in the direction $\left[\begin{smallmatrix} -2 \\ 1 \end{smallmatrix}\right]$

13. $f(x, y) = 2xy^3 + x^2 y$ at $(1, -1)$ in the direction $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$

14. $f(x, y) = \dfrac{1}{\sqrt{x^2 + y^2}}$ at $(0, 1)$ in the direction $\left[\begin{smallmatrix} 4 \\ -1 \end{smallmatrix}\right]$

In Problems 15–18, compute the directional derivative of $f(x, y)$ at the point P in the direction of the point Q.

15. $f(x, y) = 2x^2 y - 3x$, $P = (2, 1)$, $Q = (3, 2)$

16. $f(x, y) = 4xy + y^2$, $P = (-1, 1)$, $Q = (3, 2)$

17. $f(x, y) = \sqrt{xy - 2x^2}$, $P = (1, 6)$, $Q = (3, 1)$

18. $f(x, y) = e^{x-y}$, $P = (2, 2)$, $Q = (1, -1)$

19. In what direction does $f(x, y) = 3xy - x^2$ increase most rapidly at $(-1, 1)$?

20. In what direction does $f(x, y) = e^x \cos y$ increase most rapidly at $(0, \pi/2)$?

21. In what direction does $f(x, y) = \sqrt{x^2 - y^2}$ increase most rapidly at $(5, 3)$?

22. In what direction does $f(x, y) = \ln(x^2 + y^2)$ increase most rapidly at $(1, 1)$?

23. Find a unit vector that is normal to the level curve of the function

$$f(x, y) = 3x + 4y$$

at the point $(-1, 1)$.

24. Find a unit vector that is normal to the level curve of the function

$$f(x, y) = x^2 + \frac{y^2}{9}$$

at the point $(1, 3)$.

25. Find a unit vector that is normal to the level curve of the function

$$f(x, y) = x^2 - y^3$$

at the point $(1, 3)$.

26. Find a unit vector that is normal to the level curve of the function

$$f(x, y) = xy$$

at the point $(2, 3)$.

27. Chemotaxis Chemotaxis is the chemically directed movement of organisms up a concentration gradient—that is, in the direction in which the concentration increases most rapidly. The slime mold *Dictyostelium discoideum* exhibits this phenomenon. Single-celled amoebas of this species move up the concentration gradient of a chemical called cyclic adenosine monophosphate (AMP). Suppose the concentration of cyclic AMP at the point (x, y) in the x–y plane is given by

$$f(x, y) = \frac{4}{\sqrt{x^2 + y^2 + 1}}$$

If you place an amoeba at the point (3, 1) in the x–y plane, determine in which direction the amoeba will move if its movement is directed by chemotaxis.

28. Given a function $f(x, y)$ that is defined and differentiable on an open ball containing the point (x_0, y_0), show that the function f **decreases** most rapidly in the direction of $-\nabla f(x_0, y_0)$.

29. Suppose an organism moves down a sloped surface along the steepest line of descent, i.e., the direction in which the surface decreases most rapidly. If the surface is given by

$$f(x, y) = x^2 - y^2$$

find the direction in which the organism will move at the point (2, 3).

30. Chemotaxis A bacterium swimming in still water with sugar concentration $c(x, y)$ will tend to swim in the direction in which $c(x, y)$ is increasing most rapidly. Suppose a bacterium finds itself in an environment where the sugar concentration is shown by the level curves in Figure 10.58.

Use the level curves shown in the figure to sketch the direction that a bacterium located at each of the points A, B, and C would swim in.

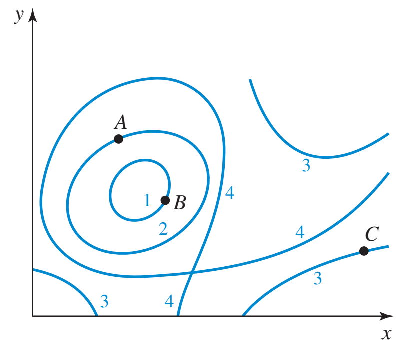


Figure 10.58 Level curves of $c(x, y)$ for Problem 30.

10.7 Maximization and Minimization of Functions

10.7.1 Local Maxima and Minima

In Section 5.1, we introduced local extrema for functions of one variable. Local extrema can also be defined for functions of more than one independent variable; here, we will restrict our discussion to functions of two variables. Recall that we denoted by $B_\delta(x_0, y_0)$ the open disk with radius δ centered at (x_0, y_0). The following definition, with which you should compare the corresponding definition in Section 5.1, extends the notion of local extrema to functions of two variables:

Definition A function $f(x, y)$ defined on a set $D \subset \mathbf{R}^2$ has a **local** (or **relative**) **maximum** at a point (x_0, y_0) if there exists a $\delta > 0$ such that

$$f(x, y) \le f(x_0, y_0) \quad \text{for all } (x, y) \in B_\delta(x_0, y_0) \cap D$$

A function $f(x, y)$ defined on a set $D \subset \mathbf{R}^2$ has a **local** (or **relative**) **minimum** at a point (x_0, y_0) if there exists a $\delta > 0$ such that

$$f(x, y) \ge f(x_0, y_0) \quad \text{for all } (x, y) \in B_\delta(x_0, y_0) \cap D$$

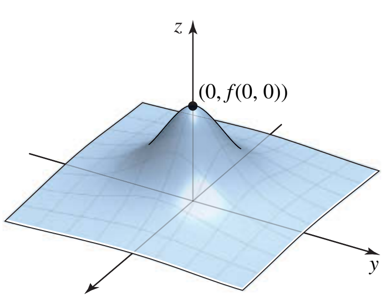


Figure 10.59 The graph of a function $f(x, y)$ with a local maximum at (0, 0).

Informally, a local maximum (local minimum) is a point that is higher (lower) than all nearby points. We can define **global** (or **absolute**) extrema as well: If the inequalities in the definition hold for all $(x, y) \in D$, then f has an global maximum (minimum) at (x_0, y_0). Figure 10.59 shows an example of a function of two variables with a local maximum at (0, 0).

How can we find local extrema? Recall that in the single-variable case, a horizontal tangent line at a point on the graph of a differentiable function is a necessary condition for the point to be a local extremum (Fermat's theorem). We can generalize this statement to functions of more than one variable: Looking at Figure 10.59, we see that the tangent plane at the local extremum will be horizontal. The equation of a horizontal tangent plane on the graph of a differentiable function $f(x, y)$ at (x_0, y_0) is

$$z = f(x_0, y_0)$$ Right hand side is constant

Comparing this equation with the general form of a tangent plane (Section 10.4), we see that both $\partial f/\partial x$ and $\partial f/\partial y$ are equal to 0 at (x_0, y_0); in other words, $\nabla f(x_0, y_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Another way to see why the gradient of f needs to vanish at a local extremum is to argue that if the gradient were nonzero at (x_0, y_0), then the function would increase in the direction of the gradient and decrease in the opposite direction, so we could not be at a local extremum. We thus have the following criterion:

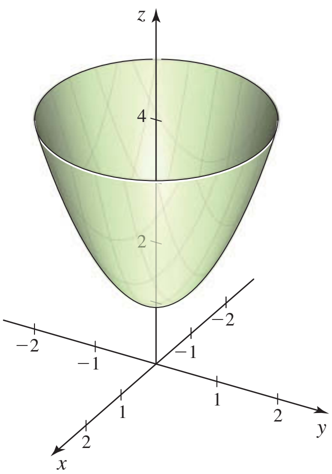


Figure 10.60 The graph of the function $f(x, y) = x^2 + y^2 + 1$.

Criterion for a Local Extremum

If $f(x, y)$ has a local extremum at (x_0, y_0) and if f is differentiable at (x_0, y_0), then

$$\nabla f(x_0, y_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{10.25}$$

A point (x_0, y_0) that satisfies (10.25) is called a **critical point**; points where $f(x, y)$ is not differentiable are also called critical points. As in Section 5.1, (10.25) only identifies *candidates* for local extrema. Other critical points are also just candidates for local extrema. A further investigation is then needed to determine whether a candidate is indeed a local extremum.

EXAMPLE 1 Figure 10.60 shows the graph of the differentiable function $f(x, y) = x^2 + y^2 + 1$, $(x, y) \in \mathbf{R}^2$. We see that $f(x, y)$ has a local minimum at $(0, 0)$. Show that $\nabla f(0, 0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and determine the equation of the tangent plane at $(0, 0)$.

Solution We compute

$$\nabla f(x, y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

Evaluating $\nabla f(0, 0)$ at $(0, 0)$, we find that

$$\nabla f(0, 0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The equation of the tangent plane at $(0, 0)$ is given by

$$z = f(0, 0) + (x - 0) f_x(0, 0) + (y - 0) f_y(0, 0)$$

Since $f(0, 0) = 1$, $f_x(0, 0) = 0$, and $f_y(0, 0) = 0$, the equation of the tangent plane at $(0, 0)$ is $z = 1$, which shows that the tangent plane is horizontal. •

EXAMPLE 2 Find all critical points of

$$f(x, y) = x^2 + y^2 + xy, \quad (x, y) \in \mathbf{R}^2$$

Solution Since the function $f(x, y)$ is differentiable in $\mathbf{R}^2$, the only critical points are points that satisfy $\nabla f(x, y) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Now,

$$\nabla f(x, y) = \begin{bmatrix} 2x + y \\ 2y + x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

We need to solve the system of linear equations

$$2x + y = 0$$
$$x + 2y = 0$$

It follows from the first equation that $y = -2x$. Substituting this into the second equation yields

$$x + 2(-2x) = 0, \qquad \text{or} \qquad -3x = 0, \qquad \text{or} \qquad x = 0$$

and, therefore, $y = 0$. The function thus has one critical point: $(0, 0)$. •

We now give a sufficient condition that will allow us to determine whether a candidate for a local extremum is indeed a local extremum and, if so, whether it is a local maximum or a local minimum. The proof of this condition is beyond the scope of this book, but you should memorize it and learn how to apply it.

Recall that in the case of a function of one variable we obtained the following sufficient condition for twice-differentiable functions: If $f'(x_0) = 0$ and $f''(x_0) > 0$, then $f(x)$ has a local minimum (or a local maximum if $f''(x_0) < 0$) at $x = x_0$. In the multivariable case, there is an analogous condition involving second partial derivatives.

Second Derivative Test for Identifying Local Minima and Maxima Suppose the second partial derivatives of f are continuous in a disk centered at (x_0, y_0). Suppose also that $\nabla f(x_0, y_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Define

$$D = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - \left(f_{xy}(x_0, y_0)\right)^2$$

1. If $D > 0$ and $f_{xx}(x_0, y_0) > 0$, then f has a local minimum at (x_0, y_0).
2. If $D > 0$ and $f_{xx}(x_0, y_0) < 0$, then f has a local maximum at (x_0, y_0).
3. If $D < 0$, then f does not have a local extremum at (x_0, y_0). The point (x_0, y_0) is then called a **saddle point**.

In all other cases, the test is inconclusive. We can compare this with the situation for functions of a single variable. If $f'(x_0) = 0$ and $f''(x_0) = 0$, then we cannot determine whether x_0 is a local minimum, a local maximum, or neither. We now return to Example 2 and determine whether $(0, 0)$ is a local extremum.

EXAMPLE 2 **(continued)** Determine whether the critical point $(0, 0)$ of $f(x, y) = x^2 + y^2 + xy$ in Example 2 is a local maximum or a local minimum.

Solution We need to find all second partial derivatives. Since

$$\frac{\partial f(x, y)}{\partial x} = 2x + y \qquad \text{and} \qquad \frac{\partial f(x, y)}{\partial y} = x + 2y$$

we have

$$\frac{\partial^2 f}{\partial x^2} = 2, \quad \frac{\partial^2 f}{\partial y^2} = 2, \quad \frac{\partial^2 f}{\partial x \partial y} = 1$$

Hence,

$$D = f_{xx} f_{yy} - f_{xy}^2 = (2)(2) - (1)^2 = 3 > 0$$

Since $D > 0$ and $f_{xx} > 0$, we conclude that $(0, 0)$ is a local minimum, which is confirmed when we use a computer to plot the graph of $f(x, y)$, in Figure 10.61. •

EXAMPLE 3 Find all local extrema of

$$f(x, y) = 3xy - x^3 - y^3, \quad (x, y) \in \mathbf{R}^2$$

and classify them according to whether each is a local maximum, a local minimum, or neither.

Solution The function $f(x, y)$ is differentiable on its domain. The critical points thus satisfy

$$\nabla f(x, y) = \begin{bmatrix} 3y - 3x^2 \\ 3x - 3y^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So to find the critical point we must solve the system of equations:

$$y - x^2 = 0 \qquad (R_1)$$

$$x - y^2 = 0 \qquad (R_2)$$

These equations are not linear in x and y, so we cannot use the methods from Chapter 9 to solve them. However, we can use one equation to eliminate one of the variables from the other. For example, (R_1) tells us that $y = x^2$. So substituting for y in (R_2) we then have:

$$x - (x^2)^2 = x - x^4 = 0$$

$$x(1 - x^3) = 0$$

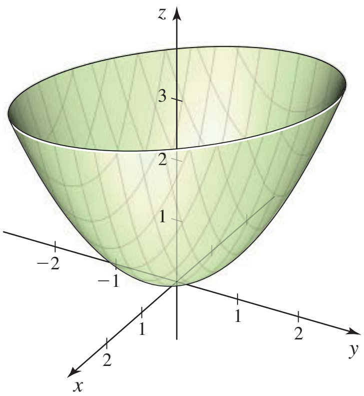


Figure 10.61 The graph of the function $f(x, y) = x^2 + y^2 + xy$.

for which $x = 0$ and $x = 1$ are both solutions. Then (R_1) yields $y = x^2$, so the set of equations has the solutions $(0, 0)$ and $(1, 1)$. Now,

$$\frac{\partial^2 f(x, y)}{\partial x^2} = -6x, \quad \frac{\partial^2 f(x, y)}{\partial y^2} = -6y, \quad \frac{\partial^2 f(x, y)}{\partial x \partial y} = 3$$

Therefore,

$$D = f_{xx} f_{yy} - (f_{xy})^2 = 36xy - 9$$

At $(1, 1)$, $D = 36 - 9 > 0$. Since $f_{xx}(1, 1) = -6 < 0$, $f(x, y)$ has a local maximum at $(1, 1)$. At $(0, 0)$, $D = -9 < 0$. The critical point $(0, 0)$ is neither a local maximum nor a local minimum, since $D < 0$ (Figure 10.62). •

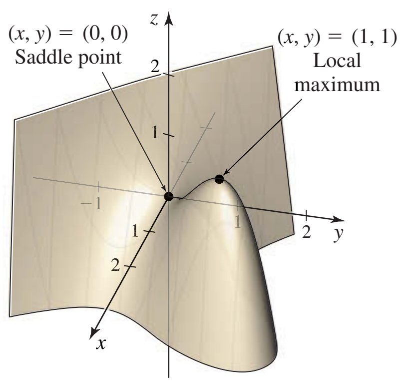


Figure 10.62 The graph of the function $f(x, y) = 3xy - x^3 - y^3$, with critical points marked.

A Sufficient Condition Based on Eigenvalues (Optional). We will now give a sufficient condition that is phrased in terms of eigenvalues to determine whether a candidate for a local extremum is indeed a local extremum and, if so, what type (i.e., local maximum or local minimum). To motivate this condition, we will look at the following three functions defined for $(x, y) \in \mathbf{R}^2$:

$$f_1(x, y) = x^2 + y^2, \qquad f_2(x, y) = x^2 - y^2, \qquad f_3(x, y) = -x^2 - y^2$$

These functions are illustrated in Figures 10.63 through 10.65.

Computing $\nabla f_i(x, y)$, $i = 1, 2$, and 3, we find that

$$\nabla f_1(x, y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}, \qquad \nabla f_2(x, y) = \begin{bmatrix} 2x \\ -2y \end{bmatrix}, \qquad \nabla f_3(x, y) = \begin{bmatrix} -2x \\ -2y \end{bmatrix}$$

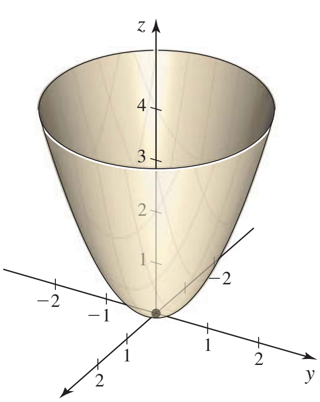


Figure 10.63 The graph of the function $f_1(x, y) = x^2 + y^2$.

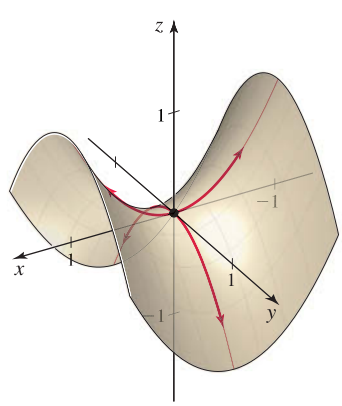


Figure 10.64 The graph of the function $f_2(x, y) = x^2 - y^2$.

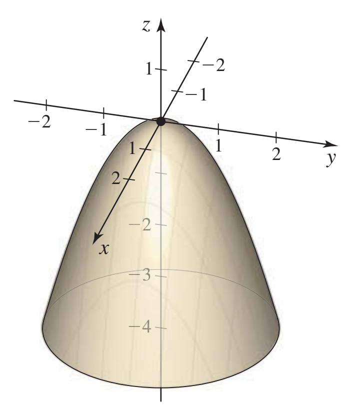


Figure 10.65 The graph of the function $f_3(x, y) = -x^2 - y^2$.

$(0, 0)$ is a candidate for a local extremum for all three functions, since $\nabla f_1(0, 0) = \nabla f_2(0, 0) = \nabla f_3(0, 0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

The analogue of a second derivative of a function of one variable for a function of two variables with continuous second partial derivatives is the following second-derivative matrix, called the **Hessian matrix**:

$$\text{Hess } f(x, y) = \begin{bmatrix} \frac{\partial^2 f(x,y)}{\partial x^2} & \frac{\partial^2 f(x,y)}{\partial y \partial x} \\ \frac{\partial^2 f(x,y)}{\partial x \partial y} & \frac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix}$$

Computing this matrix for the functions $f_i(x, y)$, $i = 1, 2,$ and 3, we obtain

$$\text{Hess } f_1(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \qquad \text{Hess } f_2(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix},$$
$$\text{Hess } f_3(x, y) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

It turns out (through the proof is beyond the scope of this book) that the eigenvalues of the second-derivative matrix provide a sufficient condition for determining whether a critical point is a local maximum, a local minimum, or neither. The following holds:

Eigenvalue Test for Identifying Local Minima and Maxima Suppose the second partial derivatives of f are continuous in a disk centered at (x_0, y_0). Suppose also that $\nabla f(x_0, y_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Then

1. If the two eigenvalues of the second-derivative matrix Hess $f(x_0, y_0)$ at (x_0, y_0) are positive, then f has a local minimum at (x_0, y_0).
2. If the two eigenvalues of Hess $f(x_0, y_0)$ are negative, then f has a local maximum at (x_0, y_0).
3. If the two eigenvalues of Hess $f(x_0, y_0)$ are of opposite signs, then f does not have a local extremum at (x_0, y_0). The point (x_0, y_0) is then called a **saddle point**.

In all other cases, the test is inconclusive. Returning to our example, we need to evaluate Hess $f_i(x, y)$, $i = 1, 2,$ and 3, at $(0, 0)$. In fact in all three cases the Hessian matrices are independent of (x, y), so Hess $f_i(0, 0) =$ Hess $f_i(x, y)$. We can read off the eigenvalues of each of the second-derivative matrices evaluated at $(0, 0)$, since they are in diagonal form (see Section 9.3). The eigenvalues of Hess $f_1(0, 0)$ are both 2; hence, $f_1(0, 0)$ is a local minimum, which agrees with the graph in Figure 10.63. The eigenvalues of Hess $f_2(0, 0)$ are 2 and -2, and we conclude that $f_2(0, 0)$ is not a local extremum. (See Figure 10.64.) The eigenvalues of Hess $f_3(0, 0)$ are both -2, and we conclude that $f_3(0, 0)$ is a local maximum. (See Figure 10.65.)

$f_2(x, y)$ does not have a local extremum at $(x, y) = (0, 0)$. From the graph in Figure 10.64 we can see that an observer who leaves the point $(x, y) = (0, 0)$ and follows the surface that is given by the function will see f_2 increase if they walk in the direction of the x-axis, and decrease if they walk in the direction of the y-axis. The graph resembles a saddle (or alternatively, the shape of a Pringles™ potato chip) near $(0, 0)$, which is why such critical points are called saddle points.

We assumed in the second-derivative criterion that all second partial derivatives are continuous in a disk centered at (x_0, y_0). This assumption implies that

$$\frac{\partial^2 f(x, y)}{\partial x \, \partial y} = \frac{\partial^2 f(x, y)}{\partial y \, \partial x}$$

That is, the off-diagonal elements of Hess $f(x, y)$ are identical and the Hessian matrix is of the form $\begin{bmatrix} a & c \\ c & b \end{bmatrix}$. Such a matrix is called **symmetric**. We can show that the eigenvalues of a symmetric matrix are always real. (See Problem 34.) This fact has an important consequence: If the second partial derivatives of f are continuous in a disk centered at (x_0, y_0), then the eigenvalues of Hess $f(x_0, y_0)$ are both real. Provided that neither eigenvalue is equal to zero, one of the three cases in our second-derivative criterion occurs, thereby allowing us to settle the question whether the candidate (x_0, y_0) for which $\nabla f(x_0, y_0) = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$ is a local extremum and, if so, of what type it is. If one or both eigenvalues are equal to zero, we cannot say anything about the nature of the critical point on the basis of the Hessian matrix. This case is discussed in Problem 11.

EXAMPLE 4 Find the local extrema of

$$f(x, y) = 2x^2 - xy + y^4, \quad (x, y) \in \mathbf{R}^2$$

Solution We compute

$$\nabla f(x, y) = \begin{bmatrix} 4x - y \\ -x + 4y^3 \end{bmatrix}$$

Setting both partial derivatives equal to 0, we find that

$$4x - y = 0 \quad (R_1) \qquad \text{and} \qquad -x + 4y^3 = 0 \quad (R_2)$$

It follows from (R_1) that $x = y/4$. Substituting this into (R_2), we obtain

$$-\frac{y}{4} + 4y^3 = 0$$
$$-\frac{y}{4}(1 - 16y^2) = 0$$

yielding

$$y_1 = 0, \quad y_2 = \frac{1}{4}, \quad \text{and} \quad y_3 = -\frac{1}{4}.$$

The corresponding x-values are

$$x_1 = 0, \quad x_2 = \frac{1}{16}, \quad \text{and} \quad x_3 = -\frac{1}{16} \qquad x = y/4$$

Since ∇f is defined for all $(x, y) \in \mathbf{R}^2$, there are no other critical points. The three candidates for local extrema are thus

$$(0, 0), \quad \left(\frac{1}{16}, \frac{1}{4}\right), \quad \text{and} \quad \left(-\frac{1}{16}, -\frac{1}{4}\right)$$

The Hessian matrix is of the form

$$\text{Hess}\, f(x, y) = \begin{bmatrix} 4 & -1 \\ -1 & 12y^2 \end{bmatrix}$$

We evaluate the Hessian matrix at each candidate and compute its eigenvalues:

(i) Hess $f(0, 0) = \begin{bmatrix} 4 & -1 \\ -1 & 0 \end{bmatrix}$

The eigenvalues satisfy

$$\det \begin{bmatrix} 4 - \lambda & -1 \\ -1 & -\lambda \end{bmatrix} = \lambda(\lambda - 4) - 1 = \lambda^2 - 4\lambda - 1 = 0$$

Thus,

$$\lambda_{1,2} = \frac{4 \pm \sqrt{16 + 4}}{2} = 2 \pm \sqrt{5} \approx \begin{cases} 4.2361 \\ -0.2361 \end{cases}$$

implying that f has a saddle point at $(0, 0)$.

(ii) Hess $f\left(\frac{1}{16}, \frac{1}{4}\right) = \begin{bmatrix} 4 & -1 \\ -1 & \frac{3}{4} \end{bmatrix}$

The eigenvalues satisfy

$$\det\begin{bmatrix} 4-\lambda & -1 \\ -1 & \frac{3}{4}-\lambda \end{bmatrix} = (4-\lambda)(\frac{3}{4}-\lambda) - 1 = 3 - 4\lambda - \frac{3}{4}\lambda + \lambda^2 - 1$$
$$= \lambda^2 - \frac{19}{4}\lambda + 2 = 0$$

Thus,

$$\lambda_{1,2} = \frac{\frac{19}{4} \pm \sqrt{\frac{361}{16} - 8}}{2} = \frac{19}{8} \pm \frac{1}{8}\sqrt{233} \approx \begin{cases} 4.2830 \\ 0.4670 \end{cases}$$

implying that f has a local minimum at $(\frac{1}{16}, \frac{1}{4})$.

(iii) Hess $f\left(-\frac{1}{16}, -\frac{1}{4}\right) = \begin{bmatrix} 4 & -1 \\ -1 & \frac{3}{4} \end{bmatrix}$

This is the same matrix as that for $(\frac{1}{16}, \frac{1}{4})$ [i.e., case (ii)]. We thus conclude that f has a local minimum at $(-\frac{1}{16}, -\frac{1}{4})$ as well.

The graph of $f(x, y)$ is illustrated in Figure 10.66. •

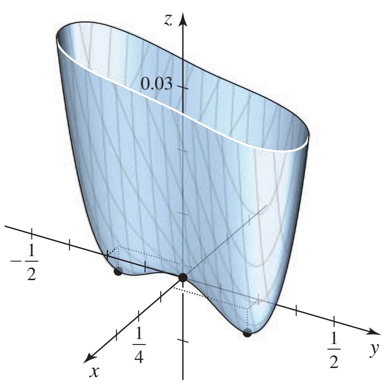


Figure 10.66 The graph of the function $f(x, y) = 2x^2 - xy + y^4$.

As you can see from Example 4 finding the eigenvalues of the Hessian matrix can be time consuming. There is another criterion which follows from the relationship that expresses the determinant and the trace of a 2×2 matrix in terms of the eigenvalues of the matrix. Recall that if A is a 2×2 matrix with eigenvalues λ_1 and λ_2, then $\det A = \lambda_1\lambda_2$ and $\operatorname{tr} A = \lambda_1 + \lambda_2$. If the eigenvalues of A are both real (as is the case for a symmetric matrix), and if $\det A > 0$, then either both λ_1 and λ_2 are positive or both are negative. If, in addition, $\operatorname{tr} A > 0$, then both λ_1 and λ_2 are positive.

Trace and Determinant Test for Identifying Local Minima and Local Maxima
Suppose the second partial derivatives of f are continuous in a disk centered at (x_0, y_0). Suppose also that $\nabla f(x_0, y_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Then

1. If $\det \text{Hess } f(x_0, y_0) > 0$ and $\operatorname{tr} \text{Hess } f(x_0, y_0) > 0$, then f has a local minimum at (x_0, y_0).
2. If $\det \text{Hess } f(x_0, y_0) > 0$ and $\operatorname{tr} \text{Hess } f(x_0, y_0) < 0$, then f has a local maximum at (x_0, y_0).
3. If $\det \text{Hess } f(x_0, y_0) < 0$, then (x_0, y_0) is not a local extremum; instead, (x_0, y_0) is a saddle point.

Recall that if one of the eigenvalues of Hess $f(x_0, y_0)$ is equal to 0 [or, equivalently, if $\det \text{Hess } f(x_0, y_0) = 0$], then we cannot say anything about the nature of the critical point on the basis of the Hessian matrix. (Such a case is explored in Problem 11.)

EXAMPLE 5 Find and classify the critical points of

$$f(x, y) = x^3 - 4xy + y, \quad (x, y) \in \mathbf{R}^2$$

Solution We find that

$$\nabla f(x, y) = \begin{bmatrix} 3x^2 - 4y \\ -4x + 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

when

$$3x^2 - 4y = 0 \quad (R_1) \qquad \text{and} \qquad -4x + 1 = 0 \quad (R_2)$$

(R_2) yields $x = 1/4$. Substituting this value into the (R_1) yields

$$\frac{3}{16} - 4y = 0$$

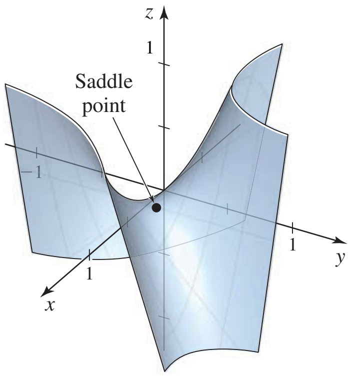


Figure 10.67 The graph of the function $f(x, y) = x^3 - 4xy + y$.

or $y = 3/64$. Since f is differentiable for all $(x, y) \in \mathbf{R}^2$, there is only one critical point: $(\frac{1}{4}, \frac{3}{64})$. To classify the critical point, we compute

$$\text{Hess } f(x, y) = \begin{bmatrix} 6x & -4 \\ -4 & 0 \end{bmatrix}$$

Evaluating this matrix at the critical point, we find that

$$\text{Hess } f\left(\frac{1}{4}, \frac{3}{64}\right) = \begin{bmatrix} \frac{3}{2} & -4 \\ -4 & 0 \end{bmatrix}$$

Since $\det \text{Hess } f(\frac{1}{4}, \frac{3}{64}) = -16 < 0$, we conclude that $f(x, y)$ has a saddle point at $(\frac{1}{4}, \frac{3}{64})$. (See Figure 10.67 for a graph of the function.) •

In each of the above examples the function $f(x, y)$ is differentiable over the entire of $\mathbf{R}^2$. So to find the critical points we only needed to examine points at which $\nabla f = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. However, we must also remember that points at which ∇f is undefined are also critical points, and therefore may also be candidate local minima or maxima, as the next example shows.

EXAMPLE 6 Find and classify the critical points of $f(x, y) = \sqrt{x^2 + y^2}$, $(x, y) \in \mathbf{R}^2$.

Solution We find that

$$\nabla f(x, y) = \begin{bmatrix} \frac{2x}{2\sqrt{x^2+y^2}} \\ \\ \frac{2y}{2\sqrt{x^2+y^2}} \end{bmatrix} = \frac{1}{\sqrt{x^2 + y^2}} \begin{bmatrix} x \\ y \end{bmatrix}, \qquad (x, y) \neq (0, 0)$$

Since the gradient of f is undefined at $(0, 0)$, the point $(0, 0)$ is a critical point. There are no other critical points, because $\nabla f(x, y) \neq \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$ for all $(x, y) \neq (0, 0)$. Now, $f(x, y) > 0$ for $(x, y) \neq (0, 0)$ and $f(x, y) = 0$ for $(x, y) = (0, 0)$. Therefore, $f(x, y)$ has a local minimum at $(0, 0)$. (See Figure 10.68.) Note that we cannot use the Hessian to decide whether $(0, 0)$ is a local extremum and, if so, of what type it is, since the theorem requires that the gradient be zero at the critical point, but here the gradient is undefined at $(0, 0)$. [The Hessian is also not defined at $(0, 0)$.] •

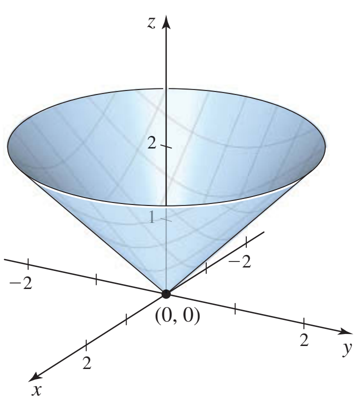


Figure 10.68 The graph of the function $f(x, y) = \sqrt{x^2 + y^2}$.

10.7.2 Global Extrema

We now turn our discussion to global extrema. Recall that, for functions of one variable, the extreme-value theorem guarantees the existence of global extrema for functions defined on a *closed* interval. The analogue of closed intervals in the two-dimensional plane is a **closed** set; similarly, the analogue of an open interval is an **open** set.

To define these concepts, we start with a set $D \subset \mathbf{R}^2$. A point (x, y) is called an **interior point** of D if there exists a $\delta > 0$ such that the disk centered at (x, y) with radius δ is contained in D—that is, if $B_\delta(x, y) \subset D$. (See Figure 10.69a.) A point (x, y) is a **boundary point** of D if every disk centered at (x, y) contains both points in D and points not in D; the boundary point (x, y) need not be contained in D. (See Figure 10.69b.) The **interior** of D consists of all interior points of D; the **boundary** of D consists of all boundary points of D. A set $D \subset \mathbf{R}^2$ is **open** if it consists only of interior points; a set $D \subset \mathbf{R}^2$ is **closed** if it contains all its boundary points as well as its interior points. (See Figure 10.70.)

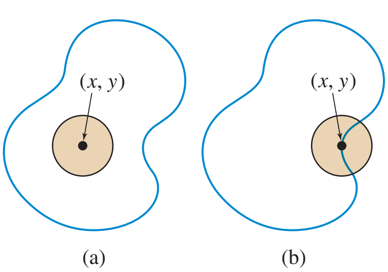


Figure 10.69 The point (x, y) on the left is an interior point; the point (x, y) on the right is a boundary point.

Most of the time, the domains of our functions will be rectangles or disks. Figure 10.71 illustrates the concepts we just learned on the unit disk. We start with the open unit disk $\{(x, y) : x^2 + y^2 < 1\}$ (Figure 10.71a). Every point in this set is an interior point. The unit circle $\{(x, y) : x^2 + y^2 = 1\}$ is the boundary of the *open unit disk*

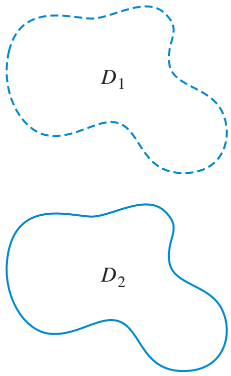


Figure 10.70 The set D_1 on the top is open. The set D_2 on the bottom is closed; the solid line is the boundary.

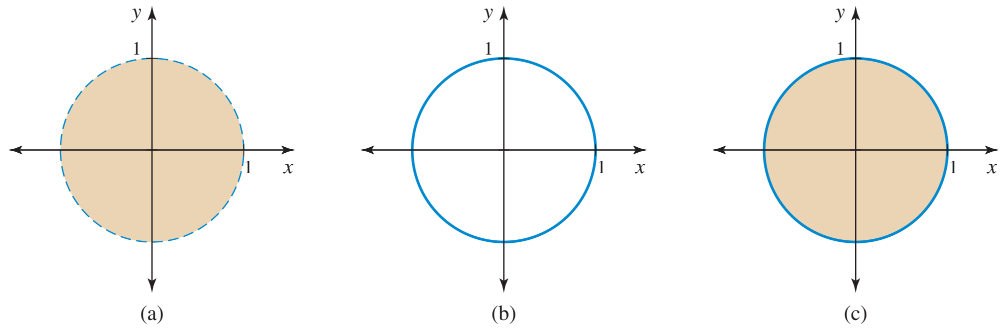


Figure 10.71 The set on the left is the open unit disk. The solid line in the middle figure is the boundary of the unit disk. The set on the right combines the open disk and its boundary; it is the closed unit disk.

(Figure 10.71b). If we combine the open disk and its boundary, we obtain the *closed unit disk* $\{(x, y) : x^2 + y^2 \le 1\}$ (Figure 10.71c).

To formulate the Extreme-Value Theorem in $\mathbf{R}^2$, we also need the notion of a **bounded** set. A set is bounded if it can be completely contained within a disk.

> **Theorem Extreme-Value Theorem in $\mathbf{R}^2$** If f is continuous on a closed and bounded set $D \subset \mathbf{R}^2$, then f has both a global maximum and a global minimum on D.

The conditions for checking that a set $D \subset \mathbf{R}^2$ is both closed and bounded are somewhat difficult to check, and the focus of this section is really on finding extrema, so we offer the following practical guidance

> **Closed and Bounded Rectangles and Disks.**
>
> The rectangle $\{(x, y) : \quad a \le x \le b, c \le y \le d\}$ $\quad a, b, c, d, \in \mathbf{R}$
>
> and the disk $\{(x, y) : \quad (x - a)^2 + (y - b)^2 \le R^2\}$ $\quad a, b, R \in \mathbf{R}$
>
> are both closed and bounded sets

Global extrema can occur in the interior of D or on the boundary of D. We already discussed how to find candidates for local extrema in the interior. To find local extrema on the boundary, it is useful to think of $f(x, y)$ restricted to the boundary of D. This restriction often allows us to write the function that is so restricted as a function of just one variable; we can then use the tools of single-variable calculus to find all candidates for local extrema on the boundary of D. (See Example 7.) To find global extrema for continuous functions defined on a closed and bounded set, we thus proceed as follows:

> 1. Determine all candidates for local extrema in the interior of D.
> 2. Determine all candidates for local extrema on the boundary of D.
> 3. Select the global maximum and the global minimum from the set of points determined in steps 1 and 2.

EXAMPLE 7 Find the global extrema of

$$f(x, y) = x^2 - 3y + y^2, \quad -1 \le x \le 1, 0 \le y \le 2.$$

Solution The function is defined on a closed and bounded rectangle and is continuous. The extreme-value theorem thus guarantees the existence of global extrema. We begin with finding critical points in the interior of the domain,

$$\nabla f(x, y) = \begin{bmatrix} 2x \\ -3 + 2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

when $x = 0$ and $y = 3/2$. The point $(0, 3/2)$ is in the interior of the domain of f and is thus a critical point with $f(0, 3/2) = -2.25$. There are no other critical points in the interior of the domain of f.

Next, we need to check the boundary values. (See Figure 10.72.) We start with the line segment C_1, which connects the points $(-1, 0)$ and $(1, 0)$ on the x-axis. On C_1, $y = 0$. Hence, on C_1, f is of the form

$$f(x, 0) = x^2, \quad -1 \le x \le 1$$

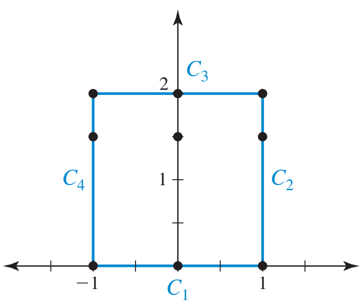


Figure 10.72 The domain of $f(x, y)$ in Example 7. The boundary consists of the line segments C_1, C_2, C_3, and C_4. The points are candidate local extrema.

By restricting $f(x, y)$ to the curve $y = 0$, we obtained a function of just one variable. Using single-variable calculus, we find that $f'(x, 0) = 2x = 0$ for $x = 0$. The critical point on C_1 is thus $(0, 0)$, with $f(0, 0) = 0$; in addition, there are the two endpoints $(-1, 0)$, with $f(-1, 0) = 1$, and $(1, 0)$, with $f(1, 0) = 1$.

On C_2, we have $x = 1$, which yields $f(1, y) = 1 - 3y + y^2$, $0 \le y \le 2$, which is again a function of just one variable. Now, $f'(1, y) = -3 + 2y = 0$ for $y = 3/2$. Hence, we find a candidate at $(1, 3/2)$, with $f(1, 3/2) = -1.25$; other candidates are the endpoints $(1, 0)$, with $f(1, 0) = 1$, and $(1, 2)$, with $f(1, 2) = -1$.

On C_3, we have $y = 2$, yielding $f(x, 2) = x^2 - 2$. Thus, $f'(x, 2) = 2x = 0$ for $x = 0$, giving the critical point $(0, 2)$, with $f(0, 2) = -2$. Other candidates are the endpoints $(-1, 2)$, with $f(-1, 2) = -1$, and $(1, 2)$, with $f(1, 2) = -1$.

On C_4, $x = -1$ and $f(-1, y) = 1 - 3y + y^2$, which is the same as on C_2. We thus have the additional candidate extrema $(-1, 3/2)$, with $f(-1, 3/2) = -1.25$; $(-1, 0)$, with $f(-1, 0) = 1$; and $(-1, 2)$, with $f(-1, 2) = -1$.

Comparing all the values of $f(x, y)$ at the candidate points (see Figure 10.72 and the table that follows), we find that the global minimum is $f(0, 3/2) = -2.25$ and the global maxima are $f(-1, 0) = 1$ and $f(1, 0) = 1$.

(x, y)	$(0, 3/2)$	$(0, 0)$	$(-1, 0)$	$(1, 0)$	$(1, 3/2)$	$(1, 2)$	$(0, 2)$	$(-1, 2)$	$(-1, 3/2)$
$f(x, y)$	-2.25	0	1	1	-1.25	-1	-2	-1	-1.25

EXAMPLE 8 Find the global maxima and minima of $f(x, y) = x^2 + y^2 - 2x + 4$ on the disk $D = \{(x, y) : x^2 + y^2 \le 4\}$.

Solution The function is defined on a closed and bounded disk and is continuous. The Extreme-Value Theorem thus guarantees global extrema. We begin with finding critical points in the interior of the domain,

$$\nabla f(x, y) = \begin{bmatrix} 2x - 2 \\ 2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

when $x = 1$ and $y = 0$. Since $x^2 + y^2 = 1 \le 4$, the point $(1, 0)$ is in the interior of the domain of f and is thus a critical point, with $f(1, 0) = 3$. There are no other critical points in the interior of the domain of f.

Next, we seek extrema on the boundary of the domain: the circle $x^2 + y^2 = 4$. The circle is centered at the origin $(0, 0)$ and has radius 2. We need a mathematical description of the circle in terms of a function of just one variable so that we can use single-variable calculus to identify extrema on the boundary. Toward that end, every

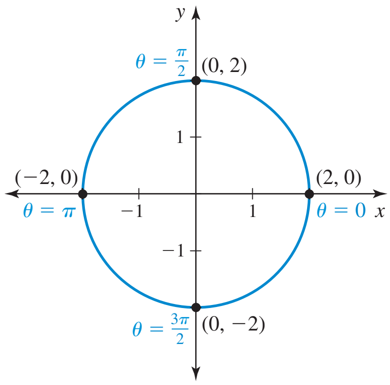


Figure 10.73 $x = 2\cos\theta$, $y = 2\sin\theta$, $0 \le \theta \le 2\pi$ parameterizes the circle $x^2 + y^2 = 4$

point (x, y) on the circle can be written as

$$x = 2\cos\theta$$
$$y = 2\sin\theta$$

for $0 \le \theta < 2\pi$. This is called a **parameterization** of the circle (see Figure 10.73).

On this circle,

$$\begin{aligned} f(x, y) &= x^2 + y^2 - 2x + 4 \\ &= 4\cos^2\theta + 4\sin^2\theta - 4\cos\theta + 4 \\ &= 4 - 4\cos\theta + 4 = 8 - 4\cos\theta \qquad \sin^2\theta + \cos^2\theta = 1 \end{aligned}$$

To find maxima and minima of the single-valued function $g(\theta) = 8 - 4\cos\theta$, we need to differentiate $g(\theta)$:

$$g'(\theta) = 4\sin\theta$$

Then we solve $g'(\theta) = 0$ in $[0, 2\pi)$. We find the two angles $\theta = 0$ and $\theta = \pi$. Now,

$$g(0) = 8 - 4 = 4 \qquad \text{and} \qquad g(\pi) = 8 + 4 = 12$$

The maximum on the boundary is at $\theta = \pi$, which corresponds to the point $(-2, 0)$. The minimum on the boundary is at $\theta = 0$, which corresponds to the point $(2, 0)$.

We compile all of our candidate global extrema into a table

(x, y)	$(1, 0)$	$(-2, 0)$	$(2, 0)$
$f(x, y)$	3	12	4

From the table we see that the global minimum is in the interior at $(1, 0)$ and the global maximum is on the boundary at $(-2, 0)$ (Figure 10.74 shows how $f(x, y)$ can be visualized using level curves). •

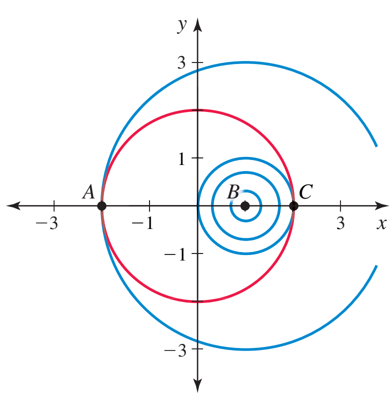


Figure 10.74 Shown are the level curves of the function $f(x, y) = x^2 + y^2 - 2x + 4$ [in blue for the values $f(x, y) = c$ with $c = 3.1, 3.5, 4$, and 12] and the boundary of the disk $x^2 + y^2 \le 4$ (in red). The point $B = (1, 0)$ is the local minimum in the interior of the disk $x^2 + y^2 \le 4$, the point $A = (-2, 0)$ is the local maximum on the boundary of the disk, and the point $C = (2, 0)$ is the other candidate extremum on the boundary of the disk.

We conclude this subsection with an application.

EXAMPLE 9 Determine the values of three nonnegative numbers whose sum is 90 and whose product is maximal.

Solution We denote the three numbers by x, y, and z, respectively. Then $x + y + z = 90$. Now, their product is xyz, and since $z = 90 - x - y$, we can write the product as $xyz = xy(90 - x - y)$. Our goal is to maximize this product. We define the function

$$f(x, y) = xy(90 - x - y),$$

Since x, y, and z are nonnegative numbers and their sum is equal to 90, the domain is the set $\{(x, y), x \ge 0, y \ge 0, x + y \le 90\}$. The first two conditions ($x \ge 0$ and $y \ge 0$) come from requiring that x and y both be non-negative. z also must be non-negative,

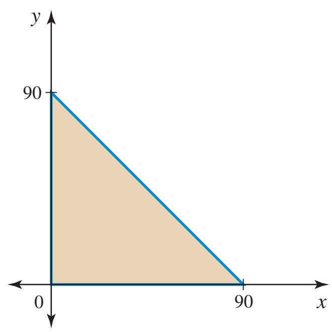


Figure 10.75 The domain of the function $f(x, y) = xy(90 - x - y)$ in Example 9.

meaning that $90 - x - y \geq 0$, which gives rise to our third condition $x + y \leq 90$. The domain on which f is defined is the triangular region bounded by the lines $x = 0$, $y = 0$, and $y = 90 - x$. We show this triangular domain in Figure 10.75.

We need to find (x, y) so that $f(x, y)$ is maximal. Now,

$$\nabla f(x, y) = \begin{bmatrix} 90y - 2xy - y^2 \\ 90x - x^2 - 2xy \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

when

$$y(90 - 2x - y) = 0 \quad (R_1) \qquad \text{and} \qquad x(90 - 2y - x) = 0 \quad (R_2)$$

Solutions with $x = 0$ or $y = 0$ are points on the boundary. To find solutions in the interior of the domain, we need to solve

$$\begin{aligned} 2x + \ \ y &= 90 \qquad (R_3) \\ x + 2y &= 90 \qquad (R_4) \end{aligned}$$

(R_3) follows from (R_1) assuming $y \neq 0$, and (R_4) follows from (R_2), assuming $x \neq 0$. (R_3) and (R_4) are linear equations, so we may solve the system using the methods of Chapter 9.

$$\begin{aligned} (R_3): \quad 2x + y &= 90 \qquad (R_5) \\ 2(R_4) - (R_3): \quad 3y &= 90 \qquad (R_6) \qquad \text{Eliminate } x \text{ from } (R_4) \end{aligned}$$

So $y = 30$ and $x = \frac{1}{2}(90 - y) = 30$, yielding the candidate $(30, 30)$, which is in the interior of the domain, with $f(30, 30) = 30^3 = 27{,}000$. There are no other candidates for local extrema in the interior of D.

The function $f(x, y)$ is continuous on a closed and bounded set, namely the triangle with corners $(0, 0)$, $(0, 90)$, and $(90, 0)$. The Extreme-Value Theorem guarantees that f has a global maximum on the domain. We see that $f(x, y)$ takes on the value 0 on the boundary of the domain since either $x = 0$, $y = 0$, or $z = 0$ on each of the line segments that makes up the boundary. Comparing the values of $f(x, y)$ on the boundary of D with the value at the candidate point $(30, 30)$, we conclude that the function $f(x, y)$ has the global maximum at the interior candidate $(30, 30)$.

The product xyz is therefore maximal when $x = y = z = 30$. •

10.7.3 Extrema with Constraints

A number of studies have shown that, in butterflies which lay their eggs singly, egg size decreases with maternal age. ***Begon and Parker (1986)*** proposed a mathematical model to explain this decline in egg size in terms of a maternal strategy that would optimize reproductive fitness. The main assumptions of their model are that all resources necessary for egg production are gathered before eggs are laid and that clutch size is fixed (e.g., a single egg per clutch). Begon and Parker assume that butterflies seek to maximize the fitness of their offspring, which is affected by egg size.

Mathematically, the problem can be phrased in terms of finding the maximum of a function that describes total offspring **fitness** (i.e., the number of offspring that survive) in terms of egg size per clutch under the constraint that the total amount of reproductive resources are fixed before reproduction starts. This type of problem falls into the category of finding extrema with constraints.

To have a concrete example at hand when we discuss how to find such extrema, let's take the case of a female that has at most two clutches, each of size n, during her lifetime. We denote egg size of the first clutch by x_1, and egg size of the second clutch by x_2; we also assume that all eggs in the same clutch have the same size. If the total amount of resources available for reproduction is R, then the constraint can be written as

$$nx_1 + nx_2 = R \tag{10.26}$$

[In order for the units to agree in (10.26), assume that both egg size and amount of resources are measured in the same units, e.g., calories.] We define a function $\rho(x)$ that describes egg fitness (that is the likelihood that the offspring in the egg survives) as a function of egg size x. If p_i is the probability that the female survives to lay her ith clutch ($i = 1, 2$), then the total offspring fitness is:

$$f(x_1, x_2) = p_1 n\rho(x_1) + p_2 n\rho(x_2)$$

Fitness is the number of offspring that survive

How is this function derived? We say that the fitness of one egg from clutch i is $\rho(x_i)$. So the number of offspring from that clutch that survive is equal to:

$$\begin{matrix}\text{Probability mother} \\ \text{survives to lay clutch } i\end{matrix} \times \begin{matrix}\text{Number of eggs} \\ \text{in clutch}\end{matrix} \times \begin{matrix}\text{Likelihood each} \\ \text{offspring survives}\end{matrix}$$

$$= p_i \times n \times \rho(x_i)$$

Summing the fitnesses of both clutches yields the function $f(x_1, x_2)$. To find the optimal maternal strategy we need to find extrema of the function $f(x_1, x_2)$ under the constraint $nx_1 + nx_2 = R$ (i.e., that also satisfy Equation (10.26)).

Finding extrema with constraints involves two functions: one describing the constraint, the other the function we wish to maximize i.e., f. All of the constraints in this section will be of the form

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

For instance, the constraint (10.26) can be written as $g(x_1, x_2) = nx_1 + nx_2 - R = 0$.

We can illustrate the constraint $g(x, y) = 0$ as a set of points in the x–y plane: $\{(x, y) : g(x, y) = 0\}$. These will typically be curves of the sort shown in Figure 10.76. That is, the set of all points that obey the constraint lie on a level curve of the function $g(x, y)$. Now to proceed we must recall from Section 10.6 that, if $g(x, y)$ is a differentiable function, then at each point along the level curve the gradient $\nabla g(x, y)$ will be perpendicular to the level curve.

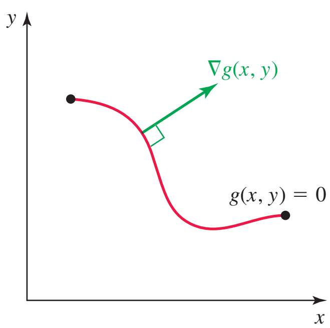


Figure 10.76 A constraint curve.

We want to find the extrema of the function $f(x, y)$ under the constraint $g(x, y) = 0$. Finding extrema with constraints amounts to restricting the function $f(x, y)$ to the constraint curve and seeking its extrema there. The constraint $g(x, y) = 0$ defines a set of points (x, y) in the x–y plane. Using level curves for $f(x, y)$, we can represent $f(x, y)$ in the x–y plane. We can then graph both the level curves of $f(x, y)$ and the constraint $g(x, y) = 0$ in the same two-dimensional coordinate system. (See Figure 10.77.) To make the discussion that follows more concrete, assume that $c_1 < c_2 < c_3 < c_4$. (Other cases will be discussed in Problems 49 and 50.) A particular value $f(x, y) = c$ can be attained only if the level curve $f(x, y) = c$ intersects with the level curve $g(x, y) = 0$. Imagine walking along the curve $g(x, y) = 0$. According to Figure 10.77 if we walk along the level curve from left to right (i.e. in the direction shown by the arrow in the figure), the value of $f(x, y)$ that we would observe initially increases through c_1, c_2 until it reaches c_3, then it starts to decrease, going back through c_2, c_1 and so on (see Figure 10.78). So, along the curve $g(x, y) = 0$, the function f has a local extremum (in this case, a local maximum) at P.

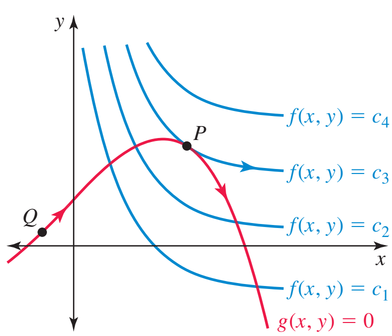


Figure 10.77 Level curves and constraints.

What characterizes the point P? The level curve through P and the constraint curve touch each other at P; that is, they both have the same tangent line. Equivalently, the normal to the tangent of f must be parallel to the normal to the tangent of g. But the gradient of f at the point P is perpendicular to the level curve through P, and the gradient of g at P is perpendicular to graph of $g(x, y) = 0$, so if the normal directions are parallel, then ∇g and ∇f are parallel.

To state the theorem more generally, we need to require that $\nabla g(x, y) \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ at P. Denoting the coordinates of P by (x_0, y_0), we can then formulate the result as Lagrange's theorem:

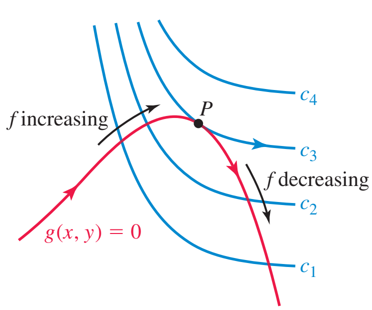


Figure 10.78 Value of $f(x, y)$ measured by an observer who walks along the curve $g(x, y) = 0$.

Lagrange's Theorem Assume that f and g have continuous first partial derivatives and that $f(x, y)$ has an extremum at (x_0, y_0) subject to the constraint $g(x, y) = 0$. If $\nabla g(x_0, y_0) \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, then there exists a number λ such that

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0) \tag{10.27}$$

The number λ is called a **Lagrange multiplier**. Using Lagrange multipliers to find candidates for extrema subject to a constraint is called the **method of Lagrange multipliers**. The condition (10.27) is a necessary condition. In the next example, we illustrate how to use Lagrange multipliers to find extrema subject to constraints.

EXAMPLE 10 Find all extrema of $f(x, y) = e^{-xy}$, subject to the constraint $x^2 + 4y^2 = 1$.

Solution We define $g(x, y) = x^2 + 4y^2 - 1$. Then the constraint is of the form $g(x, y) = 0$. Using the method of Lagrange multipliers, we are looking for (x, y) and λ such that

$$\nabla f(x, y) = \lambda \nabla g(x, y) \qquad \text{and} \qquad g(x, y) = 0$$

This translates into the set of equations

$$-ye^{-xy} = 2\lambda x \quad (R_1), \qquad -xe^{-xy} = 8\lambda y \quad (R_2), \qquad \text{and} \quad x^2 + 4y^2 = 1 \quad (R_3)$$

since

$$\nabla f(x, y) = \begin{bmatrix} -ye^{-xy} \\ -xe^{-xy} \end{bmatrix} \qquad \text{and} \qquad \nabla g(x, y) = \begin{bmatrix} 2x \\ 8y \end{bmatrix}.$$

We can eliminate λ from the first two equations. (Multiply (R_1) by $4y$ and (R_2) by x, and take the difference of the two equations.) We then find that

$$-4y^2e^{-xy} + x^2e^{-xy} = 0$$

Simplifying yields $e^{-xy}(x^2 - 4y^2) = 0$. Since $e^{-xy} \neq 0$, we obtain $x^2 - 4y^2 = 0$. Combining this with the constraint equation, we get the system

$$x^2 - 4y^2 = 0$$
$$x^2 + 4y^2 = 1$$

We leave the first equation and eliminate y from the second equation by adding the two equations. We then obtain

$$x^2 - 4y^2 = 0$$
$$2x^2 = 1$$

Thus, $x^2 = 1/2$ and $4y^2 = x^2 = 1/2$, implying that $y^2 = 1/8$. Simultaneously solving $x^2 = 1/2$ and $y^2 = 1/8$ gives the candidates

$$\left(\sqrt{\frac{1}{2}}, \frac{1}{2}\sqrt{\frac{1}{2}}\right), \quad \left(\sqrt{\frac{1}{2}}, -\frac{1}{2}\sqrt{\frac{1}{2}}\right), \quad \left(-\sqrt{\frac{1}{2}}, \frac{1}{2}\sqrt{\frac{1}{2}}\right), \quad \left(-\sqrt{\frac{1}{2}}, -\frac{1}{2}\sqrt{\frac{1}{2}}\right)$$

with

$$f\left(\sqrt{\frac{1}{2}}, \frac{1}{2}\sqrt{\frac{1}{2}}\right) = e^{-1/4}, \quad f\left(\sqrt{\frac{1}{2}}, -\frac{1}{2}\sqrt{\frac{1}{2}}\right) = e^{1/4},$$

$$f\left(-\sqrt{\frac{1}{2}}, \frac{1}{2}\sqrt{\frac{1}{2}}\right) = e^{1/4}, \quad f\left(-\sqrt{\frac{1}{2}}, -\frac{1}{2}\sqrt{\frac{1}{2}}\right) = e^{-1/4}$$

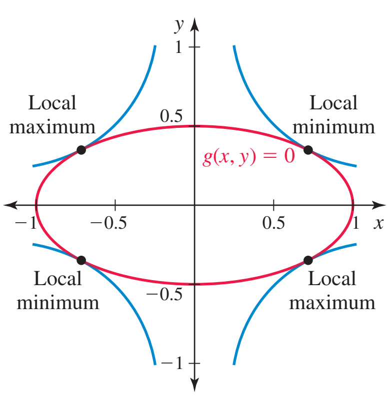


Figure 10.79 The level curves of $f(x, y)$ (blue) that touches the constraint curve $g(x, y) = 0$ in Example 10.

The extreme-value theorem applies to the constraint curve because that curve is closed and bounded; we can then conclude that maxima and minima exist on this curve, and we can select them from among our candidates. The maxima are

$$\left(-\sqrt{\frac{1}{2}}, \frac{1}{2}\sqrt{\frac{1}{2}}\right) \quad \text{and} \quad \left(\sqrt{\frac{1}{2}}, -\frac{1}{2}\sqrt{\frac{1}{2}}\right)$$

the minima are

$$\left(\sqrt{\frac{1}{2}}, \frac{1}{2}\sqrt{\frac{1}{2}}\right) \quad \text{and} \quad \left(-\sqrt{\frac{1}{2}}, -\frac{1}{2}\sqrt{\frac{1}{2}}\right)$$

(See Figure 10.79.)

Looking back at Example 10, we notice that we did not have to compute the actual value of λ to find extrema. This will typically be the case.

In the statement of the result, we mentioned that the condition $\nabla f = \lambda \nabla g$ is a necessary condition. This means that finding (x_0, y_0) so that $\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$ only identifies candidates for local extrema. The next example illustrates this point.

EXAMPLE 11 Use Lagrange multipliers to identify candidates for local extrema of

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

subject to the constraint $y - x^3 = 0$, and show that there is one such candidate that turns out not to be a local extremum. Furthermore, show that the function $f(x, y)$ subject to the constraint $y - x^3 = 0$ has no global extrema.

Solution We define $g(x, y) = y - x^3$. Then

$$\nabla f(x, y) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \nabla g(x, y) = \begin{bmatrix} -3x^2 \\ 1 \end{bmatrix}.$$

With $y - x^3 = 0$, we obtain

$$0 = -3\lambda x^2 \quad \text{and} \quad 1 = \lambda \quad \text{and} \quad y = x^3.$$

Eliminating λ, we find $x = 0$ and thus $y = 0$. We claim that $(0, 0)$ is *not* a local extremum. The easiest way to see this is to find out what $f(x, y)$ looks like along the constraint curve $y = x^3$. If we use $y = x^3$ to substitute y in the function $f(x, y)$, we obtain a single-variable function $h(x) = x^3$, $x \in \mathbf{R}$. We know from single-variable calculus that $h(x) = x^3$ has no local extrema on $\mathbf{R}$ even though, since $h'(x) = 0$ for $x = 0$, there is a candidate for a local extremum at $x = 0$. Because $\lim_{x \to \infty} h(x) = \infty$ and $\lim_{x \to -\infty} h(x) = -\infty$, the function $f(x, y)$ subject to the constraint $y - x^3 = 0$ has no global extrema. •

The method of Lagrange multipliers only identifies candidates for local extrema, and as we saw in the previous example, these candidates may not turn out to be local extrema. Just finding candidates, however, is often good enough if we are interested in global extrema. This scenario is illustrated in the next example.

EXAMPLE 12 Suppose you wish to enclose a rectangular plot. You have 1600 ft of fencing. Using that material, what are the dimensions of the plot that will have the largest area? (See Figure 10.80.)

Solution We wish to maximize

$$A = xy$$

subject to the constraint $2x + 2y = 1600$. We define

$$f(x, y) = xy \qquad \text{and} \qquad g(x, y) = 2x + 2y - 1600 = 0.$$

Then

$$\nabla f(x, y) = \begin{bmatrix} y \\ x \end{bmatrix} \qquad \text{and} \qquad \nabla g(x, y) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

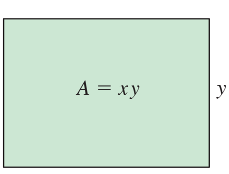


Figure 10.80 The rectangular plot in Example 12.

Using Lagrange multipliers, we need to find (x, y) and λ such that

$$\nabla f(x, y) = \lambda \nabla g(x, y) \qquad \text{and} \qquad 2x + 2y - 1600 = 0.$$

This yields the system of equations

$$y = 2\lambda \qquad x = 2\lambda \qquad x + y = 800.$$

Eliminating λ from the first two equations, we conclude that $x = y$ and thus $2x = 800$ or $x = 400$. For physical reasons, x and y can take only nonnegative values. The constraint thus restricts x and y to the line segment $y = 800 - x$, $0 \le x \le 800$. To see that $f(400, 400)$ gives us a maximum, we compare $f(400, 400)$ with the values at the endpoints of the line segment describing the constraint: $f(0, 800)$ and $f(800, 0)$. Since $f(400, 400) = 160,000$ and $f(0, 800) = f(800, 0) = 0$, $f(400, 400)$ is indeed the global maximum. •

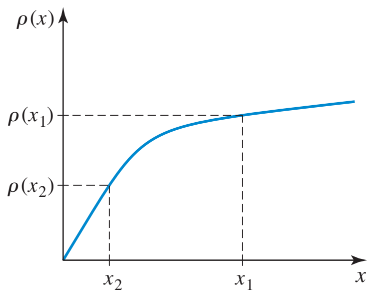


Figure 10.81 The function $\rho(x)$.

You probably remember the type of problem presented in Example 12 from Section 5.4. The method of Lagrange multipliers provides another method for solving the problems we discussed in that section. The method of Lagrange multipliers is more general than the method we learned there; it can be used even if we cannot solve the constraint for either x or y to eliminate one of the two variables, as we did in Section 5.4.

As the last example in this subsection, we return to the motivating example at the beginning of the subsection. There, we wished to maximize $f(x_1, x_2) = p_1 n\rho(x_1) + p_2 n\rho(x_2)$ subject to the constraint $nx_1 + nx_2 = R$. We now make the additional assumption that $\rho(x)$ increases at a decelerating rate (i.e., $p'(x) > 0$ and $p''(x) < 0$) and satisfies $\rho(0) = 0$ (see Figure 10.81), implying that there is a diminishing return to increasing egg size.

EXAMPLE 13 **Optimal Egg Size** Assume that x_1 and x_2 are nonnegative. Maximize

$$f(x_1, x_2) = p_1 n\rho(x_1) + p_2 n\rho(x_2)$$

subject to the constraint $nx_1 + nx_2 = R$, and show that egg size should decline with maternal age.

Solution We find that

$$\nabla f(x_1, x_2) = \begin{bmatrix} p_1 n\rho'(x_1) \\ p_2 n\rho'(x_2) \end{bmatrix} \qquad \text{and} \qquad \nabla g(x_1, x_2) = \begin{bmatrix} n \\ n \end{bmatrix}$$

Thus, we need to find (x_1, x_2) and λ such that

$$p_1 n\rho'(x_1) = n\lambda \quad (R_1) \qquad p_2 n\rho'(x_2) = n\lambda \quad (R_2) \qquad nx_1 + nx_2 = R \quad (R_3)$$

Eliminating λ from (R_1) and (R_2), we obtain

$$p_1\rho'(x_1) = p_2\rho'(x_2) \tag{10.28}$$

Now, the constraint curve $nx_1 + nx_2 = R$ is a straight line. For biological reasons, we require that both x_1 and x_2 be nonnegative. Therefore, the constraint curve is the line segment with endpoints $(R/n, 0)$ and $(0, R/n)$ (see Figure 10.82).

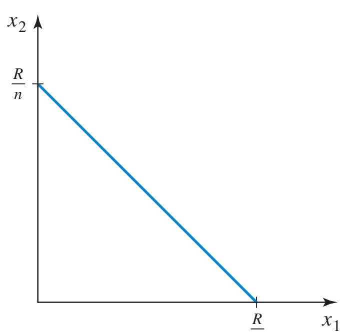


Figure 10.82 Constraint curve for Example 13.

On this line segment, we can show that there is at most one point (x_1, x_2) that satisfies (10.28). (We must also consider the possibility that no point on the line satisfies (10.28).) To do so, we use the equation $x_2 = R/n - x_1$ to eliminate x_2 and substitute the result into (10.28), yielding:

$$\frac{p_1}{p_2} = \frac{\rho'(R/n - x_1)}{\rho'(x_1)} \tag{10.29}$$

Since $\rho'(x) > 0$ and $\rho''(x) < 0$, we have

$$\frac{d}{dx_1}\frac{\rho'(R/n - x_1)}{\rho'(x_1)} = \frac{-\rho''(R/n - x_1)\rho'(x_1) - \rho'(R/n - x_1)\rho''(x_1)}{[\rho'(x_1)]^2} > 0.$$

Since the right-hand side of (10.29) is an increasing function it follows that there is at most one value of x_1 such that (10.29) [and hence (10.28)] holds.

We thus have the following situation: If there is a point (x_1, x_2) that satisfies (10.28) and lies on that line segment, then there are three candidates for global extrema, namely, (x_1, x_2), $(R/n, 0)$, and $(0, R/n)$; otherwise, there are only the two endpoints $(R/n, 0)$ and $(0, R/n)$. We need to select the global maximum from this set of candidates.

Now, $p_1 > p_2$ because the probability of being alive at a later age is smaller than at an earlier age. So $f(R/n, 0) > f(0, R/n)$. Thus, the global maximum cannot occur at the endpoint $(0, R/n)$. This leaves us with only two candidates for the global extremum: $(R/n, 0)$ and (if it exists) the unique point (x_1, x_2) that satisfies Equation (10.28). We claim that if the point (x_1, x_2) exists it is the global maximum; otherwise, the endpoint $(R/n, 0)$ is the global maximum.

How can we see this? Since $x_2 = R/n - x_1$, we can write the fitness function f as a function of x_1 alone and determine where it is increasing and where it is decreasing. On the constraint curve the fitness function is:

$$h(x_1) = f(x_1, R/n - x_1) = np_1\rho(x_1) + np_2\rho(R/n - x_1)$$

Differentiating with respect to x_1 yields

$$h'(x_1) = np_1\rho'(x_1) - np_2\rho'(R/n - x_1)$$

$$h''(x) = np_1\rho''(x_1) + np_2\rho''(R/n - x_1) < 0 \qquad \rho''(x) < 0 \text{ for all } x$$

so h' is a decreasing function (equivalently, h is concave downward) and $h'(0) > 0$ because $p_1 > p_2$ and $\rho'(0) > \rho'(R/n)$. [Recall that $\rho'(x)$ is increasing at a decelerating rate.] Moreover $h'(R/n) = np_1\rho'(R/n) - np_2\rho'(0)$. This quantity will be negative if and only if:

$$\frac{\rho'(R/n)}{\rho'(0)} < \frac{p_2}{p_1} \tag{10.30}$$

So if (10.30) holds, then $h'(x_1)$ goes from positive to negative for some value of x_1 between 0 and $h'(x_1) = 0$ only at a point satisfying (10.29). So there is a unique local maximum between 0 and R/n. Conversely, if (10.30) does not hold, then $h'(x_1)$ cannot change sign between $x_1 = 0$ and $x_1 = R/n$, because for all $0 \leq x_1 \leq R/n : h(x_1) \geq h(R/n) > 0$.

Thus the maximum of $h(x_1)$ occurs at the value of x_1 that solves (10.28) if (10.30) holds and at $x_1 = R/n$ if (10.30) does not hold.

To show that egg size should decline, we study (10.28) again using the assumption that $\rho(x)$ is a function with a diminishing return; that is, $\rho(x)$ is increasing at a decelerating rate. If a solution to this equation exists, then since $p_1 > p_2$ (the probability of being alive at a later age is smaller than at an earlier age), it follows that if (10.30) holds, then $\rho'(x_1) < \rho'(x_2)$ and therefore $x_1 > x_2$ (see Figure 10.81), which implies that egg size should decline. Otherwise the optimal strategy is at the endpoint $(R/n, 0)$, then the egg size in the first clutch is R/n and in the second clutch it is 0, implying that egg size should decline in this case as well. •

10.7.4 Least-Squares Data Fitting

This subsection requires you to be comfortable using Σ-notation, which was first introduced in Section 2.2.

A particularly useful application for minimization of functions is to fit a mathematical model to experimental measurements. Throughout this book we have introduced many mathematical models, as well as discussed methods for solving the equations that are derived for each model. A model is useful only when it can be compared with experimental measurements. Often our model contains coefficients that are not independently measured, but can be adjusted to ensure that the model agrees as closely as possible with real experimental measurements. A full study of fitting requires us to make use of the mathematical ideas of probability. We will therefore return to the topic of fitting in Section 12.7.

For example, suppose that you are a plant biologist interested in studying the effect of light intensity upon the growth rate of bean plants. You measure growth rate in plants that are grown in different light intensities. The data that you collect is presented in the following table:

light intensity (x)	0	1	2	4	6	10
growth rate (y)	1.72	2.09	1.94	2.91	3.31	5.09

More generally, suppose that we have an independent variable that we can control in our experiments (in this case our independent variable is the light intensity). Let's call this variable x. A second variable, y, depends on x (in this case y is the growth rate).

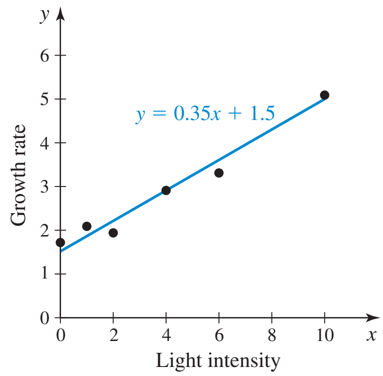


Figure 10.83 Growth rate, y, against light intensity x. The blue line is a hand-drawn line fitting the data.

To understand how y depends on x we run a sequence of experiments using different values of x; call this sequence $\{x_i\} i = 1, 2, \ldots, n$. We measure the resulting values of y; call these values $\{y_i\}$. Here $n = 6$ because we have six measurements.

When we plot the (x_i, y_i) data from our plant growth experiment, we see that y_i seems to depend linearly on x_i (see Figure 10.83).

That is, it appears that:

$$y_i = mx_i + c$$

for some coefficients m and c. We would like to estimate m and c from the data. But inevitably there is also some scatter in the data—and it is impossible to draw a single straight line that goes through all of the data points. Instead, the best that we can do is to find values for m and c that make the line $y = mx + c$ pass close to all (or as many as possible) of the data points. In previous examples we have encountered where such fitting was necessary; we did it by eye. For example, by moving a straight edge around in the data shown in Figure 10.83, we arrive at a line that seems quite close to all of the data:

$$y = 0.35x + 1.5$$

Fitting data by eye can take a lot of time, and can produce inconsistent results. For example, different people may find different coefficients when trying to fit the same data. Is there a way to find the coefficients m and c mathematically?

To start, we need to quantify how well our model fits the data. For each value of x_i our model can be used to predict a value for the dependent variable; that is, we predict:

$$Y_i = mx_i + c$$

Using our guess for the values of m and c ($m = 0.35$, $c = 1.5$) we add another row to the table showing our predictions for Y_i:

light intensity (x_i)	**0**	**1**	**2**	**4**	**6**	**10**
growth rate (y_i)	1.72	2.09	1.94	2.91	3.31	5.09
predicted (Y_i)	1.50	1.85	2.20	2.90	3.60	5.00
error ($e_i = Y_i - y_i$)	−0.22	−0.24	0.26	−0.01	0.29	−0.09

We then add another row to our table representing the error in each of our predictions. That is how far the predicted value of Y_i is away from the measured value y_i. This error is given by $e_i = Y_i - y_i$. For our model to fit the data well we would like each of the errors to be as small as possible in magnitude. So we need a way to measure the total error from all of our measurements. One possibility is to sum up all of the squares of the errors e_i^2, i.e., define a total sum-of-squares-error by $\sum_{i=1}^{6} e_i^2$. For the given values of m and c the total sum-of-squares error is

$$\sum_{i=1}^{6} e_i^2 = (-0.22)^2 + (-0.24)^2 + (0.26)^2 + (-0.01)^2 + (0.29)^2 + (-0.09)^2$$
$$= 0.27$$

$\sum_{i=1}^{6} e_i^2$ will be small only if each of the terms in the sum (i.e., each of the e_i^2) is small, meaning that all predicted y_i are close to the real measured values y_i. Why did we need to square the errors before adding them? If we sum the errors e_i, then $\sum_{i=1}^{6} e_i$ being small does not necessarily mean that each of the e_i is small in magnitude. If some of the errors are large and positive and others are large and negative, then their sum may still be small.

Different values of m and c will lead to different errors, e_i, and thus to different values for the sum-of-squares error $\sum_i e_i^2$. In other words, we can define a function $f(m, c) = \sum_i e_i^2$. Our goal is to find the values of m and c that minimize, f. To see how to minimize, let's first write out the sum to make the dependence upon m and c clear:

$$f(m, c) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - y_i)^2 = \sum_{i=1}^{n} (mx_i + c - y_i)^2 \quad \text{Consider general case of } n \text{ terms}$$

Now at local extrema:

$$0 = \frac{\partial f}{\partial m} = \sum_{i=1}^{n} \frac{\partial}{\partial m}(mx_i + c - y_i)^2$$
$$= \sum_{i=1}^{n} 2x_i(mx_i + c - y_i) \quad \text{Differentiate each term in the sum}$$

Or

$$0 = 2m\sum_{i=1}^{n} x_i^2 + 2c\sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} x_i y_i$$

which may be reorganized into the form:

$$\sum_{i=1}^{n} x_i y_i = m\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i \tag{10.31}$$

Similarly:

$$0 = \frac{\partial f}{\partial c} = \sum_{i=1}^{n} \frac{\partial}{\partial c}(mx_i + c - y_i)^2$$
$$= \sum_{i=1}^{n} 2(mx_i + c - y_i)$$

Or

$$0 = 2m\sum_{i=1}^{n} x_i + 2cn - 2\sum_{i=1}^{n} y_i \quad \textstyle\sum_{i=1}^{n} c = cn$$

which may be rearranged into the form:

$$\sum_{i=1}^{n} y_i = m\sum_{i=1}^{n} x_i + c\,n \tag{10.32}$$

We therefore have two linear equations to solve for m and c: (10.31) and (10.32). All of the coefficients on these equations can be calculated from the data. The only unknowns are m and c. We can solve these equations using the methods from Chapter 9. We will solve them by writing them in matrix form:

$$\begin{bmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} y_i \end{bmatrix}$$

So:

$$\begin{bmatrix} m \\ c \end{bmatrix} = \frac{1}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \begin{bmatrix} n & -\sum_{i=1}^{n} x_i \\ -\sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} y_i \end{bmatrix}$$

Or writing out the components

$$m = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{10.33}$$

and

$$c = \frac{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} x_i y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{10.34}$$

We call these values for m and c the **least squares estimates** for m and c because they come from minimizing the sum-of-squares error. Remember that local extrema are only candidates for global extrema. We can show that the point (m, c) is a local minimum by calculating the Hessian matrix, but doing so is a little beyond the tools developed in this text. Instead let's note that the value of f is bounded below by 0 (since f is the sum of positive numbers), so f must have some minimum value. Either this minimum value of f is attained at some local extremum point (which must be a local minimum, then) or it is only approached as either $m \to \infty$ or $c \to \infty$, or both. But if either m or c become infinitely large (whether positive or negative) we see that f will grow without bound, since each of the summands e_i will diverge. So the global minimum value must be attained at a local extremum. The only possible local extremum is given by (10.33) and (10.34), so these values for (m, c) really are the global minimum point.

EXAMPLE 14 **Plant Growth** Calculate the least squares estimates for m and c using the plant growth data provided in the introduction to this section.

Solution To find m and c we need to substitute for each of the sums appearing in Equations (10.33) and (10.34). For these data:

$$n = 6$$

$$\sum_{i=1}^{n} x_i = 0 + 1 + 2 + 4 + 6 + 10 = 23$$

$$\sum_{i=1}^{n} x_i^2 = 0^2 + 1^2 + 2^2 + 4^2 + 6^2 + 10^2 = 157$$

$$\sum_{i=1}^{n} y_i = 1.72 + 2.09 + 1.94 + 2.91 + 3.31 + 5.09 = 17.06$$

and

$$\sum_{i=1}^{n} x_i y_i = (0)(1.72) + (1)(2.09) + (2)(1.94) + (4)(2.91) + (6)(3.31) + (10)(5.09)$$
$$= 88.37.$$

So the least squares estimates are (to 3 decimal places):

$$m = \frac{(6)(88.37) - (23)(17.06)}{(6)(157) - (23)^2} = 0.334$$

and

$$c = \frac{(157)(17.06) - (23)(88.37)}{(6)(157) - (23)^2} = 1.564$$

Our best-fit by eye gave $m = 0.35$ and $c = 1.5$.

Equations (10.33) and (10.34) can be used to estimate the slope and intercept coefficients (i.e., m and c) for any data that can be manipulated into a linear form as the next example shows.

EXAMPLE 15 **Population Growth** A population of bacteria growing in a flask multiplies exponentially; that is, the size of the population at time t is believed to be given by a model:

$$N(t) = N_0 e^{rt}$$

where N_0 and r are coefficients. Use a least squares method to estimate the coefficients that fit this model to the following data:

time, t	0	1	2	3	4	8	10
number of cells, N	31	71	133	237	246	2385	6125

Solution The method that we have developed only works to estimate the coefficients in a linear model. As a plot of N against t shows, the population size here appears to grow exponentially (see Figure 10.84(a)).

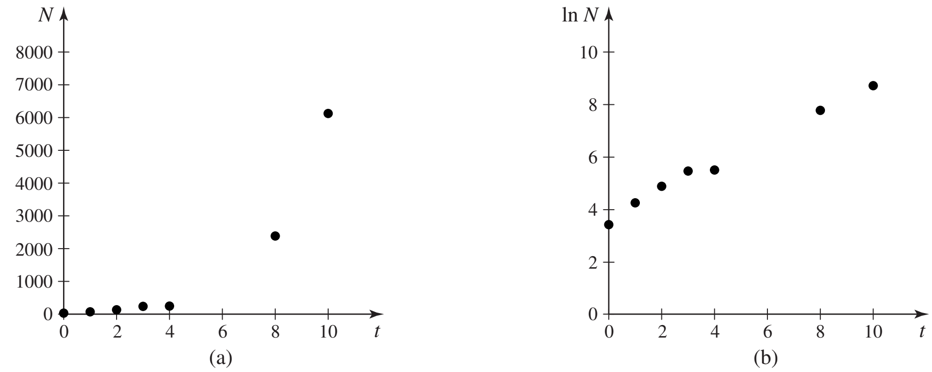


Figure 10.84 (a) N plotted against t and (b) $\ln N$ plotted against t for the data in Example 15.

However, we can transform our data, using the methods from Chapter 1, to turn our model into a linear model. Specifically, if we plot $\ln N$ against t, then according to the model:

$$\ln N = rt + \ln N_0$$

That is, there is a linear relationship between t and $\ln N$. We therefore add another row to our table, giving the values for $\ln N_i$ rather than N_i:

$\boldsymbol{x_i = t_i}$	0	1	2	3	4	8	10
$\boldsymbol{N_i}$	31	71	133	237	246	2385	6125
$\boldsymbol{y_i = \ln N_i}$	3.43	4.26	4.89	5.47	5.51	7.78	8.72

So when $\ln N$ is plotted against t we see, as anticipated, a linear relationship (see Figure 10.84(b)). We can let $x_i = t_i$ and $y_i = \ln N_i$ and use Equations (10.33) and (10.34) to estimate the slope and intercept parameters. To do so we must first calculate:

$$n = 7$$

$$\sum_{i=1}^{n} x_i = 0 + 1 + 2 + 3 + \cdots + 10 = 28$$

$$\sum_{i=1}^{n} x_i^2 = 0^2 + 1^2 + 2^2 + 3^2 + \cdots + 10^2 = 194$$

$$\sum_{i=1}^{n} y_i = 3.43 + 4.26 + 4.89 + \cdots + 8.72 = 40.06$$

$$\sum_{i=1}^{n} x_i y_i = (0)(3.43) + (1)(4.26) + (2)(4.89) + \cdots + (10)(8.72) = 201.89$$

So to 3 decimal points of accuracy:

$$r = m = \frac{(7)(201.89) - (28)(40.06)}{(7)(194) - (28)^2} = 0.508$$

and

$$\ln N_0 = c = \frac{(194) \cdot (40.06) - (28)(201.89)}{(7)(194) - (28)^2}$$
$$= 3.691$$

which implies that

$$N_0 = e^{3.691} = 40.07.$$

When these coefficients are used in the exponential growth model (i.e., in $N(t) = N_0 e^{rt}$), then the model fits the exponential measurements quite well (see Figure 10.85).

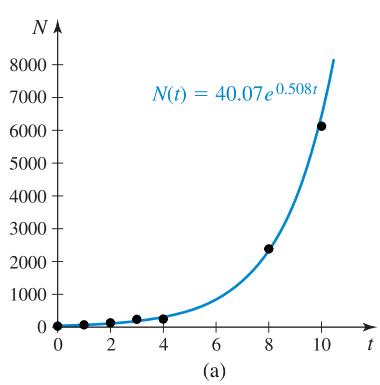


Figure 10.85 Comparing the fitted exponential model with experimental data in Example 15.

Section 10.7 Problems

10.7.1 and 10.7.2

In Problems 1–10, the functions are defined for all $(x, y) \in R^2$. Find all candidates for local extrema, and use the Hessian matrix to determine the type (maximum, minimum, or saddle point).

1. $f(x, y) = x^2 + y^2 + 2y$ **2.** $f(x, y) = -2x^2 - y^2 + 2x$

3. $f(x, y) = x^2y - 4x^2 - 4y$ **4.** $f(x, y) = -xy - 2y^2$

5. $f(x, y) = -2x^2 + y^2 - 6y$ **6.** $f(x, y) = x(1 + x - y)$

7. $f(x, y) = e^{-x^2-y^2}$ **8.** $f(x, y) = yxe^{-(x+y)}$

9. $f(x, y) = x \cos y$ **10.** $f(x, y) = y \sin x$

11. In this problem, we will illustrate that if one of the eigenvalues of the Hessian matrix at a point where the gradient vanishes is equal to 0, then we cannot make any statements about whether the point is a local extremum just on the basis of the Hessian matrix. Consider the following functions:

$$f_1(x, y) = x^2$$
$$f_2(x, y) = x^2 + y^3$$
$$f_3(x, y) = x^2 + y^4$$

Figures 10.86 through 10.88 show graphs of the three functions.

(a) Show that, for $i = 1, 2$, and 3,

$$\nabla f_i(0, 0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

(b) Show that, for $i = 1, 2$, and 3,

$$\text{Hess } f_i(0, 0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

and determine the eigenvalues of Hess $f_i(0, 0)$.

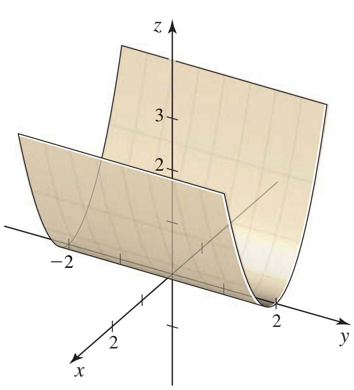


Figure 10.86 $f_1(x, y)$ in Problem 11.

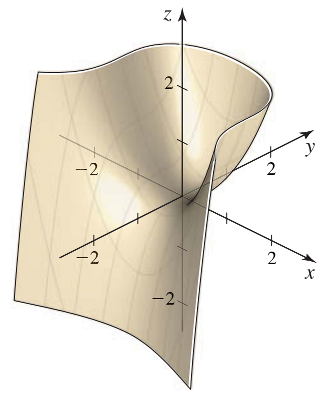


Figure 10.87 $f_2(x, y)$ in Problem 11.

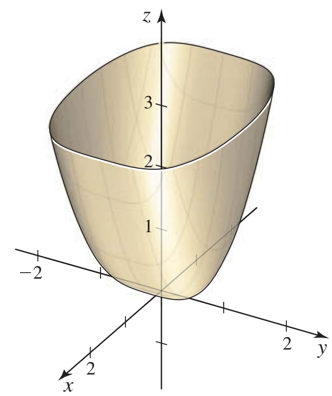


Figure 10.88 $f_3(x, y)$ in Problem 11.

(c) Since one of the eigenvalues of Hess $f_i(0, 0)$ is equal to 0, we cannot use the criterion stated in the text to determine the behavior of the three functions at $(0, 0)$. Use Figures 10.86 through 10.88 to describe what happens at $(0, 0)$ for each function.

12. Consider the function

$$f(x, y) = ax^2 + by^2$$

(a) Show that

$$\nabla f(0, 0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

(b) Find values for a and b such that **(i)** $(0, 0)$ is a local minimum, **(ii)** $(0, 0)$ is a local maximum, and **(iii)** $(0, 0)$ is a saddle point.

In Problems 13–16, the functions are defined on the rectangular domain

$$D = \{(x, y) : -1 \le x \le 1, -1 \le y \le 1\}$$

Find the global maxima and minima of f on D.

13. $f(x, y) = 2x + y$ **14.** $f(x, y) = 3 - x + 2y$

15. $f(x, y) = x^2 - y^2$ **16.** $f(x, y) = x^2 + y^2/4$

17. Find the global maxima and minima of

$$f(x, y) = x^2 + y^2 - 2x + y$$

on the set

$$D = \{(x, y) = 0 \le x \le 1, -1 \le y \le 0\}$$

18. Find the global maxima and minima of

$$f(x, y) = x^2 - y^2 + 4x + y$$

on the set

$$D = \{(x, y) = -4 \le x \le 0, 0 \le y \le 1\}$$

19. Maximize the function

$$f(x, y) = 2xy - x^2y - xy^2$$

on the triangle bounded by the line $x + y = 2$, the x-axis, and the y-axis.

20. Maximize the function

$$f(x, y) = xy(15 - 5y - 3x)$$

on the triangle bounded by the line $5y + 3x = 15$, the x-axis, and the y-axis.

21. Find the global maxima and minima of

$$f(x, y) = x^2 + y^2 + 4x - 1$$

on the disk

$$D = \{(x, y) : x^2 + y^2 \leq 9\}$$

22. Find the global maxima and minima of

$$f(x, y) = 2x^2 + y^2 - 6y + 3$$

on the disk

$$D = \{(x, y) : x^2 + y^2 \leq 16\}$$

23. Find the global maxima and minima of

$$f(x, y) = x^2 + y^2 + x - y$$

on the disk

$$D = \{(x, y) : x^2 + y^2 \leq 1\}$$

24. Find the global maxima and minima of

$$f(x, y) = x^2 + y^2 + xy - 2y$$

on the disk

$$D = \{(x, y) : x^2 + y^2 \leq 4\}$$

25. Can a continuous function of two variables have two maxima and no minima? Describe in words what the properties of such a function would be, and contrast this behavior with a function of one variable.

26. Suppose $f(x, y)$ has a horizontal tangent plane at $(0, 0)$. Can you conclude that f has a local extremum at $(0, 0)$?

27. Suppose crop yield Y depends on nitrogen (N) and phosphorus (P) concentrations as

$$Y(N, P) = NPe^{-(N+P)}$$

Find the value of (N, P) that maximizes crop yield.

28. Choose three numbers x, y, and z so that their sum is equal to 75 and their product is maximal.

29. Find the maximum volume of a rectangular closed (top, bottom, and four sides) box with surface area 48 m^2.

30. Find the maximum volume of a rectangular open (bottom and four sides, no top) box with surface area 75 m^2.

31. Find the minimum surface area of a rectangular closed (top, bottom, and four sides) box with volume 64 m^3.

32. Find the minimum surface area of a rectangular open (bottom and four sides, no top) box with volume 256 m^3.

33. The distance between the origin $(0, 0, 0)$ and the point (x, y, z) is

$$\sqrt{x^2 + y^2 + z^2}$$

Find the minimum distance between the origin and the plane $x + y + z = 1$. (*Hint*: Minimize the squared distance between the origin and the plane.)

34. Given the symmetric matrix

$$A = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$$

where a, b, and c are real numbers, show that the eigenvalues of A are real. (*Hint*: Compute the eigenvalues.)

35. Species Diversity A frequently used measure of the diversity of an habitat is the Shannon index:

$$H = -\sum_{i=1}^{n} p_i \ln p_i$$

where p_i is equal to the proportion organisms in the area that are species i, $i = 1, 2, \ldots, n$, and n is the total number of species in the study area. Assume that a habitat harbors three species with relative proportions p_1, p_2, and p_3.

(a) Use the fact that $p_1 + p_2 + p_3 = 1$ to show that H is of the form

$$H(p_1, p_2) = -p_1 \ln p_1 - p_2 \ln p_2 - (1 - p_1 - p_2) \ln(1 - p_1 - p_2)$$

and that the domain of $H(p_1, p_2)$ is the triangular set in the p_1–p_2 plane bounded by the lines $p_1 = 0$, $p_2 = 0$, and $p_1 + p_2 = 1$.

(b) Show that H attains its global maximum when $p_1 = p_2 = p_3 = 1/3$. (*Hint*: You may assume $0 \ln 0 = 0$).

10.7.3

In Problems 36–45, use Lagrange multipliers to find the maxima and minima of the functions under the given constraints.

36. $f(x, y) = 2x - y$; $x^2 + y^2 = 5$

37. $f(x, y) = 4x^2 + y$; $x^2 + y^2 = 1$

38. $f(x, y) = xy$; $x + y = 4$

39. $f(x, y) = xy$; $x^2 + 4y^2 = 1$

40. $f(x, y) = x^2 - y^2$; $2x + y = 1$

41. $f(x, y) = x^2 + y^2$; $3x - 2y = 4$

42. $f(x, y) = xy^2$; $x^2 - y = 0$

43. $f(x, y) = x^2y$; $x^2 + 3y = 1$

44. $f(x, y) = x^2y^2$; $2x - 3y = 4$

45. $f(x, y) = x^2y^2$; $x^2 + y^2 = 1$

46. Let

$$f(x, y) = x + y \quad (x, y) \in \mathbf{R}^2$$

with constraint function $xy = 1$.

(a) Use Lagrange multipliers to find all local extrema.

(b) Are there global extrema?

47. Let

$$f(x, y) = x + y$$

with constraint function

$$\frac{1}{x} + \frac{1}{y} = 1, \;\; x \neq 0, y \neq 0$$

(a) Use Lagrange multipliers to find all local extrema.

(b) Are there global extrema?

48. Let

$$f(x, y) = xy, \quad (x, y) \in \mathbf{R}^2$$

with constraint function $y - x^2 = 0$.

(a) Use Lagrange multipliers to find candidates for local extrema.

(b) Use the constraint $y - x^2 = 0$ to reduce $f(x, y)$ to a single-variable function, and then use this function to show that $f(x, y)$ has no local extrema on the constraint curve.

49. Explain why $f(x, y)$ has a local extremum at the point P in Figure 10.77 under the constraint $g(x, y) = 0$ if $c_1 > c_2 > c_3 > c_4$.

50. Assume $c_1 < c_2$ and $c_2 > c_3 > c_4$.

(a) Explain why $f(x, y)$ has a local extremum at the point P in Figure 10.77 under the constraint $g(x, y) = 0$.

(b) Explain why there must be additional local extrema on the curve $g(x, y) = 0$ [*Hint*: there must be some maximum value of $f(x, y)$ that is reached between crossing the $f(x, y) = c_1$ level curve and reaching the point P]

(c) Explain why the local extremum that you identify in (b) does not violate Lagrange's theorem.

51. In the introductory example, we discussed how egg size depends on maternal age. Assume now that the total amount of resources available is 10 (in appropriate units), the number of eggs per clutch is 3, the number of clutches is 2, and the egg size in clutch number i is denoted by x_i.

(a) Find the constraint function.

(b) Suppose the fitness function is given by

$$f(x_1, x_2) = \frac{3}{2}\rho(x_1) + \frac{1}{4}\rho(x_2)$$

where $\rho(x) = \frac{2x}{5+x}$. Find the optimal egg sizes for clutch 1 and clutch 2 under the constraint in (a).

52. In the introductory example in this subsection, we discussed how egg size depends on maternal age. Assume now that the fitness function is given by

$$f(x_1, x_2) = \frac{5}{3}\rho(x_1) + \frac{5}{6}\rho(x_2)$$

with

$$\rho(x) = \frac{3x}{4+x}$$

The constraint function is given by

$$5x_1 + 5x_2 = 7$$

(a) Compare the given functions with the corresponding ones in the text, and identify the parameters n, p_1, p_2, and R from the text.

(b) Solve the constraint function for x_2 and substitute your expression for x_2 into the function f. This then yields a function of one variable. Find the domain of this single-variable function and use single-variable calculus to determine optimal egg sizes for clutch 1 and clutch 2.

53. Species Diversity Let's revisit the Shannon diversity index from Problem 35. Suppose a habitat contains n different species of organism with a proportion p_1 of organisms being from species 1, a proportion p_2 being from species 2, and so on. Then the Shannon diversity index is defined to be

$$H = -\sum_{i=1}^{n} p_i \ln p_i$$

(a) Explain why $\sum_{i=1}^{n} p_i = 1$, and $0 \le p_i \le 1$ for each $i = 1, 2, \ldots, n$.

(b) Assume that Lagrange's theorem holds for functions of more than two variables. Then at local extrema:

$$\frac{\partial H}{\partial p_j} = \lambda \frac{\partial g}{\partial p_j} \quad \text{where} \quad g(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{n} p_i - 1.$$

For some constant λ, and for each of $j = 1, 2, \ldots, n$. Show then that at the local extremum point, $p_1 = p_2 = p_3 = \cdots = p_n = \frac{1}{n}$.

(c) In part (b) we showed that there is a candidate local extremum with $(p_1, p_2, \ldots, p_n) = (\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$. The Shannon diversity for this extremum is

$$\begin{aligned} H\left(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}\right) &= -\sum_{i=1}^{n} \left(\frac{1}{n}\right) \ln\left(\frac{1}{n}\right) \\ &= \ln n \end{aligned}$$

But it is possible that the global maximum lies on the boundaries of the domain of H.

We will not completely rule out all boundary parts, but let's consider the following boundaries:

(i) If $p_1 = 1$, show that $H = 0$ (You may assume that $0 \ln 0 = 0$)

(ii) If $p_1 = 0$, then

$$H(0, p_2, p_3, \ldots, p_n) = -\sum_{i=2}^{n} p_i \ln p_i$$

and $\sum_{i=2}^{n} p_i = 1$. Assuming that $p_i \neq 0$ and $p_i \neq 1$ for all $i > 1$, show that the only candidate local extremum that satisfies this constraint has $H = \ln(n-1)$.

Since $\ln(n-1) < \ln(n)$, we see that the local extremum identified in (b) is in fact the global maximum; that is, diversity is maximal when all species are present in the same proportions.

10.7.4

54. Model Fitting We argued based on the behavior of f as m or $c \to \infty$ that the local extremum identified in Equations (10.33) and (10.34) is the global minimum point. In this problem we will calculate the Hessian matrix.

(a) Show that the Hessian matrix is:

$$\text{Hess } f(m, c) = \begin{bmatrix} 2\sum_{i=1}^{n} x_i^2 & 2\sum_{i=1}^{n} x_i \\ 2\sum_{i=1}^{n} x_i & 2n \end{bmatrix}$$

(b) Explain why the trace of Hess $f(m, c)$ is positive.

(c) Write out the expression for det Hess $f(m, c)$.

(d) It can be shown that:

$$n\sum_{i=1}^{n} x_i^2 \ge \left(\sum_{i=1}^{n} x_i\right)^2$$

for all possible sequences $\{x_i\}$

We will not attempt to prove this result in the most general case. Instead we will prove it in the case where $n = 3$ (the proof for general n follows on similar lines).

(i) If $n = 3$, show that:

$$\begin{aligned} n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2 &= 3(x_1^2 + x_2^2 + x_3^2) - (x_1 + x_2 + x_3)^2 \\ &= (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_1)^2. \end{aligned}$$

(ii) Then explain why $n\sum_{i=1}^{n} x_i^2 \ge \left(\sum_{i=1}^{n} x_i\right)^2$ if $n = 3$. Under what circumstances does equality occur?

(e) Using your answers from parts (b) and (d), deduce whether the local extremum is a local minimum or local maximum.

55. Plant Biomass The amount of plant biomass (that is, weight of living plant matter), y, produced in a particular soil patch is thought to depend linearly on the amount of nitrogen added to the soil. Let x represent the amount of added nitrogen. Suppose you measure the following data by adding different amounts of nitrogen

x_i	0	0.5	1	1	4	6
y_i	0.79	1.62	5.54	8.14	5.00	24.62

Assuming that $y_i = mx_i + c$, find least squares estimated for the coefficients m and c.

56. Hyena Bite Strength The bite strength of juvenile hyenas is thought to increase linearly with an animal's age. Suppose you measure the following data on bite strength, y_i, as a function of animal age, x_i:

x_i	1	2	5	10	15
y_i	9.70	7.37	10.39	12.67	16.06

Assuming that $y_i = mx_i + c$, find least squares estimates for the coefficients m and c.

57. Metabolic Scaling One of the most fundamental power laws in biology is how the energy (or metabolic) needs of an animal increase as a function of its body mass. ***Kolokotrones et al. (2010)*** examined how the metabolic demands (measured in watts) of 636 species of mammals depend on their mass, M, measured in grams. Here are some representative data points from their paper:

M	5.28	22.6	121	180	608	2.99×10^3	2.68×10^4
B	0.109	0.215	0.455	0.9921	2.21	6.81	37.3

It has been hypothesized that B has a power law dependence on M

$$B = cM^a$$

for some coefficients c and a.

(a) Explain how the data can be transformed so that it may be plotted as a straight line [*Hint*: what if $\log B$ is plotted against $\log M$?]

(b) Use the method of least squares errors to estimate the coefficient a.

(c) A long-standing hypothesis, known as *Kleiber's law*, states that $a = 3/4$. Is that consistent with your estimate from (b)?

58. Hummingbird Wing Beat Frequency ***Altshuler et al. (2010)*** measured how hummingbird wing beat frequency (f, measured in wing beats per second) depends on the mass (m, measured in grams) in different birds. Here is a subset of their data:

m	3.45	5.19	5.61	7.84	8.48	21.2
f	45.1	34.7	34.5	28.0	28.7	17.7

Based on these data, the authors propose that there is a power law relationship between hummingbird body mass, m, and wing beat frequency, f, i.e.,

$$f = Cm^d$$

for some coefficients C and d.

(a) By looking at the data, would you expect $d > 0$ or $d < 0$?

(b) Explain how the data may be plotted to give a straight line relationship. [*Hint*: Plot $\log f$ against $\log m$.]

(c) Use the method of least squares to estimate the exponent d using the data above.

59. Duck Panting ***Vieth (1989)*** examined data showing how the rate of heat loss from panting ducks depends on their ambient temperature. Here is a subset of the data that she used:

ambient temperature, T (°C)	41.9	42.1	42.4	43.1
rate of heat loss, R (W· kg^{-1})	0.81	1.42	2.12	3.19

To these data she fit a linear model:

$$R = mT + c$$

where m and c are coefficients.

(a) Use Equations (10.33) and (10.34) to estimate m and c.

(b) According to the model, at what temperature does the rate of heat loss, R, decrease to 0?

60. Dopamines and Obesity ***Wang et al. (2001)*** wanted to test the hypothesis that obesity is associated with changes in brain physiology. They measured the availability of dopamine receptors in the brain in patients with different BMIs. Here BMI stands for body mass index, a measure of whether a person is over- or underweight. Here is a subset of Wang et al.'s data (we have suppressed units in their data).

BMI (B)	42.3	47.0	53.4	59.9
dopamine receptor availability (D)	3.05	2.91	2.39	1.84

Dopamines are associated with feelings of being rewarded by eating and Wang et al. hypothesized that dopamine receptor availability decreased as BMI increased, More specifically they hypothesized that there is a linear relationship

$$D = aB + b$$

for some coefficients a and b.

(a) Estimate the coefficients a and b using Equations (10.33) and (10.34).

(b) Do these data support the hypothesis that higher BMIs are associated with lower levels of dopamine receptor availability?

61. Population Growth A population of bacteria is hypothesized to grow exponentially, that is, the number of bacteria, N, increases with time, t, according to

$$N(t) = N_0 e^{rt}$$

where N_0 and r are coefficients. You measure the following data for the size of the population at different times:

t	0	1	2	3	6
N	15	15	20	21	30

(a) Use a least squares method to estimate the coefficients N_0 and r.

(b) When the fitted values of r and N_0 are substituted into the model, what is the predicted doubling time of the population (i.e., the time taken for the population to increase from N_0 to $2N_0$ bacteria)?

62. Radioactive Decay A radioactive isotope decays over time, following an exponential decay law. That is, the amount of isotope left at time t is predicted to be:

$$W(t) = W_0 e^{-\lambda t}$$

where W_0 and λ are both coefficients. You measure the following data on the amount of isotope left in a particular sample, W, at different times t.

t	0	0.1	0.2	0.4	0.8	1.0
W	113.2	63.7	66.0	32.1	13.1	3.89

(a) Use a least squares method to estimate the coefficients W_0 and λ.

(b) When the fitted coefficients W_0 and λ are input into the model, what is the predicted half-life of the isotope (that is, the time taken for the amount of isotope present to decay from W_0 to $\frac{1}{2}W_0$)?

63. Chemical Kenetics A particular chemical reaction is predicted to have *Michaelis-Menten kinetics*, meaning that the rate of reaction, r, is related to the concentration of the reacting chemical, C, by the function:

$$r = \frac{kC}{C+a}$$

where k and a are constants.

(a) Show that the reaction rate equation can be rewritten as:

$$\frac{1}{r} = \frac{a}{k} \cdot \frac{1}{C} + \frac{1}{k}$$

(b) Explain using (a) how the constants a and k could be fit from a plot of $\frac{1}{r}$ against $\frac{1}{C}$. You measure the following data for reaction rates, r, at different chemical concentrations, C.

C	0	0.1	0.5	1	2	4
r	0	0.28	1.01	1.21	1.65	1.42

Use the least squares error method to estimate k and a from these data. (*Hint*: Omit the point $(C, r) = (0, 0)$).

64. Insect Outbreak Suppose the size of an insect population, $N(t)$, grows with time t, according to the function

$$N(t) = Mte^{-mt}$$

where M and m are coefficients.

(a) Show that the model can be rewritten as:

$$\ln\left(\frac{N}{t}\right) = \ln M - mt.$$

(b) Explain how the coefficients m and M can be estimated from a plot of $\ln(N/t)$ against t.

(c) Use a least squares error method to fit M and m from the following experimental data.

t	0.1	0.3	0.5	0.8	1
N	6.11	1.64	1.00	0.196	0.0633

10.8 Diffusion

In this section we will introduce one of the most important applications of partial derivatives, which is to build partial differential equation models for biological phenomena. This is a huge topic, and so it goes well beyond the math that is covered in this book. We will show an example of this kind of modeling by deriving partial differential equations to represent diffusion, as well as showing some solutions of this equation.

Suppose that we place a sugar cube into a glass of water without stirring the water. The sugar dissolves and the sugar molecules move about randomly in the water. If we wait long enough, the sugar concentration will eventually be uniform throughout the water. This random movement of molecules is called **diffusion** and plays an important role in many processes of life. For instance, gas exchange in unicellular organisms and in many small multicellular organisms takes place by diffusion. Diffusion of chemicals is a slow process, which means that the gas that enters the organism through its surface will remain close to the surface. This limits the size and shape of organisms, unless they evolve different gas exchange mechanisms. The random swimming of some bacteria and protists also tends to scatter them diffusively.

Derivation of the One-Dimensional Diffusion Equation. We want to understand what type of microscopic description yields a diffusion equation. We assume that molecules move along the x-axis, and we denote the concentration of these molecules at x at time t by $c(x, t)$. That is, the number of molecules at time t in the interval $[x_1, x_2)$ is given by

$$N_{[x_1,x_2)}(t) = \int_{x_1}^{x_2} c(x, t)\, dx \tag{10.35}$$

Because the molecules move around, the number of molecules in a given interval changes over time. We will express this change as the difference between the net movement of molecules on the left and that on the right end of the interval. The quantity that describes this net movement is called the **flux** and is denoted by $J(x, t)$, where

$J(x,t)\,\Delta t =$ the net number of molecules crossing x from the left to the right during an interval Δt

That is, if we consider the change in the number of molecules in the interval $[x_0, x_0 + \Delta x)$ during the interval $[t, t + \Delta)$, then we have

$$N_{[x_0,x_0+\Delta x)}(t + \Delta t) - N_{[x_0,x_0+\Delta x)}(t) = \underbrace{J(x_0, t)\,\Delta t}_{\text{no. molecules entering interval from left}} - \underbrace{J(x_0 + \Delta x, t)\,\Delta t}_{\text{no. molecules entering interval to right}} \tag{10.36}$$

Dividing both sides of (10.36) by Δt and letting $\Delta t \to 0$, we find that

$$\lim_{\Delta t \to 0} \frac{N_{[x_0,x_0+\Delta x)}(t + \Delta t) - N_{[x_0,x_0+\Delta x)}(t)}{\Delta t} = J(x_0, t) - J(x_0 + \Delta x, t) \tag{10.37}$$

The left-hand side of (10.37) is equal to

$$\frac{d}{dt} N_{[x_0,x_0+\Delta x)}(t) \tag{10.38}$$

Using (10.35), we can write (10.38) as

$$\frac{d}{dt} \int_{x_0}^{x_0+\Delta x} c(x, t)\, dx$$

When $c(x, t)$ is sufficiently smooth, we can interchange differentiation and integration. (A justification of this step is beyond the scope of this course.) We find that

$$\frac{d}{dt} \int_{x_0}^{x_0+\Delta x} c(x, t)\, dx = \int_{x_0}^{x_0+\Delta x} \frac{\partial c(x, t)}{\partial t}\, dx$$

Note that the d changed into a ∂ when we moved the derivative inside the integral. Before we moved it inside, we differentiated a function that depended only on t, but once we moved it inside, we differentiated a function that depends on two variables, namely, x and t. Summarizing, we arrive at the equation

$$\int_{x_0}^{x_0+\Delta x} \frac{\partial}{\partial t} c(x, t)\, dx = J(x_0, t) - J(x_0 + \Delta x, t) \tag{10.39}$$

To obtain the diffusion equation, we divide both sides of (10.39) by Δx and take the limit as $\Delta x \to 0$. On the left-hand side of (10.39), we have

$$\lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_{x_0}^{x_0+\Delta x} \frac{\partial}{\partial t} c(x, t)\, dx = \frac{\partial c(x_0, t)}{\partial t}$$

$\int_{x_0}^{x_0+\Delta x} \frac{\partial}{\partial t} c(x, t)\, dx \approx \frac{\partial}{\partial t} c(x, t)\, \Delta x$ for small Δx

On the right-hand side, we find that

$$\lim_{\Delta x \to 0} \frac{J(x_0, t) - J(x_0 + \Delta x, t)}{\Delta x} = -\frac{\partial J(x_0, t)}{\partial x}$$

Recall definition of partial derivative.

Equating the two results, we arrive at

$$\frac{\partial c(x_0, t)}{\partial t} = -\frac{\partial J(x_0, t)}{\partial x} \tag{10.40}$$

A phenomenological law called **Fick's law** relates the flux to the change in concentration when molecules move around randomly in a solvent. Fick's law says that

$$J = -D \frac{\partial c}{\partial x} \tag{10.41}$$

where D is a positive constant called the **diffusion constant**. Equation (10.41) means that the flux is proportional to the change in concentration; the minus sign means that the net movement of molecules is from regions of high concentration to regions of low concentration. This movement agrees with our intuition: Going back to our example of sugar dissolving in water, we expect the net movement of sugar molecules to be from regions of high concentration to regions of low concentration so that ultimately the sugar concentration is uniform.

Combining (10.40) and (10.41), we arrive at the **diffusion equation**:

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} \tag{10.42}$$

The diffusion approach is ubiquitous in biology. It is used not only in the description of the random movement of molecules, but in a wide array of applications, such as the change in allele frequencies due to random genetic drift, the invasion of alien species into virgin habitat, the swimming of bacteria, the diffusion of morphogens that control growth in developing embryos, and many more phenomena. Equation (10.42) is the simplest form of an equation that incorporates diffusion. In physics, (10.42) is called the **heat equation**. It describes the diffusion of heat through a solid bar; in this case, $c(x, t)$ represents the temperature at point x at time t.

Equation (10.42) is an example of a **partial differential equation**, which is an equation that contains partial derivatives. The theory of partial differential equations is complex and well beyond the scope of this course. We will be able to discuss only some aspects of the diffusion equation.

Solving the Diffusion Equation. For most partial differential equations, it is not possible to find an analytical solution; such equations often can be solved only numerically, and even this is typically not an easy task. Fortunately, (10.42) is simple enough that we can find a solution. We claim that

$$c(x,t) = \frac{1}{\sqrt{4\pi Dt}}\exp\left[-\frac{x^2}{4Dt}\right] \tag{10.43}$$

is a solution of (10.42). As in the case of ordinary differential equations, we can check this by computing the appropriate derivatives. For the left-hand side of (10.42), we need the first partial derivative of $c(x, t)$ with respect to t. We find that

$$\begin{aligned}\frac{\partial c(x,t)}{\partial t} &= \frac{\partial}{\partial t}\left(\frac{1}{\sqrt{4\pi Dt}}\exp\left[-\frac{x^2}{4Dt}\right]\right)\\ &= -\frac{1}{2}\frac{4\pi D}{(4\pi Dt)^{3/2}}\exp\left[-\frac{x^2}{4Dt}\right] + \frac{1}{\sqrt{4\pi Dt}}\exp\left[-\frac{x^2}{4Dt}\right]\frac{x^2}{4Dt^2}\\ &= \exp\left[-\frac{x^2}{4Dt}\right]\left\{\frac{x^2}{4Dt^2\sqrt{4\pi Dt}} - \frac{2\pi D}{4\pi Dt\sqrt{4\pi Dt}}\right\}\\ &= \frac{1}{2t\sqrt{4\pi Dt}}\exp\left[-\frac{x^2}{4Dt}\right]\left\{\frac{x^2}{2Dt} - 1\right\}.\end{aligned} \tag{10.44}$$

On the right-hand side of (10.42), we need the second partial derivative of $c(x, t)$ with respect to x. We find that

$$\begin{aligned}\frac{\partial c(x,t)}{\partial x} &= \frac{\partial}{\partial x}\left(\frac{1}{\sqrt{4\pi Dt}}\exp\left[-\frac{x^2}{4Dt}\right]\right)\\ &= \frac{1}{\sqrt{4\pi Dt}}\exp\left[-\frac{x^2}{4Dt}\right]\left(-\frac{2x}{4Dt}\right)\end{aligned}$$

and then

$$\begin{aligned}\frac{\partial^2 c(x,t)}{\partial x^2} &= \frac{-1}{\sqrt{4\pi Dt}}\frac{1}{2Dt}\exp\left[-\frac{x^2}{4Dt}\right]\left\{1 - x\frac{2x}{4Dt}\right\}\\ &= \frac{1}{2Dt\sqrt{4\pi Dt}}\exp\left[-\frac{x^2}{4Dt}\right]\left\{\frac{x^2}{2Dt} - 1\right\}.\end{aligned} \tag{10.45}$$

Since substituting into the left-hand and right-hand sides of the equation produces the same function, we see that (10.43) satisfies (10.42).

The function in (10.43) is called the **Gaussian density**. Figure 10.89 shows $c(x, t)$ for $t = 1, 2$, and 4; it shows that the concentration $c(x, t)$ becomes more uniform as time goes on.

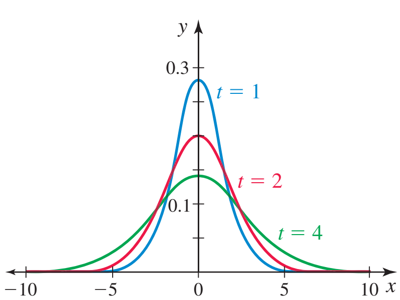


Figure 10.89 The solution of the diffusion equation at different times.

Diffusion is a very slow process. The diffusion constant D measures how quickly it proceeds. The larger D, the faster the concentration spreads out, and we can show that within t units of time, the bulk of the molecules spread over a region of length of order $\sqrt{t}$.

To give an idea of how slow diffusion is, here are a few examples taken from ***Yeargers, Shonkwiler, and Herod (1996)***. Oxygen in blood at 20°C has a diffusion constant of $10^{-5}\ \mathrm{cm}^2/\mathrm{s}$, which means that it takes an oxygen molecule roughly 500 seconds to cross a distance of 1 mm by diffusion alone. Ribonuclease (an enzyme that hydrolyzes ribonucleic acid) in water at 20°C has a diffusion constant of $1.1 \times 10^{-6}\ \mathrm{cm}^2/\mathrm{s}$, which means that ribonuclease takes roughly 4672 seconds (or 1 hr, 18 min) to cross a distance of 1 mm by diffusion alone. These examples illustrate why organisms frequently rely on other active mechanisms such as the flow of blood in arteries and veins to transport molecules.

The diffusion equation (10.42) can be generalized to higher dimensions. In that case, (10.40) becomes

$$\frac{\partial c}{\partial t} = -\nabla J \tag{10.46}$$

and (10.41) becomes

$$J = -D\nabla c \tag{10.47}$$

Combining (10.46) and (10.47), we find that

$$\frac{\partial c}{\partial t} = D\nabla \cdot (\nabla c)$$

where $\nabla \cdot (\nabla c)$ is to be interpreted as a dot product. That is, if $\mathbf{x} = (x_1, x_2, x_3) \in \mathbf{R}^3$, $t \in \mathbf{R}$, then

$$\frac{\partial c}{\partial t} = D\left(\frac{\partial^2 c}{\partial x_1^2} + \frac{\partial^2 c}{\partial x_2^2} + \frac{\partial^2 c}{\partial x_3^2}\right)$$

As a shorthand notation, we define

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$$

where $\triangle$ is called the **Laplace operator**. We then write

$$\frac{\partial c}{\partial t} = D\,\triangle c$$

Section 10.8 Problems

10.8

1. Show that

$$c(x,t) = \frac{1}{\sqrt{8\pi t}} \exp\left[-\frac{x^2}{8t}\right]$$

solves

$$\frac{\partial c(x,t)}{\partial t} = 2\frac{\partial^2 c(x,t)}{\partial x^2}$$

2. Show that

$$c(x,t) = \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{x^2}{2t}\right]$$

solves

$$\frac{\partial c(x,t)}{\partial t} = \frac{1}{2}\frac{\partial^2 c(x,t)}{\partial x^2}$$

3. A solution of

$$\frac{\partial c(x,t)}{\partial t} = D\frac{\partial^2 c(x,t)}{\partial x^2}$$

is the function

$$c(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left[-\frac{x^2}{4Dt}\right]$$

for $x \in \mathbf{R}$ and $t > 0$.

(a) Show that, as a function of x for fixed values of $t > 0$, $c(x,t)$ is **(i)** positive for all $x \in \mathbf{R}$, **(ii)** is increasing for $x < 0$ and decreasing for $x > 0$, **(iii)** has a local maximum at $x = 0$, and **(iv)** has inflection points at $x = \pm\sqrt{2Dt}$.

(b) Graph $c(x,t)$ as a function of x when $D = 1$ for $t = 0.01$, $t = 0.1$, and $t = 1$.

4. A solution of

$$\frac{\partial c(x,t)}{\partial t} = D\frac{\partial^2 c(x,t)}{\partial x^2}$$

is the function

$$c(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left[-\frac{x^2}{4Dt}\right]$$

for $x \in \mathbf{R}$ and $t > 0$.

(a) Show that a local maximum of $c(x, t)$ occurs at $x = 0$ for fixed t.

(b) Show that $c(0, t)$, $t > 0$, is a decreasing function of t.

(c) Find

$$\lim_{t \to 0^+} c(x, t)$$

when $x = 0$ and when $x \neq 0$.

(d) Use the fact that

$$\int_{-\infty}^{\infty} e^{-u^2/2}\, du = \sqrt{2\pi}$$

to show that, for $t > 0$,

$$\int_{-\infty}^{\infty} c(x, t)\, dx = 1$$

(e) The function $c(x, t)$ can be interpreted as the concentration of a substance diffusing in space. Explain the meaning of

$$\int_{-\infty}^{\infty} c(x, t)\, dx = 1$$

and use your results in (c) and (d) to explain why this means that initially (i.e., at $t = 0$) the entire amount of the substance was released at the origin.

5. A chemical diffuses in a container that occupies the interval $0 \leq x \leq 1$. The concentration of the chemical at time t and at a point x is given by the diffusion equation:

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}$$

(a) Suppose that the chemical is allowed to diffuse through the entire container until the concentration reaches an equilibrium value where c does not change any more with time, that is, $\partial c/\partial t = 0$. Suppose that chemical that touches the walls of the container is removed so that

$$c(0, t) = c(1, t) = 0.$$

The steady state concentration of chemical will be a function $C(x)$ with

$$0 = D\frac{d^2C}{dx^2} \text{ for } x \in (0, 1)$$

and $C(0) = C(1) = 0$.

Show that $C(x) = 0$ satisfies this differential equation and the constraints as the points $x = 0$ and $x = 1$.

(b) Now suppose that chemical is added to the container by a reaction that occurs at the wall $x = 0$. This reaction keeps the concentration of chemical at this wall equal to $c(0, t) = 1$. Under these conditions the steady state distribution of chemical will obey a differential equation:

$$0 = D\frac{d^2C}{dx^2} \text{ for } x \in (0, 1)$$

with $C(0) = 1$ and $C(1) = 0$. Show that $C(x) = 1 - x$ satisfies both the differential equation and the boundary conditions at $x = 0$ and $x = 1$.

(c) Notice that the steady state distributions in (a) and (b) do not depend on D. Can you explain why?

6. Morphogenesis Most embryos start out as lumps of cells. Cells in these lumps are initially undifferentiated—that is, they start out all in the same state. Over time cells then commit to different functions, e.g., to becoming legs, eyes, and so on. To do this chemicals called *morphogens* are distributed unequally through the embryo, allowing each cell to tell where in the embryo it is located. How are unequal distributions of morphogens achieved?

One model for how morphogens can be distributed through the embryo is that morphogens are continuously produced at one end (also called pole) of the embryo. From there they diffuse through the embryo. As the morphogens diffuse, they are constantly broken down by the cells in the embryo.

First let's ignore the process of morphogen degradation, and focus only on diffusion. We will assume that the pole at which the morphogen is produced is located at $x = 0$; and for simplicity's sake the cell occupies the interval $x \geq 0$. Then our partial differential equation model for the distribution of morphogen becomes:

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} \quad \text{for} \quad x > 0$$

with

$$-D\frac{\partial c}{\partial x}(0, t) = Q \quad \text{and} \quad c(x, t) \to 0 \text{ as } x \to \infty$$

where Q is the rate of morphogen production.

(a) Let's try to find a steady state distribution of morphogen. That is, we will assume that over time the morphogen concentration reaches some state that does not change with time, i.e., the concentration is given by a function $C(x)$. Then $C(x)$ will satisfy the partial differential equation if and only if:

$$0 = D\frac{d^2C}{dx^2} \quad \text{for} \quad x > 0 \tag{10.48}$$

$$-DC'(0) = Q \quad \text{and} \quad C(x) \to 0 \text{ as } x \to \infty$$

Show that there is no function $C(x)$ that satisfies this differential equation. [*Hint*: Start by integrating once (10.48) to find dC/dx and then again to find $C(x)$, then try to impose the constraints at $x = 0$, and as $x \to \infty$ on your solution.]

(b) Now let's incorporate morphogen degradation into our model. We will assume that the breakdown of morphogen has first order kinetics (see Section 5.9 for a discussion of the different kinds of kinetics that chemical reactions may have). This means that in one unit of time a fraction r of the morphogen contained in each region of the embryo is degraded. Then our partial differential equation must be altered to:

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} - rc \quad \text{for} \quad x > 0 \tag{10.49}$$

with

$$-D\frac{\partial c}{\partial x}(0, t) = Q \quad \text{and} \quad c(x, t) \to 0 \text{ as } x \to \infty$$

Show that this partial differential equation does have a steady state solution of the form:

$$C(x) = Q\sqrt{\frac{1}{Dr}}\exp\left(-\sqrt{\frac{r}{D}}x\right)$$

That is, check that this function $C(x)$ satisfies both the steady state form of (10.49) as well as the constraints at $x = 0$ and as $x \to \infty$.

7. The two-dimensional diffusion equation

$$\frac{\partial n(\mathbf{r}, t)}{\partial t} = D\left(\frac{\partial^2 n(\mathbf{r}, t)}{\partial x^2} + \frac{\partial^2 n(\mathbf{r}, t)}{\partial y^2}\right) \tag{10.50}$$

where $n(\mathbf{r}, t)$, $\mathbf{r} = (x, y)$, denotes the population density at the point $\mathbf{r} = (x, y)$ in the plane at time t, can be used to describe the spread of organisms. Assume that a large number of insects are released at time 0 at the point $(0, 0)$. Furthermore, assume that, at later times, the density of these insects can be described by the diffusion equation (10.50). Show that

$$n(x, y, t) = \frac{n_0}{4\pi Dt}\exp\left[-\frac{x^2 + y^2}{4Dt}\right]$$

satisfies (10.50).

10.9 Systems of Recurrence Equations*

10.9.1 A Biological Example

About 14% of all insect species are estimated to belong to a group of insects called *parasitoids*. These are insects that lay their eggs on, in, or near the body of another arthropod, which serves as a host for the developing parasitoids. The developing parasitoid larvae then feed on the host. Parasitoids play an important role in biological control. For example *Trichogramma* wasps parasitize insect eggs. These wasps are reared in factories for subsequent release to the field. Every year, millions of hectares of agricultural land are treated with released *Trichogramma* wasps, for instance, to protect sugar cane from sugarcane borers, in China, or to protect cornfields from the European corn borer in western Europe. Another successful example of biological control of an insect pest is the parasitoid wasp *Aphytis melinus*, which regulates red scale disease, which damages citrus trees in California.

The importance of parasitoids in pest control stimulated both experiments and mathematical modeling work. Theoretical studies of host–parasitoid interactions go back to ***Thompson (1924)*** and ***Nicholson and Bailey (1935)***. Nicholson and Bailey introduced discrete-generation, host–parasitoid models. Specifically if N_t and P_t denote the population sizes of, respectively, susceptible hosts and searching adult female parasitoids at time t then:

$$N_{t+1} = bN_t e^{-aP_t}$$

$$P_{t+1} = cN_t[1 - e^{-aP_t}]$$

for $t = 0, 1, 2, \ldots$. The parameter b is interpreted as the *net growth parameter*. The key difference between these models and the models of population growth that were introduced in Section 2.3 is that we now have two populations that depend on each other. The productive rates of the host depend on the number of parasitoids, since parasitoids destroy hosts. But the reproductive rate of parasitoids depends on hosts because parasitoids depend on hosts for food. So our model takes the form of two recurrence equations in which N_{t+1} depends on both N_t and P_t, while P_{t+1} depends on both N_t and P_t also. We see from the first equation that hosts grow exponentially in the absence of parasitoids ($P_t = 0$). The term e^{-aP_t} is the fraction of hosts that are *not* parasitized (and thus $1 - e^{-aP_t}$ is the fraction of hosts that are parasitized) at generation t. Parasitized hosts produce parasitoids. The parameter c is equal to the number of parasitoids produced per parasitized host. Only nonparasitized hosts produce offspring in each time step.

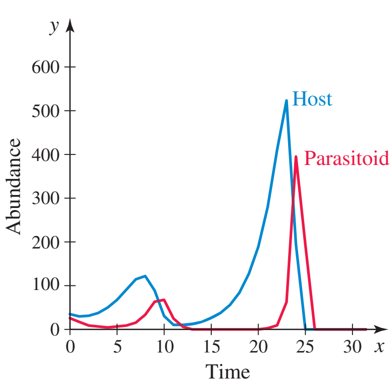


Figure 10.90 A numerical simulation of the Nicholson–Bailey equation with $a = 0.023$, $b = 1.5$, and $c = 1$. The population sizes oscillate until both host and parasitoid become extinct.

A numerical simulation (Figure 10.90) of the Nicholson–Bailey equation shows that population sizes oscillate with increasing amplitude until the host becomes extinct, followed by the extinction of the parasitoid. It is also possible that the parasitoid becomes extinct, followed by an exponential increase of the host.

The behavior of the model disagrees with most empirical studies (although some laboratory experiments have produced such unstable behavior). The model has since been modified in a number of ways to stabilize the dynamics. One such attempt is called the *negative binomial model* ***(Griffiths, 1969; May, 1978)***, in which

$$N_{t+1} = bN_t\left(1 + \frac{aP_t}{k}\right)^{-k}$$

$$P_{t+1} = cN_t\left[1 - \left(1 + \frac{aP_t}{k}\right)^{-k}\right]$$

The form of this set of equations is quite similar to that of the Nicholson–Bailey equation, and the parameters b and c have the same interpretation as before. The main (and crucial) difference is the term $\left(1 + \frac{aP_t}{k}\right)^{-k}$, which replaces the term e^{-aP_t} in the

* You should study Section 5.7 before studying this section.

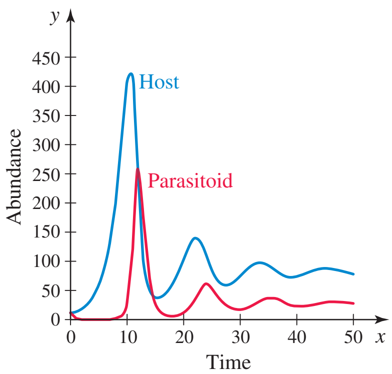


Figure 10.91 A numerical simulation of the negative binomial model with $a = 0.023$, $b = 1.5$, $c = 1$, and $k = 0.5$. The choices for the parameters a, b, and c are the same as in Figure 10.90. The choice for k stabilizes the host–parasitoid interactions, and coexistence occurs.

Nicholson–Bailey model. It has the same interpretation, though, in both equations, denoting the fraction of hosts that escape parasitism.

For some choices of parameter in the numerical simulation of the negative binomial model (Figure 10.91) host and parasitoids equilibrate so that they both have positive abundances. (We call this *coexistence*.)

The two host–parasitoid examples just give a glimpse of the possible behavior of multispecies interactions that are modeled by discrete-generation recurrence equations. In what follows, we will concentrate on coexistence in two-species, discrete-time models. The analysis will parallel our discussion of recurrence equations in Section 5.7, where we examined the equilibria and stability of single-species, discrete-time models of the form

$$x_{t+1} = f(x_t)$$

There, we found that point equilibria satisfy the equation

$$x^* = f(x^*)$$

and that such point equilibria are locally stable if $|f'(x^*)| < 1$. We obtained this condition by linearizing $f(x)$ about the equilibrium x^*.

We will see in this section that point equilibria satisfy a similar condition in two-species models and that the same strategy of linearizing about the equilibrium will yield an analogous condition of local stability in two-species models. We begin our discussion by studying linear recurrence equations.

10.9.2 Equilibria and Stability in Systems of Linear Recurrence Equations

Linear recurrence equations are of the form

$$x_1(t+1) = a_{11}x_1(t) + a_{12}x_2(t) \tag{10.51}$$

$$x_2(t+1) = a_{21}x_1(t) + a_{22}x_2(t) \tag{10.52}$$

where $t = 0, 1, 2, \ldots$. This set of equations can be written in matrix form

$$\underbrace{\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix}}_{\mathbf{x}(t+1)} = \underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}}_{\mathbf{x}(t)}$$

which shows that linear recurrence equations are *linear maps*, which we discussed in Section 9.3. First note that if $\mathbf{x}(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, then $\mathbf{x}(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ for all $t = 1, 2, 3, \ldots$. We call $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ a (**point**) **equilibrium**. More generally, a point equilibrium satisfies the equation

$$\mathbf{x}^* = A\mathbf{x}^*$$

We see that $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a point equilibrium of the linear system $\mathbf{x}(t+1) = A\mathbf{x}(t)$. We will now investigate what happens when $\mathbf{x}(0) \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

In Section 9.3, we learned how to compute an explicit formula for $\mathbf{x}(t)$ for a given initial condition $\mathbf{x}(0)$. We found that if A has two real and distinct eigenvalues λ_1 and λ_2, then we can write any vector $\mathbf{x}(0)$ as a linear combination of its eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ (corresponding to λ_1 and λ_2, respectively); that is,

$$\mathbf{x}(0) = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$$

where c_1 and c_2 are real numbers. Using this representation of $\mathbf{x}(0)$, we found that

$$\mathbf{x}(t) = c_1\lambda_1^t\mathbf{u}_1 + c_2\lambda_2^t\mathbf{u}_2 \tag{10.53}$$

which we can use to say something about the *long-term behavior* of $\mathbf{x}(t)$, or $\lim_{t\to\infty}\mathbf{x}(t)$. Let's return to the question of what happens to $\mathbf{x}(t)$ when $\mathbf{x}(t) \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. We see from the representation of $\mathbf{x}(t)$ in (10.53) that if $|\lambda_1| < 1$ and $|\lambda_2| < 1$, then $\lim_{t\to\infty}\mathbf{x}(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, regardless of $\mathbf{x}(0)$. In this case, we say that $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a **stable equilibrium**. If either $|\lambda_1| > 1$ or $|\lambda_2| > 1$, then the equilibrium $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is called **unstable**.

Criterion for $\begin{bmatrix}\mathbf{0}\\\mathbf{0}\end{bmatrix}$ to be a Stable Equilibrium The point $\begin{bmatrix}0\\0\end{bmatrix}$ is a stable equilibrium of

$$\mathbf{x}(t+1) = A\mathbf{x}(t)$$

if both eigenvalues λ_1 and λ_2 of A satisfy

$$|\lambda_1| < 1 \qquad \text{and} \qquad |\lambda_2| < 1$$

If either $|\lambda_1| > 1$ or $|\lambda_2| > 1$, $\begin{bmatrix}0\\0\end{bmatrix}$ is an unstable equilibrium.

In the preceding criterion, we no longer require the eigenvalues of A to be real and distinct. However, it is beyond the scope of this book to show the criterion for general λ_1 and λ_2.

EXAMPLE 1 Show that $\begin{bmatrix}0\\0\end{bmatrix}$ is an equilibrium of

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} -0.4 & 0.2 \\ -0.3 & 0.1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

and determine its stability.

Solution To check that $\begin{bmatrix}0\\0\end{bmatrix}$ is an equilibrium, we need to show that $\begin{bmatrix}0\\0\end{bmatrix}$ satisfies

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -0.4 & 0.2 \\ -0.3 & 0.1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

But this is true, since $A\begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$ for any matrix A with constant entries. To determine the stability of A, we need to find the eigenvalues of A. That is, we need to solve

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} -0.4-\lambda & 0.2 \\ -0.3 & 0.1-\lambda \end{bmatrix} \\ &= (-0.4-\lambda)(0.1-\lambda) + (0.2)(0.3) \\ &= \lambda^2 + 0.3\lambda + 0.02 = 0 \end{aligned}$$

The solutions are

$$\lambda_{1,2} = \frac{-0.3 \pm \sqrt{0.09 - 0.08}}{2} = \frac{-0.3 \pm 0.1}{2}$$

Hence, $\lambda_1 = -0.1$ and $\lambda_2 = -0.2$. Since $|\lambda_1| = |-0.1| < 1$ and $|\lambda_2| = |-0.2| < 1$, it follows that $\begin{bmatrix}0\\0\end{bmatrix}$ is stable. •

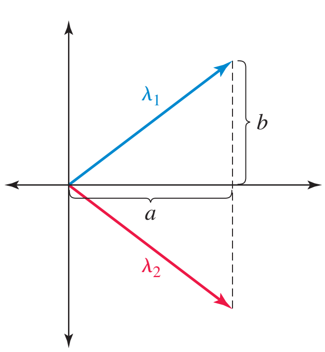


Figure 10.92 A graphical illustration of the identity $|\lambda_1|^2 = |\lambda_2|^2 = \lambda_1\lambda_2$.

When the eigenvalues λ_1 and λ_2 are complex conjugate, the criterion for stability can be simplified. Specifically, if λ_1 and λ_2 are complex conjugates, then

$$|\lambda_1|^2 = |\lambda_2|^2 = \lambda_1\lambda_2 \tag{10.54}$$

This identity is not obvious at first sight, but can be demonstrated when we graph λ_1 and λ_2 and determine their absolute values. Let

$$\lambda_1 = a + ib \qquad \text{and} \qquad \lambda_2 = a - ib$$

be the two complex conjugate eigenvalues of A. In Figure 10.92, we draw λ_1 and λ_2 when both a and b are positive. An application of the Pythagorean theorem shows that

$$|\lambda_1|^2 = a^2 + b^2 \qquad \text{and} \qquad |\lambda_2|^2 = a^2 + b^2 \tag{10.55}$$

Algebraically, we find that

$$\lambda_1\lambda_2 = (a+ib)(a-ib) = a^2 - (ib)^2 = a^2 - i^2b^2 = a^2 + b^2$$

since $i^2 = -1$. Combining the preceding equation for $\lambda_1\lambda_2$ with (10.55) proves (10.54), which we use in the next example.

EXAMPLE 2 Show that the equilibrium $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ of

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} -0.3 & -0.5 \\ 0.7 & 0.15 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

is stable.

Solution To test the stability of $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, we need to solve

$$\det \begin{bmatrix} -0.3-\lambda & -0.5 \\ 0.7 & 0.15-\lambda \end{bmatrix} = 0$$

This amounts to solving the quadratic equation

$$(-0.3-\lambda)(0.15-\lambda) + (0.5)(0.7) = 0$$

or

$$\lambda^2 + 0.15\lambda + 0.305 = 0 \tag{10.56}$$

Since the discriminant $(0.15)^2 - (4)(1)(0.305) = -1.1975 < 0$, it follows that the two solutions λ_1 and λ_2 of (10.56) are complex conjugates. Without computing λ_1 and λ_2, we can check whether $|\lambda_1|$ and $|\lambda_2|$ are both less than 1, since $|\lambda_1|^2 = |\lambda_2|^2 = \lambda_1\lambda_2 = \det A$. Now,

$$\begin{aligned} \det A &= \begin{bmatrix} -0.3 & -0.5 \\ 0.7 & 0.15 \end{bmatrix} \\ &= (-0.3)(0.15) - (0.7)(-0.5) = 0.305 < 1 \end{aligned}$$

Therefore, $|\lambda_1| < 1$ and $|\lambda_2| < 1$, and $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a stable equilibrium. •

10.9.3 Equilibria and Stability of Nonlinear Systems of Recurrence Equations

We saw examples of nonlinear systems of recurrence equations at the beginning of this section: the Nicholson–Bailey equation and the negative binomial model. The general form of a system of two nonlinear recurrence equations is

$$x_1(t+1) = F(x_1(t), x_2(t)) \tag{10.57}$$

$$x_2(t+1) = G(x_1(t), x_2(t)) \tag{10.58}$$

where F and G are (nonlinear) functions of the two variables x_1 and x_2 and $t = 0, 1, 2, \ldots$ is the independent variable that denotes time.

As in the previous subsection, we will be interested here in equilibria and their stability. We say that the point (x_1^*, x_2^*) is a (**point**) **equilibrium** of the system (10.57) and (10.58) if x_1^* and x_2^* simultaneously satisfy the two equations

$$\begin{aligned} x_1^* &= F(x_1^*, x_2^*) \\ x_2^* &= G(x_1^*, x_2^*) \end{aligned}$$

EXAMPLE 3 Find all equilibria of

$$x_1(t+1) = 2x_1(t)[1 - x_1(t)]$$
$$x_2(t+1) = x_1(t)[1 - x_2(t)].$$

Solution To find the equilibria, we need to solve

$$x_1 = 2x_1(1 - x_1)$$
$$x_2 = x_1(1 - x_2)$$

Multiplying out and rearranging terms yields

$$2x_1^2 - x_1 = 0$$
$$x_2 + x_1 x_2 - x_1 = 0$$

The first equation has the solution $x_1 = 0$ or $\frac{1}{2}$. Solving the second equation for x_2, we find that

$$x_2(1 + x_1) = x_1, \qquad \text{or} \qquad x_2 = \frac{x_1}{1 + x_1}$$

If $x_1 = 0$, then $x_2 = 0$; if $x_1 = 1/2$, then $x_2 = (1/2)/(3/2) = 1/3$. Summarizing our results, we found two point equilibria:

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1/2 \\ 1/3 \end{bmatrix}$$

•

EXAMPLE 4 Find all biologically relevant equilibria of the Nicholson–Bailey model:

$$N_{t+1} = bN_t e^{-aP_t}$$
$$P_{t+1} = cN_t[1 - e^{-aP_t}].$$

Solution To find the equilibria in question, we need to solve

$$N = bNe^{-aP} \tag{10.59}$$
$$P = cN[1 - e^{-aP}] \tag{10.60}$$

The first equation is satisfied for $N = 0$. If we substitute $N = 0$ into the second equation, we obtain $P = 0$. The system thus has the trivial equilibrium $(N^*, P^*) = (0, 0)$, which corresponds to the state when both host and parasitoid are absent.

If $N \neq 0$, then we can cancel N in equation (10.59), resulting in

$$e^{aP^*} = b, \qquad \text{or} \qquad P^* = \frac{1}{a} \ln b$$

We see that, in order for P^* to be positive (this is required to be a biologically reasonable nontrivial equilibrium), we require b to be greater than 1. Now, using $e^{aP^*} = b$, we find that

$$P^* = cN^* \left[1 - \frac{1}{b}\right]$$

This equation reduces to

$$N^* = \frac{P^*}{c[1 - 1/b]} = \frac{\ln b}{ac[1 - 1/b]} = \frac{b}{b-1} \frac{1}{ac} \ln b$$

when $P^* = \frac{1}{a} \ln b$. We see that, for $b > 1$, $N^* > 0$. We conclude that, in addition to the trivial equilibrium $(N^*, P^*) = (0, 0)$, if $b > 1$, then there exists a biologically reasonable, nontrivial equilibrium—that is, an equilibrium in which both the host and the parasitoid densities are positive. This equilibrium is given by

$$N^* = \frac{b}{b-1} \frac{1}{ac} \ln b \qquad \text{and} \qquad P^* = \frac{1}{a} \ln b$$

•

To determine the stability of point equilibria, we proceed in the same way as in the single-species case: We linearize about the equilibria and use the linearized system and what we learned about linear maps in Chapter 9 to derive an analytical condition for local stability. Here is how this works: We start with the general system of recurrence equations

$$x_1(t+1) = F(x_1(t), x_2(t)) \tag{10.61}$$

$$x_2(t+1) = G(x_1(t), x_2(t)) \tag{10.62}$$

and assume that it has a point equilibrium (x_1^*, x_2^*), which simultaneously satisfies

$$x_1^* = F(x_1^*, x_2^*) \qquad \text{and} \qquad x_2^* = G(x_1^*, x_2^*)$$

To linearize about (x_1^*, x_2^*), we write

$$x_1(t) = x_1^* + X_1(t) \qquad \text{and} \qquad x_2(t) = x_2^* + X_2(t)$$

where we interpret $X_1(t)$ and $X_2(t)$ as small perturbations, just as in Chapter 5, where we discussed stability of equilibria in recurrence equations, or in Chapter 8, where we discussed the stability of equilibria in differential equations. Now linearizing $F(x_1(t), x_2(t))$ and $G(x_1(t), x_2(t))$ about the equilibrium (x_1^*, x_2^*), we find that

$$F(x_1(t), x_2(t)) \simeq F(x_1^*, x_2^*) + \left(\frac{\partial F}{\partial X_1}\right)^* X_1(t) + \left(\frac{\partial F}{\partial X_2}\right)^* X_2(t)$$

$$G(x_1(t), x_2(t)) \simeq G(x_1^*, x_2^*) + \left(\frac{\partial G}{\partial X_1}\right)^* X_1(t) + \left(\frac{\partial G}{\partial X_2}\right)^* X_2(t)$$

where $(\cdot)^*$ means that we evaluate the expression in the parentheses at the equilibrium (x_1^*, x_2^*).

With $x_1(t) = x_1^* + X_1(t)$ and $x_2(t) = x_2^* + X_2(t)$, we find that

$$x_1^* + X_1(t+1) \approx \underbrace{F(x_1^*, x_2^*)}_{x_1^*} + \left(\frac{\partial F}{\partial x_1}\right)^* X_1(t) + \left(\frac{\partial F}{\partial x_2}\right)^* X_2(t)$$

$$x_2^* + X_2(t+1) \approx \underbrace{G(x_1^*, x_2^*)}_{x_2^*} + \left(\frac{\partial G}{\partial x_1}\right)^* X_1(t) + \left(\frac{\partial G}{\partial x_2}\right)^* X_2(t)$$

Canceling x_1^* from the first equation and x_2^* from the second equation and writing the resulting approximation in matrix form, we obtain

$$\begin{bmatrix} X_1(t+1) \\ X_2(t+1) \end{bmatrix} \approx \begin{bmatrix} \left(\frac{\partial F}{\partial x_1}\right)^* & \left(\frac{\partial F}{\partial x_2}\right)^* \\ \left(\frac{\partial G}{\partial x_1}\right)^* & \left(\frac{\partial G}{\partial x_2}\right)^* \end{bmatrix} \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} \tag{10.63}$$

We recognize the 2×2 matrix as the Jacobi matrix of the vector-valued function $\begin{bmatrix} F(x_1, x_2) \\ G(x_1, x_2) \end{bmatrix}$. The right-hand side of (10.63) is a linear map of the form $A\mathbf{x}$, where A is a 2×2 matrix and $\mathbf{x} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ is a 2×1 vector. In the previous subsection, we found that a linear map

$$\mathbf{x}_{t+1} = A\mathbf{x}_t$$

has the equilibrium $(0, 0)$, which is stable if the absolute values of the two eigenvalues of A are each less than 1. This is the criterion we need to determine the stability of the equilibrium (x_1^*, x_2^*) of the system (10.61) and (10.62).

Criterion for Stability of Point Equilibria The point equilibrium $\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix}$ of the system (10.61) and (10.62) is locally stable if the two eigenvalues λ_1 and λ_2 of the Jacobi matrix

$$\begin{bmatrix} \left(\frac{\partial F}{\partial x_1}\right)^* & \left(\frac{\partial F}{\partial x_2}\right)^* \\ \left(\frac{\partial G}{\partial x_1}\right)^* & \left(\frac{\partial G}{\partial x_2}\right)^* \end{bmatrix}$$

$\left(\frac{\partial F}{\partial x_1}\right)^* = \frac{\partial F}{\partial x_1}(x_1^*, x_2^*)$

evaluated at (x_1^*, x_2^*) satisfy

$$|\lambda_1| < 1 \qquad \text{and} \qquad |\lambda_2| < 1$$

If $|\lambda_1| > 1$ or $|\lambda_2| > 1$, the point equilibrium $\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix}$ is an unstable equilibrium.

Note that, as in Chapters 5 and 8, the stability analysis is only a local analysis and we must say that an equilibrium is *locally* stable; the local analysis does not reveal anything about global stability.

EXAMPLE 5 Discuss the stability of the equilibria of the system in Example 3.

Solution In Example 3,

$$F(x_1, x_2) = 2x_1(1 - x_1)$$
$$G(x_1, x_2) = x_1(1 - x_2)$$

The Jacobi matrix is

$$J(x_1, x_2) = \begin{bmatrix} 2 - 4x_1 & 0 \\ 1 - x_2 & -x_1 \end{bmatrix}$$

Evaluating $J(x_1, x_2)$ at the equilibrium $(0, 0)$, we find that

$$J(0, 0) = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}$$

which has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 0$. Since $|\lambda_1| > 1$, it follows that $(0, 0)$ is an unstable equilibrium.

Evaluating $J(x_1, x_2)$ at the equilibrium $(1/2, 1/3)$, we obtain

$$J(1/2, 1/3) = \begin{bmatrix} 0 & 0 \\ 2/3 & -1/2 \end{bmatrix}$$

which has eigenvalues $\lambda_1 = 0$ and $\lambda_2 = -1/2$. Since $|\lambda_1| < 1$ and $|\lambda_2| < 1$, it follows that $(1/2, 1/3)$ is a locally stable equilibrium. •

EXAMPLE 6 Show that the nontrivial equilibrium of the Nicholson–Bailey equation is unstable.

Solution The Nicholson–Bailey equation is of the form

$$F(N, P) = bNe^{-aP}$$
$$G(N, P) = cN[1 - e^{-aP}]$$

To find the Jacobi matrix evaluated at the nontrivial equilibrium, we differentiate F and G and then evaluate the derivatives at the equilibrium

$$(N^*, P^*) = \left(\frac{b}{(b-1)}\frac{1}{ac}\ln b, \frac{1}{a}\ln b\right)$$

which we computed in Example 3:

$$\left(\frac{\partial F}{\partial N}\right)^* = be^{-aP}\Big|_{(N^*, P^*)} = 1 \qquad \left(\tfrac{\partial F}{\partial N}\right)^* = \tfrac{\partial F}{\partial N}(N^*, P^*)$$

$$\left(\frac{\partial F}{\partial P}\right)^* = -abNe^{-aP}\Big|_{(N^*, P^*)} = -aN^* \qquad be^{-aP^*} = 1$$

$$\left(\frac{\partial G}{\partial N}\right)^* = c[1 - e^{-aP}]\Big|_{(N^*, P^*)} = c\left[1 - \frac{1}{b}\right] \qquad e^{-aP^*} = \tfrac{1}{b}$$

$$\left(\frac{\partial G}{\partial P}\right)^* = caNe^{-aP}\Big|_{(N^*, P^*)} = acN^*\frac{1}{b}$$

The Jacobi matrix evaluated at (N^*, P^*) is then

$$J(N^*, P^*) = \begin{bmatrix} 1 & -aN^* \\ c\left[1 - \frac{1}{b}\right] & acN^*\frac{1}{b} \end{bmatrix}$$

Instead of computing the eigenvalues explicitly, we will first show that the two eigenvalues of the matrix J are complex conjugate if $b > 1$. (This was the condition we found in Example 4 that guaranteed a biologically reasonable, nontrivial equilibrium.) The eigenvalues of J satisfy the equation $\det(J - \lambda I) = 0$; that is,

$$(1 - \lambda)\left(acN^*\frac{1}{b} - \lambda\right) + acN^*\left(1 - \frac{1}{b}\right) = 0$$

which simplifies to

$$\lambda^2 - \left(1 + \frac{ac}{b}N^*\right)\lambda + acN^* = 0$$

The solutions of this equation are complex conjugate if the discriminant

$$\left(1 + \frac{ac}{b}N^*\right)^2 - 4acN^* < 0$$

With $N^* = \frac{b}{b-1}\frac{1}{ac}\ln b$, the discriminant is

$$f(b) = \left(1 + \frac{\ln b}{b-1}\right)^2 - \frac{4b}{b-1}\ln b$$

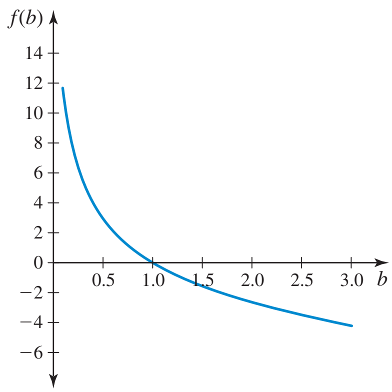


Figure 10.93 The graph of $f(b)$ confirms that $f(b) < 0$ for $b > 1$.

This function depends only on b. Graphing $f(b)$ (see Figure 10.93) shows that $f(b) < 0$ for $b > 1$, thus confirming that the two eigenvalues of J are complex conjugate if $b > 1$.

When we discussed linear systems of recurrence equations, we derived the identity

$$|\lambda_1|^2 = |\lambda_2|^2 = \lambda_1\lambda_2 = \det J$$

The determinant of J is given by

$$\det J = acN^*\frac{1}{b} + acN^*\left(1 - \frac{1}{b}\right) = acN^* = \frac{b\ln b}{b-1}$$

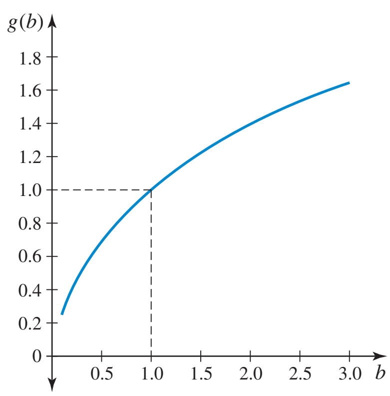


Figure 10.94 The graph of $g(b)$ confirms that $g(b) > 1$ for $b > 1$.

Graphing $g(b) = \frac{b\ln b}{b-1}$ as a function of b (see Figure 10.94), we see that $g(b) > 1$ for $b > 1$, from which we conclude that, for $b > 1$,

$$|\lambda_1|^2 = |\lambda_2|^2 = \lambda_1\lambda_2 > 1$$

implying that the nontrivial equilibrium is unstable. •

Section 10.9 Problems

10.9.1

T ***Problems 1–6 refer to the Nicholson–Bailey host–parasitoid model. Problems 1, 2, 5, and 6 are best done with the help of a spreadsheet, but can also be done with a calculator. Nicholson and Bailey introduced the discrete-generation host–parasitoid model of the form***

$$\boldsymbol{N_{t+1} = bN_t e^{-aP_t}}$$

$$\boldsymbol{P_{t+1} = cN_t[1 - e^{-aP_t}]}$$

for $\boldsymbol{t = 0, 1, 2, \ldots}$.

1. Evaluate the Nicholson–Bailey model for the first 10 generations when $a = 0.02$, $c = 2$, and $b = 2.1$. For the initial host density, choose $N_0 = 5$, and for the initial parasitoid density, choose $P_0 = 0$.

2. Evaluate the Nicholson–Bailey model for the first 10 generations when $a = 0.02$, $c = 3$, and $b = 0.5$. For the initial host density, choose $N_0 = 15$, and for the initial parasitoid density, choose $P_0 = 0$.

3. Show that when the initial parasitoid density is $P_0 = 0$, the Nicholson–Bailey model reduces to

$$N_{t+1} = bN_t$$

With N_0 denoting the initial host density, find an expression for N_t in terms of N_0 and the parameter b.

4. When the initial parasitoid density is $P_0 = 0$, the Nicholson–Bailey model reduces to

$$N_{t+1} = bN_t$$

as shown in the previous problem. For which values of b is the host density increasing if $N_0 > 0$? For which values of b is it decreasing? (Assume that $b > 0$.)

5. Evaluate the Nicholson–Bailey model for the first 15 generations when $a = 0.02$, $c = 3$, and $b = 1.5$. For the initial host density, choose $N_0 = 5$, and for the initial parasitoid density, choose $P_0 = 3$.

6. Evaluate the Nicholson–Bailey model for the first 25 generations when $a = 0.02$, $c = 3$, and $b = 1.5$. For the initial host density, choose $N_0 = 10$, and for the initial parasitoid density, choose $P_0 = 6$.

T ***Problems 7–12 refer to the negative binomial host–parasitoid model. Problems 7, 8, 11, and 12 are best done with the help of a spreadsheet, but can also be done with a calculator. The negative binomial model is a discrete-generation host–parasitoid model of the form***

$$N_{t+1} = bN_t\left(1+\frac{aP_t}{k}\right)^{-k}$$

$$P_{t+1} = cN_t\left[1-\left(1+\frac{aP_t}{k}\right)^{-k}\right]$$

for $t = 0, 1, 2, \ldots$.

7. Evaluate the negative binomial model for the first 10 generations when $a = 0.02$, $c = 3$, $k = 0.75$, and $b = 2.1$. For the initial host density, choose $N_0 = 5$, and for the initial parasitoid density, choose $P_0 = 0$.

8. Evaluate the negative binomial model for the first 10 generations when $a = 0.02$, $c = 3$, $k = 0.75$, and $b = 0.8$. For the initial host density, choose $N_0 = 10$, and for the initial parasitoid density, choose $P_0 = 0$.

9. Show that when the initial parasitoid density is $P_0 = 0$, the negative binomial model reduces to

$$N_{t+1} = bN_t$$

With N_0 denoting the initial host density, find an expression for N_t in terms of N_0 and the parameter b.

10. When the initial parasitoid density is $P_0 = 0$, the negative binomial model reduces to

$$N_{t+1} = bN_t$$

as shown in the previous problem. For which values of b is the host density increasing if $N_0 > 0$? For which values of b is it decreasing? (Assume that $b > 0$.)

11. Evaluate the negative binomial model for the first 25 generations when $a = 0.02$, $c = 3$, $k = 0.75$, and $b = 1.5$. For the initial host density, choose $N_0 = 100$, and for the initial parasitoid density, choose $P_0 = 30$.

12. Evaluate the negative binomial model for the first 25 generations when $a = 0.02$, $c = 3$, $k = 0.75$, and $b = 0.5$. For the initial host density, choose $N_0 = 100$, and for the initial parasitoid density, choose $P_0 = 30$.

T **13.** In the Nicholson–Bailey model, the fraction of hosts escaping parasitism is given by

$$f(P) = e^{-aP}$$

(a) Graph $f(P)$ as a function of P for $a = 0.1$ and $a = 0.01$.

(b) For a given value of P, how are the chances of escaping parasitism affected by increasing a?

T **14.** In the negative binomial model, the fraction of hosts escaping parasitism is given by

$$f(P) = \left(1+\frac{aP}{k}\right)^{-k}$$

(a) Graph $f(P)$ as a function of P for $a = 0.1$ and $a = 0.01$ when $k = 0.75$.

(b) For $k = 0.75$ and a given value of P, how are the chances of escaping parasitism affected by increasing a?

T **15.** In the negative binomial model, the fraction of hosts escaping parasitism is given by

$$f(P) = \left(1+\frac{aP}{k}\right)^{-k}$$

(a) Graph $f(P)$ as a function of P for $k = 0.75$ and $k = 0.5$ when $a = 0.02$.

(b) For $a = 0.02$ and a given value of P, how are the chances of escaping parasitism affected by increasing k?

16. The negative binomial model can be reduced to the Nicholson–Bailey model by letting the parameter k in the negative binomial model go to infinity. Show that

$$\lim_{k\to\infty}\left(1+\frac{aP}{k}\right)^{-k} = e^{-aP}$$

(*Hint*: Use l'Hôpital's rule.)

10.9.2

17. Show that $\begin{bmatrix}0\\0\end{bmatrix}$ is an equilibrium of

$$\begin{bmatrix}x_1(t+1)\\x_2(t+1)\end{bmatrix} = \begin{bmatrix}-0.6 & 0\\-0.3 & 0.3\end{bmatrix}\begin{bmatrix}x_1(t)\\x_2(t)\end{bmatrix}$$

and determine its stability.

18. Show that $\begin{bmatrix}0\\0\end{bmatrix}$ is an equilibrium of

$$\begin{bmatrix}x_1(t+1)\\x_2(t+1)\end{bmatrix} = \begin{bmatrix}-0.4 & 0.3\\0 & -0.8\end{bmatrix}\begin{bmatrix}x_1(t)\\x_2(t)\end{bmatrix}$$

and determine its stability.

19. Show that $\begin{bmatrix}0\\0\end{bmatrix}$ is an equilibrium of

$$\begin{bmatrix}x_1(t+1)\\x_2(t+1)\end{bmatrix} = \begin{bmatrix}-1.6 & 0\\-0.5 & 0.1\end{bmatrix}\begin{bmatrix}x_1(t)\\x_2(t)\end{bmatrix}$$

and determine its stability.

20. Show that $\begin{bmatrix}0\\0\end{bmatrix}$ is an equilibrium of

$$\begin{bmatrix}x_1(t+1)\\x_2(t+1)\end{bmatrix} = \begin{bmatrix}0.1 & 0.3\\0.1 & -1.8\end{bmatrix}\begin{bmatrix}x_1(t)\\x_2(t)\end{bmatrix}$$

and determine its stability.

21. Show that $\begin{bmatrix}0\\0\end{bmatrix}$ is an equilibrium of

$$\begin{bmatrix}x_1(t+1)\\x_2(t+1)\end{bmatrix} = \begin{bmatrix}1 & 2\\1 & 2\end{bmatrix}\begin{bmatrix}x_1(t)\\x_2(t)\end{bmatrix}$$

and determine its stability.

22. Show that $\begin{bmatrix}0\\0\end{bmatrix}$ is an equilibrium of

$$\begin{bmatrix}x_1(t+1)\\x_2(t+1)\end{bmatrix} = \begin{bmatrix}1.5 & 0.2\\0.08 & 0\end{bmatrix}\begin{bmatrix}x_1(t)\\x_2(t)\end{bmatrix}$$

and determine its stability.

23. Show that the equilibrium $\begin{bmatrix}0\\0\end{bmatrix}$ of

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} -0.2 & -0.4 \\ 0.6 & 0.1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

is stable.

24. Show that the equilibrium $\begin{bmatrix}0\\0\end{bmatrix}$ of

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} 0.2 & 0.3 \\ -0.5 & -0.4 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

is stable.

25. Show that the equilibrium $\begin{bmatrix}0\\0\end{bmatrix}$ of

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} 4.2 & -3.4 \\ 2.4 & -1.1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

is unstable.

26. Show that the equilibrium $\begin{bmatrix}0\\0\end{bmatrix}$ of

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} 2 & -4 \\ 5 & -2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

is unstable.

10.9.3

27. Show that the equilibrium $\begin{bmatrix}0\\0\end{bmatrix}$ of

$$x_1(t+1) = \frac{x_2(t)}{4(1+(x_1(t))^2)}$$

$$x_2(t+1) = \frac{3x_1(t)}{1+(x_2(t))^2}$$

is locally stable.

28. Show that the equilibrium $\begin{bmatrix}0\\0\end{bmatrix}$ of

$$x_1(t+1) = \frac{3x_2(t)}{2(1+(x_1(t))^2)}$$

$$x_2(t+1) = \frac{2x_1(t)}{1+(x_2(t))^2}$$

is unstable.

29. Show that the equilibrium $\begin{bmatrix}0\\0\end{bmatrix}$ of

$$x_1(t+1) = x_2(t)$$

$$x_2(t+1) = \frac{3x_2(t) - 2x_1(t)}{3 + x_1(t)}$$

is locally stable.

30. Show that, for any $a > 1$, the equilibrium $\begin{bmatrix}0\\0\end{bmatrix}$ of

$$x_1(t+1) = x_2(t)$$

$$x_2(t+1) = \frac{ax_2(t) - (a-1)x_1(t)}{a + x_1(t)}$$

is locally stable.

31. Show that $\begin{bmatrix}0\\0\end{bmatrix}$ is an equilibrium point of

$$x_1(t+1) = ax_2(t)$$

$$x_2(t+1) = x_1(t) - \cos(x_2(t)) + 1$$

Assume that $a > 0$. For which values of a is $\begin{bmatrix}0\\0\end{bmatrix}$ locally stable?

32. Show that $\begin{bmatrix}0\\0\end{bmatrix}$ and $\begin{bmatrix}\pi\\\pi\end{bmatrix}$ are equilibria of

$$x_1(t+1) = x_2(t)$$

$$x_2(t+1) = \sin(x_2(t)) + x_1(t)$$

and analyze their stability.

33. Find all nonnegative equilibria of

$$x_1(t+1) = x_2(t)$$

$$x_2(t+1) = \frac{x_1(t)}{4} - \frac{x_2(t)}{4} + (x_2(t))^2$$

and analyze their stability.

34. Find all nonnegative equilibria of

$$x_1(t+1) = x_2(t)$$

$$x_2(t+1) = \frac{1}{2}x_1(t) + \frac{1}{3}x_2(t) - (x_2(t))^2$$

and analyze their stability.

35. For which values of a is the equilibrium $\begin{bmatrix}0\\0\end{bmatrix}$ of

$$x_1(t+1) = \frac{ax_2(t)}{1+(x_1(t))^2}$$

$$x_2(t+1) = \frac{2x_1(t)}{1+(x_2(t))^2}$$

locally stable?

36. For which values of a is the equilibrium $\begin{bmatrix}0\\0\end{bmatrix}$ of

$$x_1(t+1) = x_2(t)$$

$$x_2(t+1) = \frac{-1}{2}x_1(t) + ax_2(t) - (x_2(t))^2$$

locally stable?

37. Denote by $x_1(t)$ the number of juveniles, and by $x_2(t)$ the number of adults, at time t. Assume that $x_1(t)$ and $x_2(t)$ evolve according to

$$x_1(t+1) = x_2(t)$$

$$x_2(t+1) = \frac{1}{2}x_1(t) + rx_2(t) - (x_2(t))^2$$

(a) Show that if $r > 1/2$, there exists an equilibrium $\begin{bmatrix}x_1^*\\x_2^*\end{bmatrix}$ with $x_1^* > 0$ and $x_2^* > 0$. Find x_1^* and x_2^*.

(b) Determine the stability of the equilibrium found in (a) when $r > 1/2$.

38. Host-Parasitoid Interactions Find all biologically relevant equilibria of the Nicholson–Bailey model

$$N_{t+1} = 2N_t e^{-0.2P_t}$$

$$P_{t+1} = N_t\left[1 - e^{-0.2P_t}\right]$$

and analyze their stability.

39. Host-Parasitoid Interactions Find all biologically relevant equilibria of the Nicholson–Bailey model

$$N_{t+1} = 4N_t e^{-0.1P_t}$$

$$P_{t+1} = N_t\left[1 - e^{-0.1P_t}\right]$$

and analyze their stability.

40. **Host-Parasitoid Interactions** Find all biologically relevant equilibria of the negative binomial host–parasitoid model

$$N_{t+1} = 4N_t \left(1 + \frac{0.01P_t}{2}\right)^{-2}$$

$$P_{t+1} = N_t \left[1 - \left(1 + \frac{0.01P_t}{2}\right)^{-2}\right]$$

and analyze their stability.

41. **Host-Parasitoid Interactions** Find all biologically relevant equilibria of the negative binomial host–parasitoid model

$$N_{t+1} = 4N_t \left(1 + \frac{0.01P_t}{0.5}\right)^{-0.5}$$

$$P_{t+1} = N_t \left[1 - \left(1 + \frac{0.01P_t}{0.5}\right)^{-0.5}\right]$$

and analyze their stability.

Chapter 10 Review

Key Terms

Discuss the following definitions and concepts:

1. Real-valued function
2. Function of two variables
3. Surface
4. Heat map
5. Level curve
6. Limit
7. Limit laws
8. Continuity
9. Partial derivative
10. Geometric interpretation of a partial derivative
11. Mixed-derivative theorem
12. Tangent plane
13. Differentiability
14. Differentiability and continuity
15. Sufficient condition for differentiability
16. Standard linear approximation, tangent plane approximation
17. Vector-valued function
18. Jacobi matrix, derivative matrix
19. Chain rule
20. Implicit differentiation
21. Directional derivative
22. Gradient
23. Local extrema
24. Sufficient condition for finding local extrema
25. Hessian matrix
26. Global extrema
27. The Extreme-Value Theorem
28. Critical points
29. Lagrange multiplier
30. Sum of squares error
31. Least squares error fitting method
32. Diffusion
33. Systems of recurrence equations
34. Point equilibria and their stability
35. Nicholson–Bailey equation

Review Problems

For each of the functions given in Problem 1–8, calculate the specified partial derivative.

1. $f(x, y) = xy + y^2 + 2; \quad \dfrac{\partial f}{\partial y}$

2. $g(x, y) = \dfrac{xy}{x + y}; \quad \dfrac{\partial f}{\partial x}$

3. $f(x, y) = \sin(x + y); \quad \dfrac{\partial^2 f}{\partial x^2}$

4. $f(x, y) = \exp(-x^2 - y^2); \quad \dfrac{\partial f}{\partial y}$

5. $h(r, s) = r + se^{-rs}; \quad \dfrac{\partial h}{\partial r}$

6. $g(x, y) = \dfrac{x + y^2}{x^2 + y}; \quad \dfrac{\partial g}{\partial x}$

7. $h(x, y) = (x + y^2) - (x^2 + y); \quad \dfrac{\partial^2 h}{\partial x\, \partial y}$

8. $f(r, s) = (r + s)\cos(s + r); \quad \dfrac{\partial^2 f}{\partial r^2}$

9. **Germination** Suppose that you conduct an experiment to measure the germination success of seeds of a certain plant as a function of temperature and humidity. You find that seeds don't germinate at all when the humidity is too low, regardless of temperature. Germination success is highest for intermediate values of temperature. At all temperatures seeds tend to germinate better when you increase humidity levels. Use the preceding information to sketch a graph of germination success as a function of temperature for different levels of humidity. Also, sketch the graph of germination success as a function of humidity for different temperature values.

10. **Plant Physiology** ***Gaastra (1959)*** measured the effects of atmospheric CO_2 levels on CO_2 fixation in sugar beet leaves at various light levels. He found that increasing CO_2 at fixed light levels increases the fixation rate and that increasing light levels at fixed atmospheric CO_2 concentration also increased fixation. If $F(A, I)$ denotes the fixation rate as a function of atmospheric CO_2 concentration (A) and light intensity (I), determine the signs of $\partial F/\partial A$ and $\partial F/\partial I$.

11. **Plant Ecology** In ***Burke and Grime (1996)***, a long-term field experiment in a limestone grassland was described.

(a) One of the experiments related total area covered by *indigenous* species to fertility and disturbance gradients. The experiment was designed so that the two variables (fertility and disturbance) could be altered independently. Burke and Grime found that the area covered by indigenous species generally increased with the amount of fertilizer added and decreased with the intensity of a disturbance. If $A_i(F, D)$ denotes the area covered by indigenous species as a function of the amount of fertilizer added (F) and the intensity of disturbance (D), determine the signs of $\partial A_i/\partial F$ and $\partial A_i/\partial D$ for Burke and Grime's experiment.

(b) In another experiment, Burke and Grime related the total area covered by *introduced* species to fertility and disturbance gradients. Let $A_e(F, D)$ denote the area covered by introduced species as a function of the amount of fertilizer added (F) and the intensity of disturbance (D). Burke and Grime found that

$$\frac{\partial A_e}{\partial F} > 0$$

and

$$\frac{\partial A_e}{\partial D} > 0$$

Explain in words what this means.

12. Water Content You measure the amount of water stored in soil in different locations within a plot of land. Your measurements can be visualized using a contour plot (see Figure 10.95) showing water content, w, as a function of x and y within the plot.

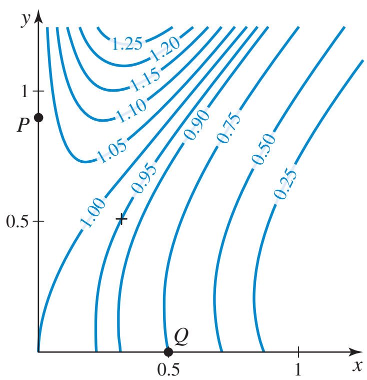


Figure 10.95 Soil water content in Problem 12. Note that level curves do not come from evenly spaced values of c.

(a) Make a sketch showing how water content varies with x on an x-transect that starts at the point P; that is, sketch the function $w(x, 0.9)$.

(b) Then make a sketch showing how water content varies with y on a y-transect that starts at the point Q; that is, sketch the function $w(0.5, y)$.

(c) A particular type of plant can only grow in soil in which $w \geq 0.9$. Make a sketch showing the regions in this plot where the plant could be expected to grow.

(d) Within soil, water often flows in the direction in which the water content of soil decreases most rapidly. Assume that water is introduced into the soil at the point marked by a +. Draw the direction in which we expect the water to flow at this point.

13. Find the Jacobi matrix for the function:

$$\mathbf{f}(x, y) = \begin{bmatrix} x^2 - y \\ x^3 - y^2 \end{bmatrix}.$$

14. Find a linear approximation to

$$\mathbf{f}(x, y) = \begin{bmatrix} 2xy^2 \\ \frac{x}{y} \end{bmatrix}$$

at $(1, 1)$.

15. Mark–Recapture Experiment We can compute the average distance traveled by foraging animals at time t, denoted by r_{avg}. We find that

$$r_{\text{avg}} = \sqrt{\pi D t} \tag{10.64}$$

(a) Graph r_{avg} as a function of D for $t = 0.1$, $t = 1$, and $t = 5$. Describe in words how an increase in D affects the average radius the animals travel.

(b) Show that

$$D = \frac{(r_{\text{avg}})^2}{\pi t} \tag{10.65}$$

(c) Equation (10.65) can be used to measure D, the diffusion constant, from field data of mark–recapture experiments, taken from ***Kareiva (1983),*** as follows: Marked animals are released from the release site and then recaptured after a certain amount of time t from the time of release. The distance of the recaptured animals from the release site is measured.

If N denotes the total number of recaptured animals, d_i denotes the distance of the ith recaptured animal from the release site, and t is the time between release and recapture, use (10.65) to explain why

$$D = \frac{1}{\pi t}\left(\frac{1}{N}\sum_{i=1}^{N} d_i\right)^2$$

can be used to measure D from field data. (Note that the time between release and recapture is the same for each individual in this study.)

In Problems 16–19 find the linear approximation for each function $f(x, y)$ at the specified point.

16. $f(x, y) = x^2 + xy - 2x$ at $(1, -1)$

17. $f(x, y) = \dfrac{xy}{1 + x^2}$ at $(0, 1)$

18. $f(x, y) = \exp(-x^2 - y^2)$ at $(0, 0)$

19. $f(x, y) = (x^2 + y)(x - y)$ at $(1, -1)$

20. Probability Model Fitting One of the most common uses of optimization methods is to calculate parameters that can be used to fit probability models to real, measured data. You are studying a wild population of California condors. You want to estimate in each year what fraction of condors die (p_1), what fraction migrate out of the study region (p_2), and what fraction do not die and also remain in the study region (p_3).

(a) Explain why we expect $p_1 + p_2 + p_3 = 1$.

(b) Suppose that out of a total of N birds the following numbers of birds die, migrate away, or neither.

No dying	No migrating away	Neither
2	3	5

We can show (see Chapter 12) that the likelihood of seeing these data from a probability model is:

$$L(p_1, p_2, p_3) = \frac{2!3!5!}{10!} p_1^2 p_2^3 p_3^5$$

Or using the result from part (a) to eliminate p_3:

$$L(p_1, p_2) = \frac{2!3!5!}{10!} p_1^2 p_2^3 (1 - p_1 - p_2)^5$$

What is the domain of the function L? [*Hint*: All probabilities must be between 0 and 1.]

(c) We want to find values of p_1, p_2 that maximize the likelihood L. However, it is a little easier if, instead of maximizing L, we maximize the function $f(p_1, p_2) = \ln(L(p_1, p_2))$ (since $\ln L$

increases as L increases, maximizing f is equivalent to maximizing L).

Now

$$f(p_1, p_2) = \ln L(p_1, p_2)$$
$$= \ln\left[\frac{2!3!5!}{10!}\right] + 2\ln p_1 + 3\ln p_2 + 5\ln(1 - p_1 - p_2)$$

Explain why the global maximum of ln L has to occur in the interior of the domain that you found in part (b).

(d) Show that there is a critical point with

$$p_1 = \frac{1}{5}, p_2 = \frac{3}{10}$$

Are there any other critical points in the interior of the domain?

(e) By calculating Hessian $f\left(\frac{1}{5}, \frac{3}{10}\right)$ show that this critical point is a local maximum.

Since $p_1 = \frac{1}{5}$ and $p_2 = \frac{3}{10}$ (and $p_3 = 1 - p_1 - p_2 = \frac{1}{2}$) maximize the likelihood of the model producing the measured data, we call these values the *maximum likelihood estimates* for p_1, p_2, and p_3.

CHAPTER

11

Systems of Differential Equations

This chapter develops the theory of systems of differential equations. Specifically, we will learn how to

- solve systems of linear differential equations;
- analyze and classify point equilibria of linear and nonlinear systems of differential equations; and
- employ systems of differential equations to model phenomena and processes in biology.

In Chapters 5 and 8 we learned how to build and analyze mathematical models for biological phenomena. These models took the form of differential equations. For example, if there are N organisms in a particular population at time t then our model for population growth may be a differential equation of the form:

$$\frac{dN}{dt} = g(t, N),$$

for population growth. The function $g(t, N)$ represents the growth rate and depends both on time t and on the size of the population because the rate of births depends on the number of organisms present.

However, in models of ecosystems, it can be hard to isolate different types of organisms. For example consider an ecosystem made up of two species: an herbivore, and the plant that it feeds on. Suppose that at time t there are $h(t)$ herbivores and $p(t)$ plants. Then for the reasons discussed above we expect that the rate of growth of the herbivore population will depend on h. But the amount of resources that herbivores can devote to reproduction will depend on how many plants they eat. So the growth rate will also depend on the number of plants, p. That is, the growth rate of the herbivore population will in general be a function of both h and p, as well as potentially depending on time also, so:

$$\frac{dh}{dt} = f(t, h, p) \tag{11.1}$$

for some function $f(t, h, p)$.

We cannot solve this differential equation unless we know $p(t)$. But the growth or decrease in the plant population is also impacted by the number of herbivores. If the herbivore population is large, then high rates of herbivory will cause the plant population to decline. Conversely, if the herbivore population is low, then the plant population will grow faster. So the rate of growth of the plant population depends on the number of herbivores. The growth of the plant population will therefore be given by a differential equation:

$$\frac{dp}{dt} = g(t, h, p) \tag{11.2}$$

for some function $g(t, h, p)$. Equations (11.1) and (11.2) together constitute a system of differential equations. Because one equation cannot be solved independently of the other equation, the techniques that we learned in Chapter 8 for solving differential equations cannot be used here. Indeed in most cases it is not possible to solve the differential equations analytically, i.e., we cannot derive explicit functions $h(t)$ and $p(t)$ that solve equations (11.1) and (11.2). Graphical methods, however, will enable us to analyze the differential equations without solving them explicitly, just as they did in Section 8.2.

We start by giving a more formal definition of what a system of differential equations is. Suppose that we are given a set of variables $x_1, x_2, \ldots, x_n$, each depending on an independent variable, say, t, so that $x_1 = x_1(t), x_2 = x_2(t), \ldots, x_n = x_n(t)$. $x_1, x_2, \ldots, x_n$ could, for example, represent the number of organisms present in an ecosystem from each of n different species. Suppose also that the dynamics of the variables are linked by differential equations of the form

$$\begin{aligned}\frac{dx_1}{dt} &= g_1(t, x_1, x_2, \ldots, x_n)\\ \frac{dx_2}{dt} &= g_2(t, x_1, x_2, \ldots, x_n)\\ &\vdots\\ \frac{dx_n}{dt} &= g_n(t, x_1, x_2, \ldots, x_n)\end{aligned} \tag{11.3}$$

This set of equations is called a **system of differential equations**. On the left-hand side are the derivatives of $x_i(t)$ with respect to t. On the right-hand side of each equation is a function g_i that depends on the variables $x_1, x_2, \ldots, x_n$ and on t. We will first look at the case when the functions g_i are linear in the variables $x_1, x_2, \ldots, x_n$ where linear is meant in the same sense as it was used in Chapter 9, i.e., for $i = 1, 2, \ldots, n$,

$$g_i(t, x_1, x_2, \ldots, x_n) = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$$

for some set of constants $a_{i1}, a_{i2}, \ldots, a_{in}$

We can write the linear system in matrix form as

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \tag{11.4}$$

where

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t)\\ x_2(t)\\ \vdots\\ x_n(t)\end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ a_{n1} & a_{n2} & \ldots & a_{nn}\end{bmatrix}.$$

Equation (11.4) is an example of a **system of linear first-order differential equations** (first order because only first derivatives occur). Equation (11.4) is called a **homogeneous linear first-order system with constant coefficients**. Another type of differential, called inhomogeneous, has an additional function $\mathbf{f}(t)$ on the right-hand side of (11.4). Systems of linear differential equations also can have coefficients, a_{ij}, that depend on t. Both of these scenarios are outside the scope of this book. Since none of the functions, g_i, depends on t, the system is **autonomous**. (We encountered autonomous systems in Section 8.1.) Note that A is a square matrix.

In Section 11.1, we will present some of the theory for systems of the form (11.4). In Section 11.2, we will discuss some applications of linear systems. Section 11.3 is devoted to the theory of nonlinear systems, Section 11.4 to applications of nonlinear systems.

11.1 Linear Systems: Theory

In this section, we will analyze homogeneous, linear first-order systems with constant coefficients—that is, systems of the form

$$\begin{aligned}
\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
\frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
&\vdots \\
\frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n
\end{aligned} \tag{11.5}$$

where the variables $x_1, x_2, \ldots, x_n$ are functions of t and the parameters a_{ij}, $1 \le i, j \le n$, are constants. We can write this system in matrix form as

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t)$$

where

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} \qquad A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

Most of the time we will restrict ourselves to the case $n = 2$.

EXAMPLE 1 Write

$$\begin{aligned}
\frac{dx_1}{dt} &= 4x_1 - 2x_2 \\
\frac{dx_2}{dt} &= -3x_1 + x_2
\end{aligned}$$

in matrix notation.

Solution We write $x_1 = x_1(t)$ and $x_2 = x_2(t)$. Using the rules for matrix multiplication, we find that

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 4x_1 - 2x_2 \\ -3x_1 + x_2 \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

That is,

$$\frac{d\mathbf{x}(t)}{dt} = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} \mathbf{x}(t)$$

First, we will be concerned with solutions of (11.5): A **solution** is an ordered n-tuple of functions $(x_1(t), x_2(t), \ldots, x_n(t))$ that satisfies all n equations in (11.5). Second, we will discuss equilibria: An **equilibrium** is a point $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)$ such that $A\hat{\mathbf{x}} = \mathbf{0}$. We begin with a graphical approach to visualizing solutions.

11.1.1 The Vector Field

We sketch solutions graphically in the x_1–x_2 plane with the help of **vector fields**. Consider

$$\begin{aligned}
\frac{dx_1}{dt} &= x_1 - 2x_2 \\
\frac{dx_2}{dt} &= x_2
\end{aligned} \tag{11.6}$$

Imagine now that you are standing at a point (x_1, x_2) in the x_1–x_2 plane and the system (11.6) determines your future location. Where should you go next? The two differential equations tell you how your coordinates will change. To give a concrete example, look at the point $(2, -1)$. At this point the x_1 coordinate changes at a rate

$$\frac{dx_1}{dt} = x_1 - 2x_2 = 2 - 2(-1) = 4,$$

and the x_2 coordinate changes at a rate

$$\frac{dx_2}{dt} = x_2 = -1.$$

We claim that you move along a curve whose tangent line at the point $(2, -1)$ has slope

$$\frac{dx_2}{dx_1} = \frac{dx_2/dt}{dx_1/dt} = -\frac{1}{4}$$

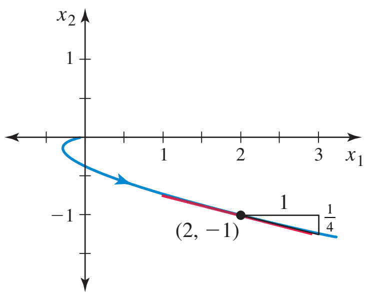


Figure 11.1 Solution curve through $(2, -1)$ with tangent line. The arrow on the solution curve shows the direction of the solution as t increases.

A solution of (11.6) that starts at time 0 at the point $(x_1(0), x_2(0))$ is given by points of the form $(x_1(t), x_2(t))$, $t \geq 0$, that satisfy (11.6), i.e., by a curve in the x_1–x_2 plane. An example of a solution curve is shown in Figure 11.1. At each point on the solution curve, we can draw a tangent line whose slope is $\frac{dx_2}{dx_1}$. At each point we can calculate $\frac{dx_2}{dx_1}$ from $\frac{dx_2}{dt} / \frac{dx_1}{dt}$. The tangent line at $(2, -1)$, for which we computed the slope (namely, $-1/4$), is also drawn in Figure 11.1.

We can draw tangent lines at each point (x_1, x_2) in the x_1–x_2 plane. Knowing all the tangent lines then allows us to sketch the solution curve passing through any point in the $x_1 - x_2$ plane. This is done by assigning each point (x_1, x_2) in the x_1–x_2 plane a vector $\begin{bmatrix} dx_1/dt \\ dx_2/dt \end{bmatrix}$, which has the property that it is tangential to the solution curve that passes through the point (x_1, x_2) and it points in the direction of the solution.

In our example, at a point (x_1, x_2) the vector is of the form $\begin{bmatrix} x_1 - 2x_2 \\ x_2 \end{bmatrix}$ and the slope of the solution curve that goes through (x_1, x_2) is

$$\frac{dx_2}{dx_1} = \frac{dx_2/dt}{dx_1/dt} = \frac{x_2}{x_1 - 2x_2}$$

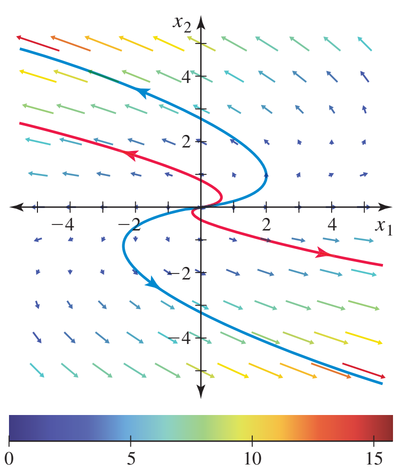


Figure 11.2 The vector field of the system (11.6) together with some solution curves.

The collection of these vectors is called a **vector field** (it is also called a **slope field** or a **direction field**). Each vector of the vector field is called a **direction vector**. Since (11.6) is an autonomous system, the direction vector depends only on the location of the point (x_1, x_2) and not on t. This property implies that the vector field is the same for all times t.

The vector field for (11.6) is shown in Figure 11.2. [The figure also contains four solution curves that were computer generated; each curve starts at a different point very close to the origin (0, 0).] The length of the direction vector at a given point tells us how quickly the solution curve passes through the point; the length is proportional to $\sqrt{\left(\frac{dx_1}{dt}\right)^2 + \left(\frac{dx_2}{dt}\right)^2}$. In Figure 11.2, as a way of emphasizing how $\sqrt{\left(\frac{dx_1}{dt}\right)^2 + \left(\frac{dx_2}{dt}\right)^2}$ varies between different points (x_1, x_2), we color code each arrow by the value of $\sqrt{\left(\frac{dx_1}{dt}\right)^2 + \left(\frac{dx_2}{dt}\right)^2}$. It is not in general necessary to color code vector fields, and we will not do so in future examples. If we are interested only in the direction of the solution curves, we can indicate the direction by small line segments (as shown in Figure 11.3), which often results in a less cluttered picture. Like Figure 11.2, Figure 11.3 contains four solution curves that were computer generated; note that the direction vectors are always tangent to the curves. We can therefore use the vector field to sketch solution curves by drawing curves in such a way that the direction vectors are always tangent to the curve.

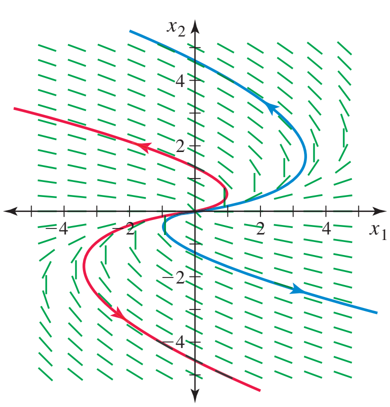


Figure 11.3 The vector field of the system (11.6) with some solution curves, where the direction vectors are not drawn to scale.

The point (0, 0) is special: When we compute the direction vector at (0, 0), we find that

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

That is, if we start at this point, neither $x_1(t)$ nor $x_2(t)$ will change. We call such points **equilibria**. We will discuss their significance in Subsection 11.1.3.

11.1.2 Solving Linear Systems

Specific Solutions. Consider the following differential equation:

$$\frac{dx}{dt} = ax$$

This is a linear, first-order differential equation with a constant coefficient. [That is, the equation is of the form (11.5) with $n = 1$]. We can find a solution by integrating after separating variables (as we learned in Chapter 8). All solutions are of the form

$$x(t) = ce^{at}$$

where c is a constant that depends on the initial condition. An initial condition picks out a specific solution among the set of solutions. For instance, if $x(0) = x_0$, then $c = x_0$.

We will now show that exponential functions are also solutions of *systems* of linear differential equations. We restrict our discussion to systems of two differential equations. Consider the system

$$\frac{dx_1}{dt} = a_{11}x_1(t) + a_{12}x_2(t) \tag{11.7}$$

$$\frac{dx_2}{dt} = a_{21}x_1(t) + a_{22}x_2(t) \tag{11.8}$$

which, in matrix form, is written as

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \tag{11.9}$$

A solution of (11.9) is a vector-valued function. As in the example presented at the beginning of this subsection (i.e., $dx/dt = ax$), we will find that (11.9) admits a collection of solutions, and if we choose an initial condition, a particular solution will be picked out. We will get to initial conditions later; let's first see what the solutions look like.

We claim that the vector-valued function

$$\mathbf{x}(t) = \begin{bmatrix} u_1 e^{\lambda t} \\ u_2 e^{\lambda t} \end{bmatrix} = e^{\lambda t} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{11.10}$$

where λ, u_1, and u_2 are constants, is a solution of (11.9) for an appropriate choice of values for λ, u_1, and u_2. To see how we must choose these values, we differentiate $\mathbf{x}(t)$ in (11.10):

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} u_1 \lambda e^{\lambda t} \\ u_2 \lambda e^{\lambda t} \end{bmatrix} = \lambda e^{\lambda t} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{11.11}$$

Using (11.11) for the left-hand side of (11.9) and (11.10) for the right-hand side, we find that $x(t)$ solves (11.9) if:

$$\underbrace{\lambda e^{\lambda t} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}}_{\frac{d\mathbf{x}(t)}{dt}} = \underbrace{A e^{\lambda t} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}}_{A\mathbf{x}(t)}$$

or, after dividing both sides by $e^{\lambda t}$,

$$\lambda \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = A \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

This last expression should remind you of eigenvalues and eigenvectors that we encountered in Section 9.3: We have shown that $x(t)$ solves (11.9) if $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is an eigenvector corresponding to an eigenvalue λ of A.

In this subsection, we will look only at differential equations of the form (11.9) for which the eigenvalues of A are both real and distinct. We will discuss complex eigenvalues in the next subsection. (We will not discuss the case when both eigenvalues are identical.) We use the following system to illustrate how to find specific solutions of a system of the form (11.9):

$$\begin{aligned}\frac{dx_1}{dt} &= 2x_1 - 2x_2 \\ \frac{dx_2}{dt} &= 2x_1 - 3x_2\end{aligned} \tag{11.12}$$

1. *Finding eigenvalues* The coefficient matrix of (11.12) is given by

$$A = \begin{bmatrix} 2 & -2 \\ 2 & -3 \end{bmatrix}$$

We first determine the eigenvalues and corresponding eigenvectors of A. To find the eigenvalues of A, we must solve

$$\det(A - \lambda I) = 0$$

That is,

$$\begin{aligned}\det\begin{bmatrix} 2-\lambda & -2 \\ 2 & -3-\lambda \end{bmatrix} &= (2-\lambda)(-3-\lambda) + 4 \\ &= \lambda^2 + \lambda - 2 = (\lambda - 1)(\lambda + 2) = 0\end{aligned}$$

which has solutions

$$\lambda_1 = 1 \qquad \text{and} \qquad \lambda_2 = -2.$$

2. *A solution corresponding to the eigenvalue λ_1* To find an eigenvector $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ corresponding to $\lambda_1 = 1$, we solve

$$A\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \lambda_1 \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad \text{with } \lambda_1 = 1$$

We find

$$\begin{bmatrix} 2u_1 - 2u_2 \\ 2u_1 - 3u_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

which can be written as

$$\begin{aligned} u_1 - 2u_2 &= 0 \\ 2u_1 - 4u_2 &= 0 \end{aligned}$$

Both equations reduce to the same equation, namely,

$$u_1 = 2u_2$$

Setting $u_2 = 1$, for instance, we obtain $u_1 = 2$. An eigenvector corresponding to $\lambda_1 = 1$ is then

$$\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(Recall that any nonzero multiple of $\mathbf{u}$ will also be an eigenvector with eigenvalue $\lambda_1 = 1$.)

We claim that

$$\mathbf{x}(t) = e^t \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

solves (11.12) Let's check by substituting for $x(t)$ in the left- and right-hand sides of (11.12). On the left-hand side we find that

$$\frac{d\mathbf{x}}{dt} = e^t \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \mathbf{x}(t)$$

While on the right-hand side:

$$A\mathbf{x}(t) = Ae^t \begin{bmatrix} 2 \\ 1 \end{bmatrix} = e^t A \begin{bmatrix} 2 \\ 1 \end{bmatrix} = e^t \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \mathbf{x}(t) \qquad A\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Therefore,

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \qquad \text{Both sides evaluate to } x(t)$$

and we conclude that $\mathbf{x}(t) = e^t\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is indeed a solution of (11.12).

3. ***A solution corresponding to the eigenvalue*** $\boldsymbol{\lambda_2}$ We can now repeat the same steps for the eigenvalue $\lambda_2 = -2$. To find an eigenvector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ corresponding to the eigenvalue $\lambda_2 = -2$, we solve

$$\begin{bmatrix} 2v_1 - 2v_2 \\ 2v_1 - 3v_2 \end{bmatrix} = -2\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

which can be written as

$$4v_1 - 2v_2 = 0$$

$$2v_1 - v_2 = 0.$$

The two equations reduce to the same equation, namely,

$$2v_1 = v_2.$$

Setting $v_1 = 1$, we find that $v_2 = 2$. An eigenvector corresponding to $\lambda_2 = -2$ is then

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

We set

$$\mathbf{x}(t) = e^{-2t}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

and check that

$$\frac{d\mathbf{x}}{dt} = -2e^{-2t}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = -2\mathbf{x}(t)$$

and

$$A\mathbf{x}(t) = Ae^{-2t}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = e^{-2t}A\begin{bmatrix} 1 \\ 2 \end{bmatrix} = -2e^{-2t}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = -2\mathbf{x}(t). \qquad A\begin{bmatrix} 1 \\ 2 \end{bmatrix} = -2\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

This shows that $\mathbf{x}(t) = e^{-2t}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is also a solution of (11.12).

4. ***The vector field*** Consider the two particular solutions we just found. We wish to illustrate them in the corresponding vector field. The eigenvectors of A will turn out to correspond to special lines on the vector field plot. Note that any multiple of $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector of A. So there is an infinite line of points (given by the equation $x_1 - 2x_2 = 0$) whose position vectors are eigenvectors of A, with eigenvalue 1. For any of these points, $\frac{d\mathbf{x}}{dt} = \mathbf{x}$ so the slope of the vector field at any of these points is:

$$\frac{dx_2}{dx_1} = \frac{\frac{dx_2}{dt}}{\frac{dx_1}{dt}} = \frac{x_2}{x_1}$$

which is the same as the slope of the line of eigenvectors. In fact, for any system $\frac{d\mathbf{x}}{dt} = A\mathbf{x}$, if (x_1, x_2) lies on a line of eigenvectors with eigenvalue λ, then:

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x} = \lambda\mathbf{x}$$

$$\text{implies that} \quad \frac{dx_2}{dx_1} = \frac{dx_2/dt}{dx_1/dt} = \frac{\lambda x_2}{\lambda x_1} = \frac{x_2}{x_1}.$$

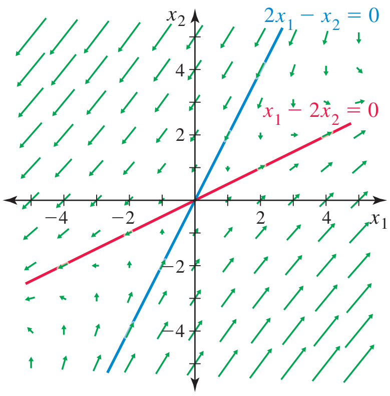


Figure 11.4 The vector field of the system (11.12) with the lines in the direction of the eigenvectors.

That is, the vector field has the same slope as the line of eigenvectors. In other words, the direction vectors point along the line of eigenvectors. In Figure 11.4, we show the vector field together with the two lines in the direction of the eigenvectors.

If $\mathbf{x}(0)$ is a point on one of the lines defined by the eigenvectors, then the solution $\mathbf{x}(t)$ will remain on that line at *all* later times. The location of $\mathbf{x}(t)$ on the line will change with time but $\mathbf{x}(t)$ cannot leave the line of eigenvectors. If the corresponding eigenvalue is positive, the solution will move away from the origin; if the eigenvalue is negative, it will move toward the origin, as can be seen from the direction of the direction vectors. The solid line in Figure 11.4 ($2x_1 - x_2 = 0$) corresponds to the eigenvalue $\lambda_2 = -2$, and we see that the direction vectors on this line point toward the origin. The dashed line in Figure 11.4 ($x_1 - 2x_2 = 0$) corresponds to the eigenvalue $\lambda_2 = 1$, and we see that the direction vectors on this line point away from the origin.

The General Solution. Suppose that $\mathbf{y}(t)$ and $\mathbf{z}(t)$ are two solutions of

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \tag{11.13}$$

That is,

$$\frac{d\mathbf{y}}{dt} = A\mathbf{y}(t) \qquad \text{and} \qquad \frac{d\mathbf{z}}{dt} = A\mathbf{z}(t)$$

Then the linear combination

$$c_1\mathbf{y}(t) + c_2\mathbf{z}(t) \qquad \text{c_1 and c_2 are any constants.} \tag{11.14}$$

also solves (11.13). This can be seen as follows: First, note that

$$\frac{d}{dt}\left[c_1\mathbf{y}(t) + c_2\mathbf{z}(t)\right] = c_1\frac{d\mathbf{y}}{dt} + c_2\frac{d\mathbf{z}}{dt}$$

But since $\mathbf{y}(t)$ and $\mathbf{z}(t)$ are both solutions of (11.13), it follows that

$$c_1\frac{d\mathbf{y}}{dt} + c_2\frac{d\mathbf{z}}{dt} = c_1A\mathbf{y}(t) + c_2A\mathbf{z}(t) \tag{11.15}$$

We can write the right-hand side of (11.15) as:

$$A\left[c_1\mathbf{y}(t) + c_2\mathbf{z}(t)\right] \qquad \text{Use properties of matrix multiplication}$$

Summarizing, we find that

$$\frac{d}{dt}[\underbrace{c_1\mathbf{y}(t) + c_2\mathbf{z}(t)}_{\mathbf{x}(t)}] = A[\underbrace{c_1\mathbf{y}(t) + c_2\mathbf{z}(t)}_{\mathbf{x}(t)}]$$

which shows that the linear combination $c_1\mathbf{y}(t) + c_2\mathbf{z}(t)$ is also a solution.

Combining solutions, as in (11.14), illustrates the important **superposition principle**:

Superposition Principle Suppose that

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{11.16}$$

If

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} \quad \text{and} \quad \mathbf{z}(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}$$

are solutions of (11.16), then

$$\mathbf{x}(t) = c_1\mathbf{y}(t) + c_2\mathbf{z}(t)$$

is also a solution of (11.16).

We just saw that if $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is an eigenvector corresponding to the real eigenvalue λ of A, then $e^{\lambda t}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ solves (11.13). If we have two real and distinct eigenvalues λ_1 and λ_2 with corresponding eigenvectors $\mathbf{u}$ and $\mathbf{v}$, then, setting $\mathbf{y}(t) = e^{\lambda_1 t}\mathbf{u}$ and $\mathbf{z}(t) = e^{\lambda_2 t}\mathbf{v}$ and using the superposition principle, we find that the linear combination

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{u} + c_2 e^{\lambda_2 t}\mathbf{v} \tag{11.17}$$

where c_1 and c_2 are constants, is also a solution of (11.13). The constants c_1 and c_2 depend on the initial conditions. We can show that every solution of (11.4) can be written in the form (11.17); we therefore call a solution of the form (11.17) the **general solution**. (The situation is more complicated when A has repeated eigenvalues—that is, when $\lambda_1 = \lambda_2$. We do not give the general solution for that case here, but will discuss two such examples in Problems 27 and 28.)

The General Solution Let

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \tag{11.18}$$

where A is a 2×2 matrix with two real and distinct eigenvalues λ_1 and λ_2 with corresponding eigenvectors $\mathbf{u}$ and $\mathbf{v}$. Then

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{u} + c_2 e^{\lambda_2 t}\mathbf{v} \tag{11.19}$$

is the general solution of (11.18). The constants c_1 and c_2 depend on the initial condition.

We will now check that (11.19) is indeed a solution of (11.18). To do so, we differentiate (11.19) with respect to t:

$$\frac{d\mathbf{x}}{dt} = \lambda_1 c_1 e^{\lambda_1 t}\mathbf{u} + \lambda_2 c_2 e^{\lambda_2 t}\mathbf{v} \qquad \text{Left hand side of (11.18)}$$

Since $\mathbf{u}$ and $\mathbf{v}$ are eigenvectors corresponding to λ_1 and λ_2, it follows that

$$\begin{aligned} A\mathbf{x}(t) &= A(c_1 e^{\lambda_1 t}\mathbf{u} + c_2 e^{\lambda_2 t}\mathbf{v}) \qquad \text{Right hand side of (11.18)} \\ &= c_1 e^{\lambda_1 t} A\mathbf{u} + c_2 e^{\lambda_2 t} A\mathbf{v} \\ &= c_1 e^{\lambda_1 t}\lambda_1 \mathbf{u} + c_2 e^{\lambda_2 t}\lambda_2 \mathbf{v} \end{aligned}$$

Since the left-hand and right-hand sides evaluate to the same vector function,

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t).$$

Going back to the example at the beginning of this subsection, we find that the general solution of (11.12), or

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} 2 & -2 \\ 2 & -3 \end{bmatrix}\mathbf{x}(t)$$

is

$$\mathbf{x}(t) = c_1 e^{t}\begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 e^{-2t}\begin{bmatrix} 1 \\ 2 \end{bmatrix} \tag{11.20}$$

5. *An initial condition for (11.12)* Suppose we know that, at time 0,

$$\mathbf{x}(0) = \begin{bmatrix} -1 \\ 4 \end{bmatrix} \tag{11.21}$$

holds for (11.12). Then we can determine the constants c_1 and c_2 in (11.20) so that $\mathbf{x}(t)$ satisfies the initial condition (11.21), i.e., so that

$$\mathbf{x}(0) = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$$

To find c_1 and c_2, we must solve the system of linear equations

$$\begin{aligned} 2c_1 + c_2 &= -1 \qquad (R_1) \\ c_1 + 2c_2 &= 4 \qquad (R_2) \end{aligned}$$

Eliminating c_1 in the second equation yields

$$\begin{aligned} (R_1){:}\quad 2c_1 + c_2 &= -1 \qquad (R_3) \\ 2(R_2) - (R_1){:}\quad 3c_2 &= 9 \qquad (R_4) \end{aligned}$$

Hence, $c_2 = 3$, and therefore on substituting into (R_3):

$$2c_1 = -1 - c_2 = -1 - 3 = -4 \qquad \text{or} \qquad c_1 = -2.$$

Thus,

$$\mathbf{x}(t) = -2e^t \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 3e^{-2t} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

or, to summarize, (writing the vector equation out as separate equations):

$$x_1(t) = -4e^t + 3e^{-2t}$$

$$x_2(t) = -2e^t + 6e^{-2t}$$

solves the system

$$\frac{dx_1}{dt} = 2x_1 - 2x_2$$

$$\frac{dx_2}{dt} = 2x_1 - 3x_2$$

with initial condition $x_1(0) = -1$ and $x_2(0) = 4$.

We give one more example of an initial-value problem that illustrates all at once the different steps that must be carried out in order to obtain a solution in the case when A has two distinct and real eigenvalues.

EXAMPLE 2 Solve

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix} \mathbf{x}(t) \tag{11.22}$$

with the initial condition

$$\mathbf{x}(0) = \begin{bmatrix} 3 \\ -1 \end{bmatrix}. \tag{11.23}$$

Solution The first step is to find the eigenvalues and the corresponding eigenvectors. To find the eigenvalues, we must solve

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 2-\lambda & -3 \\ 1 & -2-\lambda \end{bmatrix} = (2-\lambda)(-2-\lambda) + 3 \\ &= \lambda^2 - 1 = 0 \end{aligned}$$

which has solutions

$$\lambda_1 = 1 \qquad \text{and} \qquad \lambda_2 = -1$$

The eigenvector $\mathbf{u}$ corresponding to the eigenvalue $\lambda_1 = 1$ satisfies

$$\begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Rewriting this matrix equation as a system of equations, we obtain

$$\begin{aligned} 2u_1 - 3u_2 &= u_1 \\ u_1 - 2u_2 &= u_2. \end{aligned}$$

These two equations reduce to the same equation, namely,

$$u_1 - 3u_2 = 0.$$

If we set $u_2 = 1$, then $u_1 = 3$, and therefore, $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_1 = 1$.

The eigenvector $\mathbf{v}$ corresponding to the eigenvalue $\lambda_2 = -1$ satisfies

$$\begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = - \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Rewriting this as a system of equations, we find that

$$\begin{aligned} 2v_1 - 3v_2 &= -v_1 \\ v_1 - 2v_2 &= -v_2. \end{aligned}$$

These two equations reduce to the same equation, namely,

$$v_1 - v_2 = 0.$$

If we set $v_1 = 1$, then $v_2 = 1$, and therefore, $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_2 = -1$.

The general solution of (11.22) is thus

$$\mathbf{x}(t) = c_1 e^t \begin{bmatrix} 3 \\ 1 \end{bmatrix} + c_2 e^{-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

where c_1 and c_2 are constants. The initial condition (11.23) will allow us to determine the constants c_1 and c_2:

$$\mathbf{x}(0) = c_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

That is, c_1 and c_2 satisfy

$$\begin{aligned} 3c_1 + c_2 &= 3 \qquad (R_1) \\ c_1 + c_2 &= -1 \qquad (R_2) \end{aligned}$$

We solve this system by the standard elimination method:

$$\begin{array}{rrl} (R_1) & 3c_1 + c_2 = 3 & (R_3) \\ (R_1) - 3(R_2) & -2c_2 = 6 & (R_4) \end{array}$$

Hence, $c_2 = -3$ and $3c_1 + (-3) = 3$, which implies $c_1 = 2$. The solution of (11.22) that satisfies the initial condition (11.23) is therefore

$$\mathbf{x}(t) = 2e^t \begin{bmatrix} 3 \\ 1 \end{bmatrix} - 3e^{-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

which can also be written in the form

$$\begin{aligned} x_1(t) &= 6e^t - 3e^{-t} \\ x_2(t) &= 2e^t - 3e^{-t} \end{aligned}$$

The vector field, two lines in the direction of the two eigenvectors, and the solution curve, are all shown in Figure 11.5.

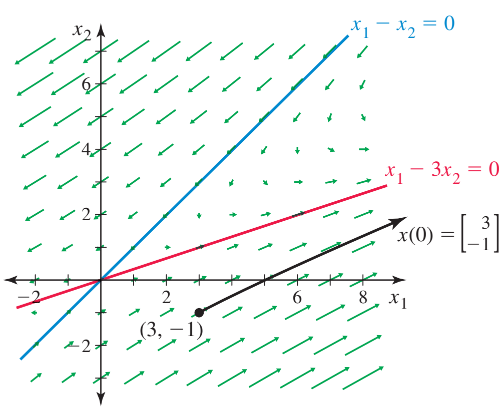


Figure 11.5 The vector field of the system (11.22) with the lines in the direction of the eigenvectors. We also show the solution with initial condition $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$.

You might have noticed that all initial conditions we have discussed thus far have been formulated at time $t = 0$. This is a natural choice for an initial condition; however, we could have chosen any other time—for instance, $t = 1$. Suppose that in Example 2 the initial condition had been

$$\mathbf{x}(1) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \tag{11.24}$$

Then the general solution would still be

$$\mathbf{x}(t) = c_1 e^t \begin{bmatrix} 3 \\ 1 \end{bmatrix} + c_2 e^{-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

but the constants c_1 and c_2 would now satisfy

$$3ec_1 + e^{-1}c_2 = 2 \qquad (R_1)$$

$$ec_1 + e^{-1}c_2 = 1 \qquad (R_2)$$

We solve this system by the standard elimination method:

$$\begin{array}{lrcl} (R_1): & 3ec_1 + e^{-1}c_2 & = & 2 \qquad (R_3) \\ (R_1) - 3(R_2): & -2e^{-1}c_2 & = & -1 \qquad (R_4) \end{array}$$

Hence, $c_2 = \frac{e}{2}$, and by substituting in to (R_3) we find $c_1 = \frac{1}{2e}$. The solution satisfying (11.24) is then

$$\mathbf{x}(t) = \frac{1}{2}e^{t-1} \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \frac{1}{2}e^{1-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

11.1.3 Equilibria and Stability

The concepts of equilibria and stability, two concepts that we initially encountered in Section 8.2, can be extended to systems of differential equations. We will restrict ourselves to the case

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \tag{11.25}$$

with

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{11.26}$$

We say that a point

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\mathbf{x}}_2 \end{bmatrix}$$

is an equilibrium of (11.25) if

$$A\hat{\mathbf{x}} = \mathbf{0}$$

that is, if $\hat{\mathbf{x}}$ is a point at which $\frac{d\mathbf{x}}{dt} = 0$. If we start the solution of a system of differential equations at an equilibrium point, it will remain there at all later times.

To find equilibria of (11.25), we must solve $A\mathbf{x} = \mathbf{0}$. We see immediately that $\hat{\mathbf{x}} = \mathbf{0}$ solves $A\mathbf{x} = \mathbf{0}$. It follows from results in Subsection 9.2.3 that if $\det A \neq 0$, then $(0, 0)$ is the only equilibrium of (11.25). If $\det A = 0$, then there will be other equilibria.

As in Chapter 8, the characteristic property of an equilibrium is that if we start a system in equilibrium, it will stay there at all future times. This does *not* mean that if the system is in equilibrium and is **perturbed** by a small amount (i.e., the solution is moved to a nearby point), it will return to the equilibrium. Whether or not a solution will return to an equilibrium after a small perturbation is addressed by the **stability** of the equilibrium. We saw in the previous subsection that, in the case when A has two real and distinct eigenvalues, the solution of (11.25) is given by

$$\mathbf{x}(t) = c_1 e^{\lambda_1 t}\mathbf{u} + c_2 e^{\lambda_2 t}\mathbf{v} \tag{11.27}$$

where $\mathbf{u}$ and $\mathbf{v}$ are the eigenvectors corresponding to the eigenvalues λ_1 and λ_2 of the matrix A and the constants c_1 and c_2 depend on the initial condition. Knowing the solution will allow us to study the behavior of the solution (11.27) as $t \to \infty$ and thus address the question of stability, at least when the eigenvalues are real and distinct.

Since A is a 2×2 matrix and all entries of A are real, the eigenvalues of A are either both real or both complex conjugate. We will treat these two cases separately. To simplify our discussion, we again assume that A has two distinct eigenvalues.

The equilibria of (11.25) can be found by solving

$$A\mathbf{x} = \mathbf{0} \tag{11.28}$$

If $\det A \neq 0$, then (11.28) has only one solution, namely, the trivial solution $(0, 0)$. (See Subsection 9.2.3.) Since $\det A = \lambda_1\lambda_2$, where λ_1 and λ_2 are the eigenvalues of A, $\det A \neq 0$ if and only if λ_1 and λ_2 are both nonzero.

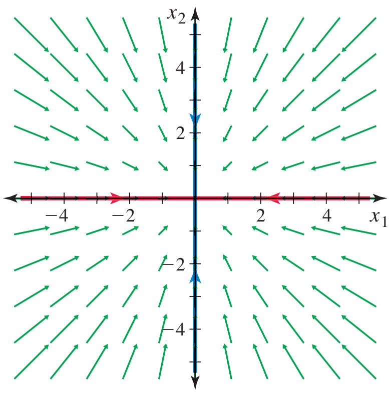


Figure 11.6 The vector field of a linear system where both eigenvalues are negative, together with the lines in the direction of the eigenvectors.

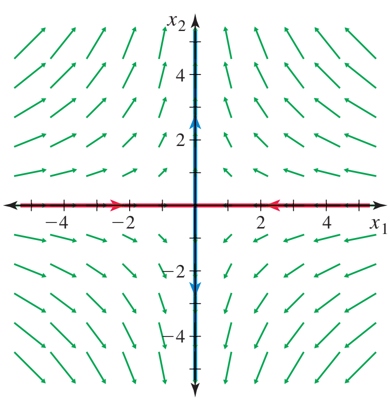


Figure 11.7 The vector field of a linear system where both eigenvalues are of opposite sign, together with the lines in the direction of the eigenvectors.

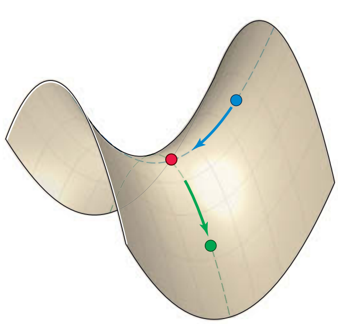


Figure 11.8 A saddle point has both stable and unstable directions.

Case 1: A has two distinct real nonzero eigenvalues. In this case, the equation $A\mathbf{x} = \mathbf{0}$ has only one solution, namely $(0, 0)$, and thus $(0, 0)$ is the only equilibrium. The general solution of (11.25) is given by (11.27), and we can therefore study the behavior of the solution of (11.25) directly by investigating (11.27). We are interested in determining

$$\lim_{t\to\infty} \mathbf{x}(t) = \lim_{t\to\infty} \left[c_1 e^{\lambda_1 t}\mathbf{u} + c_2 e^{\lambda_2 t}\mathbf{v}\right] \tag{11.29}$$

The behavior of $\mathbf{x}(t)$ is determined by $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$. Recall that

$$\lim_{t\to\infty} e^{\lambda_i t} = \begin{cases} 0 & \text{if } \lambda_i < 0 \\ \infty & \text{if } \lambda_i > 0 \end{cases}$$

We distinguish the following three categories:

1. Both eigenvalues are negative: $\lambda_1 < 0$ and $\lambda_2 < 0$.
2. The eigenvalues are of opposite signs: $\lambda_1 < 0$ and $\lambda_2 > 0$ or $\lambda_1 > 0$ and $\lambda_2 < 0$. We will assume without loss of generality that $\lambda_1 > 0 > \lambda_2$.
3. Both eigenvalues are positive: $\lambda_1 > 0$ and $\lambda_2 > 0$.

Representative vector fields for each of the three categories are shown in Figures 11.6 through 11.9. We will discuss each category separately.

Category 1. $\boldsymbol{\lambda_1 < 0}$ and $\boldsymbol{\lambda_2 < 0}$ When both eigenvalues are negative, we conclude from (11.29) that

$$\lim_{t\to\infty} \mathbf{x}(t) = \mathbf{0}$$

regardless of $(x_1(0), x_2(0))$. We say that the equilibrium $(0, 0)$ is **globally stable**, since the solution will approach the equilibrium $(0, 0)$ regardless of the starting point (and not just from nearby points). We call $(0, 0)$ a **sink** or a **stable node**. A vector field for this case is shown in Figure 11.6; the shape of the field explains why we call the equilibrium $(0, 0)$ a sink: All solutions "flow" into the origin.

The system of differential equations that gave rise to the vector field in Figure 11.6 is

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \quad \text{with } A = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$$

The eigenvalues of A are $\lambda_1 = -2$ and $\lambda_2 = -1$. Both are negative. The two straight lines in the figure are the lines in the directions of the two eigenvectors. We see that, starting at a point on either straight line, the solution will approach the equilibrium $(0, 0)$ along the straight line. Furthermore, we see from the vector field that, starting from any other point, the solution will approach the equilibrium $(0, 0)$, as we concluded from the general solution and the fact that both eigenvalues are negative.

Category 2. $\boldsymbol{\lambda_1 > 0 > \lambda_2}$ When the eigenvalues are of opposite signs, we see from (11.29) that the component of the solution associated with the negative eigenvalue goes to 0 as $t \to \infty$ and the component associated with the positive eigenvalue goes to infinity. That is, unless we start in the direction of the eigenvector associated with the negative eigenvalue, the solution will not converge to the equilibrium

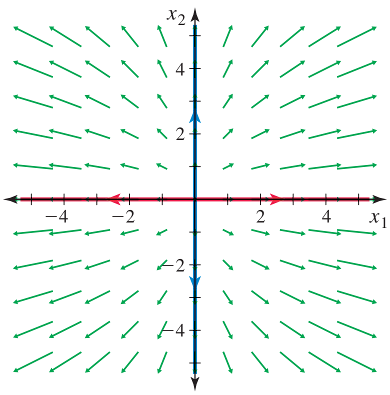


Figure 11.9 The vector field of a linear system where both eigenvalues are positive, together with the lines in the direction of the eigenvectors.

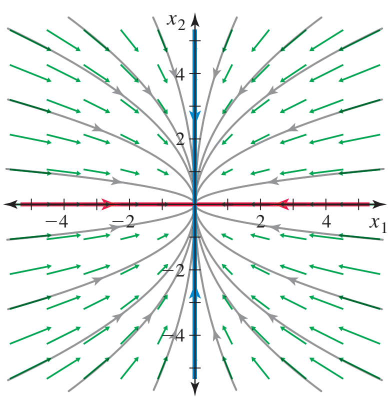


Figure 11.10 Solution curves for a linear system where both eigenvalues are negative.

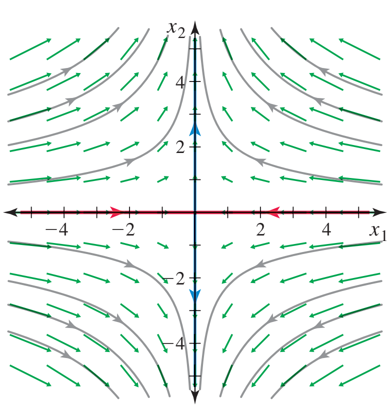


Figure 11.11 Solution curves for a linear system where one eigenvalue is positive and the other is negative.

$(0, 0)$. We say that the equilibrium $(0, 0)$ is **unstable** and call $(0, 0)$ a **saddle point**. A vector field for this case is shown in Figure 11.7. Why is this type of equilibrium called a saddle point? Look back at our study of local extrema in Section 10.7. There we identified a saddle point to be a critical point that is neither a local minimum nor a local maximum. Instead, in one direction the surface is increasing away from the critical point, while in another direction it is decreasing (see Figure 11.8) rather like the saddle of a horse. Imagine a ball rolling along this surface. If the ball starts exactly at the equilibrium (like the red ball in Figure 11.8) it will remain at rest at the equilibrium. If it is started along one of the increasing directions (like the blue ball in Figure 11.8), it will roll back toward the equilibrium point, behaving as if the equilibrium were stable. However, if it is started along one of the decreasing directions (like the green ball in Figure 11.8), then it will roll away from the equilibrium.

The system of differential equations that gave rise to the vector field in Figure 11.7 is

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \quad \text{with } A = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}.$$

The eigenvalues of A are $\lambda_1 = -2$ and $\lambda_2 = 1$. The two straight lines in Figure 11.7 are the lines in the directions of the two eigenvectors. If started at a point on the straight line corresponding to the negative eigenvalue (the horizontal line), the solution will approach the equilibrium $(0, 0)$ along the straight line. Starting at a point on the straight line corresponding to the positive eigenvalue (the vertical line), the solution will move away from the equilibrium $(0, 0)$ along the straight line. If started from any other point, the solution will eventually move away from the equilibrium $(0, 0)$.

In sum, we see from the vector field that the solution can approach $(0, 0)$ from only one direction (the direction of the eigenvector associated with the negative eigenvalue); it eventually moves away from $(0, 0)$ if it is started anywhere else.

Category 3. $\lambda_1 > 0$ and $\lambda_2 > 0$ Finally, if both eigenvalues are positive, we see from (11.29) that the solution will not converge to $(0, 0)$ unless we start at $(0, 0)$. We say that the equilibrium $(0, 0)$ is unstable, and we call $(0, 0)$ a **source** or an **unstable node**. A vector field for this case is shown in Figure 11.9; the shape of the field explains why we call this equilibrium a source.

The system of differential equations that gave rise to this vector field is

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \quad \text{with } A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

The eigenvalues of A are $\lambda_1 = 2$ and $\lambda_2 = 1$, which are both positive. The two straight lines in Figure 11.9 are the lines in the directions of the two eigenvectors. If started at a point on either straight line, the solution will move away from the equilibrium $(0, 0)$ along the straight line. If started from any other point, the solution will move away from the equilibrium $(0, 0)$.

In addition to using vector fields to visualize the behaviors of solutions of systems of differential equations, we may also same of the plot solution curves $(x_1(t), x_2(t))$ starting from different initial conditions to visualize whether an equilibrium point is stable or unstable. Figures 11.10 to 11.12 show some of the solution curves derived from the vector fields plotted in Figures 11.6 to 11.9, respectively.

In all cases, as we previously noted, if the initial condition lies on one of the lines of eigenvectors of A, then the solution will stay on this line of eigenvectors for all times. Equivalently the lines of eigenvectors are solution curves.

11.1.4 Systems with Complex Conjugate Eigenvalues

So far we have only considered linear systems in which the matrix A has real eigenvalues. However, it is also possible, even when all of the components of A are real, that A will have complex conjugate eigenvalues. We will not solve the system when A has

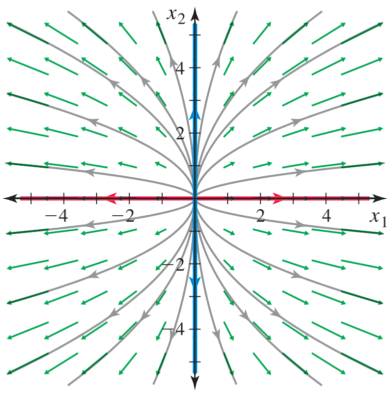


Figure 11.12 Solution curves for a linear system where both eigenvalues are positive.

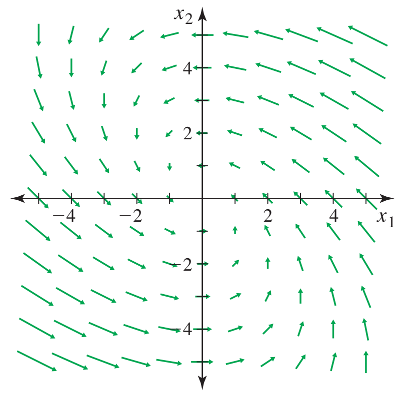


Figure 11.13 Some of the solution curves for the stable spiral equilibrium point given by (11.30).

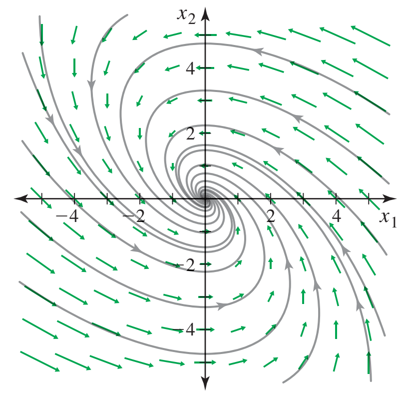


Figure 11.14 The vector field for the system with matrix A in (11.30).

complex conjugate eigenvalues. Instead, we will look at some examples to see what typical vector fields look like. We will see that the stability of the equilibrium point is determined by whether the real part of the eigenvalue is positive, negative, or zero. Since the eigenvalues are a complex conjugate pair they have the same real parts.

Category 1. $\mathbf{Re}(\lambda_1), \mathbf{Re}(\lambda_2) < 0$ Let

$$A = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \tag{11.30}$$

Then

$$\det(A - \lambda I) = \det \begin{bmatrix} -1-\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = (-1-\lambda)(-\lambda) + 1$$
$$= \lambda^2 + \lambda + 1 = 0$$

which gives

$$\lambda_{1,2} = \frac{-1 \pm \sqrt{1-4}}{2} = -\frac{1}{2} \pm \frac{i}{2}\sqrt{3}$$

Both eigenvalues are complex and form a conjugate pair. (Note that the real parts of both eigenvalues are negative.) To see what the solutions of

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t)$$

look like, we graph the vector field and some solutions $x_1(t)$ and $x_2(t)$ against t in Figures 11.13 and 11.15, respectively.

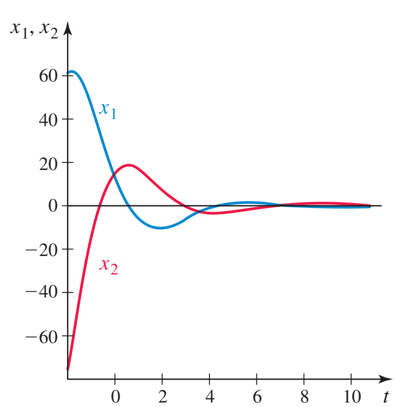


Figure 11.15 The solutions for (11.30).

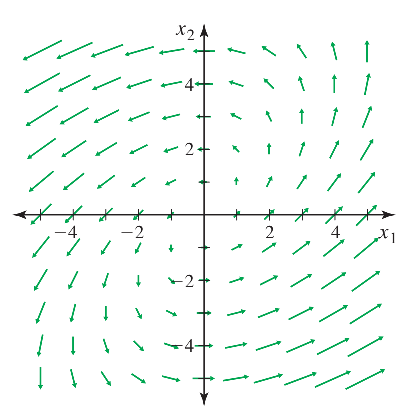


Figure 11.16 The vector field for the system with matrix B in (11.31).

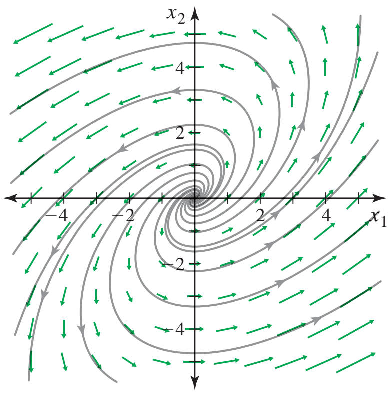


Figure 11.17 Some solution curves for the unstable spiral point given by (11.31).

We see from the vector field that, starting from any point other than $(0, 0)$, solutions spiral into the equilibrium $(0, 0)$. (In Figure 11.14 we show some sample solution curves for this system.) For this reason, the equilibrium $(0, 0)$ is called a **stable spiral**. When we plot solutions as functions of time, they show oscillations, as illustrated in Figure 11.15. The amplitude of the oscillations decreases over time. We therefore call the oscillations **damped**.

The oscillations are caused by the imaginary part of the eigenvalues; the damping of the oscillations is caused by the negative real part of the eigenvalues. Before we explain any further, we give two more examples, one in which the complex conjugate eigenvalues have positive real parts, the other in which the eigenvalues are purely imaginary.

Category 2. $\mathrm{Re}(\lambda_1), \mathrm{Re}(\lambda_2) > 0$ Let

$$B = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \tag{11.31}$$

Then

$$\det(B - \lambda I) = \det \begin{bmatrix} 1-\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = (1-\lambda)(-\lambda) + 1$$
$$= \lambda^2 - \lambda + 1 = 0$$

which has solutions

$$\lambda_{1,2} = \frac{1 \pm \sqrt{1-4}}{2} = \frac{1}{2} \pm \frac{i}{2}\sqrt{3}$$

Both eigenvalues of B are complex and form a complex conjugate pair, but their real parts are positive. To see what the solutions of

$$\frac{d\mathbf{x}}{dt} = B\mathbf{x}(t)$$

look like, we graph the vector field and solution curves in Figures 11.16 and 11.17, respectively.

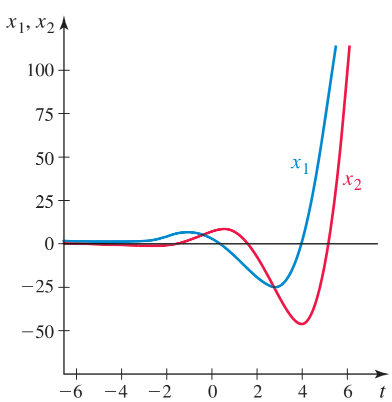


Figure 11.18 The solutions for (11.31).

We see from the vector field that, starting from any point other than $(0, 0)$, the solutions spiral out from the equilibrium $(0, 0)$. For this reason, we call the equilibrium $(0, 0)$ an **unstable spiral**. When we plot solutions as functions of time, as in Figure 11.18, we see that the solutions show oscillations as before, but now their amplitudes are increasing and we call these oscillations unstable. The oscillations are again caused by the imaginary parts of the eigenvalues; the increase in amplitude is caused by the positive real parts.

Category 3. $\mathrm{Re}(\lambda_1), \mathrm{Re}(\lambda_2) = 0$ Here is an example where both eigenvalues are purely imaginary. Let

$$C = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{11.32}$$

Then

$$\det(C - \lambda I) = \det \begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 + 1 = 0$$

which gives

$$\lambda_{1,2} = \pm i$$

Both eigenvalues of C are complex and form a complex conjugate pair, but they are purely imaginary. (Their real parts are equal to 0.) To see what the solutions of

$$\frac{d\mathbf{x}}{dt} = C\mathbf{x}(t)$$

look like, we graph both the vector field and solution curves in Figures 11.19 and 11.20.

Looking at the solutions $x_1(t)$ and $x_2(t)$ as functions of t in Figure 11.21, we see that the solutions oscillate as in the previous two examples, but this time the amplitude does not change with time. Looking at the solution curves in Figure 11.20, we see that solutions spiral around the equilibrium $(0, 0)$, but since the amplitude of the solutions does not change, the solutions neither approach nor move away

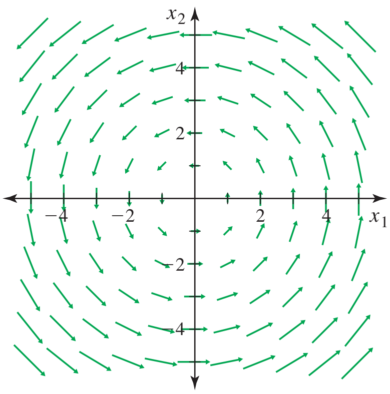


Figure 11.19 The vector field for the system with matrix C in (11.32).

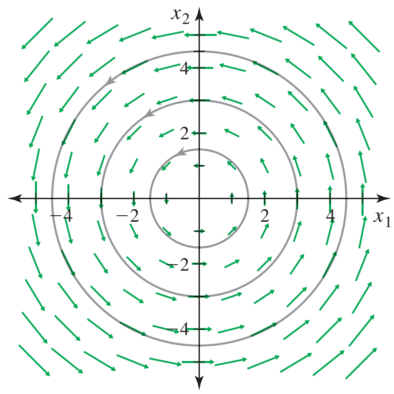


Figure 11.20 Some of the solution curves for the center point given by the matrix in (11.32).

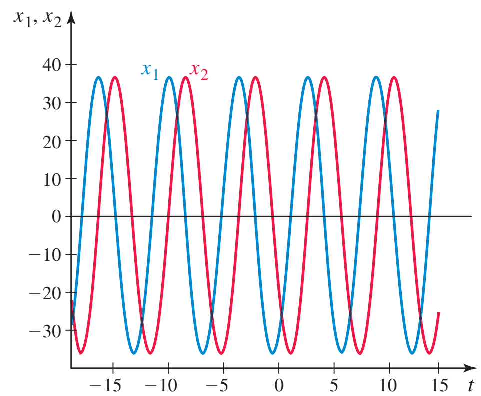


Figure 11.21 The solutions for (11.32).

from the equilibrium. The equilibrium (0, 0) is called a **neutral spiral** or a **center**. We can show that the solutions form closed curves. (We will analyze the solution curves more closely in Example 3.)

Where Do the Oscillations Come From? We can show that the solutions of (11.25) are given by (11.27), regardless of whether the eigenvalues are real or complex, as long as the eigenvalues are distinct. If the eigenvalues are complex, then the solution contains terms of the form $e^{\lambda_1 t}$, where λ_1 is a complex number.

We can split λ_1 into real and imaginary parts, where a and b are both real:

$$\lambda_1 = a + ib$$

The number a is called the *real part* of z, the number b is the *imaginary part* of λ_1.

To understand what $e^{\lambda_1 t}$ means when λ_1 is complex, we write

$$e^{\lambda_1 t} = e^{(a+ib)t} = e^{at}e^{ibt}$$

The term e^{at} is a real number. If $a > 0$, then e^{at} will grow exponentially, while if $a < 0$, then e^{at} will decay exponentially. Hence, since $a = Re(\lambda_1)$, the solutions grow or decay depending on the sign of $Re(\lambda_1)$, as we already claimed. The term e^{ibt} is an exponential with a purely imaginary exponent. We have not encountered this kind of number before. The following formula, which we cannot prove here, explains its meaning:

Euler's Formula

$$e^{ibt} = \cos bt + i \sin bt$$

Both the sine and the cosine functions show oscillations; this is the reason systems of differential equations with complex eigenvalues have solutions that oscillate. The combination of an exponentially growing or decaying part and an oscillating part is what produces the spirals we saw in categories 1 and 2. To understand why the solutions are closed curves when λ_1 and λ_2 are purely imaginary we will consider a specific example in which the system of equations can be solved to find the solution curves.

EXAMPLE 3 Find the equation of the solution curves for the system of equations:

$$\frac{dx_1}{dt} = -x_2$$

$$\frac{dx_2}{dt} = x_1.$$

Solution The associated matrix of coefficients is

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

which is the same as the matrix C that we saw in (11.32). We therefore know that

$$\lambda_1 = i \quad \text{and} \quad \lambda_2 = -i.$$

However, in this instance knowing the eigenvalues does not give us much information about the shape of the solution curves. Instead we will make use of the fact that at each point (x_1, x_2) the slope of the solution curve is equal to:

$$\frac{dx_2}{dx_1} = \frac{dx_2/dt}{dx_1/dt} = \frac{x_1}{(-x_2)} = -\frac{x_1}{x_2} \tag{11.33}$$

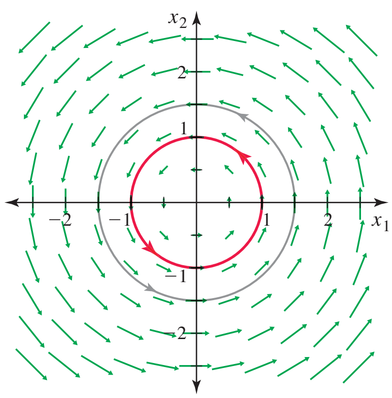


Figure 11.22 Solution curves trajectory for Example 3. The solution curve that passes through $(x_1(0), x_2(0)) = (1, 0)$ is shown in red.

Equation (11.33) is separable, and we therefore solve it using the methods of Section 8.1.

$$\int x_2\,dx_2 = -\int x_1\,dx_1 \qquad \text{Separate variables and integrate}$$

$$\frac{1}{2}x_2^2 = -\frac{1}{2}x_1^2 + C_1$$

or

$$x_1^2 + x_2^2 = 2C_1$$

Notice that C_1 must be nonnegative (there are no points (x_1, x_2) with $x_1^2 + x_2^2 < 0$). So we can redefine our constant on the right-hand side to be R^2. We have therefore shown that the solution curves consist of points (x_1, x_2) that solve the implicit equation $x_1^2 + x_2^2 = R^2$.

We recognize this equation from Subsection 1.2.3 as the equation of a circle of radius R centered at $(0, 0)$. Different initial conditions $x_1(0), x_2(0)$ give different values for the constant R, but all solution curves follow a circular path around $(0, 0)$. Some solution curves are shown in Figure 11.22, along with the vector field of the equation system. •

In Example 3 we were able to draw the solution curves of the equation system without solving for $x_1(t)$ and $x_2(t)$. Suppose that we have initial conditions $x_1(0) = 1$ and $x_2(0) = 0$. Then, although directly deriving the solution is beyond the scope of this book, we can check that the functions

$$x_1(t) = \cos t \quad \text{and} \quad x_2(t) = \sin t$$

solve the system of differential equations (11.33).

We find that

$$\frac{dx_1}{dt} = -\sin t = -x_2(t)$$

and

$$\frac{dx_2}{dt} = \cos t = x_1(t)$$

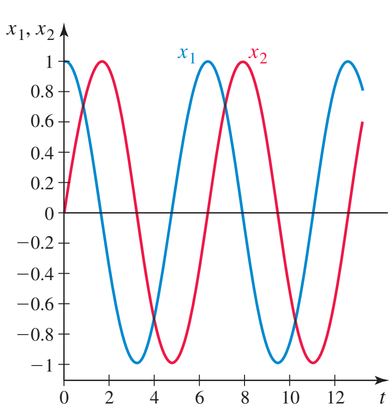


Figure 11.23 The solution for Example 3 as functions of t with initial condition $x_1(0) = 1$ and $x_2(0) = 0$.

Also $x_1(0) = \cos 0 = 1$ and $x_2(0) = \sin 0 = 0$ so the functions satisfy the initial conditions. The solution is shown in Figure 11.23. As noted above, the corresponding solution curve in the x_1–x_2 plane is a circle. Indeed since

$$x_1^2 + x_2^2 = \cos^2 t + \sin^2 t = 1$$

the solution curve is a circle of radius 1 and is shown in Figure 11.22.

We can see from either Figure 11.22 or from Figure 11.23, in which $x_1(t)$ and $x_2(t)$ are plotted as functions of t that the solutions $x_1(t)$ and $x_2(t)$ are both periodic. Figure 11.23 shows sustained oscillations; that is, the amplitudes of $x_1(t)$ and $x_2(t)$ do not change with time. In Figure 11.22, we see that all solutions form closed curves in the x_1–x_2 plane, indicating that the solutions are periodic.

11.1.5 Summary of the Theory of Linear Systems

We close this section by bringing together all of the kinds of equilibria that were introduced in this section, and by giving a straightforward rule for classifying the equilibrium $(0, 0)$ for a linear system

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \quad \text{with } A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Recall that

The trace of A is $\text{trace}(A) = a_{11} + a_{22}$.

The determinant of A is $\det(A) = a_{11}a_{22} - a_{12}a_{21}$.

We denote the trace of A by τ and the determinant of A by Δ. The eigenvalues of A are found by solving

$$\begin{aligned}\det(A - \lambda I) &= \det \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} \\ &= (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} \\ &= \lambda^2 - \tau\lambda + \Delta = 0\end{aligned}$$

where we used $\tau = a_{11} + a_{22}$ and $\Delta = a_{11}a_{22} - a_{12}a_{21}$. The last equation gives

$$\lambda_1 = \frac{\tau + \sqrt{\tau^2 - 4\Delta}}{2} \qquad \text{and} \qquad \lambda_2 = \frac{\tau - \sqrt{\tau^2 - 4\Delta}}{2} \tag{11.34}$$

So since λ_1, λ_2 are real if and only if $\tau^2 - 4\Delta \geq 0$:

If $\tau^2 > 4\Delta$, both eigenvalues are real and distinct.
If $\tau^2 < 4\Delta$, the eigenvalues are complex conjugates.

When both eigenvalues are real and distinct—that is, when $\tau^2 > 4\Delta$—we can distinguish the following three cases:

1. $\Delta < 0$: $\lambda_1 > 0$, $\lambda_2 < 0$ (saddle point)
2. $\Delta > 0$, $\tau < 0$: $\lambda_1 < 0$, $\lambda_2 < 0$ (sink, or stable node)
3. $\Delta > 0$, $\tau > 0$: $\lambda_1 > 0$, $\lambda_2 > 0$ (source, or unstable node)

When both eigenvalues are complex conjugates—that is, when $\tau^2 < 4\Delta$—we can distinguish the following three cases:

1. $\tau < 0$: Both eigenvalues have negative real parts (stable spiral).
2. $\tau > 0$: Both eigenvalues have positive real parts (unstable spiral).
3. $\tau = 0$: Both eigenvalues are purely imaginary (center).

We can summarize all this graphically in the τ–Δ plane as shown in Figure 11.24. The parabola $4\Delta = \tau^2$ is the boundary line between oscillatory and nonoscillatory behavior. The line $\tau = 0$ divides the stable and the unstable regions. The line $\Delta = 0$ separates the saddle point from the node regions. Puzzling out where a matrix A lies on the diagram in Figure 11.24 gives a quick way to determine what kind of equilibrium $(0, 0)$ is.

Each line in the diagram corresponds to a special kind of equilibrium. We have already discussed the line $\tau = 0$ and $\Delta = 0$. Equilibria on this line are centers.

The line $\tau^2 = 4\Delta$ divides stable nodes from stable spirals (if $\tau < 0$) and unstable nodes from unstable spirals (if $\tau > 0$). On this line, because $\tau^2 = 4\Delta$, (11.34) tells us that $\lambda_1 = \lambda_2 = \frac{\tau}{2}$; i.e., the eigenvalues of A are not distinct. We will not discuss this situation in this book, because it requires us to go deeper into the theory of matrices than we were able to go in Chapter 9.

The line $\Delta = 0$ divides stable nodes from saddles (if $\tau < 0$) and unstable nodes from saddles (if $\tau > 0$). On this line one eigenvalue is equal to 0. As long as the other

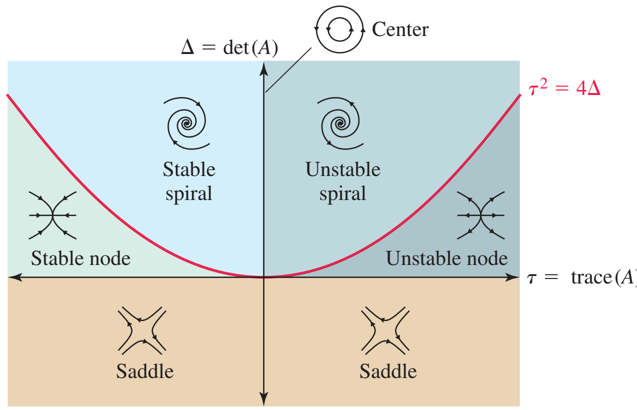


Figure 11.24 The stability behavior of a system of two linear, homogeneous differential equations with constant coefficients.

eigenvalue is not equal to 0, both eigenvalues are again distinct and the solution is of the form (11.27). However, in this case there are equilibria other than (0, 0). We will discuss two examples in Problems 67 and 68 and another example in Subsection 11.2.1.

Section 11.1 Problems

11.1.1

In Problems 1–4, write each system of differential equations in matrix form.

1. $\dfrac{dx_1}{dt} = 2x_1 + 3x_2$

$\dfrac{dx_2}{dt} = -x_1 + x_2$

2. $\dfrac{dx_1}{dt} = x_1 + x_2$

$\dfrac{dx_2}{dt} = -x_1$

3. $\dfrac{dx_1}{dt} = x_3 - 2x_1$

$\dfrac{dx_2}{dt} = -x_1 + x_3$

$\dfrac{dx_3}{dt} = x_1 + x_2 + x_3$

4. $\dfrac{dx_1}{dt} = 2x_2 - 3x_1 - x_3$

$\dfrac{dx_2}{dt} = x_2 - 2x_1$

$\dfrac{dx_3}{dt} = 5x_1 + x_3$

5. Consider

$$\frac{dx_1}{dt} = -x_1 + 2x_2$$
$$\frac{dx_2}{dt} = x_1$$

Determine the direction vectors associated with the following points in the x_1–x_2 plane, and graph the direction vectors in the x_1–x_2 plane: (1, 0), (0, 1), (−1, 0), (0, −1), (1, 1), (0, 0), and (−1, 1).

6. Consider

$$\frac{dx_1}{dt} = 2x_1 - x_2$$
$$\frac{dx_2}{dt} = -x_2$$

Determine the direction vectors associated with the following points in the x_1–x_2 plane, and graph the direction vectors in the x_1–x_2 plane: (1, 0), (0, 1), (−1, 0), (0, −1), (1, 1), (0, 0), and (−2, −2).

7. Consider

$$\frac{dx_1}{dt} = x_1 + x_2$$
$$\frac{dx_2}{dt} = 3x_1 - x_2.$$

Determine the direction vectors associated with the following points in the x_1–x_2 plane, and graph the direction vectors in the x_1–x_2 plane: (1, 0), (0, 1), (−1, 0), (0, −1), (−1, −1), (0, 0), and (1, 2).

8. Consider

$$\frac{dx_1}{dt} = x_2$$
$$\frac{dx_2}{dt} = x_1 + x_2$$

Determine the direction vectors associated with the following points in the x_1–x_2 plane, and graph the direction vectors in the x_1–x_2 plane: (1, 0), (0, 1), (−1, 0), (0, −1), (1, 1), (0, 0), and (−2, −2).

9. In Figures 11.25 through 11.28, vector fields are given. Each of the following systems of differential equations corresponds to exactly one of the vector fields. Match the systems to the appropriate figures.

(a) $\dfrac{dx_1}{dt} = -2x_1 + x_2$

$\dfrac{dx_2}{dt} = x_1 + 2x_2$

(b) $\dfrac{dx_1}{dt} = 2x_1 + x_2$

$\dfrac{dx_2}{dt} = x_1 + x_2$

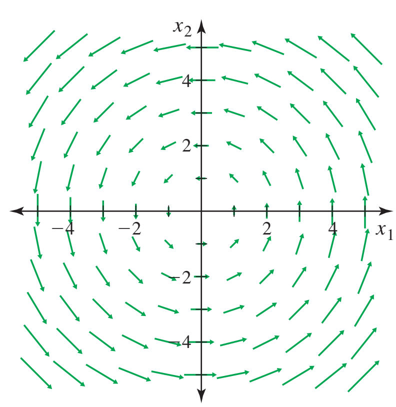


Figure 11.25 See Problem 9.

(c) $\dfrac{dx_1}{dt} = x_1$

$\dfrac{dx_2}{dt} = x_2$

(d) $\dfrac{dx_1}{dt} = -x_2$

$\dfrac{dx_2}{dt} = x_1$

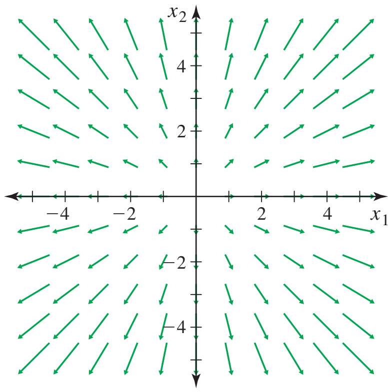


Figure 11.26 See Problem 9.

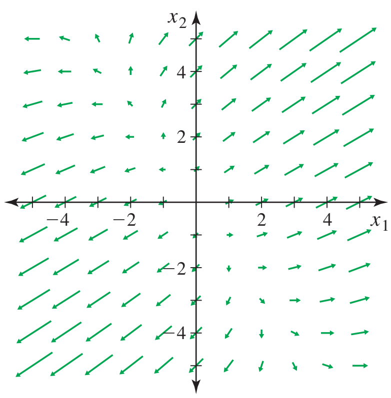


Figure 11.27 See Problem 9.

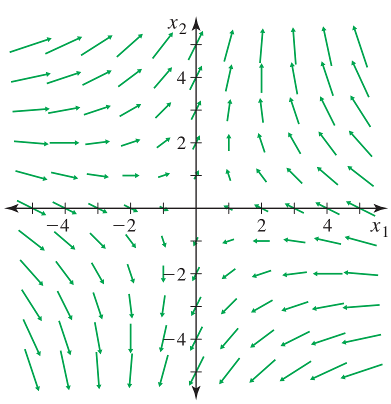


Figure 11.28 See Problem 9.

10. The vector field of

$$\frac{dx_1}{dt} = x_1 + 3x_2$$

$$\frac{dx_2}{dt} = 2x_1 + 3x_2$$

is given in Figure 11.29. Sketch the solution curve that goes through the point $(1, -1)$.

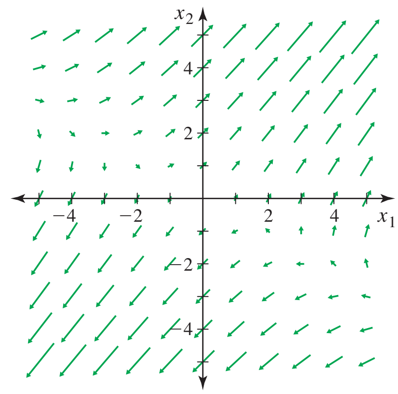


Figure 11.29 See Problem 10.

11. The vector field of

$$\frac{dx_1}{dt} = 2x_1 + 3x_2$$

$$\frac{dx_2}{dt} = -x_1 + x_2$$

is given in Figure 11.30. Sketch the solution curve that goes through the point $(-2, 0)$.

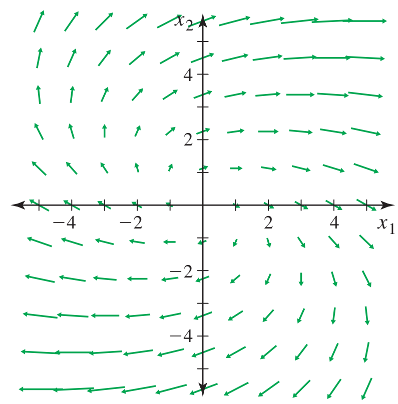


Figure 11.30 See Problem 11.

12. The vector field of

$$\frac{dx_1}{dt} = -x_1 - x_2$$

$$\frac{dx_2}{dt} = -2x_2$$

is given in Figure 11.31. Sketch the solution curve that goes through the point $(3, 4)$.

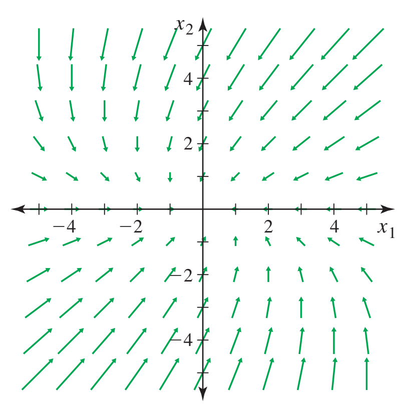


Figure 11.31 See Problem 12.

11.1.2

In Problems 13–18, find the general solution of each system of differential equations and sketch the lines in the direction of the eigenvectors. Indicate on each line of eigenvectors the direction in which the solution would move if it starts on that line.

13. (Figure 11.32)

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

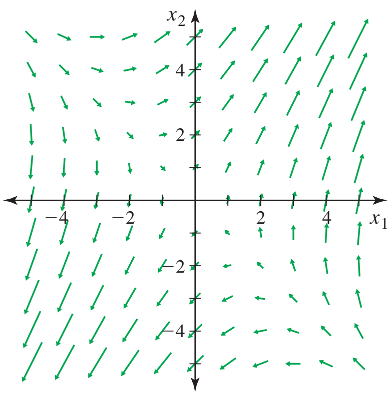


Figure 11.32 See Problem 13.

14. (Figure 11.33)

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

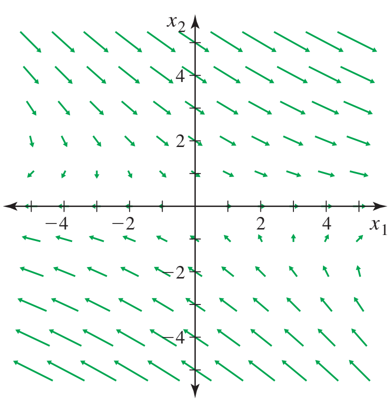


Figure 11.33 See Problem 14.

15. (Figure 11.34)

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} -5 & -9 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

16. (Figure 11.35)

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} -5 & 3 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

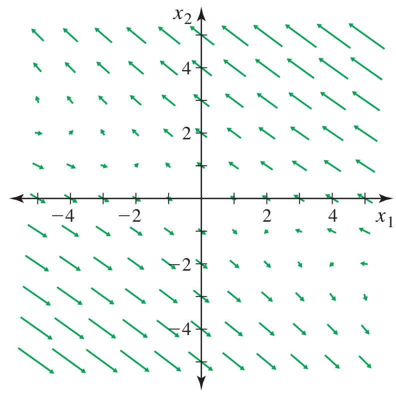


Figure 11.34 See Problem 15.

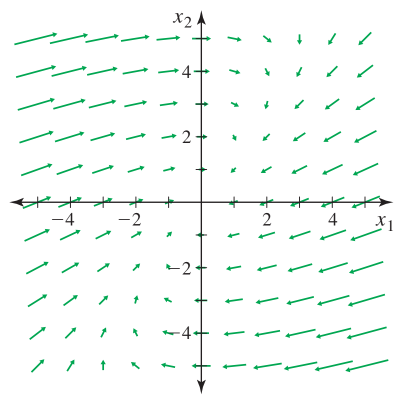


Figure 11.35 See Problem 16.

17. (Figure 11.36)

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

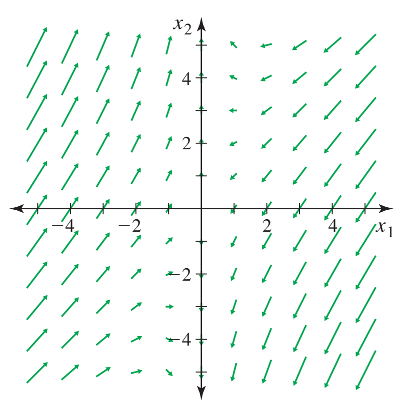


Figure 11.36 See Problem 17.

18. (Figure 11.37)

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 7 & 4 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

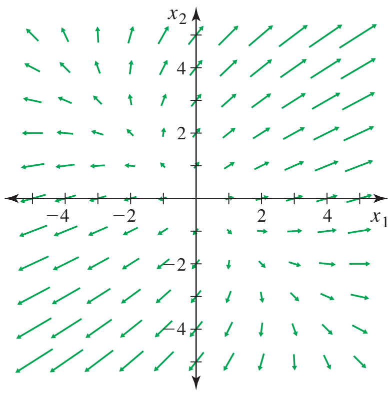


Figure 11.37 See Problem 18.

In Problems 19–26, solve the given initial-value problem.

19. $$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$
with $x_1(0) = -5$ and $x_2(0) = 5$.

20. $$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$
with $x_1(0) = 2$ and $x_2(0) = -1$.

21. $$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$
with $x_1(0) = 1$ and $x_2(0) = 1$.

22. $$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$
with $x_1(0) = -1$ and $x_2(0) = -2$.

23. $$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 4 & -7 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$
with $x_1(0) = 13$ and $x_2(0) = 3$.

24. $$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} -3 & 4 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$
with $x_1(0) = 1$ and $x_2(0) = 2$.

25. $$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 4 & 7 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$
with $x_1(0) = -1$ and $x_2(0) = -2$.

26. $$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$
with $x_1(0) = -3$ and $x_2(0) = 1$.

In Problems 27 and 28, we discuss the case of repeated eigenvalues.

27. Let
$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{11.35}$$

(a) Show that
$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
has the repeated eigenvalues $\lambda_1 = \lambda_2 = 1$.

(b) Show that $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are eigenvectors of A and that any vector $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ can be written as
$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

(c) Show that
$$\mathbf{x}(t) = c_1 e^t \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 e^t \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
is a solution of (11.35) that satisfies the initial condition $x_1(0) = c_1$ and $x_2(0) = c_2$.

28. Let
$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{11.36}$$

(a) Show that
$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$
has the repeated eigenvalues $\lambda_1 = \lambda_2 = 1$.

(b) Show that every eigenvector of A is of the form
$$c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
where c_1 is a real number different from 0.

(c) Show that
$$\mathbf{x}_1(t) = e^t \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
is a solution of (11.36).

(d) Show that
$$\mathbf{x}_2(t) = te^t \begin{bmatrix} 1 \\ 0 \end{bmatrix} + e^t \begin{bmatrix} 0 \\ 0.5 \end{bmatrix}$$
is a solution of (11.36).

(e) Show that
$$\mathbf{x}(t) = c_1 \mathbf{x}_1(t) + c_2 \mathbf{x}_2(t)$$
is a solution of (11.36) for any pair of constants c_1 and c_2. (In fact this is the general solution of the system.)

11.1.3

In Problems 29–42, we consider differential equations of the form
$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t)$$
where
$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$
The eigenvalues of A will be real, distinct, and nonzero. Analyze the stability of the equilibrium (0, 0), and classify the equilibrium according to whether it is a sink, a source, or a saddle point.

29. $A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$

30. $A = \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$

31. $A = \begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix}$

32. $A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$

33. $A = \begin{bmatrix} -4 & 2 \\ -5 & 3 \end{bmatrix}$ **34.** $A = \begin{bmatrix} 2 & 4 \\ 2 & -2 \end{bmatrix}$

35. $A = \begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix}$ **36.** $A = \begin{bmatrix} -1 & 3 \\ 2 & 4 \end{bmatrix}$

37. $A = \begin{bmatrix} -3 & -1 \\ 1 & -6 \end{bmatrix}$ **38.** $A = \begin{bmatrix} -3 & 1 \\ 1 & -2 \end{bmatrix}$

39. $A = \begin{bmatrix} 0 & -2 \\ 1 & -3 \end{bmatrix}$ **40.** $A = \begin{bmatrix} 1 & -2 \\ 1 & -3 \end{bmatrix}$

41. $A = \begin{bmatrix} -2 & -3 \\ 1 & 3 \end{bmatrix}$ **42.** $A = \begin{bmatrix} 4 & -1 \\ 5 & -1 \end{bmatrix}$

11.1.4

In Problems 43–56, we consider differential equations of the form

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t)$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

The eigenvalues of A will be complex conjugates. Analyze the stability of the equilibrium (0, 0), and classify the equilibrium according to whether it is a stable spiral, an unstable spiral, or a center.

43. $A = \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix}$ **44.** $A = \begin{bmatrix} -1 & -5 \\ 4 & -3 \end{bmatrix}$

45. $A = \begin{bmatrix} -2 & 4 \\ -3 & -2 \end{bmatrix}$ **46.** $A = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$

47. $A = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$ **48.** $A = \begin{bmatrix} 2 & -4 \\ 2 & -3 \end{bmatrix}$

49. $A = \begin{bmatrix} 4 & 5 \\ -3 & -3 \end{bmatrix}$ **50.** $A = \begin{bmatrix} 1 & -4 \\ 1 & -1 \end{bmatrix}$

51. $A = \begin{bmatrix} -1 & 1 \\ -3 & 1 \end{bmatrix}$ **52.** $A = \begin{bmatrix} 3 & -2 \\ 1 & 3 \end{bmatrix}$

53. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ **54.** $A = \begin{bmatrix} 0 & -3 \\ 2 & 2 \end{bmatrix}$

55. $A = \begin{bmatrix} 1 & 2 \\ -5 & -3 \end{bmatrix}$ **56.** $A = \begin{bmatrix} 2 & -3 \\ 3 & -2 \end{bmatrix}$

11.1.5

In Problems 57–66, we consider differential equations of the form

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t)$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Analyze the stability of the equilibrium (0, 0), and classify the equilibrium.

57. $A = \begin{bmatrix} -1 & -2 \\ 1 & 3 \end{bmatrix}$ **58.** $A = \begin{bmatrix} -2 & 2 \\ -4 & 3 \end{bmatrix}$

59. $A = \begin{bmatrix} -1 & -1 \\ 5 & -3 \end{bmatrix}$ **60.** $A = \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}$

61. $A = \begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix}$ **62.** $A = \begin{bmatrix} -1 & 5 \\ -3 & 1 \end{bmatrix}$

63. $A = \begin{bmatrix} -2 & 3 \\ 1 & -4 \end{bmatrix}$ **64.** $A = \begin{bmatrix} -2 & -7 \\ 1 & 2 \end{bmatrix}$

65. $A = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}$ **66.** $A = \begin{bmatrix} 2 & 2 \\ 3 & -2 \end{bmatrix}$

67. The following system has two distinct real eigenvalues, but one eigenvalue is equal to 0:

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}\mathbf{x}(t),\ x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{11.37}$$

(a) Find both eigenvalues and the associated eigenvectors.

(b) From the general solution of (11.27) find $x_1(t)$ and $x_2(t)$.

(c) The vector field is shown in Figure 11.38. Sketch the lines corresponding to the eigenvectors. Compute dx_2/dx_1, and conclude that all direction vectors are parallel to the line of eigenvectors corresponding to the nonzero eigenvalue. Describe in words how solutions starting at different points behave.

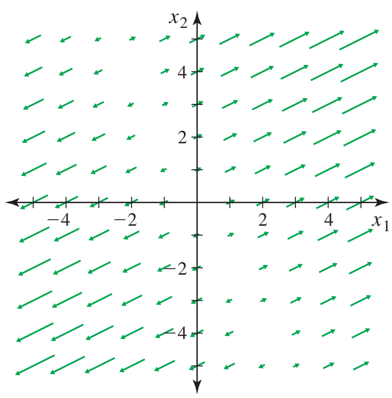


Figure 11.38 See Problem 67.

68. The following system has two distinct real eigenvalues, but one eigenvalue is equal to 0:

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}\mathbf{x}(t) \tag{11.38}$$

(a) Find both eigenvalues and the associated eigenvectors.

(b) Use the general solution (11.27) to find $x_1(t)$ and $x_2(t)$.

(c) The vector field is shown in Figure 11.39. Sketch the lines corresponding to the eigenvectors. Compute dx_2/dx_1, and conclude that all direction vectors are parallel to the line of eigenvectors corresponding to the nonzero eigenvalue. Describe in words how solutions starting at different points behave.

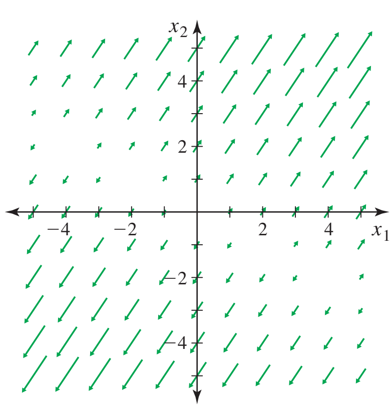


Figure 11.39 See Problem 68.

11.2 Linear Systems: Applications

11.2.1 Two-Compartment Models

Compartment models (which we encountered in Chapter 8) describe flow between compartments, such as nutrient flow between lakes or the flow of a radioactive tracer between different parts of an organism. In many situations, the resulting model is a system of linear differential equations. In Chapter 8 we analyzed two kinds of compartment model. In Section 8.3 we used a graphical method to analyze the flow into and out of a single compartment, while in Section 8.4 we analyzed flows between two compartments. Two-compartment models are often used to predict the passage of medication through a patient's body. To solve the system of equations describing flows between two compartments we had to use integrating factors, and the math was quite complicated. Even so, we still encountered situations where we could not solve the models explicitly. We will now show that the methods introduced in the previous section offer an easier way of analyzing two-compartment models.

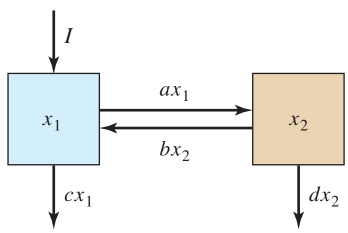


Figure 11.40 A schematic description of a general two-compartment model.

We will consider a general two-compartment model that can be described by a system of two linear differential equations. A schematic description of the model is given in Figure 11.40.

We denote by $x_1(t)$ the amount of matter in compartment 1 at time t and by $x_2(t)$ the amount of matter in compartment 2 at time t. To have a concrete example in mind, think of $x_1(t)$ and $x_2(t)$ as the amount of water in each of the two compartments, respectively. The direction of the flow of matter and the rates at which matter flows are shown in Figure 11.40. We see that matter enters compartment 1 at the constant rate I. Matter moves from compartment 1 to compartment 2 at rate ax_1 if x_1 is the amount of matter in compartment 1, that is, in one unit of time a fraction a of the matter in compartment 1 is transferred to compartment 2. Matter in compartment 1 is lost at rate cx_1 that is, in one unit of time a fraction c of the matter in compartment 1 is lost. In addition, matter flows from compartment 2 to compartment 1 at rate bx_2 if x_2 is the amount of matter in compartment 2. Matter in compartment 2 is lost at rate dx_2; there is no external input into compartment 2. The constants I, a, b, c, and d are all nonnegative.

To derive a system of equations describing the flow of matter between the two compartments let's start, as we did in Section 8.3, by writing word equations describing how matter is added to and removed from each compartment.

$$\begin{array}{c}\text{Rate of change of}\\ \text{matter in}\\ \text{compartment 1}\end{array} = \begin{array}{c}\text{Rate at which}\\ \text{matter enters}\\ \text{compartment 1}\\ \text{from inflow}\end{array} - \begin{array}{c}\text{Rate at which}\\ \text{matter is}\\ \text{transferred to}\\ \text{compartment 2}\end{array} - \begin{array}{c}\text{Rate at which}\\ \text{matter is lost}\end{array} + \begin{array}{c}\text{Rate at which}\\ \text{matter is}\\ \text{transferred from}\\ \text{compartment 2}\end{array}$$

Writing each term out mathematically:

$$\frac{dx_1}{dt} = I - ax_1 - cx_1 + bx_2 \tag{11.39}$$

or:

$$\frac{dx_1}{dt} = I - (a+c)x_1 + bx_2$$ Simplify right-hand side.

Similarly, for the matter in compartment 2,

$$\begin{array}{c}\text{Rate of change of}\\ \text{matter in}\\ \text{compartment 2}\end{array} = \begin{array}{c}\text{Rate at which}\\ \text{matter is}\\ \text{transferred to}\\ \text{compartment 1}\end{array} - \begin{array}{c}\text{Rate at which}\\ \text{matter is lost}\end{array} + \begin{array}{c}\text{Rate at which}\\ \text{matter is}\\ \text{transferred from}\\ \text{compartment 1}\end{array}$$

or, when the terms are all written out mathematically:

$$\frac{dx_2}{dt} = -bx_2 - dx_2 + ax_1$$

or: (11.40)

$$\frac{dx_2}{dt} = ax_1 - (b+d)x_2$$

There is a big difference between this system of equations and the systems that we studied in the previous section; namely, if we attempt to write the system of equations out using matrix and vector notation, i.e., set $\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$, we obtain:

$$\frac{d\mathbf{x}}{dt} = \underbrace{\begin{bmatrix} -(a+c) & b \\ a & -(b+d) \end{bmatrix}}_{A} \mathbf{x} + \begin{bmatrix} I \\ 0 \end{bmatrix} \tag{11.41}$$

Notice that in addition to the term $A\mathbf{x}$ we have a vector on the right-hand side. The terms in this vector cannot be absorbed into the matrix A because they are not a linear map in $[x_1, x_2]'$.

Equations of the type (11.41) are called **inhomogeneous** equations because of the presence of this extra term (if the right-hand side of the equation does not contain these terms it is called **homogeneous**).

The methods we learned in Section 11.1. can only be used for homogeneous equations, i.e., equations of the form $\frac{d\mathbf{x}}{dt} = A\mathbf{x}$. In order for these methods to be applied to (11.41) we must assume that $I = 0$, so that there is no inflow of matter into the system. Now there are many situations where, for example, constant inflow is an important part of the system being modeled. For example, when modeling the flow of nutrients between soil and plants we may need to include an input term, like I, to represent the addition of nutrients to the soil at a constant rate by microorganisms living in the soil. We will discuss how to extend our analysis to inhomogeneous equations in Section 11.3. You may be able to guess how the system behaves when $I = 0$: Either some matter is continually lost, so one or both compartments empty out, or no matter is lost, so at least one compartment will contain matter. We will discuss both cases.

When $I = 0$, (11.40) reduces to the linear system

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \quad \text{with} \quad A = \begin{bmatrix} -(a+c) & b \\ a & -(b+d) \end{bmatrix} \tag{11.42}$$

Because we are assuming there is some flow of matter in the system, we assume that at least one of the parameters a, b, c, and d is positive.

To find the eigenvalues of A, we compute

$$\begin{aligned} &\det \begin{bmatrix} -(a+c)-\lambda & b \\ a & -(b+d)-\lambda \end{bmatrix} \\ &= (a+c+\lambda)(b+d+\lambda) - ab \\ &= \lambda^2 + (a+b+c+d)\lambda + (a+c)(b+d) - ab \\ &= \lambda^2 - \tau\lambda + \Delta = 0 \end{aligned}$$

where τ is the trace of A and Δ is the determinant of A:

$$\tau = -(a+b+c+d) < 0 \qquad \text{At least one of } a, b, c, d \text{ is positive.}$$

$$\Delta = (a+c)(b+d) - ab = ad + bc + cd \geq 0$$

The eigenvalues λ_1 and λ_2 are the solutions of $\lambda^2 - \tau\lambda + \Delta = 0$:

$$\lambda_{1,2} = \frac{\tau \pm \sqrt{\tau^2 - 4\Delta}}{2}$$

To check whether the eigenvalues are real, we simplify the expression under the square root:

$$\begin{aligned} \tau^2 - 4\Delta &= (a+b+c+d)^2 - 4(a+c)(b+d) + 4ab \\ &= [(a+c) - (b+d)]^2 + 4ab \geq 0 \end{aligned}$$

So both eigenvalues are real.

Since $\Delta \geq 0$ and $\tau^2 - 4\Delta \geq 0$, it follows that $\sqrt{\tau^2 - 4\Delta} \leq |\tau|$. Since, in addition, $\tau < 0$, we conclude that both eigenvalues are less than or equal to 0.

$\Delta > 0$: When $\Delta = ad+bc+cd > 0$, then both eigenvalues will be strictly negative; then $(0, 0)$ will be the only equilibrium and will be a stable node. For $\Delta > 0$, at least one of the terms in $ad + bc + cd$ must be positive; that is, either

$$a, d > 0 \qquad \text{or} \qquad b, c > 0 \qquad \text{or} \qquad c, d > 0$$

If $ad > 0$ then $a > 0$ and $d > 0$.

To see what this means, we return to Figure 11.40. When a and d are both positive, the matter in compartment 1 can move to compartment 2, and matter in compartment 2 can leave the system. Matter then flows out of the system through compartment 2. Similarly, when both b and c are positive, matter from both compartments can leave the system through compartment 1. If c and d are both positive, then matter can leave both compartments. In any of these three cases, all matter will eventually leave the system, implying that $(0, 0)$ is a stable equilibrium.

$\Delta = 0$: Then

$$\lambda_1 = \tau = -(a + b + c + d) \qquad \text{and} \qquad \lambda_2 = 0$$

Recall from Section 11.1 that if one eigenvalue is 0, then there will be multiple equilibria (i.e., not just (0,0))

Since we assume that at least one of the four parameters a, b, c, and d is positive, it follows that $\lambda_1 < 0$; thus, the eigenvalues are distinct. To have $\Delta = 0$, one of the following must hold:

$$c = a = 0 \qquad \text{or} \qquad d = b = 0 \qquad \text{or} \qquad c = d = 0$$

If $c = a = 0$, then matter gets stuck in compartment 1; if $d = b = 0$, then matter gets stuck in compartment 2; if $c = d = 0$, no matter will ever leave the system. In all three cases there are multiple equilibria, as the theory from Section 11.1 predicts. For example, in the first case, since matter cannot leave the first compartment, $(x_1, x_2) = (C, 0)$ is an equilibrium for absolutely any value of C. In the second case (matter is trapped in compartment 2), $(x_1, x_2) = (0, C)$ is an equilibrium for any value of C. In the third case ($c = d = 0$), we can say even more. Since matter neither enters nor leaves the system, the total amount of matter in both compartments, which is $x_1 + x_2$, must be constant. We therefore call $x_1(t) + x_2(t)$ a **conserved quantity**, which is the general term for a quantity that does not depend on t. The same conclusion can be reached by a mathematical argument, since, if $c = d = 0$, then (11.40) becomes

$$\frac{dx_1}{dt} = -ax_1 + bx_2$$

$$\frac{dx_2}{dt} = ax_1 - bx_2$$

Adding the two equations, we find that

$$\frac{dx_1}{dt} + \frac{dx_2}{dt} = 0$$

Since

$$\frac{dx_1}{dt} + \frac{dx_2}{dt} = \frac{d}{dt}(x_1 + x_2)$$

it follows that $x_1(t) + x_2(t)$ is a constant.

Solving the system when $c = d = 0$ We can write the solution explicitly in order to find out the equilibrium state. To determine the solution, we must compute the eigenvectors corresponding to the eigenvalues λ_1 and λ_2.

The eigenvector $\mathbf{u}$ corresponding to $\lambda_1 = \tau$ satisfies

$$\begin{bmatrix} -a & b \\ a & -b \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = -(a + b) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Writing this system out yields

$$-au_1 + bu_2 = -au_1 - bu_1$$
$$au_1 - bu_2 = -au_2 - bu_2$$

Both equations simplify to:

$$u_1 = -u_2$$

If we set $u_1 = 1$, then $u_2 = -1$, and an eigenvector corresponding to $\lambda_1 = -(a+b)$ is

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

The eigenvector $\mathbf{v}$ corresponding to $\lambda_2 = 0$ satisfies

$$\begin{bmatrix} -a & b \\ a & -b \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This matrix equation simplifies to one algebraic equation, namely,

$$-av_1 + bv_2 = 0$$

If we set $v_1 = b$, then $v_2 = a$, and an eigenvector corresponding to $\lambda_2 = 0$ is

$$\mathbf{v} = \begin{bmatrix} b \\ a \end{bmatrix}$$

The general solution when $\Delta = 0$ is, therefore,

$$\mathbf{x}(t) = c_1 e^{-(a+b)t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} b \\ a \end{bmatrix}$$

and so:

$$\lim_{t\to\infty} \mathbf{x}(t) = c_2 \begin{bmatrix} b \\ a \end{bmatrix}$$

At time 0,

$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} b \\ a \end{bmatrix}$$

That is,

$$c_1 + bc_2 = x_1(0) \qquad (R_1)$$

$$-c_1 + ac_2 = x_2(0) \qquad (R_2)$$

we can eliminate c_1 to find:

$$(R_1) + (R_2): \qquad (a+b)c_2 = x_1(0) + x_2(0)$$

or

$$c_2 = \frac{x_1(0) + x_2(0)}{a+b}.$$

Recall that the sum $x_1(t) + x_2(t)$ is a constant for all $t \geq 0$; we set $x_1(0) + x_2(0) = K$. Then

$$\lim_{t\to\infty} x_1(t) = K\frac{b}{a+b} \qquad \text{and} \qquad \lim_{t\to\infty} x_2(t) = K\frac{a}{a+b}$$

Together, these two limit equations mean that compartment 1 will eventually contain a fraction $b/(a+b)$ of the total amount of matter, and compartment 2 will contain a fraction $a/(a+b)$ of the total amount.

EXAMPLE 1 Find the system of differential equations corresponding to the compartment diagram shown in Figure 11.41, and analyze the stability of the equilibrium $(0, 0)$.

Solution If time, t, is measured in units of hours then the compartment model is described by the following system of differential equations:

$$\frac{dx_1}{dt} = -(0.1 + 0.2)x_1 + 0.5x_2$$

$$\frac{dx_2}{dt} = 0.2x_1 - 0.5x_2$$

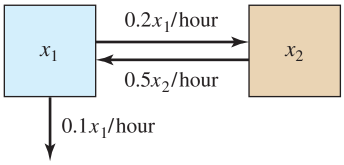


Figure 11.41 The compartment diagram for Example 1.

In matrix notation, this system is equal to

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} -0.3 & 0.5 \\ 0.2 & -0.5 \end{bmatrix} \mathbf{x}(t)$$

To investigate the stability of (0, 0), we find the eigenvalues of the matrix describing the system. That is, we solve

$$\begin{aligned} \det \begin{bmatrix} -0.3-\lambda & 0.5 \\ 0.2 & -0.5-\lambda \end{bmatrix} &= (-0.3-\lambda)(-0.5-\lambda) - (0.2)(0.5) \\ &= (0.3)(0.5) + (0.3+0.5)\lambda + \lambda^2 - (0.2)(0.5) \\ &= \lambda^2 + 0.8\lambda + 0.05 = 0 \end{aligned}$$

We obtain

$$\begin{aligned} \lambda_{1,2} &= \frac{-0.8 \pm \sqrt{0.64 - 0.2}}{2} \\ &= \begin{cases} -0.4 + \frac{1}{2}\sqrt{0.44} \approx -0.068 \\ -0.4 - \frac{1}{2}\sqrt{0.44} \approx -0.732 \end{cases} \end{aligned}$$

We find that both eigenvalues are negative. Therefore, the equilibrium (0, 0) is a stable node. •

EXAMPLE 2 Given the system of differential equations

$$\frac{dx_1}{dt} = -0.7x_1 + 0.2x_2$$

$$\frac{dx_2}{dt} = 0.3x_1 - 0.2x_2$$

determine the parameters for the compartment diagram in Figure 11.40.

Solution The general compartment diagram describing a linear system with two states is shown in Figure 11.40. Comparing the diagram and the system of equations with Equations (11.40) and (11.41), we find that $I = 0$ and

$$\begin{aligned} a + c &= 0.7 \\ b &= 0.2 \\ a &= 0.3 \\ b + d &= 0.2 \end{aligned}$$

Solving this system of equations, we conclude that

$$a = 0.3, \qquad b = 0.2, \qquad c = 0.4, \qquad \text{and} \qquad d = 0$$ •

Compartment models are used in pharmacology to study how drug concentrations change within a patient's body. If the drug is administered in a single dose and we study its passage through the body after the drug has been administered, then there is no input into the patient's body for $t > 0$. The initial dose is modeled through the initial conditions $x_1(0)$ and $x_2(0)$. The techniques from this section can be used to analyze the model.

EXAMPLE 3 A drug is administered to a person in a single dose. We assume that the drug does not accumulate in body tissue, but is filtered from the blood by the kidneys which then pass the drug into the urine. We denote the amount of drug in the body at time t by $x_1(t)$ and in the urine at time t by $x_2(t)$. Initially,

$$x_1(0) = K \qquad \text{and} \qquad x_2(0) = 0$$

Suppose a fraction k_1 of the drug is filtered out by the kidneys in each unit of time. Then the movement of the drug between the body and the urine is modeled by

$$\frac{dx_1}{dt} = -k_1 x_1(t) \tag{11.43}$$

$$\frac{dx_2}{dt} = k_1 x_1(t) \tag{11.44}$$

Solve for $x_1(t)$ and $x_2(t)$.

Solution The system of differential equations (11.43) and (11.44) can be written in matrix form as:

$$\frac{d\mathbf{x}}{dt} = \underbrace{\begin{bmatrix} -k_1 & 0 \\ k_1 & 0 \end{bmatrix}}_{A} \mathbf{x} \qquad \mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

The eigenvalues of the matrix A satisfy

$$\det(A - \lambda I) = 0$$
$$(-k_1 - \lambda)(-\lambda) = 0$$

so $\lambda_1 = -k_1$ and $\lambda_2 = 0$. We can show using the methods from Chapter 9 that the eigenvector corresponding to $\lambda_1 = -k_1$ is $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, while the eigenvector for $\lambda_2 = 0$ is $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, so the general solution of the equation is:

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-k_1 t} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

for some pair of constants c_1 and c_2. As usual we determine constants by imposing the initial conclusions, $\mathbf{x}(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} K \\ 0 \end{bmatrix}$, so $c_1 = K$, $-c_1 + c_2 = 0$, from which we can read off $c_1 = K$, $c_2 = K$.

The solution is then

$$x_1(t) = Ke^{-k_1 t}$$
$$x_2(t) = K(1 - e^{-k_1 t})$$

so as $t \to \infty$, $x_1(t) \to 0$ and $x_2(t) \to K$; that is, all of the drug eventually ends up in the patient's urine. •

11.2.2 A Mathematical Model for Love

This light-hearted but charming example originates with ***Strogatz (2015)***. We set out to model the love between two people; let's call them Romeo and Juliet (to avoid gender stereotyping, you might prefer to call them Romina and Julio, or Romina and Juliet, or Romeo and Julio). Let R represent the amount of affection that Romeo shows Juliet. (One of the reasons that this is a light-hearted model is that it is hard to imagine how this amount of affection might be measured; number of love letters per week, perhaps?) A large positive value of R means that Romeo adores Juliet. Conversely, a large negative value of R means that he loathes her. Let the variable J, which can also be either positive or negative, measure the amount of affection that Juliet shows Romeo.

Both $R(t)$ and $J(t)$ are functions of time and will depend on the interactions between the two lovers. We will explore some possible scenarios below.

EXAMPLE 4 **Hot and Cold Romeo** Let's imagine a situation where Juliet reciprocates Romeo's love. That is, her love for Romeo grows if he shows her affection, but decreases if he treats her poorly. Romeo, on the other hand, backs off if Juliet shows him affection, but becomes interested in her again if she backs off from him.

How might we model Romeo and Juliet's love? We want J to increase if R is positive (to reflect Juliet's love for Romeo growing when he shows her affection) and to decrease if R is negative. One such model would be:

$$\frac{dJ}{dt} = aR \tag{11.45}$$

where a is a positive constant. To represent Romeo's response to this affection we model the change of R in time by

$$\frac{dR}{dt} = -bJ \tag{11.46}$$

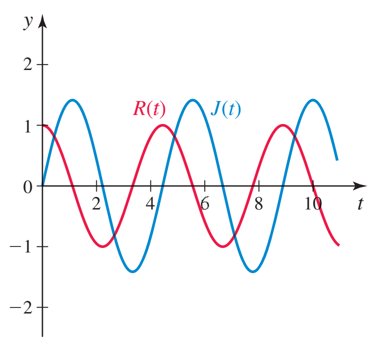


Figure 11.42 Solutions of (11.45) and (11.46) for $a = 2, b = 1$.

where b is another positive constant. As required we then have $\frac{dR}{dt} < 0$ when $J > 0$ (so $R(t)$ decreases, i.e., Romeo *backs off* from Juliet, when she shows him affection), and $\frac{dR}{dt} > 0$ when $J < 0$ (Romeo's affection for Juliet *grows* when she dislikes him). We can analyze the system of equations (11.45) and (11.46) without needing to specify the constants a and b. Written in matrix form, the system of equations is:

$$\frac{d}{dt}\begin{bmatrix} J \\ R \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & a \\ -b & 0 \end{bmatrix}}_{A}\begin{bmatrix} J \\ R \end{bmatrix}$$

Since a, b are both positive constants, the matrix A has eigenvalues $\pm i\sqrt{ab}$, and so $(0, 0)$ is a center in the JR-plane. The solution curves are therefore closed. When plotted as functions of t, $J(t)$ and $R(t)$ oscillate, as shown in Figure 11.42.

In this model Romeo and Juliet are doomed to an endless cycle of affection and then loathing. First they love each other ($R, J > 0$), then Romeo hates Juliet though she loves him ($R < 0, J > 0$), then the hatred becomes mutual ($R, J < 0$), then Romeo loves Juliet though she hates him ($R > 0, J < 0$), and finally they both love each other again. •

EXAMPLE 5 **Cautious Lovers** We now consider a case of cautious lovers. Suppose that, as in Example 4, Juliet's affection for Romeo increases if he shows affection toward her (that is $dJ/dt > 0$ if $R > 0$), and decreases if he treats her poorly (so $dJ/dt < 0$ if $R < 0$). However, Juliet is also cautious: If she finds herself loving Romeo when he is indifferent to her (i.e., $R = 0$), her affection will decrease over time. Conversely, if Romeo is indifferent to Juliet, but she hates him, then her self-consciousness will cause her hatred to vanish over time. We can model Juliet's cautiousness by adding another term to our model from Equation (11.45)

$$\frac{dJ}{dt} = \underbrace{-cJ}_{\text{caution term}} + aR \tag{11.47}$$

where a and c are positive constants.

The aR term is identical to the term in (11.45) and represents the growth of Juliet's affection for Romeo when he shows love toward her. But if Romeo is indifferent to Juliet (i.e., $R = 0$), then (11.47) implies that $\frac{dJ}{dt} = -cJ$, so $J(t)$ will decay exponentially to 0 over time. If Romeo is indifferent to Juliet, then Juliet will put her affection and energy to different uses (maybe she will foster dogs, or learn Portuguese) and then she will become indifferent to Romeo over time. Let's assume that Romeo behaves in the same way toward Juliet, so:

$$\frac{dR}{dt} = aJ - cR \tag{11.48}$$

The trajectory of Romeo and Juliet's romance predicted by Equations (11.47) and (11.48) will depend on the values of the constants a and c. First we write the system in

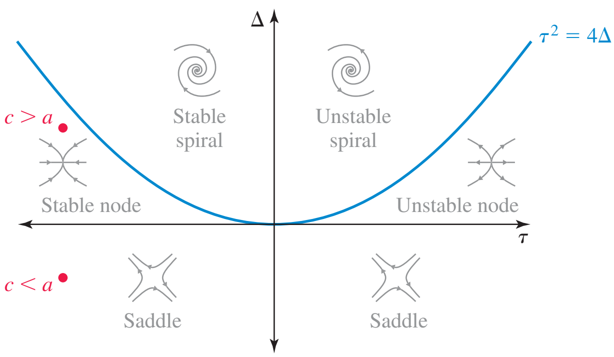


Figure 11.43 Where Equations (11.47) and (11.48) are placed on a (Δ, τ) plot depends on the values of a and c.

matrix form:

$$\frac{d}{dt}\begin{bmatrix} J \\ R \end{bmatrix} = \underbrace{\begin{bmatrix} -c & a \\ a & -c \end{bmatrix}}_{A}\begin{bmatrix} J \\ R \end{bmatrix}$$

For this matrix, A $\tau = \text{trace}\,(A) = -2c$ while $\Delta = \det(A) = c^2 - a^2$.

Since $\tau^2 - 4\Delta = 4c^2 - 4(c^2 - a^2) = 4a^2 > 0$, while $\tau < 0$, the equilibrium at $(0, 0)$ always lies below the line $\tau^2 = 4\Delta$ and on the left of the Δ-axis on a plot of Δ and τ (such as Figure 11.43). Now, if $c > a$, then $\Delta > 0$, whereas if $c < a$, then $\Delta < 0$. Figure 11.43 summarizes how different values for c and a are placed on the (τ, Δ) plot. So if $c > a$, $(0, 0)$ is a stable node; some sample solution curves are then shown in Figure 11.44. In this case J and R both converge to 0 as $t \to \infty$; that is, the lovers' caution means that they inevitably become indifferent to each other over time.

Conversely, if $c < a$, then $(0, 0)$ is a saddle node; however, as a plot of the solution curves shows, this may lead either to $R \to \infty$ and $J \to \infty$ (mutual love) or $R \to -\infty$ and $J \to -\infty$ (mutual hatred); see Figure 11.44.

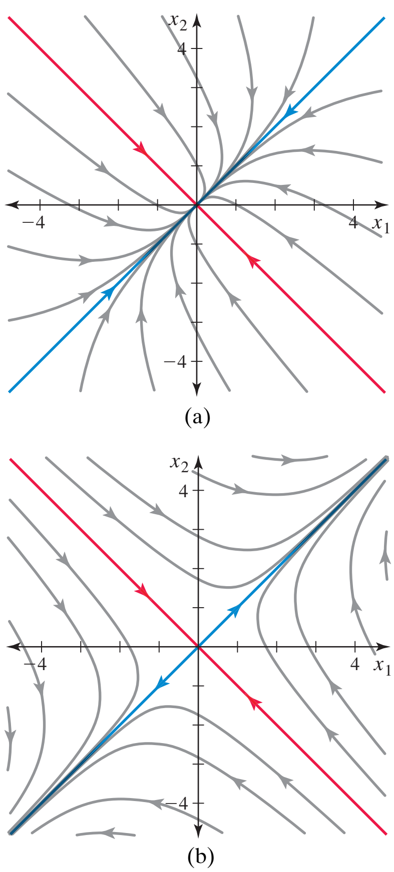


Figure 11.44 Solution curves of Equations (11.47) and (11.48) when (a) $a < c$; $a = 1, c = 2$. (b) $a > c$; $a = 2, c = 1$.

11.2.3 The Harmonic Oscillator

Consider a particle moving along the x-axis. We assume that the particle experiences a force that is proportional to the distance to the origin and that the direction of this force always points toward the origin. If $x(t)$ is the location of the particle at time t, then the second derivative of $x(t)$ denotes the acceleration of the particle, and we find that

$$\frac{d^2x}{dt^2} = -kx \tag{11.49}$$

for some $k > 0$. This is a **second-order differential equation**, because the derivative of the highest-order derivative in the equation is of order 2. We will use this example to show how a second-order differential equation can be transformed into a system of first-order differential equations. To do so, we set

$$\frac{dx}{dt} = v(t)$$

Then

$$\frac{dv}{dt} = \frac{d^2x}{dt^2}$$

and, hence,

$$\frac{dv}{dt} = -kx$$

We thus obtain the following system of first-order differential equations:

$$\frac{dx}{dt} = v$$

$$\frac{dv}{dt} = -kx$$

or

$$\frac{d}{dt}\begin{bmatrix} x \\ v \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -k & 0 \end{bmatrix}\begin{bmatrix} x \\ v \end{bmatrix}$$

If we denote the matrix by A, then

$$\text{tr}\,A = 0 \qquad \text{and} \qquad \det A = k > 0$$

which implies that the eigenvalues of A are complex conjugates with real parts equal to 0. We find that

$$\det(A - \lambda I) = \det\begin{bmatrix} -\lambda & 1 \\ -k & -\lambda \end{bmatrix} = \lambda^2 + k = 0$$

Hence,

$$\lambda_1 = i\sqrt{k} \qquad \text{and} \qquad \lambda_2 = -i\sqrt{k}$$

Thus, we expect the equilibrium $(0, 0)$ to be a neutral spiral (or center) and the solutions to exhibit oscillations whose amplitudes do not change with time.

We can solve (11.49) directly: Since

$$\frac{d}{dt}\sin(at) = a\cos(at)$$

$$\frac{d}{dt}\cos(at) = -a\sin(at)$$

it follows that

$$\frac{d^2}{dt^2}\sin(at) = -a^2\sin(at)$$

and

$$\frac{d^2}{dt^2}\cos(at) = -a^2\cos(at)$$

If we set $a = \sqrt{k}$, we see that $\cos(\sqrt{k}t)$ and $\sin(\sqrt{k}t)$ solve (11.49). Using the superposition principle, we therefore obtain the solution of (11.49) as

$$x(t) = c_1\sin(\sqrt{k}t) + c_2\cos(\sqrt{k}t)$$

To determine the constants c_1 and c_2, we must have an initial condition. If we assume, for instance, that

$$x(0) = 0 \qquad \text{and} \qquad v(0) = v_0 \tag{11.50}$$

then

$$0 = c_2$$

Since $v(t) = dx/dt$, we have

$$v(t) = c_1\sqrt{k}\cos(\sqrt{k}t) - c_2\sqrt{k}\sin(\sqrt{k}t)$$

and, therefore,

$$v(0) = c_1\sqrt{k} = v_0$$

which implies that

$$c_1 = \frac{v_0}{\sqrt{k}}$$

Hence, the solution of (11.49) that satisfies the initial condition (11.50) is given by

$$x(t) = \frac{v_0}{\sqrt{k}}\sin(\sqrt{k}t)$$

The harmonic oscillator is quite important in physics. It describes, for instance, the motion of a particle bouncing at the end of a spring.

Section 11.2 Problems

11.2.1

In Problems 1–8, determine the system of differential equations corresponding to each compartment model and analyze the stability of the equilibrium $(0, 0)$. The parameters have the same meaning as in Figure 11.40.

1. $a = 0.4, b = 1.2, c = 0.3, d = 0$

2. $a = 0.5, b = 0.1, c = 0.05, d = 0.02$

3. $a = 2.5, b = 0.7, c = 0, d = 0.1$

4. $a = 0, b = 0.1, c = 0, d = 0.3$

5. $a = 1.7, b = 0.6, c = 0.1, d = 0.3$

6. $a = 0.2, b = 0.1, c = 0, d = 0$

7. $a = 0.1, b = 0.5, c = 0.5, d = 0.1$

8. $a = 0.2, b = 0, c = 0.5, d = 0$

In Problems 9–18, find the corresponding compartment diagram for each system of differential equations.

9.
$$\frac{dx_1}{dt} = -0.4x_1 + 0.3x_2$$
$$\frac{dx_2}{dt} = 0.1x_1 - 0.5x_2$$

10.
$$\frac{dx_1}{dt} = -0.4x_1 + 0.5x_2$$
$$\frac{dx_2}{dt} = 0.2x_1 - 1.5x_2$$

11. $\dfrac{dx_1}{dt} = -0.2x_1 + 0.1x_2$

$\dfrac{dx_2}{dt} = -0.1x_2$

12. $\dfrac{dx_1}{dt} = -0.2x_1 + 1.1x_2$

$\dfrac{dx_2}{dt} = 0.2x_1 - 1.1x_2$

13. $\dfrac{dx_1}{dt} = -2.3x_1 + 1.1x_2$

$\dfrac{dx_2}{dt} = 0.2x_1 - 2.3x_2$

14. $\dfrac{dx_1}{dt} = -1.6x_1 + 0.3x_2$

$\dfrac{dx_2}{dt} = 0.1x_1 - 0.5x_2$

15. $\dfrac{dx_1}{dt} = -1.2x_1$

$\dfrac{dx_2}{dt} = 0.3x_1 - 0.2x_2$

16. $\dfrac{dx_1}{dt} = -0.2x_1 + 0.4x_2$

$\dfrac{dx_2}{dt} = 0.2x_1 - 0.4x_2$

17. $\dfrac{dx_1}{dt} = -0.2x_1$

$\dfrac{dx_2}{dt} = -0.3x_2$

18. $\dfrac{dx_1}{dt} = -x_1$

$\dfrac{dx_2}{dt} = x_1 - 0.5x_2$

19. Drug Elimination Suppose that a drug is administered to a person in a single dose, and assume that the drug does not accumulate in body tissue, but is excreted through urine. Denote the amount of drug in the body at time t by $x_1(t)$ and in the urine at time t by $x_2(t)$. If $x_1(0) = 4$ mg and $x_2(0) = 0$, find $x_1(t)$ and $x_2(t)$ if

$$\frac{dx_1}{dt} = -0.3x_1(t)$$

20. Drug Elimination Suppose that a drug is administered to a person in a single dose, and assume that the drug does not accumulate in body tissue, but is excreted through urine. Denote the amount of drug in the body at time t by $x_1(t)$ and in the urine at time t by $x_2(t)$. If $x_1(0) = 6$ mg and $x_2(0) = 0$, find a system of differential equations for $x_1(t)$ and $x_2(t)$ if it takes 20 minutes for the drug to be at one-half of its initial amount in the body.

21. Forest Disturbances Disturbances in forests (wind, fire, etc.) create gaps by killing trees. These gaps are eventually filled by new trees. We will model this process by a two-compartment model. We denote by $x_1(t)$ the area occupied by gaps and by $x_2(t)$ the area occupied by adult trees. We assume that the dynamics are given by

$$\frac{dx_1}{dt} = -0.2x_1 + 0.1x_2 \tag{11.51}$$

$$\frac{dx_2}{dt} = 0.2x_1 - 0.1x_2 \tag{11.52}$$

(a) Find the corresponding compartment diagram.

(b) Show that $x_1(t) + x_2(t)$ is a constant. Denote the constant by A and give its meaning. [*Hint*: Show that $\frac{d}{dt}(x_1 + x_2) = 0$.]

(c) Let $x_1(0)+x_2(0) = 20$. Use your answer in (b) to explain why this equation implies that $x_1(t) + x_2(t) = 20$ for all $t > 0$.

(d) Use your result in (c) to replace x_2 in (11.51) by $20 - x_1$, and show that doing so reduces the system (11.51) and (11.52) to

$$\frac{dx_1}{dt} = 2 - 0.3x_1 \tag{11.53}$$

with $x_1(t) + x_2(t) = 20$ for all $t \geq 0$.

(e) Solve the system (11.51) and (11.52), and determine what fraction of the forest is occupied by adult trees at time t when $x_1(0) = 2$ and $x_2(0) = 18$. What happens as $t \to \infty$?

22. Forest Succession Forest succession can be modeled by a three-compartment model. We assume that gaps in a forest are created by disturbances just as in Problem 21. These gaps are initially filled by fast-growing, early colonizing plants, which are then replaced by slower growing species, a process known as succession. We denote by $x_1(t)$ the total area occupied by gaps at time t, by $x_2(t)$ the total area occupied by fast growing species at time t, and by $x_3(t)$ the total area occupied by slow growing species at time t. The dynamics are given by

$$\frac{dx_1}{dt} = 0.2x_2 + x_3 - 2x_1$$

$$\frac{dx_2}{dt} = 2x_1 - 0.7x_2$$

$$\frac{dx_3}{dt} = 0.5x_2 - x_3$$

(a) Draw the corresponding compartment diagram.

(b) Show that

$$x_1(t) + x_2(t) + x_3(t) = A$$

where A is a constant, and give the meaning of A.

11.2.2

23. Frightened Romeo We will explore the situation where Juliet behaves as a cautious lover (see Example 5), but Romeo is so frightened he does not pick up her signals. In this case

$$\begin{aligned} \frac{dJ}{dt} &= -cJ + aR \\ \frac{dR}{dt} &= -dR \end{aligned} \tag{11.54}$$

where a, c, and d are all positive constants.

(a) Interpret what behavior the equation for $\frac{dR}{dt}$ models.

(b) Write (11.54) as a matrix equation and find the eigenvalues of the associated matrix.

(c) Based on your answer to (a), what happens to Romeo and Juliet's relationship as $t \to \infty$?

24. Two Hot and Cold Lovers Imagine that Romeo and Juliet both behave in the same way Romeo does in Example 4; that is, their affections are modeled by a system of equations

$$\begin{aligned} \frac{dJ}{dt} &= -bR \\ \frac{dR}{dt} &= -bJ \end{aligned} \tag{11.55}$$

where b is a positive constant.

(a) By writing the system (11.55) as a matrix equation classify the equilibrium (0, 0); i.e., is it a stable node, spiral, saddle point, and so on?

(b) Find the eigenvector directions for the equilibrium.

(c) Based on your answers to (a) and (b), what is the fate of Romeo and Juliet's relationship as $t \to \infty$?

25. Love-Struck Romeo Juliet is cautious and behaves like she does in Example 4. By contrast, Romeo becomes intoxicated by love; if he starts being even mildly fond of Juliet (i.e., $R > 0$), then his love will grow and grow over time. Their affections are modeled by a system of equations:

$$\frac{dJ}{dt} = -cJ + aR \tag{11.56}$$
$$\frac{dR}{dt} = kR$$

where a, c, and k are all positive constants.

(a) By writing the differential equation system (11.56) as a matrix equation, classify the equilibrium $(0, 0)$; i.e., identify whether it is a stable node, spiral, saddle point, or some other kind of point equilibrium.

(b) Find the eigenvector directions for the equilibrium.

(c) Based on your answers to (a) and (b), what is the fate of Romeo and Juliet's relationship as $t \to \infty$?

26. Love-Struck Romeo and Juliet Romeo and Juliet are both reckless lovers; being in love intoxicates each of them and causes their love to increase, regardless of the feelings of the other. So if Romeo starts off even mildly fond of Juliet (i.e., $R > 0$), then his love will grow regardless of her feelings. Juliet behaves similarly. Conversely, if Romeo hates Juliet, his hatred will grow by itself. We therefore anticipate that if $R > 0$, then $dR/dt > 0$, while if $R < 0$, then $dR/dt < 0$, and similarly for J. We model Romeo and Juliet's relationship by a system of differential equations.

$$\frac{dJ}{dt} = kJ \tag{11.57}$$
$$\frac{dR}{dt} = kR$$

where k is a positive constant.

(a) Write the system (11.57) as a matrix equation, and find the eigenvalues of the corresponding matrix. Why can we not use the methods from Section 11.1 to classify the equilibrium $(0, 0)$?

(b) Solve the equations (11.57) directly to find $R(t), J(t)$ if $R(0) = 1, J(0) = -1$. What is the fate of Romeo and Juliet's relationship as $t \to \infty$?

(c) Try to describe generally what happens to Romeo and Juliet's relationship as $t \to \infty$ for all possible initial conditions.

27. Love-Struck Romeo, Hot and Cold Juliet Romeo's affection for Juliet (or his hatred of her) grows, over time, with no reference to how she feels, like in Problems 25 and 26. Juliet, on the other hand, behaves like Romeo did in Example 4; she is attracted to Romeo when he spurns her, and repulsed when he loves her. We model Romeo and Juliet's relationship by a system of equations

$$\frac{dR}{dt} = kR \tag{11.58}$$
$$\frac{dJ}{dt} = -bR$$

where b and k are both positive constants.

(a) Show that when (11.58) is written in matrix form, the matrix A has 0 as one eigenvalue. What is the other eigenvalue?

(b) Because one of the eigenvalues is 0, $(0, 0)$ is not one of the types of equilibria that we studied in Section 11.1. Solve (11.58) directly to find $R(0)$ and $J(t)$, under the initial conditions $R(0) = 1$ and $J(0) = -1$. What is the fate of Romeo and Juliet's relationship as $t \to \infty$?

(c) Try to describe generally what happens to Romeo and Juliet's relationship as $t \to \infty$ for all possible initial conditions.

In Problems 28–31, based on each system of equations modeling Romeo and Juliet's relationship, describe in words how Romeo and Juliet are both behaving (you do not need to solve any of the systems).

28.
$$\frac{dJ}{dt} = 0.1R$$
$$\frac{dR}{dt} = 10J$$

29.
$$\frac{dJ}{dt} = 0.1R - 2J$$
$$\frac{dR}{dt} = J - 0.1R$$

30.
$$\frac{dJ}{dt} = J - 0.2R$$
$$\frac{dR}{dt} = J - 0.1R$$

31.
$$\frac{dJ}{dt} = R - 2J$$
$$\frac{dR}{dt} = 3R + J$$

11.2.3

32. Solve
$$\frac{d^2x}{dt^2} = -4x$$
with $x(0) = 0$ and $x'(0) = 6$.

33. Solve
$$\frac{d^2x}{dt^2} = -9x$$
with $x(0) = 0$ and $x'(0) = 12$.

34. Transform the second-order differential equation
$$\frac{d^2x}{dt^2} = 3x$$
into a system of first-order differential equations.

35. Transform the second-order differential equation
$$\frac{d^2x}{dt^2} = -\frac{1}{2}x$$
into a system of first-order differential equations.

36. Transform the second-order differential equation
$$\frac{d^2x}{dt^2} + \frac{dx}{dt} = 2x$$
into a system of first-order differential equations.

37. Transform the second-order differential equation
$$\frac{d^2x}{dt^2} - 2\frac{dx}{dt} = \frac{x}{2}$$
into a system of first-order differential equations.

11.3 Nonlinear Autonomous Systems: Theory

At the start of this chapter we motivated the study of systems of differential equations by looking at an example of an ecosystem containing both plants and herbivores. In that example the rate of change of the number of plants depended both on the number of plants and on the number of herbivores.

Generally we are interested in systems of differential equations of the form

$$\begin{aligned}\frac{dx_1}{dt} &= f_1(x_1, x_2, \ldots, x_n)\\ \frac{dx_2}{dt} &= f_2(x_1, x_2, \ldots, x_n)\\ &\vdots\\ \frac{dx_n}{dt} &= f_n(x_1, x_2, \ldots, x_n)\end{aligned} \tag{11.59}$$

where $f_i : \mathbf{R}^n \to \mathbf{R}$, for $i = 1, 2, \ldots, n$. We assume that the functions f_i, $i = 1, 2, \ldots, n$, do not explicitly depend on t; the system (11.59) is therefore called **autonomous**. We no longer assume that the functions f_i are linear, as in Section 11.1. Using vector notation, we can write the system (11.59) in the form

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x})$$

where $\mathbf{x} = [x_1, x_2, \ldots, x_n]'$, and $\mathbf{f}(\mathbf{x})$ is a vector-valued function $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^n$ with components $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, 2, \ldots, n$. The function $\mathbf{f}(\mathbf{x})$ defines a vector field, just as in the linear case.

Unless the functions f_i are linear, it is typically not possible to find explicit solutions of systems of differential equations. If we want to solve such systems, we frequently must use numerical methods. Instead of trying to find solutions, we will focus on point equilibria and their stability, just as in Section 8.2.

The definition of a point equilibrium (as given in Section 8.2) must be extended to systems of the form (11.59). We say that a point

$$\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)$$

is a **point equilibrium** (or simply **equilibrium**) of (11.59) if

$$\mathbf{f}(\hat{\mathbf{x}}) = \mathbf{0}$$

An equilibrium is also called a **critical point**. As in the linear case, at an equilibrium; $\frac{dx}{dt} = 0$, implying that if we start the solution of a system of differential equations at an equilibrium point, it will stay there for all later times.

As in the linear case, a solution might not return to an equilibrium after a small perturbation; if the solution returns to the equilibrium we call it **stable**, while if the solution does not return, then we call the equilibrium **unstable**. The theory of stability for systems of nonlinear autonomous differential equations is parallel to that in Section 8.2; there is both an analytical and a graphical approach. We will restrict our discussion to systems of two equations in two variables. (The concepts are the same when we have more than two equations, but the calculations become more involved.)

11.3.1 Analytical Approach

We start by reminding you of how we determine stability for single differential equations.

A Single Autonomous Differential Equation.

EXAMPLE 1 Find all equilibria of

$$\frac{dx}{dt} = x(1-x) \tag{11.60}$$

and analyze their stability.

Solution We developed the theory for single autonomous differential equations in Section 8.2. To find equilibria, we set

$$x(1-x) = 0$$

which yields

$$\hat{x}_1 = 0 \qquad \text{and} \qquad \hat{x}_2 = 1$$

To analyze the stability of these equilibria, we linearize the differential equation (11.60) about each equilibrium and compute the corresponding eigenvalue. We set

$$f(x) = x(1-x)$$

Then

$$f'(x) = 1 - 2x$$

$f'(\hat{x}) > 0 \Rightarrow \hat{x}$ is unstable
$f'(\hat{x}) < 0 \Rightarrow \hat{x}$ is stable

The eigenvalue associated with the equilibrium $\hat{x}_1 = 0$ is

$$\lambda_1 = f'(0) = 1 > 0$$

which implies that $\hat{x}_1 = 0$ is unstable.
The eigenvalue associated with the equilibrium $\hat{x}_2 = 1$ is

$$\lambda_2 = f'(1) = -1 < 0$$

which implies that $\hat{x}_2 = 1$ is locally stable. •

The eigenvalue corresponding to an equilibrium of the differential equation

$$\frac{dx}{dt} = f(x) \tag{11.61}$$

is the slope of the function $f(x)$ at the equilibrium value. The reason for this is discussed in detail in Section 8.2; we repeat the basic argument here. Suppose that $\hat{x}$ is an equilibrium of (11.61); that is, $f(\hat{x}) = 0$. If we perturb $\hat{x}$ slightly (i.e., if we assume $x = \hat{x} + X$ where $|X|$ is small), we can find out what happens to $\hat{x} + X$ by examining

$$\frac{dx}{dt} = \frac{d}{dt}(\hat{x} + X) = \frac{dX}{dt}$$

Since the perturbation is small, we can linearize

$$f(\hat{x} + X) \approx \underbrace{f(\hat{x})}_{=0} + f'(\hat{x})(\hat{x} + X - \hat{x}) = f'(\hat{x})X$$

so:

$$\frac{dX}{dt} \approx f'(\hat{x})X,$$

which has the approximate solution

$$X(t) \approx X(0)e^{\lambda t} \quad \text{with } \lambda = f'(\hat{x})$$

Therefore, if $f'(\hat{x}) < 0$, then $X(t) \to 0$ as $t \to \infty$ and, hence, $x(t) = \hat{x} + X(t) \to \hat{x}$ as $t \to \infty$; that is, the solution will return to the equilibrium $\hat{x}$ after a small perturbation. In this case, $\hat{x}$ is locally stable. If $f'(\hat{x}) > 0$, then $X(t)$ will not go to 0, which implies that $\hat{x}$ is unstable. The linearization of $f(x)$ thus tells us whether an equilibrium is locally stable or unstable. We will also use linearization to determine the stability of equilibria of systems of differential equations.

Systems of Two Differential Equations. We consider a pair of differential equations of the form

$$\frac{dx_1}{dt} = f_1(x_1, x_2) \tag{11.62}$$

$$\frac{dx_2}{dt} = f_2(x_1, x_2) \tag{11.63}$$

or, in vector notation,

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}) \tag{11.64}$$

where $\mathbf{x} = \mathbf{x}(t) = [x_1(t), x_2(t)]^T$ and $\mathbf{f}(\mathbf{x})$ is a vector-valued function $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x})]^T$ with $f_i : \mathbf{R}^2 \to \mathbf{R}$. An equilibrium or critical point, $\hat{\mathbf{x}}$, of (11.64) satisfies

$$\mathbf{f}(\hat{\mathbf{x}}) = \mathbf{0}$$

For any point equilibrium, we look at what happens to a small perturbation of $\hat{\mathbf{x}}$ to determine the stability of $\hat{\mathbf{x}}$. We perturb $\hat{\mathbf{x}}$; that is, we look at how $\mathbf{x} = \hat{\mathbf{x}} + \mathbf{X}$ changes under the dynamics described by (11.64) assuming that $\mathbf{x}$ is very small:

$$\frac{d}{dt}(\hat{\mathbf{x}} + \mathbf{X}) = \frac{d}{dt}\mathbf{X} = \mathbf{f}(\hat{\mathbf{x}} + \mathbf{X})$$

The linearization of $\mathbf{f}(\mathbf{x})$ about $\mathbf{x} = \hat{\mathbf{x}}$ is

$$\mathbf{f}(\mathbf{x}) \approx \underbrace{\mathbf{f}(\hat{\mathbf{x}})}_{=0} + D\mathbf{f}(\hat{\mathbf{x}})\mathbf{X} = D\mathbf{f}(\hat{\mathbf{x}})\mathbf{X}$$

See Section 10.4

The matrix $D\mathbf{f}(\hat{\mathbf{x}})$ is the **Jacobi matrix** evaluated at $\hat{\mathbf{x}}$. If we approximate $\mathbf{f}(\hat{\mathbf{x}} + \mathbf{X})$ by its linearization $D\mathbf{f}(\hat{\mathbf{x}})\mathbf{X}$, then

$$\frac{d\mathbf{X}}{dt} = D\mathbf{f}(\hat{\mathbf{x}})\mathbf{X} \tag{11.65}$$

is the linear approximation of the dynamics of the perturbation $\mathbf{X}$. Equation (11.65) looks a little complicated, but note that $D\mathbf{f}(\hat{\mathbf{x}})$ is a 2×2 matrix of constants. (11.65) is a linear system of equations, just like the systems that we studied in Section 11.1.

The eigenvalues of the matrix $D\mathbf{f}(\hat{\mathbf{x}})$ allow us to determine the nature of the equilibrium. We emphasize that this is now a *local* analysis, just as in the case of a single differential equation, since we know that the linearization (11.65) is a good approximation only as long as we are sufficiently close to the point about which we linearized.

Following the same approach as in Section 11.1 we set:

$$\Delta = \det D\mathbf{f}(\hat{\mathbf{x}}) \qquad \text{and} \qquad \tau = \operatorname{tr} D\mathbf{f}(\hat{\mathbf{x}})$$

From Δ and τ we can classify a point equilibrium using the diagram in Figure 11.45. In most cases if we zoom in on the vector field of the nonlinear system (11.64) near $\hat{\mathbf{x}}$, it is the same as the vector field for the linearized system (11.65) and similarly the solution curves for (11.64) will look locally like the solution curves for (11.65). This result is known as the *Hartman-Grobman* theorem. However, if $D\mathbf{f}(\hat{\mathbf{x}})$ lies on any of the boundaries of the regions in Figure 11.45 (i.e., if $\tau^2 = 4\Delta$, or $\tau = 0$ and $\Delta > 0$, or if

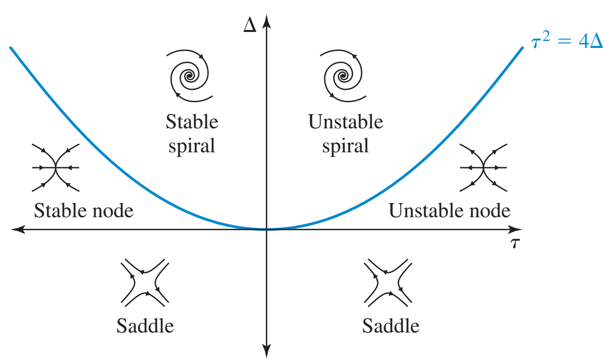


Figure 11.45 The stability behavior of a system of two autonomous equations.

$\Delta = 0$), then the linearized system and nonlinear systems may look different. In these cases the nonlinear terms cannot reasonably be neglected. There are additional methods that one can use to analyze the stability of equilibria from boundary regions, but these methods are beyond the scope of this book. You need to know that linearization cannot be trusted for these equilibria, and you may regard them as being unclassifiable for the time being.

The main challenge when identifying equilibria in nonlinear equations is that we must solve a system of equations to find all the points where $\mathbf{f}(\hat{\mathbf{x}}) = \mathbf{0}$. These equations will, in general, be nonlinear, so we cannot solve them using methods that we learned in Chapter 9. Typically we must use one of the equations to eliminate a variable; that is, we must rewrite the other equation in terms of a single variable. We may then solve the rewritten equation in a single variable. The process of solving the pair of equations is shown in Example 2.

EXAMPLE 2 Consider

$$\begin{aligned}\frac{dx_1}{dt} &= x_1 - 2x_1^2 - 2x_1x_2 \\ \frac{dx_2}{dt} &= 4x_2 - 5x_2^2 - 7x_1x_2\end{aligned} \tag{11.66}$$

(a) Find all equilibria of (11.66) and **(b)** analyze their stability.

Solution **(a)** To find equilibria, we set the right-hand sides of (11.66) equal to 0:

$$x_1 - 2x_1^2 - 2x_1x_2 = 0 \qquad \text{Two nonlinear equations} \tag{11.67}$$

$$4x_2 - 5x_2^2 - 7x_1x_2 = 0 \tag{11.68}$$

Factoring out x_1 in the first equation and x_2 in the second yields

$$x_1(1 - 2x_1 - 2x_2) = 0 \quad (R_1) \qquad \text{and} \qquad x_2(4 - 5x_2 - 7x_1) = 0 \quad (R_2)$$

That is, to solve (R_1) either

$$x_1 = 0 \qquad \text{or} \qquad 2x_1 + 2x_2 = 1$$

We then use either of these equations to eliminate x_1 from (R_2):

If $x_1 = 0$, then (R_2) implies $x_2(4 - 5x_2) = 0$, so either $x_2 = 0$ or $x_2 = 4/5$.

If $2x_1 + 2x_2 = 1$, then (R_2) implies $x_1 = \frac{1}{2} - x_2$. Then

$$\begin{aligned}x_2\left(4 - 5x_2 - 7\left(\frac{1}{2} - x_2\right)\right) &= 0 \\ = x_2\left(\frac{1}{2} + 2x_2\right) &= 0\end{aligned}$$

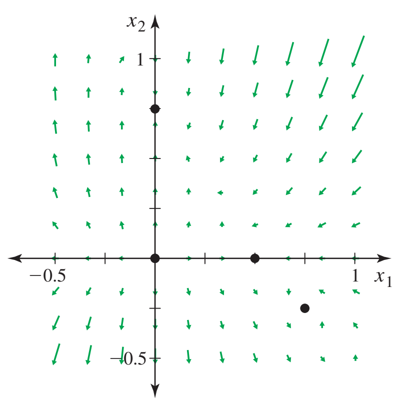


Figure 11.46 The vector field of (11.66) together with the equilibria.

so either $x_2 = 0$ or $x_2 = -\frac{1}{4}$. We can then find x_1 by substituting for x_2 in $x_1 = \frac{1}{2} - x_2$. If $x_2 = 0$, then $x_1 = \frac{1}{2} - 0 = \frac{1}{2}$, while if $x_2 = -\frac{1}{4}$, then $x_1 = \frac{1}{2} - \left(-\frac{1}{4}\right) = \frac{3}{4}$.

To summarize, there are four equilibria: at $(0, 0)$, $\left(0, \frac{4}{5}\right)$, $\left(\frac{1}{2}, 0\right)$, and $\left(\frac{3}{4}, -\frac{1}{4}\right)$.

We can illustrate all equilibria vector field of (11.66), which is displayed in Figure 11.46. The equilibria are shown as dots.

(b) To analyze the stability of the four equilibria, we compute the Jacobi matrix

$$D\mathbf{f} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix}$$

Since

$$f_1(x_1, x_2) = x_1 - 2x_1^2 - 2x_1x_2 \qquad \text{and} \qquad f_2(x_1, x_2) = 4x_2 - 5x_2^2 - 7x_1x_2,$$

we find that

$$D\mathbf{f}(x_1, x_2) = \begin{bmatrix} 1 - 4x_1 - 2x_2 & -2x_1 \\ -7x_2 & 4 - 10x_2 - 7x_1 \end{bmatrix}.$$

We will now go through the four cases and analyze each equilibrium:

(0, 0). The Jacobi matrix at (0, 0) is

$$D\mathbf{f}(0,0) = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}.$$

Since this matrix is in diagonal form, the eigenvalues are the diagonal elements, i.e., $\lambda_1 = 1$ and $\lambda_2 = 4$. Because both eigenvalues are positive, the equilibrium is unstable, so (0, 0) is an unstable node.

The linearization of the vector field about (0, 0) is displayed in Figure 11.47, where we show the vector field of

$$\frac{d\mathbf{X}}{dt} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \mathbf{X}(t)$$

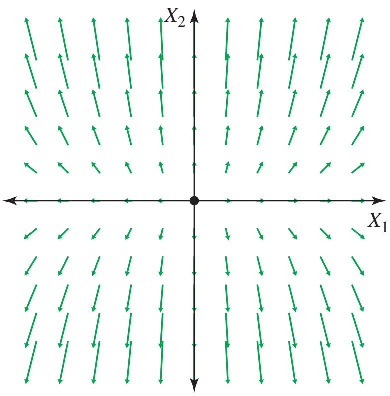


Figure 11.47 The linearization of the vector field about (0, 0).

If you now compare Figure 11.47 with the vector field of (11.66) shown in Figure 11.46, you will find that the linearized vector field and the direction field of (11.66) close to the equilibrium (0, 0) look the same.

$(0, \frac{4}{5})$. The Jacobi matrix at $(0, \frac{4}{5})$ is

$$D\mathbf{f}\left(0, \frac{4}{5}\right) = \begin{bmatrix} 1 - \frac{8}{5} & 0 \\ -\frac{28}{5} & 4 - 8 \end{bmatrix} = \begin{bmatrix} -\frac{3}{5} & 0 \\ -\frac{28}{5} & -4 \end{bmatrix}.$$

Since this matrix is in lower triangular form, the eigenvalues are the diagonal elements, i.e., $\lambda_1 = -\frac{3}{5}$ and $\lambda_2 = -4$. Because both eigenvalues are negative, $(0, \frac{4}{5})$ is locally stable. Using the same classification as in the linear case, we say that $(0, \frac{4}{5})$ is a stable node.

The linearization of the vector field about $(0, \frac{4}{5})$ is displayed in Figure 11.48. Again, the linearized vector field in Figure 11.48 looks like the vector field close to the equilibrium $(0, \frac{4}{5})$ in Figure 11.46.

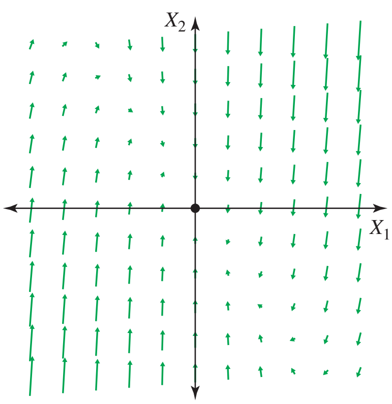


Figure 11.48 The linearization of the vector field about $(0, \frac{4}{5})$.

$(\frac{1}{2}, 0)$. The Jacobi matrix at $(\frac{1}{2}, 0)$ is

$$D\mathbf{f}\left(\frac{1}{2}, 0\right) = \begin{bmatrix} 1 - 2 & -1 \\ 0 & 4 - \frac{7}{2} \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 0 & \frac{1}{2} \end{bmatrix}$$

Since this matrix is in upper triangular form, the eigenvalues are the diagonal elements, i.e., $\lambda_1 = -1$ and $\lambda_2 = \frac{1}{2}$. Because one eigenvalue is positive and the other is negative, $(\frac{1}{2}, 0)$ is unstable. Using the same classification as in the linear case, we say that $(\frac{1}{2}, 0)$ is a saddle point.

The linearization of the vector field about $(\frac{1}{2}, 0)$ is displayed in Figure 11.49. The linearized vector field in Figure 11.49 looks like the vector field close to the equilibrium $(\frac{1}{2}, 0)$ in Figure 11.46.

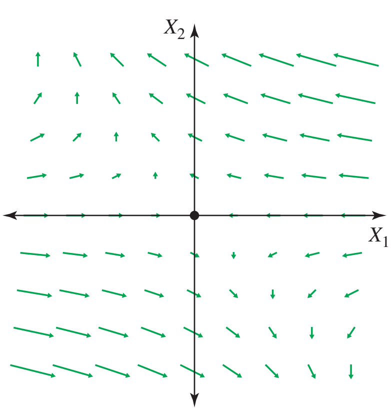


Figure 11.49 The linearization of the vector field about $(\frac{1}{2}, 0)$.

$(\frac{3}{4}, -\frac{1}{4})$. The Jacobi matrix is:

$$D\mathbf{f}\left(\frac{3}{4}, -\frac{1}{4}\right) = \begin{bmatrix} 1 - 3 + \frac{1}{2} & -\frac{3}{2} \\ \frac{7}{4} & 4 + \frac{10}{4} - \frac{21}{4} \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} & -\frac{3}{2} \\ \frac{7}{4} & \frac{5}{4} \end{bmatrix}.$$

To find the eigenvalues, we must solve

$$\det \begin{bmatrix} -\frac{3}{2} - \lambda & -\frac{3}{2} \\ \frac{7}{4} & \frac{5}{4} - \lambda \end{bmatrix} = 0.$$

Evaluating the determinant on the left-hand side and simplifying yields

$$\left(-\frac{3}{2} - \lambda\right)\left(\frac{5}{4} - \lambda\right) + \left(\frac{3}{2}\right)\left(\frac{7}{4}\right) = 0$$

$$\lambda^2 + \frac{3}{2}\lambda - \frac{5}{4}\lambda - \frac{15}{8} + \frac{21}{8} = 0$$

$$\lambda^2 + \frac{1}{4}\lambda + \frac{3}{4} = 0.$$

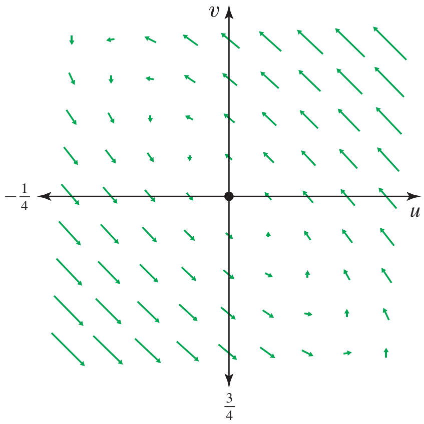


Figure 11.50 The linearization of the vector field about $(\frac{3}{4}, -\frac{1}{4})$.

Solving this quadratic equation, we find that

$$\lambda_{1,2} = \frac{-\frac{1}{4} \pm \sqrt{\frac{1}{16} - 3}}{2}$$
$$= -\frac{1}{8} \pm \frac{1}{8}\sqrt{-47} = -\frac{1}{8} \pm \frac{1}{8}i\sqrt{47}.$$

That is,

$$\lambda_1 = -\frac{1}{8} + \frac{1}{8}i\sqrt{47} \quad \text{and} \quad \lambda_2 = -\frac{1}{8} - \frac{1}{8}i\sqrt{47}.$$

The eigenvalues are complex conjugates with negative real parts. Thus, $(\frac{3}{4}, -\frac{1}{4})$ is locally stable; we expect the solutions to spiral into the equilibrium when we start close to the equilibrium.

The linearization of the vector field about $(\frac{3}{4}, -\frac{1}{4})$ is displayed in Figure 11.50. The linearized vector field in Figure 11.50 looks like the vector field close to the equilibrium $(\frac{3}{4}, -\frac{1}{4})$ in Figure 11.46. •

EXAMPLE 3 **Water Transport in Plants** Plants absorb water from soil into their roots. Water is transported from the roots through specially adapted conduits called xylem. In the leaves some fraction of the water is lost into the air by evaporation. Another fraction returns to the roots of the plant, carrying sugars that are made in the leaves and distributed through the rest of the plant or shared with fungi that are linked to the plants' roots. We will build a two-compartment model of water flows in the plant, treating the roots as one compartment, and the leaves as another compartment (see Figure 11.51).

Let's put numbers in the model, assuming water is absorbed into the roots at a constant rate of 1 g/hr. A fraction 0.1 of the water present in the roots is transported to the leaves each hour. In the leaves a fraction 0.2 of the water present is evaporated off per hour, while a fraction 0.05 of the water present is returned to the roots. Find any possible equilibria for the water present in the roots and leaves and determine whether these equilibria are stable or unstable.

Solution We use x_1 to represent the amount of water in the roots, and x_2 the amount in the leaves (both x_1 and x_2 are measured in grams). We can draw a diagram of the flows between roots as leaves (Figure 11.52) similar to the diagram we drew for Figure 11.40.

From the diagram we see that we can write a system of equations describing the flows of water in the system.

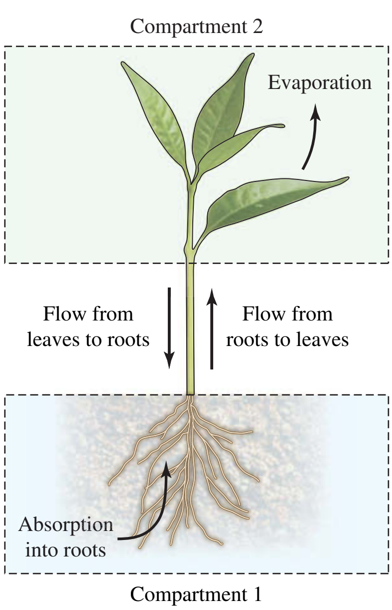


Figure 11.51 A two-compartment model for flow of water in a plant.

$$\frac{dx_1}{dt} = 1 - 0.1x_1 + 0.05x_2 \tag{11.69}$$
$$\frac{dx_2}{dt} = 0.1x_1 - 0.25x_2$$

Equation (11.69) is in the same form as Equation (11.40).

Previously we could only analyze two compartment models when there was no input term (I), making the equations homogeneous. However, the methods that we introduced for nonlinear equations also work for inhomogeneous linear equations.

First we identify all equilibria

$$\frac{dx_1}{dt} = 0 \Rightarrow 1 - 0.1x_1 + 0.05x_2 = 0$$
$$\text{or} \quad 0.1x_1 - 0.05x_2 = 1 \qquad (R_1)$$
$$\frac{dx_2}{dt} = 0 \Rightarrow 0.1x_1 - 0.25x_2 = 0 \qquad (R_2)$$

We can use (R_2) to eliminate x_2 from (R_1): $0.1x_1 - 0.25x_2 = 0$ implies $x_2 = \frac{0.1}{0.25}x_1 = 0.4x_1$.

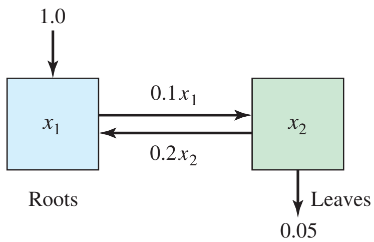


Figure 11.52 Flows of water between roots and leaves in Example 3.

Then (R_1) becomes

$$0.1x_1 - 0.05(0.4x_1) = 0.08x_1 = 1$$

or

$$x_1 = 12.5 \text{ g and } x_2 = 0.4x_1 = 5 \text{ g}$$

So there is a unique equilibrium: $(x_1, x_2) = (12.5, 5)$. Is the equilibrium stable? To find out we need to find the Jacobi matrix; since $f_1(x_1, x_2) = 1 - 0.1x_1 + 0.05x_2$ and $f_2(x_1, x_2) = 0.1x_1 - 0.25x_2$, we find:

$$D\mathbf{f} = \begin{bmatrix} -0.1 & 0.05 \\ 0.1 & -0.25 \end{bmatrix}. \qquad \text{f is linear, so } D\mathbf{f} \text{ is independent of } (x_1, x_2)$$

Now $\tau = \operatorname{tr}(D\mathbf{f}) = -0.35 < 0$, while $\Delta = \det(D\mathbf{f}) = (-0.1)(-0.25) - (0.05)(0.1) = 0.02$. Since $\tau < 0$ and $\Delta > 0$, and the equilibrium $(12.5, 5)$ is stable. In fact, because $\tau^2 - 4\Delta = (-0.35)^2 - 4 \cdot 0.02 = 0.0425 > 0$, the equilibrium is a stable node. •

11.3.2 Graphical Approach for 2 x 2 Systems

We have shown how to use linearization to understand how solutions behave near the point equilibria of a system of nonlinear equations. What other information can be gleaned from the system? In this subsection we will describe a graphical method for analyzing the behavior of solutions over the entire x_1–x_2 plane.

Suppose that

$$\frac{dx_1}{dt} = f_1(x_1, x_2)$$

$$\frac{dx_2}{dt} = f_2(x_1, x_2)$$

which in vector form is

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}).$$

The curves

$$f_1(x_1, x_2) = 0$$

$$f_2(x_1, x_2) = 0$$

are called **zero isoclines** or **null clines**, and they represent the points in the x_1–x_2 plane where either $\frac{dx_1}{dt} = 0$ or $\frac{dx_2}{dt} = 0$. This situation is illustrated in Figure 11.53 for a particular choice of f_1 and f_2. Let's assume that x_1 and x_2 are nonnegative; this restricts the discussion to the first quadrant of the x_1–x_2 plane.

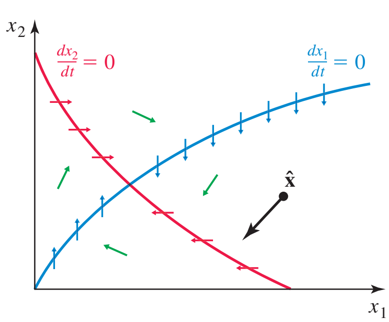


Figure 11.53 Graphical approach: zero isoclines and direction vectors.

The point where both null clines in Figure 11.53 intersect is a point equilibrium or critical point, which we call $\hat{\mathbf{x}}$. We can use the linearization method described in Subsection 11.3.1 to determine how solutions behave near this equilibrium point. What about in the rest of the x_1–x_2 plane? On the $f_1 = 0$ isoclines, $\frac{dx_1}{dt} = 0$, so on this isocline the direction vectors must point either vertically upward (if $f_2 > 0$) or vertically downward (if $f_2 < 0$). Similarly on the $f_2 = 0$ isocline, $\frac{dx_2}{dt} = 0$, so the direction vectors must point horizontally, either to the right (if $f_1 > 0$) or to the left (if $f_1 < 0$). Here is an important observation: if we find that $f_1 > 0$ at one point of an $f_2 = 0$ isocline, then, if we were to continue along the $f_2 = 0$ isocline, we would continue to find that $f_1 > 0$. In fact, $f_1 > 0$ until we reach an equilibrium. To see why, note that f_1 can only change its sign from $f_1 > 0$ to $f_1 < 0$ at a point where $f_1 = 0$. But if $f_1 = 0$ at a point on the $f_2 = 0$ isocline, that point is at equilibrium (see Figure 11.54). By the same reasoning, if a segment of the $f_2 = 0$ isocline does not contain an equilibrium, then f_1 must have the same sign over that entire segment.

Suppose we analyze the two zero-isoclines shown in Figure 11.53 and find the vector field directions on the zero-isoclines. What can we say about the vector field in regions of the x_1–x_2 plane between the zero-isoclines? Consider the point $\mathbf{x}$ shown in Figure 11.53. This point lies in a region bounded by two zero-isoclines, one with $\frac{dx_1}{dt} = 0$ and $\frac{dx_2}{dt} < 0$ (i.e., downward flow), and the other with $\frac{dx_2}{dt} = 0$ and $\frac{dx_1}{dt} < 0$ (i.e., flow to the left). We claim the direction vector at $\mathbf{x}$ is downward and to the left. In fact, at any point in this region of the x_1–x_2 plane the direction vector will be downward and to the left, and we show this by adding an arrow to the figure. We can argue similarly for the other three regions of the x_1–x_2 plane, adding arrows as shown in Figure 11.53.

How could we be so sure about the direction of the vector field at $\mathbf{x}$? (See Figure 11.53.) Consider a line that joins the point $\mathbf{x}$ to the $f_1 = 0$ isocline. As we travel along this line, $\frac{dx_2}{dt}$ can change continuously. Where the line meets the zero isocline, $\frac{dx_2}{dt} < 0$. So if $\frac{dx_2}{dt} > 0$ when we reach $\mathbf{x}$, then there must be a point on the line where $\frac{dx_2}{dt} = 0$ (since $\frac{dx_2}{dt}$ can only pass from negative to positive values by going through 0). But that cannot occur because the line does not cross a $f_2 = 0$ isocline. So we must have $\frac{dx_2}{dt} < 0$ at the point $\mathbf{x}$. Similarly we can argue that $\frac{dx_1}{dt} < 0$ at $\mathbf{x}$ by joining $\mathbf{x}$ to the $\frac{dx_2}{dt} = 0$ zero isocline. We can use the graphical approach (that is, finding the vector field on zero isoclines and then reconstructing the vector field in the regions between isoclines) to sketch vector fields for general systems of differential equations.

EXAMPLE 4 Use the graphical approach to analyze the vector field across the entire quadrant $x_1 \geq 0$, $x_2 \geq 0$, for the system of differential equations:

$$\frac{dx_1}{dt} = 5 - x_1 - x_1x_2 + 2x_2$$

$$\frac{dx_2}{dt} = x_1x_2 - 3x_2$$

Solution The zero isoclines satisfy either

$$\frac{dx_1}{dt} = 0, \quad \text{which holds if} \quad 5 - x_1 - x_1x_2 + 2x_2 = 0$$

$$\text{i.e., if} \quad x_2 = \frac{5 - x_1}{x_1 - 2} \qquad \text{solve for } x_2$$

or

$$\frac{dx_2}{dt} = 0, \quad \text{which holds if} \quad x_2(x_1 - 3) = 0,$$

i.e.,

$$\text{if} \quad x_2 = 0 \text{ or } x_1 = 3$$

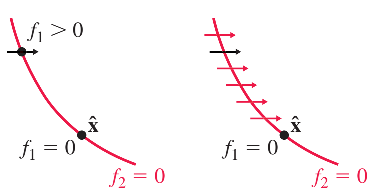


Figure 11.54 Suppose that at a point on the $f_2 = 0$ isocline, we find $f_1 > 0$. Then, if we walk along the isocline, it will continue to be the case that $f_1 > 0$, until we encounter an equilibrium.

The zero isoclines in the x_1–x_2 plane are drawn in Figure 11.55. The points where these zero isoclines intersect are the equilibria.

$$\text{If} \quad x_2 = 0, \text{ then } \frac{5 - x_1}{x_1 - 2} = 0, \text{ so } x_1 = 5$$

$$\text{whereas if} \quad x_1 = 3, \text{ then } x_2 = \frac{5 - (3)}{(3) - 2} = 2.$$

So there are two equilibria, at $(5, 0)$ and at $(3, 2)$. (Note that the two $f_2 = 0$ isoclines intersect at $(3, 0)$, but this is not an equilibrium because the $f_1 = 0$ isocline does not pass through the point.)

To understand the solution curves near the two equilibria, we study the linearized equations, starting with the Jacobi matrix:

$$D\mathbf{f}(x_1, x_2) = \begin{bmatrix} -1 - x_2 & -x_1 + 2 \\ x_2 & x_1 - 3 \end{bmatrix}.$$

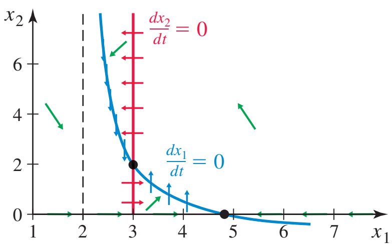


Figure 11.55 Using zero isoclines and a graphical method to sketch the vector field for Example 4.

So

$$D\mathbf{f}(5,0) = \begin{bmatrix} -1 & -3 \\ 0 & 2 \end{bmatrix}$$

This is an upper triangular matrix, so its eigenvalues are the diagonal entries; i.e., $\lambda_1 = -1$ and $\lambda_2 = 2$. Since $\lambda_1 < 0 < \lambda_2$, $(5, 0)$ is a saddle point. Meanwhile

$$D\mathbf{f}(3,2) = \begin{bmatrix} -3 & -1 \\ 2 & 0 \end{bmatrix} \qquad \tau = -3,\ \Delta = 2$$

In this case $\tau^2 - 4\Delta = (-3)^2 - 4(2) = 1 > 0$, so $(3, 2)$ is a stable node.

The zero isoclines of the system are shown in Figure 11.55. We fill in the directions of the vector field on those zero isoclines as follows.

On the $f_1 = 0$ isocline the vector field must be either vertically upward or downward, depending on whether f_2 is positive or negative. Since $f_2 = x_2(x_1 - 3)$ and $x_2 > 0$, $f_2 > 0$ (i.e., flow is upward) if $x_1 > 3$, and $f_2 < 0$ (i.e., flow is downward) if $x_1 < 3$.

Similarly, on the $f_2 = 0$ isoclines, the vector field is to the left if $f_1 < 0$ and to the right if $f_1 > 0$. On the $x_2 = 0$ isocline, $f_1 = 5 - x_1$. So $f_1 > 0$ (rightward flow) if $x_1 < 5$ and $f_1 < 0$ (leftward flow) if $x_1 > 5$. On the other $f_2 = 0$ isocline $(x_1 = 3)$, $f_1 = 5 - 3 - 3x_2 + 2x_2 = 2 - x_2$, so $f_1 > 0$ (rightward flow) if $x_2 < 2$, and $f_1 < 0$ (leftward flow) if $x_2 > 2$.

We show the direction of flow on the zero isoclines in Figure 11.55. Along with the x_1 and x_2-axes the zero isoclines divide the quadrant $x_1, x_2 \geq 0$ into four regions. Using the directions on the zero isoclines, we can add to Figure 11.55 arrows showing the vector field direction in each of these four regions. •

Section 11.3 Problems

11.3.1

In Problems 1–6, the point (0, 0) is always an equilibrium. Determine whether it is stable or unstable.

1. $\frac{dx_1}{dt} = x_1 + 2x_2 + 2x_1x_2$
$\frac{dx_2}{dt} = -x_1 + x_2 - x_1x_2$

2. $\frac{dx_1}{dt} = -x_1 + x_2 + x_1^2$
$\frac{dx_2}{dt} = x_2 + x_1^3$

3. $\frac{dx_1}{dt} = x_1 + x_1^2 - 2x_1x_2 + 3x_2$
$\frac{dx_2}{dt} = -x_1$

4. $\frac{dx_1}{dt} = 3x_1x_2 - x_1 + x_2$
$\frac{dx_2}{dt} = x_2^2 - x_1$

5. $\frac{dx_1}{dt} = \ln(1 + x_1 + x_2)$
$\frac{dx_2}{dt} = x_1 - x_2$

6. $\frac{dx_1}{dt} = -2\sin x_1$
$\frac{dx_2}{dt} = -x_2e^{x_1}$

In Problems 7–12, find all equilibria of each system of differential equations and determine the stability of each equilibrium.

7. $\frac{dx_1}{dt} = -x_1 + 2x_1(1 - x_1)$
$\frac{dx_2}{dt} = -x_2 + 5x_2(1 - x_1 - x_2)$

8. $\frac{dx_1}{dt} = 2x_1 - x_1^2 - 2x_2x_1$
$\frac{dx_2}{dt} = x_2 - 2x_2^2 - x_1x_2$

9. $\frac{dx_1}{dt} = 4x_1(1 - x_1) - 2x_1x_2$
$\frac{dx_2}{dt} = x_2(2 - x_2) - x_2$

10. $\frac{dx_1}{dt} = 2x_1(5 - x_1 - x_2)$
$\frac{dx_2}{dt} = 3x_2(7 - 3x_1 - x_2)$

11. $\frac{dx_1}{dt} = x_1x_2 - 2x_2$
$\frac{dx_2}{dt} = x_1 + x_2$

12. $\frac{dx_1}{dt} = x_1 - x_2$
$\frac{dx_2}{dt} = x_1x_2 - x_2$

13. Assume that $a > 0$. Find all point equilibria of

$$\frac{dx_1}{dt} = x_2(x_1 - a)$$
$$\frac{dx_2}{dt} = x_2^2 - x_1$$

and characterize their stability.

14. Assume that $a > 0$. Find all point equilibria of

$$\frac{dx_1}{dt} = 1 - ax_1x_2$$
$$\frac{dx_2}{dt} = ax_1x_2 - x_2$$

and characterize their stability.

15. **Nitrogen Flow in Plants** Plants depend on microbial partners (bacteria and fungi) within soil to provide them with nitrogen (in return, the plant supplies its microbial partners with sugars). We will build a two-compartment model for the available nitrogen in the system. Measure time, t, in days. Let x_1 represent the amount of nitrogen held by soil fungi, and let x_2 represent the amount of nitrogen held by the plant. Assume that in one day the fungi gain (from the atmosphere) 1 g of nitrogen. Assume also that in one day they pass 50% of their free nitrogen to the plant, and they also use up 30% of this nitrogen for growth. In the plant, 50% of the nitrogen is used up (e.g., for growth) each day, but no nitrogen is returned to the fungal partner.

(a) Draw a compartment diagram (like Figure 11.40) showing the flows of nitrogen between fungi and plant.

(b) Write down a system of differential equations for $x_1(t)$ and $x_2(t)$.

(c) Find all point equilibria for the system of differential equations that you wrote down in (b). Are these equilibria stable or unstable?

16. Insulin Pump An insulin pump is used to treat type I diabetes by continuously infusing insulin into the fat in a patient's abdomen or thigh. We will model this flow by a two-compartment model. We identify the fat into which insulin is pumped as the first compartment, and the patient's blood as the second compartment. Assume that the pump infuses 0.5 IU of insulin into the fat each hour. In one hour 10% of this insulin is eliminated from the fat (i.e., passes from the body without entering the patient's blood), and 70% is absorbed into blood. In the patient's blood 80% of insulin is metabolized (i.e., used up) by the patient's tissues each hour.

(a) Draw a compartment diagram (like Figure 11.40) showing the flow of insulin between fat and blood.

(b) Write down a system of differential equations for $x_1(t)$ and $x_2(t)$.

(c) Find all equilibria for the system of differential equations that you wrote down in (b). Are these equilibria stable or unstable?

11.3.2

In Problems 17–22, assume $x_1, x_2 \geq 0$. The figure corresponding to each problem shows the zero isoclines of a system of differential equations ($f_1 = 0$ isoclines are shown in blue and $f_2 = 0$ isoclines in red). Each figure also includes arrows showing the direction of the vector field on some portions of each zero isocline as well as in some of the regions between the zero isoclines. Use this information to complete the plot, i.e., to add arrows showing the vector field direction on every zero isocline and in all of the regions of the x_1–x_2 plane.

17. See Figure 11.56.

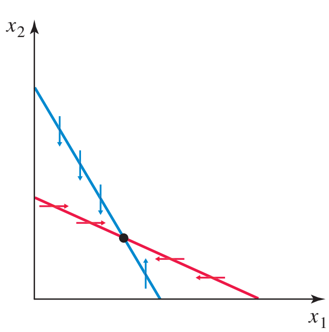


Figure 11.56 See Problem 17.

18. See Figure 11.57.

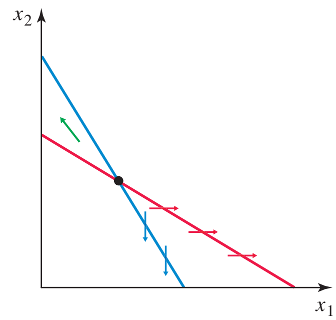


Figure 11.57 See Problem 18.

19. See Figure 11.58.

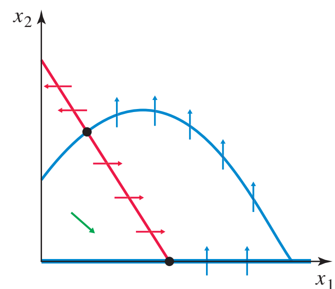


Figure 11.58 See Problem 19.

20. See Figure 11.59.

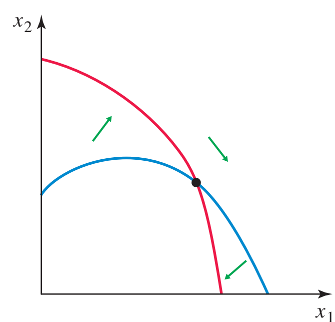


Figure 11.59 See Problem 20.

21. See Figure 11.60.

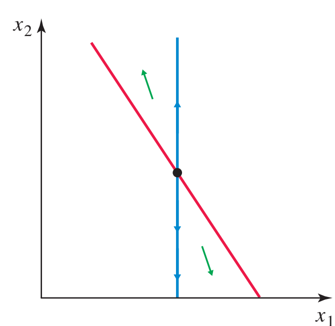


Figure 11.60 See Problem 21.

22. See Figure 11.61.

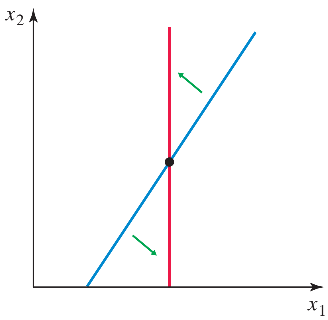


Figure 11.61 See Problem 23.

23. Assume that

$$\frac{dx_1}{dt} = x_1(10 - 2x_1 - x_2)$$

$$\frac{dx_2}{dt} = x_2(10 - x_1 - 2x_2)$$

(a) Graph the zero isoclines.

(b) Find all equilibria and classify them, by linearizing the system near each equilibrium.

(c) Draw the directions of the vector field on the zero isoclines, and in the regions between the zero isoclines.

24. Assume that

$$\frac{dx_1}{dt} = x_1(10 - x_1 - 2x_2)$$

$$\frac{dx_2}{dt} = x_2(10 - 2x_1 - x_2)$$

(a) Graph the zero isoclines.

(b) Find all equilibria and classify them, by linearizing the system near each equilibrium.

(c) Draw the directions of the vector field on the zero isoclines, and in the regions between the zero isoclines.

25. Let

$$\frac{dx_1}{dt} = x_1(2 - x_1) - x_1x_2$$

$$\frac{dx_2}{dt} = x_1x_2 - x_2$$

(a) Graph the zero isoclines.

(b) Find all equilibria and classify them, by linearizing the system near each equilibrium.

(c) Draw the directions of the vector field on the zero isoclines, and in the regions between the zero isoclines.

26. Let

$$\frac{dx_1}{dt} = x_1(2 - x_1^2) - x_1x_2$$

$$\frac{dx_2}{dt} = x_1x_2 - x_2$$

(a) Graph the zero isoclines.

(b) Find all equilibria and classify them, by linearizing the system near each equilibrium.

(c) Draw the directions of the vector field on the zero isoclines, and in the regions between the zero isoclines.

11.4 Nonlinear Systems: Lotka–Volterra Model for Interspecific Interactions

In this section we will show how the methods of Section 11.3 can be applied to an important class of ecosystem models called the Lotka–Volterra equations. These equations can be used to model the interactions between two different species, for example, two species competing for a limited resource, or a predator species that feeds on a prey species.

11.4.1 Competition

Imagine two species of plants growing together in the same plot. They both use similar resources: light, water, and nutrients and **compete** for these resources. Competition may result in reduced fecundity or reduced survivorship (or both).

The Lotka–Volterra model of interspecific competition incorporates density-dependent effects of competition in a way that extends the logistic equation to the case of two species. To describe it, we denote the population size of species 1 at time t by $N_1(t)$ and that of species 2 at time t by $N_2(t)$. Each species grows according to the logistic equation when the other species is absent. We denote their respective carrying capacities by K_1 and K_2, and their respective intrinsic rates of growth by r_1 and r_2. We assume that K_1, K_2, r_1, and r_2 are positive. Then, if there were no interactions between species, the population sizes would grow according to equations:

$$\frac{dN_1}{dt} = r_1N_1\left(1 - \frac{N_1}{K_1}\right)$$

$$\frac{dN_2}{dt} = r_2N_2\left(1 - \frac{N_2}{K_2}\right)$$

Without interactions, each species would grow to its carrying capacity; species 1 would grow to a size K_1 and species 2 would grow to a size K_2. Competition between the two species affects both populations. We measure the effect of species 1 on species 2 by the **competition coefficient** α_{21}; the effect of species 2 on species 1 is measured by the competition coefficient α_{12}. The Lotka–Volterra model of interspecific competition is then given by the following system of differential equations:

$$\begin{aligned} \frac{dN_1}{dt} &= r_1 N_1 \left(1 - \frac{N_1}{K_1} - \overbrace{\alpha_{12}\frac{N_2}{K_1}}^{\text{competition term}}\right) \\ \frac{dN_2}{dt} &= r_2 N_2 \left(1 - \frac{N_2}{K_2} - \underbrace{\alpha_{21}\frac{N_1}{K_2}}_{\text{competition term}}\right) \end{aligned} \tag{11.70}$$

The coefficients $r_1, r_2, K_1, K_2, \alpha_{12}$, and α_{21} are all positive. Different values of these coefficients are needed to fit the model to different species or different habitats.

Let's look at the first equation. If $N_2 = 0$, the first equation reduces to the logistic equation $dN_1/dt = r_1 N_1(1 - N_1/K_1)$, as mentioned. To understand the precise meaning of the competition coefficient α_{12}, observe that N_2 individuals of species 2 have the same effect on species 1 as $\alpha_{12}N_2$ individuals of species 1. The term $\alpha_{12}N_2$ thus converts N_2 species 2 individuals into "N_1-equivalents." For instance, set $\alpha_{12} = 0.2$ and assume that $N_2 = 20$; the effect of 20 individuals of species 2 on species 1 is the same as the effect of $(0.2)(20) = 4$ individuals of species 1 on species 1, because both $N_2 = 20$ and $N_1 = 4$ reduce the species 1 reproductive rate by the same amount. A similar interpretation can be attached to the competition coefficient α_{21} in the second equation.

We will use both zero isoclines and eigenvalues to analyze the model.

Zero Isoclines. The first step is to find the equations of the zero isoclines. First $\frac{dN_1}{dt} = 0$ implies

$$r_1 N_1 \left(1 - \frac{N_1}{K_1} - \alpha_{12}\frac{N_2}{K_1}\right) = 0$$

This equation is satisfied if either $N_1 = 0$ or

$$N_2 = \frac{K_1}{\alpha_{12}} - \frac{1}{\alpha_{12}} N_1. \tag{11.71}$$

If $\frac{dN_2}{dt} = 0$, then:

$$r_2 N_2 \left(1 - \frac{N_2}{K_2} - \alpha_{21}\frac{N_1}{K_2}\right) = 0$$

The solutions are $N_2 = 0$ and

$$N_2 = K_2 - \alpha_{21} N_1. \tag{11.72}$$

It makes sense that, if $N_1 = 0$, then $\frac{dN_1}{dt} = 0$ and if $N_2 = 0$, then $\frac{dN_2}{dt} = 0$—once a species is absent, it remains absent.

The other two isoclines, given by (11.71) and (11.72), are of particular interest because they tell us whether the two species can coexist stably. Both isoclines are straight lines in the N_1–N_2 plane, and there are four ways these isoclines can be arranged. These arrangements are illustrated in Figures 11.62 through 11.65, together with the corresponding vector fields. The solid dots in each figure are the equilibria. We see that there are always two equilibria $(K_1, 0)$ and $(0, K_2)$. At these equilibria only one species is present so they are referred to as **monoculture** equilibria. Non-trivial equilibria (that is, equilibria with both species present) occur if the lines in (11.71) and (11.72)

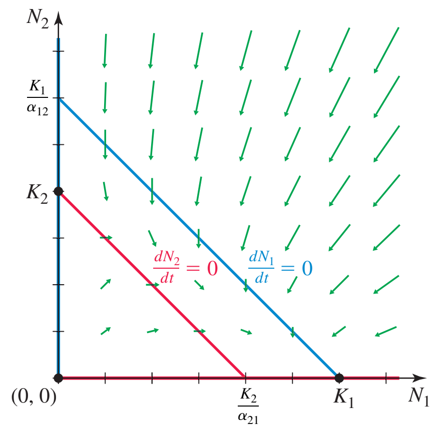


Figure 11.62 Case 1: Species 1 outcompetes species 2.

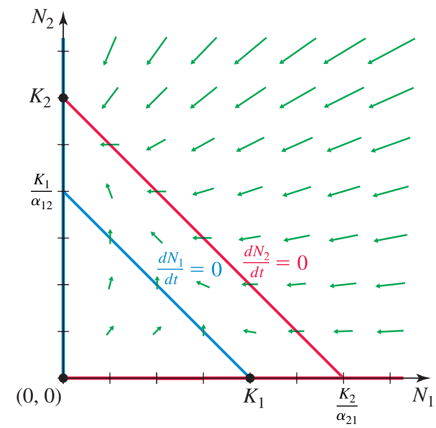


Figure 11.63 Case 2: Species 2 outcompetes species 1.

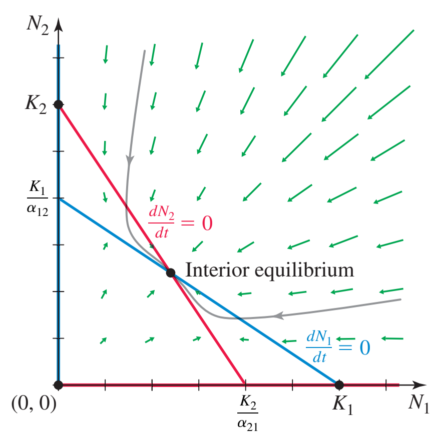


Figure 11.64 Case 3: Species 1 and 2 can coexist. All solution curves converge to the nontrivial equilibrium.

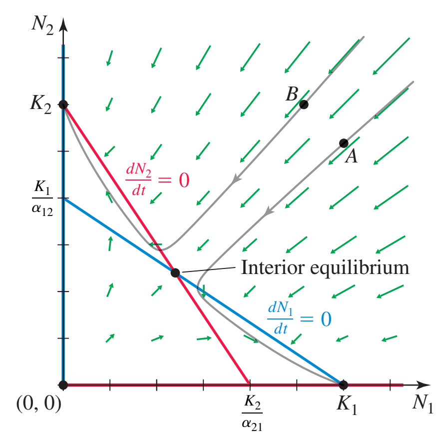


Figure 11.65 Case 4: Either species 1 or species 2 wins, depending on the initial condition.

intersect. There are four cases to consider because the $\frac{dN_1}{dt} = 0$ isocline in (11.71) intersects the N_1 axis at $N_1 = K_1$ and the N_2 axis at $N_2 = K_1/\alpha_{12}$, while the $\frac{dN_2}{dt} = 0$ isocline in (11.72) intersects these axes at $N_1 = K_2/\alpha_{21}$ and $N_2 = K_2$ respectively.

Case 1: $K_1 > \alpha_{12}K_2$ and $K_2 < \alpha_{21}K_1$ Because $K_1/\alpha_{12} > K_2$ and $K_1 > K_2/\alpha_{12}$, the $\frac{dN_1}{dt} = 0$ isocline lies above the $\frac{dN_2}{dt} = 0$ isocline. So the only points of intersection are the monoculture equilibria. From the vector field plot (Figure 11.62) we see that species 1 outcompetes species 2; i.e., unless $N_1 = 0$ initially, the populations converge to $(K_1, 0)$, meaning that species 2 becomes extinct.

Case 2: $K_1 < \alpha_{12}K_2$ and $K_2 > \alpha_{21}K_1$ This is the same as case 1, but with the roles of species 1 and 2 interchanged. Now $K_2 > K_1/\alpha_{12}$ and $K_2/\alpha_{21} > K_1$, so the $\frac{dN_2}{dt} = 0$ isocline lies above the $\frac{dN_1}{dt} = 0$ isocline, meaning that there are again only monoculture equilibria. We see from the vector field in Figure 11.63 that species 2 outcompetes species 1 and drives species 1 to extinction. The equilibrium $(K_1, 0)$ is therefore unstable, and the equilibrium $(0, K_2)$ is locally stable.

Case 3: $K_1 > \alpha_{12}K_2$ and $K_2 > \alpha_{21}K_1$ Then because $K_1/\alpha_{12} > K_2$, the $\frac{dN_2}{dt} = 0$ isocline intersects the N_2-axes above the $\frac{dN_1}{dt} = 0$ isocline. However, since $K_1 < K_2/\alpha_{21}$, the $\frac{dN_1}{dt} = 0$ isocline intersects the N_1-axis left of the $\frac{dN_2}{dt} = 0$ isocline. So these isoclines must intersect at some point in the upper quadrant, meaning that there is a non-trivial equilibrium, i.e., the two species can **coexist**. By linearization (see below) we see that the non-trivial equilibrium is stable; in fact, if N_1 and N_2 are initially non-zero, then all solution curves converge to the non-trivial equilibrium. The two monoculture equilibria $(K_1, 0)$ and $(0, K_2)$ are unstable.

Case 4: $K_1 < \alpha_{12}K_2$ and $K_2 < \alpha_{21}K_1$ Just as in case 3, the $\frac{dN_1}{dt} = 0$ and $\frac{dN_2}{dt} = 0$ isoclines intersect at a non-trivial equilibrium. However, this equilibrium is a saddle point (see its linearization below) and, hence, unstable. The outcome of competition depends on the initial densities. For instance, if the initial densities of N_1 and N_2 are given by the point A in the figure, then species 1 will win and species 2 will become extinct (following the solution curve that starts at A). If, however, the densities of N_1 and N_2 are given initially by the point B in the figure, then species 2 will win and species 1 will become extinct. Since the outcome of competition depends on initial abundances (i.e., the size of the founding populations of species 1 and 2) we refer to this scenario as **founder control**. In this scenario, both monoculture equilibria $(K_1, 0)$ and $(0, K_2)$ are locally stable.

We see from the preceding analysis that the Lotka–Volterra model allows for three possible outcomes in a two-species interaction. In cases 1 and 2 one species drives the other extinct. In case 3 coexistence is possible. In case 4, one or the other species eventually wins but which species wins depends on the initial abundances of the two species.

Interpreting the Conditions for Coexistence. We showed that the system of differential equations predicts qualitatively different behavior depending on whether K_1 is larger or smaller than $\alpha_{12}K_2$ and on whether K_2 is larger or smaller than $\alpha_{21}K_1$. Can we explain these conditions? Assuming that species 2 is initially not present, then species 1 will grow to its carrying capacity (i.e., K_1). Now imagine that once N_1 has reached its carrying capacity we add a small number of species 2 to the ecosystem. The per capita growth rate for species 2 is then:

$$r_2\left(1 - \frac{N_2}{K_2} - \alpha_{12}\frac{N_1}{K_2}\right) \approx r_2\left(1 - \alpha_{12}\frac{K_1}{K_2}\right) \qquad N_1 \approx K_1,\ N_2 \approx 0$$

This growth rate will be positive if

$$1 - \alpha_{12}\frac{K_1}{K_2} > 0 \quad \text{i.e., if} \quad K_2 > \alpha_{12}K_1.$$

If this condition is satisfied, then a small population of species 2 introduced into a group of species 1 will grow. We say that species 2 can **invade** species 1. Conversely, if $K_2 < \alpha_{12}K_1$, the small population of species 2 will decrease; species 2 cannot invade species 1.

Similarly a small population of species 1 can invade species 2 when N_1 is small and $N_2 = K_2$ if the initial per capita growth rate for species 1 is positive; i.e., if

$$r_1\left(1 - \frac{N_1}{K_1} - \alpha_{21}\frac{N_2}{K_1}\right) \approx r_1\left(1 - \alpha_{21}\frac{K_2}{K_1}\right) > 0 \qquad N_1 \approx 0, N_2 \approx K_2$$

so species 1 can invade species 2 if $K_1 > \alpha_{21}K_2$.

Let's re-examine some of the cases discussed above. In case 1, $K_1 > \alpha_{12}K_2$, so species 1 can invade species 2, but $K_2 < \alpha_{21}K_1$, so species 2 cannot invade species 1. Species 1 is therefore the *stronger* species; it drives species 2 to extinction.

In case 3, on the other hand, $K_1 > \alpha_{12}K_2$, while $K_2 > \alpha_{21}K_1$, so species 1 can invade species 2 and species 2 can invade species 1. Neither species is stronger, and both monoculture equilibria are unstable. We showed above that the two populations end up stably coexisting.

Linearization. In our graphical analysis we used the vector field to identify the stable and unstable equilibria of the model. However, looking at the vector field plots does not prove that given equilibria are indeed stable or unstable. To prove that an equilibrium is stable (or unstable) we must study the linearization of the system near the equilibrium. We will first determine all possible equilibria. Recall that if $dN_1/dt = 0$, then

$$N_1 = 0 \quad (R_1) \qquad \text{or} \qquad N_1 + \alpha_{12}N_2 = K_1 \quad (R_2)$$

Whereas $dN_2/dt = 0$, if:

$$N_2 = 0 \quad (R_3) \qquad \text{or} \qquad \alpha_{21}N_1 + N_2 = K_2 \quad (R_4)$$

There are four possible combinations:

1. (R_1) and (R_2): $(\hat{N}_1, \hat{N}_2) = (0, 0)$ corresponds to both species being absent.
2. (R_2) and (R_3): $(\hat{N}_1, \hat{N}_2) = (K_1, 0)$ corresponds to species 2 being absent and species 1 being at its carrying capacity K_1.
3. (R_1) and (R_4): $(\hat{N}_1, \hat{N}_2) = (0, K_2)$ corresponds to species 1 being absent and species 2 being at its carrying capacity K_2.
4. (R_2) and (R_4): The fourth equilibrium is obtained by simultaneously solving

$$\begin{aligned} N_1 + \alpha_{12}N_2 &= K_1 \quad (R_2) \\ \alpha_{21}N_1 + N_2 &= K_2 \quad (R_4) \end{aligned} \tag{11.73}$$

and requiring that both solutions be positive.

To analyze the stability of the equilibria, we must compute the Jacobi matrix associated with (11.70). We find that

$$D\mathbf{f}(N_1, N_2) = \begin{bmatrix} r_1 - 2\frac{r_1}{K_1}N_1 - \frac{r_1\alpha_{12}}{K_1}N_2 & -\frac{r_1\alpha_{12}}{K_1}N_1 \\ -\frac{r_2\alpha_{21}}{K_2}N_2 & r_2 - 2\frac{r_2}{K_2}N_2 - \frac{r_2\alpha_{21}}{K_2}N_1 \end{bmatrix}$$

1. At the trivial equilibrium $(\hat{N}_1, \hat{N}_2) = (0, 0)$, we

$$D\mathbf{f}(0, 0) = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}$$

which is a diagonal matrix; the eigenvalues are therefore

$$\lambda_1 = r_1 \qquad \text{and} \qquad \lambda_2 = r_2$$

Since $r_1 > 0$ and $r_2 > 0$, the equilibrium $(0, 0)$ is always unstable.

2. At the equilibrium $(\hat{N}_1, \hat{N}_2) = (K_1, 0)$, we obtain

$$D\mathbf{f}(K_1, 0) = \begin{bmatrix} -r_1 & -r_1\alpha_{12} \\ 0 & r_2\left(1 - \alpha_{21}\frac{K_1}{K_2}\right) \end{bmatrix}$$

Because $D\mathbf{f}(K_1, 0)$ is an upper triangular matrix, the eigenvalues are the diagonal entries:

$$\lambda_1 = -r_1 \qquad \text{and} \qquad \lambda_2 = r_2\left(1 - \alpha_{21}\frac{K_1}{K_2}\right)$$

Since $r_1 > 0$, it follows that $\lambda_1 < 0$. The eigenvalue $\lambda_2 < 0$ if $1 - \alpha_{21}\frac{K_1}{K_2} < 0$, i.e. when $K_2 < \alpha_{21}K_1$. Therefore, the equilibrium

$$(K_1, 0) \quad \text{is} \quad \begin{cases} \text{locally stable} & \text{for } K_2 < \alpha_{21}K_1 \\ \text{unstable} & \text{for } K_2 > \alpha_{21}K_1 \end{cases}$$

The conditions for stability are consistent with how we explained coexistence. If $K_2 > \alpha_{21}K_1$ then species 2 can invade species 1 (i.e., the monoculture equilibrium is unstable). If $K_2 < \alpha_{21}K_1$, species 2 cannot invade the monoculture equilibrium.

3. At the equilibrium $(\hat{N}_1, \hat{N}_2) = (0, K_2)$, we find that

$$D\mathbf{f}(0, K_2) = \begin{bmatrix} r_1\left(1 - \alpha_{12}\frac{K_2}{K_1}\right) & 0 \\ -r_2\alpha_{21} & -r_2 \end{bmatrix}$$

Because $D\mathbf{f}(0, K_2)$ is a lower triangular matrix, the eigenvalues are the diagonal entries:

$$\lambda_1 = r_1\left(1 - \alpha_{12}\frac{K_2}{K_1}\right) \qquad \text{and} \qquad \lambda_2 = -r_2$$

Since $r_2 > 0$, it follows that $\lambda_2 < 0$. The eigenvalue $\lambda_1 < 0$ if $1 - \alpha_{12}\frac{K_2}{K_1} < 0$, i.e. when $K_1 < \alpha_{12}K_2$. Therefore, the equilibrium

$$(0, K_2) \quad \text{is} \quad \begin{cases} \text{locally stable} & \text{for } K_1 < \alpha_{12}K_2 \\ \text{unstable} & \text{for } K_1 > \alpha_{12}K_2 \end{cases}$$

4. The fourth equilibrium can be obtained by simultaneously solving

$$N_1 + \alpha_{12}N_2 = K_1 \qquad (R_2)$$

$$\alpha_{21}N_1 + N_2 = K_2 \qquad (R_4)$$

It is possible to find the fourth equilibrium point, and also to find its stability analytically, for any values of the constants $\alpha_{12}, \alpha_{21}, r_1, r_2, K_1$, and K_2, but the details involve some complicated algebra. To avoid getting lost in the algebra, in cases 3 and 4 we will substitute in specific values for each of the constants.

Case 3 Let $r_1 = r_2 = 1$, $K_1 = 2$, $K_2 = 1$, $\alpha_{12} = 1$, and $\alpha_{21} = \frac{1}{3}$. Using these constants, $K_1 > \alpha_{12}K_2$ and $K_2 > \alpha_{21}K_1$, as is required for case 3 to be applicable. At the equilibrium representing coexistence, Equations (11.73) imply that

$$\hat{N}_1 + \hat{N}_2 = 2 \qquad (R_2)$$

$$\text{and} \qquad \frac{1}{3}\hat{N}_1 + \hat{N}_2 = 1 \qquad (R_4)$$

so

$$(R_2) - (R_4): \qquad \frac{2\hat{N}_1}{3} = 1,$$

which yields $\hat{N}_1 = \frac{3}{2}$ and $\hat{N}_2 = \frac{1}{2}$.

The Jacobi matrix is then

$$D\mathbf{f}(3/2, 1/2) = \begin{bmatrix} -3/4 & -3/4 \\ -1/6 & -1/2 \end{bmatrix} \qquad \tau = -5/4, \Delta = 1/4$$

Since $\Delta > 0$, $\tau < 0$, and $\tau^2 - 4\Delta = \frac{25}{16} - 1 = 9/16 > 0$, the equilibrium point at $(\hat{N}_1, \hat{N}_2) = (3/2, 1/2)$ is a stable node.

Case 4 Let $r_1 = r_2 = 1$, $K_1 = 2$, $K_2 = 1$, $\alpha_{12} = 5$, and $\alpha_{21} = 1$. Using these values for the constants, $K_1 < \alpha_{12}K_2$ and $K_2 < \alpha_{21}K_1$, as required for case 4 to be applicable. At the equilibrium representing coexistence, Equations (11.73) imply that:

$$\hat{N}_1 + 5\hat{N}_2 = 2 \qquad (R_2)$$

$$\hat{N}_1 + \hat{N}_2 = 1 \qquad (R_4)$$

So

$$(R_2) - (R_4) \qquad 4\hat{N}_2 = 1$$

which yields $\hat{N}_2 = 1/4$ and $\hat{N}_1 = 3/4$

$$Df(3/4.1/4) = \begin{bmatrix} -3/8 & -15/8 \\ -1/4 & -1/4 \end{bmatrix} \qquad \tau = -5/8, \Delta = -3/8$$

Because $\Delta < 0$, the equilibrium is a saddle point; in this case all solution curves converge to one of the two monoculture equilibria, but which monoculture equilibria a solution curve converges to depends on the starting values of N_1 and N_2.

11.4.2 A Predator–Prey Model

Lotka–Volterra models can also be developed for predator–prey systems. Here we designate one species to be the prey species, whose abundance is given by $N(t)$. The other species is the predator species, and its abundance is $P(t)$.

In the absence of predators we assume the prey population will grow exponentially with per capita growth rate r. On the other hand, if there are no prey present, then the predator species will die out. Let d be per capita death rate for the predators. In the absence of interactions between species, we then expect:

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

$$\text{and} \qquad \frac{dP}{dt} = -dP.$$

How do interactions between species affect the two populations? Again we modify the logistic equation to incorporate interactions. In the case of prey species, the presence of predators introduces a carrying capacity into the habitat. This carrying capacity is larger if there are fewer predators, and smaller if there are more predators. We can model this effect by introducing a term, $-aPN$, in the differential equation for $\frac{dN}{dt}$, where a is another positive coefficient. If P were simply a constant, we would then have a logistic equation with $\frac{dN}{dt} = rN\left(1 - \frac{aP}{r}N\right)$; that is, the carrying capacity for prey would be r/aP. The prey carrying capacity is larger if P is small (few predators) and smaller if P is large.

The presence of large numbers of prey boosts the reproduction rate of predators. Assume that each prey organism boosts the per capita growth rate of the predator species by some constant b (b is another positive coefficient). Then we must add a term, bNP, to the equation for dP/dt.

Putting these ingredients together we arrive at a system of differential equations

$$\begin{aligned} \frac{dN}{dt} &= rN - aPN \\ \frac{dP}{dt} &= -dP + bNP \end{aligned} \tag{11.74}$$

In these equations r, a, b, and d are all positive constants. We will treat these constants as arbitrary; using different values for these constants, the model can be fit to different species and to different habitats.

To find equilibria of (11.74), we set

$$N(r - aP) = 0$$

$$P(bN - d) = 0$$

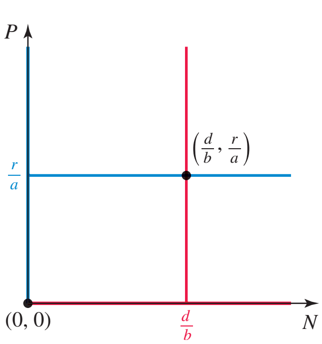


Figure 11.66 Zero isoclines.

This system yields the zero isoclines $N = 0$ and $P = r/a$ for $dN/dt = 0$ and $P = 0$ and $N = d/b$ for $dP/dt = 0$. The zero isoclines are shown in Figure 11.66. The points of intersection of the zero isoclines for N and P are equilibria. We find equilibria at:

$$(\hat{N}, \hat{P}) = (0, 0) \qquad \text{and} \qquad (\hat{N}, \hat{P}) = \left(\frac{d}{b}, \frac{r}{a}\right)$$

To analyze the stability of these two equilibria, we linearize (11.74). The Jacobi matrix is given by

$$D\mathbf{f}(N, P) = \begin{bmatrix} r - aP & -aN \\ bP & bN - d \end{bmatrix}$$

When $(\hat{N}, \hat{P}) = (0, 0)$,

$$D\mathbf{f}(0,0) = \begin{bmatrix} r & 0 \\ 0 & -d \end{bmatrix}$$

This is a diagonal matrix; hence, the eigenvalues are given by the diagonal elements

$$\lambda_1 = r > 0 \qquad \text{and} \qquad \lambda_2 = -d < 0$$

We conclude that $(0, 0)$ is unstable.

When $(\hat{N}, \hat{P}) = (\frac{d}{b}, \frac{r}{a})$,

$$D\mathbf{f}\left(\frac{d}{b}, \frac{r}{a}\right) = \begin{bmatrix} 0 & -\frac{ad}{b} \\ \frac{rb}{a} & 0 \end{bmatrix}$$

To determine the eigenvalues of $D\mathbf{f}(\frac{d}{b}, \frac{r}{a})$, we must solve

$$\det \begin{bmatrix} -\lambda & -\frac{ad}{b} \\ \frac{rb}{a} & -\lambda \end{bmatrix} = \lambda^2 + rd = 0$$

Solving this equation, we find that

$$\lambda_1 = i\sqrt{rd} \qquad \text{and} \qquad \lambda_2 = -i\sqrt{rd}$$

That is, both eigenvalues are purely imaginary. According to the linearized equations the point $(\frac{d}{b}, \frac{r}{a})$ is a center. But as we pointed out in Section 11.3, we cannot trust the linearized equations to correctly classify equilibria when the eigenvalues are purely imaginary. Fortunately, we can solve (11.74) exactly, and our derivation will enable us to show that the equilibrium truly is a center. We will give an outline of the proof skipping some of its steps.

To solve (11.74) exactly, we divide dP/dt by dN/dt:

$$\frac{dP/dt}{dN/dt} = \frac{dP}{dN} = \frac{P(bN-d)}{N(r-aP)}$$

This equation looks complicated, but it is separable. Separating variables and integrating, we obtain

$$\int \frac{r-aP}{P}\,dP = \int \frac{bN-d}{N}\,dN$$

Carrying out the integration gives

$$r\ln P - aP = bN - d\ln N + C_1$$

where C_1 is the constant of integration. Rearranging terms and exponentiating yields

$$\left(N^d e^{-bN}\right)\left(P^r e^{-aP}\right) = C \qquad C = e^{C_1}$$

We define the function

$$f(N, P) = \left(N^d e^{-bN}\right)\left(P^r e^{-aP}\right)$$

and set

$$g(N) = N^d e^{-bN} \qquad \text{and} \qquad h(P) = P^r e^{-aP}$$

so $f(N, P) = g(N)h(P)$.

Then we can show that $g(N)$ has its global maximum when $N = d/b$ and $h(P)$ has its global maximum when $P = r/a$ using the methods of Section 5.3 (this is the first step we will skip). The function $f(N, P)$ thus takes on its global maximum at the equilibrium point $(d/b, r/a)$. If the solution does not start at the equilibrium then for all points $(N(t), P(t))$ in the solution,

$$f(N(t), P(t)) = C$$

for some constant C. What this calculation means is that the solution curves are level curves of the function $f(N, P)$. The value of C for a particular level curve is found by evaluating $f(N, P)$ at the starting point of the solution curve $(N(0), P(0))$. Thereafter, the solution curve must stay on the level curve $f(N, P) = C$. Although we will not be able to demonstrate it here, these level curves are closed curves. (We show the level curves in Figure 11.67, to convince you that they are indeed closed.)

Solutions of $N(t)$ and $P(t)$ as functions of time corresponding to two of the closed curves in Figure 11.67 are shown in Figures 11.68 and 11.69, respectively. When we plot

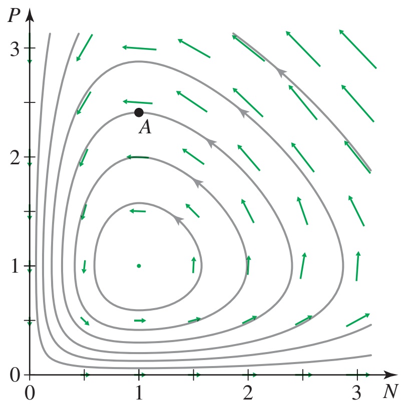


Figure 11.67 Solutions for (11.74) in the N–P plane.
[Here $a = b = r = d = 1$]

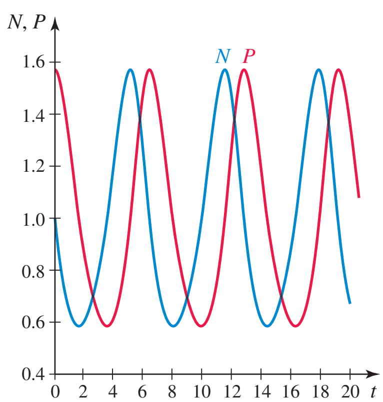


Figure 11.68 Solutions for (11.74) as functions of time.

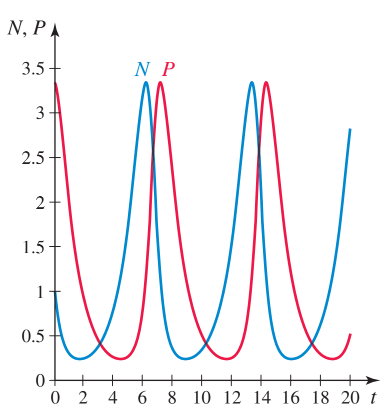


Figure 11.69 Solutions for (11.74) as functions of time.

$N(t)$ and $P(t)$ versus t, we see that the closed trajectories in the N–P plane correspond to periodic solutions for the predator and the prey. The amplitudes of the oscillations depend on the initial condition. We can understand the behaviors of the two populations as follows. Suppose our initial condition is the point marked A in Figure 11.67. At this point the number of prey is equal to $d/b = 1$ (for the values used in the figure). $N = d/b$ would be an equilibrium if we also had $P = r/a = 1$ (for the values used in the figure). But because $P > r/a$, the prey species will suffer more losses, due to predation, than occur when both populations are in equilibrium. Hence, the prey population, N, will start to decline. However, once N drops below d/b, the predators don't have enough prey to sustain their population size, so the predator population, P, also begins to decline. Eventually the number of predators drops below r/a, and the prey species can then recover. Predators continue to decline because the number of prey is still less than d/b. Once the number of prey grows past d/b, the predator species will start to recover (i.e., grow in size), but once P reaches r/a, over-predation again leads to a decline in N, and the system returns to the point A.

Notice how the sizes of the two populations in equilibrium depend on the coefficients in our model. The equilibrium prey population size is $\hat{N} = d/b$, i.e., it depends on the predator death rate, d, and the boost in predator reproduction rate, b, per prey organism but it does not depend on the prey growth rate, r, or on the coefficient a, which represents the impact of predation on the prey population. Similarly the equilibrium predator population size depends on the coefficients r and a (which show up in the equation for prey population) but not on the coefficients b and d that appear in the equation for dP/dt!

The Lotka–Volterra equations are able to partly predict the oscillations seen in real predator and prey populations. But under this model, if a small perturbation changes the value of N or P, the solution will follow a *different* closed trajectory. This property is a major drawback of the model; that if a natural population actually followed this simple model, abundances of the two species would vary irregularly because external factors would constantly shift the population from one cycle to another. If a natural population exhibits regular cycles, we would expect these cycles to be stable; that is, the population would return to the same cycle after a small perturbation. Such cycles are called **stable limit cycles**. A locally stable equilibrium that is approached by oscillations can be obtained by modifying the original Lotka–Volterra model to include prey carrying capacity (see Problems 17 through 19), but modification of the equations to produce a limit cycle is beyond the scope of this book.

Section 11.4 Problems

11.4.1

In Problems 1–4, use the graphical approach to classify the following Lotka–Volterra models of interspecific competition according to "coexistence," "founder control," "species 1 excludes species 2," or "species 2 excludes species 1."

1. $$\frac{dN_1}{dt} = 2N_1\left(1 - \frac{N_1}{10} - 0.2\frac{N_2}{10}\right)$$
$$\frac{dN_2}{dt} = 5N_2\left(1 - \frac{N_2}{15} - 0.5\frac{N_1}{15}\right)$$

2. $$\frac{dN_1}{dt} = 5N_1\left(1 - \frac{N_1}{50} - 0.2\frac{N_2}{50}\right)$$
$$\frac{dN_2}{dt} = N_2\left(1 - \frac{N_2}{30} - 1.2\frac{N_1}{30}\right)$$

3. $$\frac{dN_1}{dt} = N_1\left(1 - \frac{N_1}{20} - \frac{N_2}{5}\right)$$
$$\frac{dN_2}{dt} = 2N_2\left(1 - \frac{N_2}{15} - \frac{N_1}{3}\right)$$

4. $$\frac{dN_1}{dt} = 3N_1\left(1 - \frac{N_1}{25} - 1.2\frac{N_2}{25}\right)$$
$$\frac{dN_2}{dt} = N_2\left(1 - \frac{N_2}{30} - 0.8\frac{N_1}{30}\right)$$

In Problems 5–8, use the eigenvalue approach to analyze all equilibria of the given Lotka–Volterra models of interspecific competition.

5. $$\frac{dN_1}{dt} = 3N_1\left(1 - \frac{N_1}{18} - 1.3\frac{N_2}{18}\right)$$
$$\frac{dN_2}{dt} = 2N_2\left(1 - \frac{N_2}{20} - 0.6\frac{N_1}{20}\right)$$

6. $$\frac{dN_1}{dt} = 4N_1\left(1 - \frac{N_1}{12} - 0.3\frac{N_2}{12}\right)$$
$$\frac{dN_2}{dt} = 5N_2\left(1 - \frac{N_2}{15} - 0.2\frac{N_1}{15}\right)$$

7. $$\frac{dN_1}{dt} = N_1\left(1 - \frac{N_1}{35} - 3\frac{N_2}{35}\right)$$
$$\frac{dN_2}{dt} = 3N_2\left(1 - \frac{N_2}{40} - 4\frac{N_1}{40}\right)$$

8. $$\frac{dN_1}{dt} = N_1\left(1 - \frac{N_1}{25} - 0.1\frac{N_2}{25}\right)$$
$$\frac{dN_2}{dt} = N_2\left(1 - \frac{N_2}{28} - 1.2\frac{N_1}{28}\right)$$

9. Suppose that two species of beetles are reared together in one experiment and separately in another. When species 1 is reared alone, it reaches an equilibrium of about 200. When species 2 is reared alone, it reaches an equilibrium of about 150. When both of them are reared together, they seem to be able to coexist: Species 1 reaches an equilibrium of about 180 and species 2 reaches an equilibrium of about 80. If their densities follow the Lotka–Volterra equation of interspecific competition, find α_{12} and α_{21}.

10. Suppose that two species of beetles are reared together. Species 1 wins if there are initially 100 individuals of species 1 and 20 individuals of species 2. But species 2 wins if there are initially 20 individuals of species 1 and 100 individuals of species 2. When the beetles are reared separately, both species seem to reach an equilibrium of about 120. On the basis of this information and assuming that the densities follow the Lotka–Volterra model of interspecific competition, can you give lower bounds on α_{12} and α_{21}?

11.4.2

In Problems 11 and 12, use a graphing calculator to sketch solution curves of the given Lotka–Volterra predator–prey model in the N–P plane. That is, you should plot the level curves of the associated function $f(N, P)$.

T **11.** $$\frac{dN}{dt} = 2N - PN$$
$$\frac{dP}{dt} = 2PN - \frac{1}{2}P$$
passing through the points:

(a) $(N(0), P(0)) = (2, 2)$ **(b)** $(N(0), P(0)) = (3, 3)$

(c) $(N(0), P(0)) = (4, 4)$

T **12.** $$\frac{dN}{dt} = 3N - 2PN$$
$$\frac{dP}{dt} = PN - P$$
passing through the points:

(a) $(N(0), P(0)) = (1, 3/2)$ **(b)** $(N(0), P(0)) = (2, 2)$

(c) $(N(0), P(0)) = (3, 1)$

In Problems 13 and 14, we investigate the Lotka–Volterra predator–prey model.

13. Assume that
$$\frac{dN}{dt} = N - 4PN$$
$$\frac{dP}{dt} = 2PN - 3P$$

(a) Show that this system has two equilibria: the trivial equilibrium $(0, 0)$, and a nontrivial one in which both species have positive densities.

(b) Use the eigenvalue approach to show that the trivial equilibrium is unstable.

(c) Determine the eigenvalues corresponding to the nontrivial equilibrium. Does your analysis allow you to infer anything about the stability of this equilibrium?

(d) Use a graphing calculator to sketch curves in the N–P plane. Also, sketch solution curves of the prey and the predator densities as functions of time.

14. Assume that
$$\frac{dN}{dt} = 5N - PN$$
$$\frac{dP}{dt} = PN - P$$

(a) Show that this system has two equilibria: the trivial equilibrium $(0, 0)$, and a nontrivial one in which both species have positive densities.

(b) Use the eigenvalue approach to show that the trivial equilibrium is unstable.

(c) Determine the eigenvalues corresponding to the nontrivial equilibrium. Does your analysis allow you to infer anything about the stability of this equilibrium?

T **(d)** Use a graphing calculator to sketch curves in the N–P plane. Also, sketch solution curves of the prey and the predator densities as functions of time.

15. Biological Control Agent Assume that $N(t)$ denotes the density of an insect species at time t and $P(t)$ denotes the density of its predator at time t. The insect species is an agricultural pest, and its predator is used as a biological control agent. Their dynamics are given by the system of differential equations
$$\frac{dN}{dt} = 5N - 3PN$$
$$\frac{dP}{dt} = 2PN - P$$

(a) Explain why
$$\frac{dN}{dt} = 5N \tag{11.75}$$
describes the dynamics of the insect in the absence of the predator. Solve (11.75). Describe what happens to the insect population in the absence of the predator.

(b) Explain why introducing the insect predator into the system can help to control the density of the insect.

T **(c)** Assume that at the beginning of the growing season the insect density is 0.5 and the predator density is 2. You decide to control the insects by using an insecticide in addition to the predator. You are careful and choose an insecticide that does not harm the predator. After you spray, the insect density drops to 0.01 and the predator density remains at 2. Use a graphing calculator to investigate the long-term implications of your decision to spray the field. In particular, investigate what would have happened to the insect densities if you had decided not to spray the field, and compare your results with the insect density over time that results from your application of the insecticide.

16. Assume that $N(t)$ denotes prey density at time t and $P(t)$ denotes predator density at time t. Their dynamics are given by the system of equations

$$\frac{dN}{dt} = 4N - 2PN$$

$$\frac{dP}{dt} = PN - 3P$$

Assume that initially $N(0) = 3$ and $P(0) = 2$.

(a) If you followed this predator–prey community over time, what would you observe?

(b) Suppose that bad weather kills 90% of the prey population and 67% of the predator population. If you continued to observe this predator–prey community, what would you expect to see?

17. An unrealistic feature of the Lotka–Volterra model is that the prey exhibits unlimited growth in the absence of the predator. The model described by the following system remedies this shortcoming (in the model, we assume that the prey evolves according to logistic growth in the absence of the predator; the other features of the model are retained):

$$\begin{aligned}\frac{dN}{dt} &= 3N\left(1 - \frac{N}{10}\right) - 2PN \\ \frac{dP}{dt} &= PN - 4P\end{aligned} \tag{11.76}$$

(a) Explain why the prey evolves according to

$$\frac{dN}{dt} = 3N\left(1 - \frac{N}{10}\right) \tag{11.77}$$

in the absence of the predator. Investigate the long-term behavior of solutions to (11.77).

(b) Find all equilibria of (11.76), and use the eigenvalue approach to determine their stability.

T **(c)** Use a graphing calculator to sketch the solution curve of (11.76) in the N–P plane when $N(0) = 2$ and $P(0) = 2$. Also, sketch $N(t)$ and $P(t)$ as functions of time, starting with $N(0) = 2$ and $P(0) = 2$.

18. An unrealistic feature of the Lotka–Volterra model is that the prey exhibits unlimited growth in the absence of the predator. The model described by the following system remedies this shortcoming (in the model, we assume that the prey evolves according to logistic growth in the absence of the predator; the other features of the model are retained):

$$\begin{aligned}\frac{dN}{dt} &= N\left(1 - \frac{N}{K}\right) - 4PN \\ \frac{dP}{dt} &= PN - 5P\end{aligned} \tag{11.78}$$

Here, $K > 0$ denotes the carrying capacity of the prey in the absence of the predator. In what follows, we will investigate how the carrying capacity affects the outcome of this predator–prey interaction.

(a) Draw the zero isoclines of (11.78) for **(i)** $K = 10$ and **(ii)** $K = 3$.

(b) When $K = 10$, the zero isoclines intersect, indicating the existence of a nontrivial equilibrium. Analyze the stability of this nontrivial equilibrium.

(c) Is there a minimum carrying capacity required in order to have a nontrivial equilibrium? If yes, find it and explain what happens when the carrying capacity is below this minimum and what happens when the carrying capacity is above this minimum.

19. An unrealistic feature of the Lotka–Volterra model is that the prey exhibits unlimited growth in the absence of the predator. The model described by the following system remedies this shortcoming (in the model, we assume that the prey evolves according to logistic growth in the absence of the predator; the other features of the model are retained):

$$\begin{aligned}\frac{dN}{dt} &= N\left(1 - \frac{N}{20}\right) - 5PN \\ \frac{dP}{dt} &= 2PN - 8P\end{aligned} \tag{11.79}$$

(a) Draw the zero isoclines of (11.79).

(b) Use linearization to determine whether the nontrivial equilibrium is locally stable.

In Problems 20–24, we will analyze how a change in parameters in the modified Lotka–Volterra predator–prey model

$$\begin{aligned}\frac{dN}{dt} &= rN\left(1 - \frac{N}{K}\right) - aPN \\ \frac{dP}{dt} &= bPN - dP\end{aligned} \tag{11.80}$$

affects predator–prey interactions.

20. Find the zero isoclines of (11.80), and determine conditions under which a nontrivial equilibrium (i.e., an equilibrium in which both prey and predator have positive densities) exists.

In Problems 21–24, we use the results of Problem 20. Assume that the parameters are chosen so that a nontrivial equilibrium exists.

21. Use the results of Problem 20 to show that an increase in r (the intrinsic rate of growth of the prey) results in an increase in the predator density, but leaves the prey density unchanged.

22. Use the results of Problem 20 to show that an increase in a (the increase in prey death per predator present) reduces predator abundance, but has no effect on the equilibrium abundance of the prey.

23. Use the results of Problem 20 to show that an increase in b (the boost in predator reproductive rate per prey organism) reduces the prey equilibrium abundance and increases the predator equilibrium abundance.

24. Use the results of Problem 20 to show that an increase in K (the prey carrying capacity in the absence of the predator) increases the predator equilibrium abundance, but has no effect on the prey equilibrium abundance.

11.5 More Mathematical Models

The methods we have learned for studying systems of differential equations are very general. In this section, we will describe some more mathematical models that can be analyzed using these methods including models for ecosystem stability, the firing of neurons, enzymatic reactions, and the spread of diseases.

11.5.1 The Community Matrix

In this subsection, we consider a fairly general multispecies population model, first considered by ***Levins (1970)*** and ***May (1975)***. The goal is to determine how interactions between pairs of species can influence the stability of the equilibria of an ecosystem containing different species.

We assume an ecosystem contains two species, and the abundance of species i at time t is given by $N_i(t)$. Suppose that the following set of differential equations describes the dynamics of our two-species ecosystem:

$$\begin{aligned} \frac{dN_1}{dt} &= f_1(N_1(t), N_2(t)) \\ \frac{dN_2}{dt} &= f_2(N_1(t), N_2(t)) \end{aligned} \tag{11.81}$$

In vector notation,

$$\frac{d\mathbf{N}}{dt} = \mathbf{f}(\mathbf{N}(t)) \qquad N(t) = \begin{bmatrix} N_1(t) \\ N_2(t) \end{bmatrix} \tag{11.82}$$

To determine equilibria, we must solve the system of equations

$$\begin{aligned} f_1(N_1, N_2) &= 0 \\ f_2(N_1, N_2) &= 0 \end{aligned} \tag{11.83}$$

We assume that (11.83) has a nontrivial solution $\hat{\mathbf{N}} = (\hat{N}_1, \hat{N}_2)$ with $\hat{N}_1 > 0$ and $\hat{N}_2 > 0$. We can determine the stability of this equilibrium. To do so, we must evaluate the Jacobi matrix associated with the system (11.81) at the equilibrium $(\hat{N}_1, \hat{N}_2)$. The Jacobi matrix at the equilibrium $\hat{\mathbf{N}}$ is given by

$$D\mathbf{f}(\hat{\mathbf{N}}) = \begin{bmatrix} \frac{\partial f_1(\hat{N}_1, \hat{N}_2)}{\partial N_1} & \frac{\partial f_1(\hat{N}_1, \hat{N}_2)}{\partial N_2} \\ \frac{\partial f_2(\hat{N}_1, \hat{N}_2)}{\partial N_1} & \frac{\partial f_2(\hat{N}_1, \hat{N}_2)}{\partial N_2} \end{bmatrix}$$

This 2×2 matrix is called the **community matrix**. Its elements

$$a_{ij} = \frac{\partial f_i(\hat{N}_1, \hat{N}_2)}{\partial N_j}$$

describe the effect of species j on species i at equilibrium, because the partial derivative $\frac{\partial f_i}{\partial N_j}$ tells us how the function f_i, which describes the rate of growth of species i, changes when the abundance of species j changes.

The diagonal elements $a_{ii} = \frac{\partial f_i}{\partial N_i}$ measure the effect species i has on itself, whereas the off-diagonal elements $a_{ij} = \frac{\partial f_i}{\partial N_j}$, $i \neq j$, measure the effect species j has on species i. The signs of the elements a_{ij} thus tell us something about the pairwise effects the species in this ecosystem have on each other at equilibrium.

The quantity a_{ij} can be negative, 0, or positive. If $a_{ij} < 0$, then the growth rate of species i is decreased if species j increases its abundance; we therefore say that species j has a negative, or **inhibitory**, effect on species i. If $a_{ij} = 0$, then changes in the abundance of species j have no effect on the growth rate of species i. If $a_{ij} > 0$, then the growth rate of species i is increased if species j increases its abundance; we then say that species j has a positive, or **facilitory**, effect on species i.

We look at the possible combinations of the pair (a_{21}, a_{12}). This pair describes the interactions *between* the two species in the ecosystem. The following table lists all possible combinations:

		a_{12}		
		+	**0**	**−**
	+	$++$	$+0$	$+-$
a_{21}	**0**	$0+$	00	$0-$
	−	$-+$	-0	$--$

To interpret the table, take the pair (00), for instance. The pair (00) represents the case in which neither species has an effect on the other species at equilibrium. With another pair—for instance, (0+)—$a_{21} = 0$ and species 1 has no effect on species 2, but $a_{12} > 0$ and species 2 has a positive effect on species 1.

The case (00) is the simplest, and we will discuss it first. The community matrix in this case is

$$\begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}$$

Since the community matrix is in diagonal form, its eigenvalues are the diagonal elements a_{11} and a_{22}. Hence, the equilibrium $(\hat{N}_1, \hat{N}_2)$ is stable only if both a_{11} and a_{22} are negative. That is, if neither species has an effect on the other species, then a locally stable nontrivial equilibrium in which both species coexist exists only if they each have a negative effect on themselves; this means that each species needs to regulate its own population size.

Following ***May (1975)***, the remaining eight combinations in the table can be categorized into five biologically distinct types of interactions:

Mutualism, or symbiosis (++): Each species has a positive effect on the other.

Competition (−−): Each species has a negative effect on the other.

Commensalism (+0): One species benefits from the interaction, whereas the other is unaffected.

Amensalism (−0): One species is harmed by the interaction, whereas the other is unaffected.

Predation (+−): One species benefits, whereas the other is harmed.

We will now discuss the stability of the nontrivial equilibrium $(\hat{N}_1, \hat{N}_2)$ in all five cases. Recall that we assumed that this equilibrium exists and that both $\hat{N}_1 > 0$ and $\hat{N}_2 > 0$. The community matrix [the Jacobi matrix of (11.81) at equilibrium] is of the form

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Both eigenvalues of A have negative real parts if and only if

$$\operatorname{tr} A = a_{11} + a_{22} < 0 \qquad \text{and} \qquad \det A = a_{11}a_{22} - a_{12}a_{21} > 0$$

In what follows, we assume that

$$a_{11} < 0 \qquad \text{and} \qquad a_{22} < 0 \tag{11.84}$$

so that the first condition $\operatorname{tr} A < 0$ is automatically satisfied. This has the same interpretation as discussed in the case (00): both species have a negative effect on themselves or regulate their own population densities.

We will now go through all five cases and determine under which conditions the nontrivial equilibrium is stable:

Mutualism We assume (11.84). The sign structure of the community matrix at equilibrium in the case of mutualism is then of the form

$$A = \begin{bmatrix} - & + \\ + & - \end{bmatrix}$$

Since

$$\det A = \underbrace{a_{11}a_{22}}_{>0} - \underbrace{a_{12}a_{21}}_{>0}$$

the determinant of A may be either positive or negative. If $a_{12}a_{21}$ is sufficiently small compared with $a_{11}a_{22}$, then $\det A > 0$ and the equilibrium $(\hat{N}_1, \hat{N}_2)$ is locally stable. In other words, if the positive effects of the species on each other are sufficiently counteracted by their own population control (represented by a_{11} and a_{22}), then the equilibrium is locally stable.

Competition Again, we assume (11.84). The sign structure of the community matrix at equilibrium in the case of competition is then of the form

$$A = \begin{bmatrix} - & - \\ - & - \end{bmatrix}$$

Since

$$\det A = \underbrace{a_{11}a_{22}}_{>0} - \underbrace{a_{12}a_{21}}_{>0}$$

the determinant of A may be either positive or negative. Now, $\det A > 0$, meaning the equilibrium $(\hat{N}_1, \hat{N}_2)$ is locally stable, if the negative effects each species has on the other are smaller than the effects each species has on itself.

Commensalism and Amensalism Once more, assume (11.84). The signs in the community matrices at equilibrium are then of the form

$$A = \begin{bmatrix} - & 0 \\ + & - \end{bmatrix} \qquad \text{or} \qquad A = \begin{bmatrix} - & 0 \\ - & - \end{bmatrix}$$

In either case, the determinant is positive; therefore, the equilibrium is locally stable.

Predation Yet again, we assume (11.84). The sign structure of the community matrix at equilibrium in the case of predation is then of the form

$$A = \begin{bmatrix} - & - \\ + & - \end{bmatrix}$$

and $\det A$ is always positive. That is, provided that a nontrivial equilibrium exists, it is certainly locally stable.

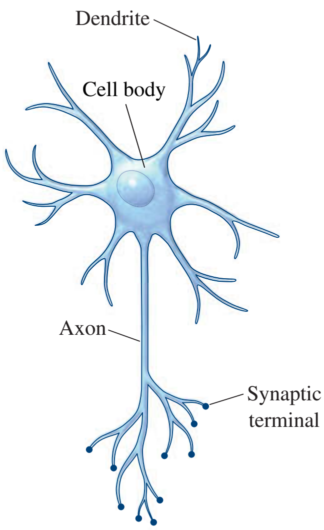


Figure 11.70 A typical vertebrate neuron.

11.5.2 Neuron Activity

Bio Info • The nervous system of an organism is a communication network that allows the rapid transmission of information between cells. The nervous system of an animal consists of nerve cells called **neurons**. A typical neuron has a cell body that contains the cell nucleus and nerve fibers. Nerve fibers that receive information are called **dendrites**, whereas those that transport information are called **axons**; the latter provide links to other neurons via **synapses**. A typical vertebrate neuron is shown in Figure 11.70.

Neurons respond to electrical stimuli, a property that is exploited by scientists studying them. When the cell body of an isolated neuron is stimulated with a very mild electrical shock, the neuron shows no response; increasing the intensity of the shock beyond a certain threshold, however, will trigger a response, namely, an impulse that travels along the axon. Increasing the intensity of the electrical shock further does not change the response. The impulse is thus an all-or-nothing response.

Let us briefly examine how a neuron works. The main players in the functioning of a neuron are sodium (Na^+) and potassium (K^+) ions. The cell membrane of a neuron is impermeable to these ions when the cell is in a resting state. In a typical neuron in its resting state, the concentration of Na^+ in the interior of the cell is about one-tenth of the extracellular concentration of Na^+ and the concentration of K^+ in the interior of the cell is about 30 times the extracellular concentration of K^+. When the neuron is in its resting state, the interior of the cell is negatively charged (at -70 mV) relative to the exterior of the cell.

When a nerve cell is stimulated, Na^+ ions rush into the cell and K^+ ions move from the inside to the outside of the cell. This movement of ions leaves the cell interior at a different voltage from when the flow of ions first started. To restore the original concentration of Na^+ and K^+ energy must be expended to run sodium and potassium pumps on the surface of the cell to pump the excess Na^+ from the interior to the exterior of the cell and to pump K^+ from the exterior to the interior of the cell.

We can think of the cell as being a **bistable** system; that is, it can be in either of two states—either with the Na^+ ions outside and K^+ ions inside, or the other way around. Both of these states are stable; we call the system **excitable**—it needs to be excited (by some external form of stimulation) in order to transition between the two states. The neuron will only change between the two states if it receives a strong enough stimulation.

Fitzhugh (1961) and ***Nagumo et al. (1962)*** derived a model that captures these effects, namely the existence of two equilibria, and the need for a large enough excitation for the system to transition between the states.

The Fitzhugh–Nagumo model is described by two variables. One variable, denoted by V, describes the voltage difference between the inside and outside of the neuron the cell surface. It is a measure of the net difference in positive or negative charge between the inside of the cell and its outside. The other variable, denoted by w, models the sodium and potassium channels which allow ions to flow into or out of the cell. The equations are

$$\begin{aligned}\frac{dV}{dt} &= -V(V-a)(V-1) - w \\ \frac{dw}{dt} &= V - cw\end{aligned} \tag{11.85}$$

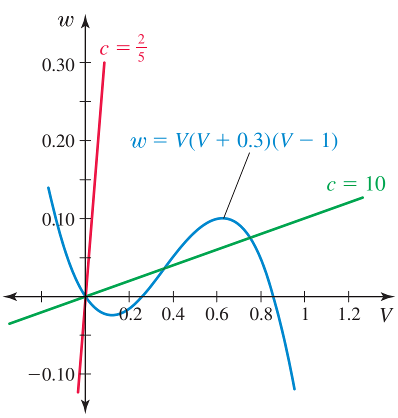


Figure 11.71 The zero isoclines of (11.79), assuming $a = 0.3$. $\frac{dw}{dt} = 0$ isoclines are drawing for two different values of the constant c; $c = 2/5$ (red) and $c = 10$ (green)

Note that these equations contain two unknown coefficients, a and c, which are constants that satisfy $0 < a < 1$, and $c > 0$.

We will analyze the system graphically to see what kinds of behaviors it can predict, and how these behaviors depend on the constants a and c. The zero isoclines of (11.85) are given by

$$w = -V(V-a)(V-1) \qquad \text{and} \qquad w = \frac{V}{c}$$

The important feature of this model is that the zero isocline $dV/dt = 0$ is a cubic function while the $\frac{dw}{dt} = 0$ isocline is a straight line. The zero isoclines are illustrated in Figure 11.71.

Increasing the value of c decreases the slope of the $\frac{dw}{dt} = 0$ isocline. If this slope is large (i.e., if c is small), there is just one equilibrium, namely, $(0, 0)$, whereas when c is sufficiently large, the line $w = V/c$ intersects the graph of $dV/dt = 0$ three times. There are therefore three equilibria, but only if c is sufficiently large.

We can analyze the stability of $(0, 0)$ by linearizing the system about this equilibrium. We find that

$$D\mathbf{f}(V, w) = \begin{bmatrix} -3V^2 + 2V + 2aV - a & -1 \\ 1 & -c \end{bmatrix}$$

Hence,

$$D\mathbf{f}(0, 0) = \begin{bmatrix} -a & -1 \\ 1 & -c \end{bmatrix}$$

Hence

$$\text{trace}\,(D\mathbf{f}(0, 0)) = -(a + c) < 0 \qquad \text{since } a, c > 0$$

and

$$\det(D\mathbf{f}(0, 0)) = ac + 1 > 0$$

So, for all values of a and c, $(0, 0)$ is a stable equilibrium. In fact, if c is small, so that $(0, 0)$ is the only equilibrium, then whatever initial conditions the neuron starts with, it returns to a resting state with $(V, w) = (0, 0)$.

What happens at larger values of c? As we can see from Figure 11.71, there are then three different equilibria. The equilibria obey:

$$\begin{aligned}-V(V-a)(V-1) &= w \\ w &= V/c\end{aligned}$$

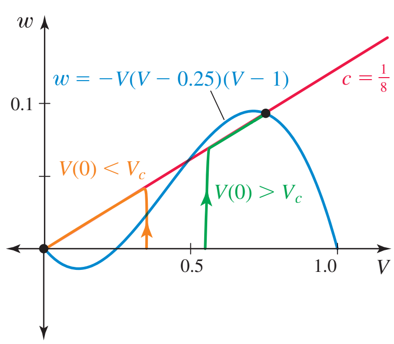


Figure 11.72 Solution curves for the Fitzhugh–Nagumo model.

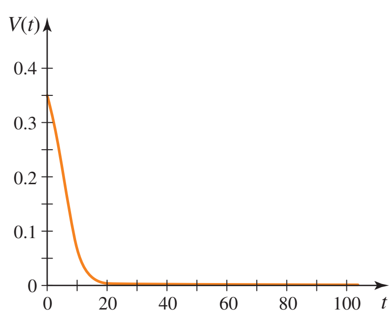


Figure 11.73 When $V(0) < V_c$, the potential V dies away.

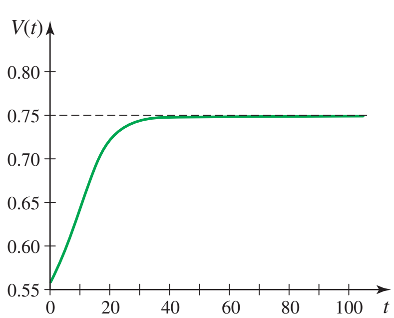


Figure 11.74 When $V(0) > V_c$, the neuron fires.

So using the second equation to eliminate w from the first equation, $V(V-a)(V-1) = -V/c$, i.e., either $V = 0$ or: $(V-a)(V-1) = -1/c$. Solving this equation for arbitrary coefficients a and c is possible, but the algebra is somewhat messy. Just as we did in Section 11.4.1, we will simplify the algebra by considering specific values for a and c. Let $a = 1/4$ and $c = 8$; then at the equilibria we either have $\hat{V} = 0$ or:

$$\left(\hat{V} - \frac{1}{4}\right)(\hat{V} - 1) = -\frac{1}{8}$$

$$\text{so} \qquad \hat{V}^2 - \frac{5\hat{V}}{4} + \frac{3}{8} = 0$$

$$\left(\hat{V} - \frac{3}{4}\right)\left(\hat{V} - \frac{1}{2}\right) = 0$$

Additional equilibria therefore occur at $\hat{V} = 1/2$ and $\hat{V} = 3/4$ with respective w-values $\hat{w} = \frac{\hat{V}}{c} = \frac{1}{16}$ and $\frac{3}{32}$. That is, the equilibria are $(\hat{V}, \hat{w}) = (0, 0)$, $(1/2, 1/16)$, and $(3/4, 3/32)$.

To analyze the stability of these equilibria, we study the Jacobi matrix. We have already shown that $(0, 0)$ is always stable. Looking at the other equilibria we find:

$$D\mathbf{f}\left(\frac{1}{2}, \frac{1}{16}\right) = \begin{bmatrix} 1/4 & -1 \\ 1 & -8 \end{bmatrix} \qquad \tau = -31/4, \Delta = -1$$

$\Delta < 0$, so the equilibrium $(\hat{V}, \hat{w}) = (1/2, 1/16)$ is a saddle point.

Meanwhile:

$$D\mathbf{f}\left(\frac{3}{4}, \frac{3}{32}\right) = \begin{bmatrix} -1/16 & -1 \\ 1 & -8 \end{bmatrix} \qquad \tau = -129/16, \Delta = 3/2$$

$\Delta > 0$ and $\tau < 0$, so the equilibrium is stable. In fact, since $\tau^2 - 4\Delta = \left(\frac{-129}{16}\right)^2 - 4\left(\frac{3}{2}\right) \approx 59 > 0$, $(\hat{V}, \hat{w}) = (3/4, 3/32)$ is a stable node.

For this value of c, the neuron has two stable equilibria, one at $(0, 0)$ and another at $(3/4, 3/32)$. We think of these equilibria as representing the two possible rest states of the neuron; that is, the resting state ($V = 0$ with Na^+ ions outside and K^+ inside) and the state after firing. The null clines (and some solution curves) for the neuron model is shown in Figure 11.72. From Figure 11.72 we see that which stable equilibria the neuron converges to will depend on its initial conditions. In particular, if the potential across the membrane is perturbed slightly from its rest state 0, then the solution curves return to $(0, 0)$; that is, the neuron returns to its resting state. A larger stimulus (in general, if $V(0) > V_c$ for some threshold V_c) will cause the neuron to "fire"; (V, w) then converges to the fired state $(3/4, 3/22)$. For these values of a and c, $V_c \approx 0.4923$. Two different solution curves, corresponding to different values of $V(0)$, are shown in Figure 11.72, and the corresponding potentials $V(t)$ are plotted in Figures 11.73 and 11.74.

Figures 11.73 and 11.74 show two solution curves $V(t)$. In either case, $w(0) = 0$. In Figure 11.75, $V(0) = 0.35 < a$, and the initial stimulus is too small and dies away quickly. In Figure 11.74, $V(0) = 0.55 > V_c$, and we observe the neuron firing. The system mimics the bistability of real neurons. Specifically, if we apply a weak stimulus (i.e., if we increase V to a value less than V_c), then V will quickly return to 0, as shown in Figure 11.73. However, if we apply a strong enough stimulus (i.e., we increase V above V_c), the solution will move away from the equilibrium point, as shown in Figure 11.74. Weakly excited neurons will return to their resting state; $V(t)$ decays back to 0. Larger excitations will cause the neuron to fire; $V(t)$ converges to the other equilibrium (i.e., $V(t) \to 3/4$, for the values of a and c used in our model).

11.5.3 Enzymatic Reactions

In this subsection we will study **enzymatic reactions**, which are ubiquitous in the living world. Enzymes are proteins that accelerate reactions by reducing the activation energy required to initiate the reaction. Enzymes are not altered by the reaction; they aid in the initial steps of the reaction by binding the reactants (called substrates), thus forming an enzyme–substrate complex that then allows the substrates to react and to form the product. This sequence of steps is illustrated in Figure 11.75.

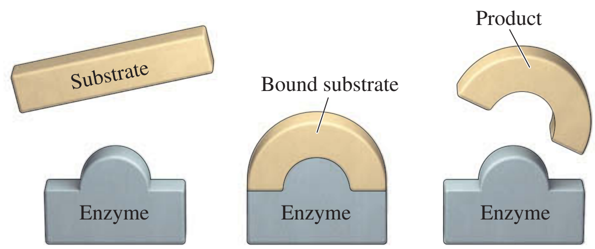


Figure 11.75 A schematic description of an enzymatic reaction.

If we denote the substrate by S, the enzyme by E, and the product of the reaction by P, then an enzymatic reaction can be described by

$$\mathrm{E} + \mathrm{S} \underset{k_-}{\overset{k_+}{\rightleftarrows}} \mathrm{ES} \xrightarrow{k_2} \mathrm{E} + \mathrm{P} \tag{11.86}$$

where k_+, k_-, and k_2 are the reaction rates in the corresponding reaction steps.

We can translate the schematic description of the reaction (11.86) into a system of differential equations. We use the following notation:

$$e(t) = [\mathrm{E}] = \text{enzyme concentration at time } t$$

$$s(t) = [\mathrm{S}] = \text{substrate concentration at time } t$$

$$c(t) = [\mathrm{ES}] = \text{concentration of enzyme–substrate complex at time } t$$

$$p(t) = [\mathrm{P}] = \text{product concentration at time } t$$

Michaelis and Menten (1913) were instrumental in the description of enzyme kinetics, through both experimental and theoretical work. On the experimental side, they developed techniques that allowed them to measure reaction rates under controlled conditions. Their experiments showed rate of the enzymatic reaction dp/dt is given by a rational function of the substrate concentration s, which they described as

$$\frac{dp}{dt} = \frac{v_m s}{K_m + s} \tag{11.87}$$

where v_m is known as the **saturation rate** and K_m is the **half-saturation constant** (i.e., if $s = K_m$, then $dp/dt = v_m/2$).

On the theoretical side, they developed a mathematical model for enzyme kinetics that predicted the observed hyperbolic relationship between the substrate concentration and the initial rate at which the product is formed.

Using the mass action laws (see Section 4.2) we start by writing down equations for each of the chemicals in (11.86)

$$\begin{aligned}
\frac{ds}{dt} &= k_-c - k_+se \\
\frac{de}{dt} &= (k_- + k_2)c - k_+se \\
\frac{dc}{dt} &= k_+se - (k_- + k_2)c \\
\frac{dp}{dt} &= k_2c
\end{aligned} \tag{11.88}$$

In what follows, we will analyze (11.88). This system of four equations is not easy to analyze, but we will simplify it, and in the end, we will obtain the **Michaelis–Menten law** for the rate of the reaction. We will also use (11.88) to illustrate that it is sometimes possible to reduce the number of equations in a system.

We claim that the system has a **conserved quantity**—that is, a quantity that does not depend on time and is therefore constant throughout the reaction. To find this quantity, note that

$$\frac{de}{dt} + \frac{dc}{dt} = 0.$$

That is,

$$\frac{d}{dt}(e + c) = 0,$$

which implies that

$$e(t) + c(t) = e_0, \tag{11.89}$$

where e_0 is a constant. We can explain (11.89) because enzyme is not used up in the reaction; so the total amount of free enzyme plus the amount appearing in enzyme-substrate complex cannot change over time. Since $e(t)+c(t)$ is constant, we say that the sum $e(t) + c(t)$ is a conserved quantity. The advantage of having conserved quantities is that if we know the initial amount of enzyme $e(0)+c(0)$ and one of the two variables $e(t)$ and $c(t)$, we can deduce the other variable since $e(0) + c(0) = e(t) + c(t)$. Since we no longer have to solve for both $e(t)$ and $c(t)$, we can reduce the number of equations from four to three.

To reduce the number of equations even further, we make another assumption. Whereas the existence of a conserved quantity followed from the system of equations, and we could have derived it without knowing the meaning of those equations, the next assumption requires understanding of the enzymatic reaction itself and cannot be deduced from the set of equations (11.88). ***Briggs and Haldane (1925)***, realized that in many reactions the rate of formation of enzyme-substrate complex balances the rate of its breakdown; that is:

$$\frac{dc}{dt} = 0$$

If we make this assumption it follows that:

$$0 = k_+ se - (k_- + k_2)c$$

which can be rewritten as

$$\frac{se}{c} = \frac{k_- + k_2}{k_+}$$

We denote this ratio by K_m; that is,

$$K_m = \frac{k_- + k_2}{k_+}$$

and, therefore,

$$\frac{se}{c} = K_m \tag{11.90}$$

Solving (11.90) for e—that is, $e = e_0 - c$—and substituting the result into (11.89), we find that

$$\frac{s(e_0 - c)}{c} = K_m$$

which, when we solve for c, yields

$$c = \frac{e_0 s}{K_m + s} \tag{11.91}$$

allowing us in turn to rewrite the equation for the rate at which the product is formed. Since $dp/dt = k_2 c$, it follows from (11.91) that

$$\frac{dp}{dt} = \frac{k_2 e_0 s}{K_m + s} \tag{11.92}$$

We can interpret the factor $k_2 e_0$ as follows. Product is formed most rapidly if all of the enzyme is complexed with the substrate, i.e., if $e = 0$ and $c = e_0$ (since $e + c = e_0$ is a conserved quantity). This implies that the rate at which the product is formed,

$dp/dt = k_2 c$, is fastest when $c = e_0$, in which case $dp/dt = k_2 e_0$. We can therefore interpret $k_2 e_0$ as the maximum rate at which this reaction can proceed. We define:

$$v_m = k_2 e_0$$

and rewrite (11.92) as

$$\frac{dp}{dt} = \frac{v_m s}{K_m + s} \tag{11.93}$$

We have therefore deduced the *Michaelis–Menten law* (11.87). We see from Equation (11.93) that the reaction rate dp/dt is limited by the availability of the enzyme. Specifically, imagine that we keep increasing the quantity of substrate present (that is, we increase s). Then (11.93) implies that the rate at which product is produced will increase, but that there are diminishing returns. In fact, product cannot be produced any faster than at the rate $v_m = k_2 e_0$. This rate is increased by increasing e_0 (i.e., by adding more enzyme). We call v_m the **saturation rate**, because as $s \to \infty$, the reaction rate saturates (i.e., has a horizontal asymptote) at v_m.

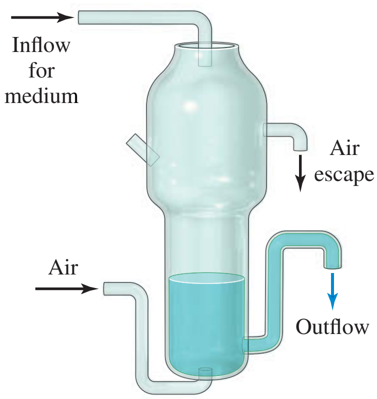


Figure 11.76 A chemostat.

11.5.4 Microbial Growth in a Chemostat

This subsection can only be studied after you have studied Subsection 11.5.3. Microbes grow by converting a substrate (nutrients from the medium they grow in) into a product (cell biomass) through enzymatic reactions. In what follows, we will investigate a mathematical model for microbial growth in a chemostat. The growth of the microbes will be limited by the availability of the substrate.

A chemostat is a growth chamber in which medium with concentration s_0 of the substrate enters the chamber at a constant rate D. Air is pumped into the chamber to mix and aerate the culture. To keep the volume in the chamber constant, the content of the chamber is removed at the same rate D as new medium enters. A sketch of a chemostat is shown in Figure 11.76.

We denote the microbial biomass at time t by $x(t)$ and the substrate concentration at time t by $s(t)$. In 1950, Jacques Lucien Monod derived the following system of differential equations to describe the growth of microbes in a chemostat:

$$\frac{ds}{dt} = D(s_0 - s) - q(s)x \tag{11.94}$$

$$\frac{dx}{dt} = Yq(s)x - Dx \tag{11.95}$$

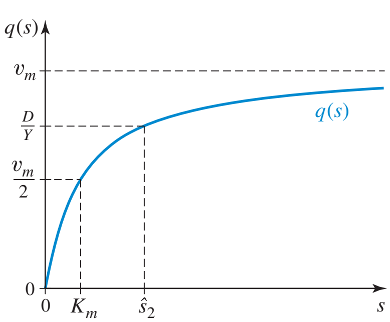


Figure 11.77 A graph of $q(s)$. We see that (11.98) can be solved for $\hat{s}_2$ if $\frac{D}{Y} < v_m$.

In these equations, $s_0 > 0$ is the substrate concentration in the entering medium, $D > 0$ is the rate at which medium flows into (and out of) the chemostat. The function $q(s)$ is the rate at which microbes digest the substrate; the argument s indicates that q depends on the substrate concentration. This substrate is converted into biomass by the microbes. Suppose for every unit of substrate they consume, they produce Y units of biomass. $Y > 0$ is called the **yield constant**.

Because the digestion of substrate depends on enzymes, it must obey the Michaelis–Menten law of chemical kinetics that we derived in Subsection 11.5.3. That is:

$$q(s) = \frac{v_m s}{K_m + s} \tag{11.96}$$

where v_m is the saturation rate and K_m is the half-saturation constant [i.e., $q(K_m) = v_m/2$]. A graph of $q(s)$ is shown in Figure 11.77.

In what follows, we will determine possible equilibria of (11.94) and (11.95) and analyze their stability. There is always the trivial equilibrium, which is obtained when substrate enters a growth chamber that is devoid of microbes [i.e., when $x(0) = 0$]. In that case, there will be no microbes at later times and, hence, $\frac{dx}{dt} = 0$ for all times $t \geq 0$. The substrate equilibrium is then found by setting $\frac{ds}{dt} = 0$ with $x = 0$ in Equation (11.94):

$$0 = D(s_0 - s)$$

This equation has the solution $s = s_0$. Hence, one equilibrium is

$$(\hat{s}_1, \hat{x}_1) = (s_0, 0) \tag{11.97}$$

To obtain a nontrivial equilibrium $(\hat{s}_2, \hat{x}_2)$, we will look for an equilibrium with $\hat{x}_2 > 0$. To find this equilibrium, we solve the simultaneous equations

$$\frac{ds}{dt} = 0 \qquad \text{and} \qquad \frac{dx}{dt} = 0$$

It follows from setting $dx/dt = 0$ in Equation (11.95) that

$$q(\hat{s}_2) = \frac{D}{Y} \tag{11.98}$$

We see from Figure 11.77 that (11.98) has a solution $\hat{s}_2 > 0$ if $0 < D/Y < v_m$. Substituting (11.96) into (11.98), we can compute $\hat{s}_2$:

$$\frac{v_m \hat{s}_2}{K_m + \hat{s}_2} = \frac{D}{Y}$$

$$\Rightarrow \qquad \hat{s}_2 = \frac{DK_m}{Yv_m - D} \tag{11.99}$$

$\hat{s}_2 > 0$, provided that $0 < D/Y < v_m$. Using (11.98) in $ds/dt = 0$, we find that

$$D(s_0 - \hat{s}_2) = \frac{D}{Y}\hat{x}_2$$

Solve (11.94) with $\frac{ds}{dt} = 0$.

or

$$\hat{x}_2 = Y(s_0 - \hat{s}_2) \tag{11.100}$$

from which we can see that $\hat{x}_2 > 0$, provided that $\hat{s}_2 < s_0$. Let's investigate this inequality a little more carefully. Since $q(s)$ is monotonic increasing in s, $\hat{s}_2 < s_0$ is equivalent to $q(\hat{s}_2) < q(s_0)$. Now $q(\hat{s}_2) = D/Y$ so the conditions for a non-trivial equilibrium to exist can be rewritten as $D/Y < v_m$ while $q(\hat{s}_2) < q(s_0)$, implies $D/Y < q(s_0)$ since $q(\hat{s}_2) = D/Y$. Although it appears that two conditions must be satisfied, we automatically have $q(s_0) < v_m$, so if the second condition is satisfied, then the first one is superfluous. Under the assumption that $0 < D/Y < v_m$ and $\hat{s}_2 < s_0$, we therefore have the nontrivial equilibrium

$$\hat{s}_2 = \frac{DK_m}{Yv_m - D}, \hat{x}_2 = Y(s_0 - \hat{s}_2) \tag{11.101}$$

There are no other equilibria.

To analyze the stability of the two equilibria (11.97) and (11.101), we find the Jacobi matrix $D\mathbf{f}$ associated with the system (11.94) and (11.95):

$$D\mathbf{f}(s, x) = \begin{bmatrix} -D - q'(s)x & -q(s) \\ Yq'(s)x & Yq(s) - D \end{bmatrix}$$

We analyze the stability of the trivial equilibrium (11.97) first:

$$D\mathbf{f}(s_0, 0) = \begin{bmatrix} -D & -q(s_0) \\ 0 & Yq(s_0) - D \end{bmatrix}$$

Since the Jacobi matrix is in upper triangular form, the eigenvalues are the diagonal elements, and we find that

$$\lambda_1 = -D < 0$$

$$\lambda_2 = Yq(s_0) - D < 0, \quad \text{provided that } \frac{D}{Y} > q(s_0)$$

Therefore, the equilibrium

$$(s_0, 0) \quad \text{is} \quad \begin{cases} \text{locally stable} & \text{if } \frac{D}{Y} > q(s_0) \\ \text{unstable} & \text{if } \frac{D}{Y} < q(s_0) \end{cases}$$

For the nontrivial equilibrium (11.101), supposing it exists, we obtain

$$A = D\mathbf{f}(\hat{s}_2, \hat{x}_2) = \begin{bmatrix} -D - q'(\hat{s}_2)\hat{x}_2 & -q(\hat{s}_2) \\ Yq'(\hat{s}_2)\hat{x}_2 & Yq(\hat{s}_2) - D \end{bmatrix}$$

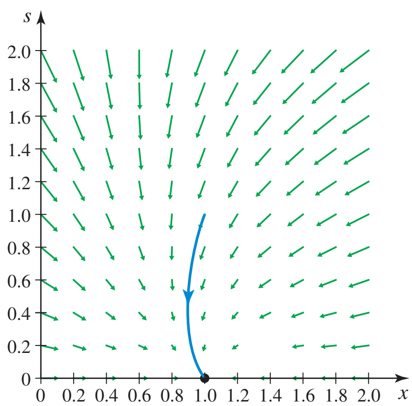


Figure 11.78 Chemostat model with $Y = 1$, $D = 1$, and $s_0 = 1$. In the reaction rate equation $v_m = 1$, $K_m = 1$. There is only one equilibrium $(\hat{s}, \hat{x})$, and our theory predicts that the microbes die out.

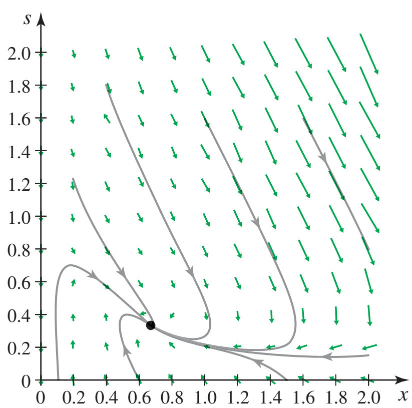


Figure 11.79 Chemostat model with $Y = 1$, $D = 1/4$, and $s_0 = 1$. In the reaction rate equation $v_m = 1$ and $K_m = 1$. Our theory shows that there is a stable non-trivial equilibrium. All solution curves converge to this equilibrium.

Using (11.98), we see that this equation simplifies to

$$A = \begin{bmatrix} -D - q'(\hat{s}_2)\hat{x}_2 & -\frac{D}{Y} \\ Y q'(\hat{s}_2)\hat{x}_2 & 0 \end{bmatrix}$$

Now,

$$\operatorname{tr} A = -D - q'(\hat{s}_2)\hat{x}_2 < 0 \qquad q'(s) > 0 \text{ since } q(s) \text{ is increasing}$$

and

$$\det A = Dq'(\hat{s}_2)\hat{x}_2 > 0$$

for $\hat{x}_2 > 0$. Therefore, if the nontrivial equilibrium exists (i.e., if both $\hat{s}_2 > 0$ and $\hat{x}_2 > 0$), then it is locally stable.

We can now summarize our results. The chemostat has two equilibria: a trivial one in which microbes are absent and, if $D < Yq(s_0)$ also a nontrivial one that allows stable microbial growth. If $D > Yq(s_0)$, then the trivial equilibrium is the only equilibrium and it is locally stable. We can interpret this inequality biologically: If D is too large (meaning the rate at which substrate and microbes is removed is too large) or if s_0 is too small (too little substrate is provided for the microbes), then the population will converge to 0; that is, all the microbes in the chemostat will go extinct. Extinction also occurs when Y is too small; i.e., if microbes are not able to convert the substrate that they digest into biomass at a fast enough rate. Conversely if $D < Yq(s_0)$, there are two equilibria; the trivial one is now unstable and the nontrivial one is the locally stable one. Stable microbial growth is therefore possible. Figures 11.78 and 11.79 show vector field plots and some solution curves in the two cases.

11.5.5 A Model for Epidemics

For our last application we return to the model of disease spread that was first introduced in Section 8.3. Mathematical models are used throughout epidemiology (the science of how diseases spread). They are used both to predict the future growth of emerging disease outbreaks like Ebola (see, for example, ***Fisman, Khoo, and Tuite, 2014***) and vaccination strategies to eliminate or control diseases like measles (see, for example, ***Keeling and Grenfell, 1997*** and ***Babad et al., 1995***). In fact the simplified models that we will present in this subsection have been shown to accurately predict the spread of many different human and animal diseases, and even the spread of computer viruses!

Just as in our previous model of infectious disease spread introduced in Section 8.3.5, we will model how a disease spreads through a population of N people (or organisms). Previously we considered a two-state model for the disease. That is, we divided the population into two classes; individuals who have the disease and individuals who do not have it. In that model after recovering from the disease, an individual is immediately returned to the population that does not have the disease, which means that they can immediately catch the disease again. This assumption may be a good model for the cold virus, which evolves quickly and has many different variants. However, for many diseases like measles, after a person recovers their immune system adapts to the disease, which makes them immune to it for some period of time (though this immunity can also be lost). To model the progression of such a disease we will use a model that was created by ***Kermack and McKendrick (1927, 1932, 1933)***.

We will divide the population into three classes: **susceptible** individuals who do not have the disease, but who could catch it; **infected** individuals who currently have (and can also transmit) the disease; and **recovered** individuals who had the disease but who have since recovered from it, and are now immune to the disease.

Denote by $S(t)$ the number of susceptible individuals, $I(t)$ the number of infected individuals, and $R(t)$ the number of individuals in the recovered class all at time t. Individuals start off in the susceptible class (S). After contact with an infected individual, they too may become infected (I). An individual will remain infected until their immune system fights off the disease. Then they enter the recovered class (R). After some time, an individual's immunity to the disease may wear off, returning them to the susceptible class. We can diagram these transitions using chemical reaction notation

$$S \to I \to R \to S$$

This kind of model is known as an SIRS model (similarly, the type of model we studied in Section 8.3 is called a SIS model).

Transitions from S to I are modeled in the same way as in Section 8.3.5. Specifically, let's assume that in one unit of time each individual comes into contact with b other individuals. The chance that a susceptible individual will become infected is proportional to the number of infected individuals that they contact. Of the b individuals that each susceptible individual encounters, a fraction $\frac{I(t)}{N}$ will be infected (similarly a fraction $\frac{S(t)}{N}$ will be susceptible, and a fraction $\frac{R(t)}{N}$ will be in the recovered class). So the rate at which susceptible individuals become infected is:

$$\begin{array}{c}\text{Rate at which}\\ \text{susceptibles are infected}\end{array} = \begin{array}{c}\text{Number of susceptible}\\ \text{individuals}\end{array} \times \begin{array}{c}\text{Likelihood that susceptible}\\ \text{individual is infected in}\\ \text{unit time}\end{array}$$

And:

$$\begin{aligned}\begin{array}{c}\text{Likelihood that susceptible}\\ \text{individual is infected in}\\ \text{unit time}\end{array} &= k \times \begin{array}{c}\text{Number of infected}\\ \text{individuals may contact}\\ \text{in unit time}\end{array}\\ &= k \times \frac{bI(t)}{N}\end{aligned}$$

Here, $k \geq 0$ is a constant of proportionality. Larger values of k make it more likely that contact with an infected individual will transmit the disease. We call k the **transmission rate**. So

$$\begin{aligned}\text{Rate of infection} &= S(t) \times \frac{kbI(t)}{N}\\ &= \frac{kb}{N} SI\end{aligned}$$

Now let's model the movements of individuals between the three classes, starting, as usual, with word equations:

$$\begin{array}{c}\text{Rate of change of}\\ \text{no. susceptibles}\end{array} = - \begin{array}{c}\text{Rate at which}\\ \text{susceptibles are}\\ \text{infected}\end{array} + \begin{array}{c}\text{Rate at which}\\ \text{recovered individuals}\\ \text{lose immunity}\end{array}$$

$$\begin{array}{c}\text{Rate of change of}\\ \text{no. infected}\end{array} = \begin{array}{c}\text{Rate at which}\\ \text{susceptibles are}\\ \text{infected}\end{array} - \begin{array}{c}\text{Rate at which}\\ \text{infected individuals}\\ \text{recover}\end{array}$$

$$\begin{array}{c}\text{Rate of change of}\\ \text{no. recovered}\end{array} = \begin{array}{c}\text{Rate at which}\\ \text{infected individuals}\\ \text{recover}\end{array} - \begin{array}{c}\text{Rate at which}\\ \text{recovered individuals}\\ \text{lose immunity}\end{array}$$

Suppose that infected individuals recover at a rate c, that is, in one unit of time a fraction c of infected individuals recover. Then

$$\begin{array}{c}\text{Rate at which}\\ \text{infected individuals}\\ \text{recover}\end{array} = cI \qquad \text{Number infected} \times \text{fraction that recover}$$

The coefficient $c \geq 0$ is a constant; we call it the **recovery rate**.

Similarly, let's assume that in one unit of time a fraction a of all recovered individuals lose immunity to the disease. Then:

$$\begin{array}{c}\text{Rate at which}\\ \text{recovered individuals}\\ \text{lose immunity}\end{array} = aR \qquad \text{No. recovered} \times \text{fraction that lose immunity}$$

We call a the **rate of immunity loss**. It is also a non-negative constant, and like the constants b, k, and c, it can be used to fit our model to different populations and to different diseases.

Putting all of the ingredients together, we derive a system of differential equations to model the spread of the disease through the population:

$$\frac{dS}{dt} = aR - \frac{kbSI}{N}$$
$$\frac{dI}{dt} = \frac{kbSI}{N} - cI$$
$$\frac{dR}{dt} = cI - aR.$$

Now unlike the previous examples that we have analyzed, our model consists of *three* separate differential equations for the three dependent variables, S, I, and R. However, we have assumed that all of the individuals must belong in exactly one of the three classes. This assumption means that:

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

In other words, $S+I+R$ is a **conserved quantity**; it does not change with time. We can check that this $S+I+R$ really is constant by using the differential equation system:

$$\begin{aligned}\frac{d}{dt}(S+I+R) &= \frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt}\\ &= \left(aR - \frac{kb}{N}SI\right) + \left(\frac{kb}{N}SI - cI\right)\\ &\quad + (cI - aR)\\ &= 0\end{aligned}$$

So if we know S and I at any point in time, we can always calculate R from

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

We can use this equation to eliminate R from the model:

$$\begin{aligned}\frac{dS}{dt} &= a(N-S-I) - \frac{kb}{N}SI\\ \frac{dI}{dt} &= \frac{kb}{N}SI - cI\end{aligned} \tag{11.102}$$

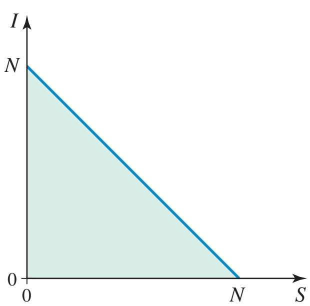


Figure 11.80 Domain for the $SIRS$ model.

To determine the domain of these equations, note that since S, I, and R all represent numbers of individuals, all three quantities must be non-negative, meaning that

$$S \geq 0, I \geq 0, \text{ and } N - S - I \geq 0 \text{ or } S + I \leq N \tag{11.103}$$

Figure 11.80 shows the domain of the differential equation system.

To analyze this model we identify its equilibria, that is, we look for values $(\hat{S}, \hat{I})$ that satisfy Equations (11.102)

$$a(N-S-I) - \frac{kb}{N}SI = 0 \qquad dS/dt = 0 \tag{11.104}$$

and

$$\frac{kb}{N}SI - cI = 0 \qquad dI/dt = 0 \tag{11.105}$$

In our initial analysis of the equations we will assume that all of the constants in our model (N, a, b, c, and k) are positive. Starting with Equation (11.105):

$$\left(\frac{kb}{N}S - c\right)I = 0$$

so either $\hat{I} = 0$ or $\hat{S} = \frac{c}{kb}N$.

Then for $\frac{dS}{dt} = 0$, either

$$\hat{I} = 0 \quad \text{and} \quad a(N - \hat{S}) = 0 \quad \text{so} \quad \hat{S} = N$$

or

$$\hat{S} = \frac{c}{kb}N \quad \text{and} \quad \hat{I} = \frac{a(N - \hat{S})}{a + \frac{kb}{N}\hat{S}}$$
$$= \frac{aN}{(a+c)}\left(1 - \frac{c}{kb}\right)$$

i.e., we have two potential equilibria, at

$$(\hat{S}, \hat{I}) = (N, 0) \quad \text{and} \quad (\hat{S}, \hat{I}) = \left(\frac{c}{kb}N, \frac{aN}{a+c}\left(1 - \frac{c}{kb}\right)\right)$$

We call these points *potential equilibria* because although they both solve (11.104) and (11.105) they may not lie in the domain of the differential equation system.

If $(\hat{S}, \hat{I}) = (N, 0)$ then $\hat{S} \geq 0, \hat{I} \geq 0$, and $\hat{S} + \hat{I} = N \leq N$, so inequalities (11.103) are certainly satisfied. For the other equilibrium $\hat{S} = \frac{c}{kb}N$, because $c, k, b, N > 0$, we certainly have $\hat{S} \geq 0$. $\hat{I} = \frac{aN}{a+c}\left(1 - \frac{c}{kb}\right)$, so $\hat{I} \geq 0$ if $1 - \frac{c}{kb} \geq 0$, i.e., if $c \leq kb$. Meanwhile:

$$\hat{S} + \hat{I} = \frac{c}{kb}N + \frac{a}{(a+c)}N\left(1 - \frac{c}{kb}\right)$$
$$= \frac{N}{(a+c)}\left(a + \frac{c^2}{kb}\right)$$

For the second equilibrium to lie in the domain of the differential equation we need $\hat{S} + \hat{I} \leq N$ or

$$\hat{S} + \hat{I} - N = \frac{N}{(a+c)}\left(a + \frac{c^2}{kb}\right) - N \leq 0$$
$$\Rightarrow \frac{Nc}{(a+c)}\left(\frac{c}{kb} - 1\right) \leq 0$$

This inequality is certainly satisfied if $c \leq kb$. Thus it is necessary and sufficient that $c \leq kb$ for the second equilibrium point to lie in the domain given by (11.103).

To analyze the stability of the equilibria we calculate the Jacobi matrix

$$\mathbf{Df}(S, I) = \begin{bmatrix} -a - \frac{kb}{N}I & -a - \frac{kb}{N}S \\ \frac{kb}{N}I & \frac{kb}{N}S - c \end{bmatrix} \qquad \mathbf{f}(S, I) = [a(N - S - I) - \tfrac{kb}{N}SI, \tfrac{kb}{N}SI - cI]'$$

For the equilibrium $(\hat{S}, \hat{I}) = (N, 0)$,

$$\mathbf{Df}(N, 0) = \begin{bmatrix} -a & -a - kb \\ 0 & kb - c \end{bmatrix}$$

The Jacobi matrix is an upper triangular matrix, so the eigenvalues may be read off from the diagonal entries:

$$\lambda_1 = -a, \quad \text{and} \quad \lambda_2 = kb - c$$

The first eigenvalue $\lambda_1 < 0$, assuming $a > 0$, while the second eigenvalue $\lambda_2 < 0$ if $c > kb$ and $\lambda_2 > 0$ if $c < kb$.

Thus, if $c > kb$, then there is only one equilibrium, at $(\hat{S}, \hat{I}) = (N, 0)$, and this equilibrium is a stable node because both eigenvalues are negative.

Figure 11.81 shows an example vector field plot and some solution curves in the case where $c > kb$. All solutions converge to $(S, I) = (N, 0)$ as $t \to \infty$. At this equilibrium $\hat{S} = N, \hat{I} = 0$, and $\hat{R} = 0$. In words, the disease disappears from the population.

If $c < kb$, then the equilibrium with $(\hat{S}, \hat{I}) = (N, 0)$ is a saddle point, and therefore unstable. However, a second equilibrium occurs if $c < kb$, namely

$$(\hat{S}, \hat{I}) = \left(\frac{c}{kb}N, \frac{aN}{a+c}\left(1 - \frac{c}{kb}\right)\right).$$

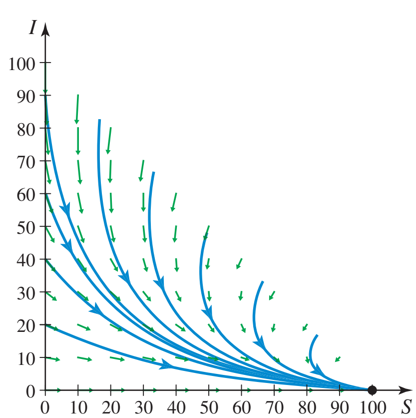


Figure 11.81 Vector field plot and solution curves for the SIRS model with constants $N = 100$, $a = k = b = 1$, $c = 2$. Because $c < kb$, all solution curves converge to $(\hat{S}, \hat{I}) = (100, 0)$.

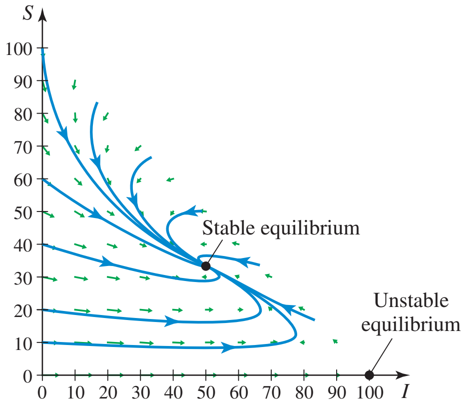


Figure 11.82 Vector field plot and solution curves for the SIRS model with constants $N = 100$, $a = k = b = 1, c = 1/2$. Because $kb > c$, the disease becomes endemic.

At this equilibrium the Jacobi matrix is

$$D\mathbf{f}\left(\frac{cN}{kb}, \frac{aN}{a+c}\left(1-\frac{c}{kb}\right)\right) = \begin{bmatrix} -a\frac{(a+kb)}{a+c} & -a-c \\ a\frac{(kb-c)}{a+c} & 0 \end{bmatrix}$$

For this equilibrium trace $(D\mathbf{f}) = -\frac{a(a+kb)}{a+c} < 0$ and $\det(D\mathbf{f}) = -(-a-c)\frac{a(kb-c)}{a+c} = a(kb-c)$. So trace $(D\mathbf{f}) < 0$ and $\det(D\mathbf{f}) = a(kb-c) > 0$, which means that the equilibrium point is stable (i.e., either a stable node or a stable spiral).

Figure 11.82 shows the vector field plot of the solution together with some solution curves when $c < kb$. All solutions converge to the stable equilibrium point with $\hat{S}, \hat{I}$ both non-zero. That is, the disease becomes **endemic**; it neither dies out nor infects everyone in the population, but remains present in the population at a stable level.

We can interpret this condition $c > kb$ as follows. Suppose that we start with a population made up almost entirely of susceptibles, with a very small number of infected individuals. Then from Equation (11.102) we deduce that:

$$\frac{dI}{dt} = \frac{kb}{N}SI - cI \approx (kb-c)I \qquad S \approx N$$

So the infected population, I, will grow or decay exponentially; growing if $c < kb$, and decaying if $kb < c$. A common way that epidemiologists write this condition is to define a new constant, R_0, called the **basic reproductive number** and defined by:

$$R_0 = \frac{kb}{c}.$$

The disease spreads if $R_0 > 1$, and dies out if $R_0 < 1$. We can go a little deeper into the dependence of R_0 on the constants that appear in our model. $R_0 < 1$ implies that $c < kb$, which will be the case if k or b is small, meaning that either each individual has few contacts with others (i.e., small b) or these contacts are not likely to transmit the disease (i.e., small k). Alternatively $R_0 < 1$ may occur if c is large, meaning that infected individuals quickly recover from the disease. In all of these situations, the disease spreads slowly and eventually disappears. So $(\hat{S}, \hat{I}) = (N, 0)$ is a stable equilibrium.

To get a clearer sense of the behaviors predicted by the model, we can use the graphical method to sketch the behaviors of the solution for different values of S and I. Although this can be done for general values of the constants a, k, b, c, and N, the details are quite involved. So we consider specific values for the constants in the model.

EXAMPLE 1 Consider the SIRS model with constants $N = 1000$, $a = 1/4$, $k = 1$, $b = 1$, and $c = 1/4$. Sketch the zero isoclines of the model in the plane, and draw the direction of the solution on the regions between these zero isoclines.

Solution For the given constants the basic reproductive number is

$$R_0 = \frac{kb}{c} = \frac{(1)(1)}{(1/4)} = 4.$$

Since $R_0 > 1$, there will be two equilibria; an unstable equilibrium corresponding to the disease not being present $(\hat{S}, \hat{I}) = (1000, 0)$, and a stable equilibrium that we find to be $(\hat{S}, \hat{I}) = (250, 375)$, representing the disease being endemic.

The corresponding Jacobi matrices are:

$$D\mathbf{f}(1000, 0) = \begin{bmatrix} -1/4 & -5/4 \\ 0 & 3/4 \end{bmatrix}$$

for which the eigenvalues are $-1/4$ and $3/4$, confirming that $(\hat{S}, \hat{I}) = (1000, 0)$ is a saddle. And:

$$D\mathbf{f}(250, 375) = \begin{bmatrix} -5/8 & -1/2 \\ 3/8 & 0 \end{bmatrix}.$$

Because $\tau < 0$ and $\Delta > 0$, this second equilibrium is stable, and because $\tau^2 - 4\Delta = -23/64 < 0$, it is, in fact, a stable spiral.

To fill in the direction of the vector field we start by reviewing the domain of the differential equation; it is $S \geq 0$, $I \geq 0$, $S + I \leq 1000$. Then we draw the zero isoclines. Writing out the differential equations, we obtain:

$$\frac{dS}{dt} = \frac{1}{4}(1000 - S - I) - \frac{1}{1000}SI$$
$$\frac{dI}{dt} = \frac{1}{1000}SI - \frac{1}{4}I.$$

Now on the $dI/dt = 0$ isoclines:

$$\frac{dI}{dt} = 0 \Rightarrow \frac{1}{1000}(S - 250)I = 0 \qquad \text{Factorize}$$

so either $I = 0$ or $S = 250$. Meanwhile on the $\frac{dS}{dt} = 0$ isocline,

$$\frac{dS}{dt} = 0 \Rightarrow \frac{1}{4}(1000 - S - I) - \frac{1}{1000}SI = 0$$

so

$$I\left(\frac{1}{4} + \frac{1}{1000}S\right) = \frac{1}{4}(1000 - S) \qquad \text{Solve for } I.$$
$$I = \frac{\frac{1}{4}(1000 - S)}{\frac{1}{4} + \frac{S}{1000}} = \frac{1000 - S}{1 + \frac{S}{250}}$$

The isocline is given by a complicated looking expression, but have courage! On the isocline, I is given by a rational function of S, and we learned how to draw the graphs of such functions in Chapter 5. Specifically we observe that the zero isocline intersects the S-axis at $S = 1000$ and the I-axis at $I = 1000$. Between these points:

$$I = \frac{1000 - S}{1 + \frac{S}{250}} < 1000 - S \qquad \text{Because } 1 + S/250 > 1$$

So $I + S < 1000$, which means that the null cline is fully contained within the domain of the differential equation system. To get the rest of the shape of the isocline, we calculate:

$$\frac{dI}{dS} = \frac{(-1)\left(1 + \frac{S}{250}\right) - (1000 - S) \cdot \frac{1}{250}}{\left(1 + \frac{S}{250}\right)^2}$$
$$= \frac{-5}{\left(1 + \frac{S}{250}\right)^2}$$

Since $\frac{dI}{dS} < 0$, the $\frac{dS}{dt} = 0$ isocline slopes downward from $(S, I) = (0, 1000)$ to $(S, I) = (1000, 0)$. Moreover, since

$$\frac{d^2I}{dS^2} = \frac{1}{25\left(1 + \frac{S}{250}\right)^3} > 0,$$

the isocline is concave upward.

In Figure 11.83 we graph the zero isoclines of the system. As our general analysis showed, there are two points of intersection between the isoclines, which are the equilibria.

On the $\frac{dI}{dt} = 0$ isoclines, the vector field must be horizontal to the left or to the right. If $I = 0$, then

$$\frac{dS}{dt} = \frac{1}{4}(1000 - S) \geq 0 \qquad S \leq 1000$$

So the vector field is rightward on the isocline $I = 0$. If $S = 250$, then $\frac{dS}{dt} = \frac{750}{4} - \frac{1}{2}I$, so the vector field is leftward ($\frac{dS}{dt} < 0$) if $I > 375$ and rightward ($\frac{dS}{dt} > 0$) if $I < 375$.

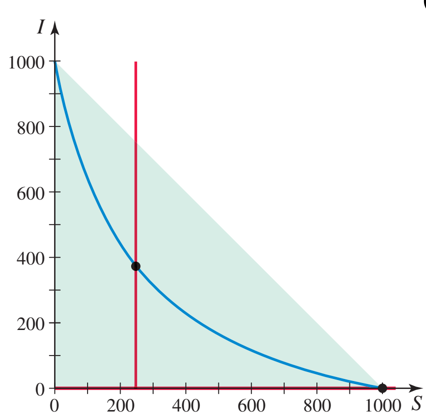


Figure 11.83 Zero isoclines for Example 1.

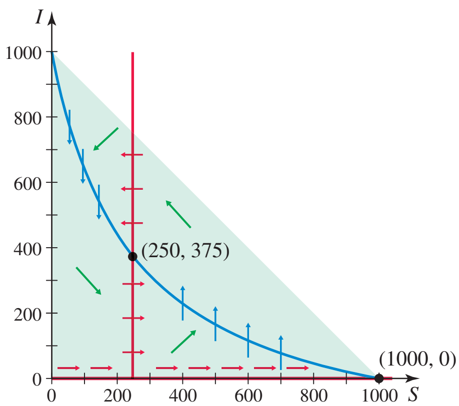


Figure 11.84 Direction of flow for Example 1.

On the $\frac{dS}{dt} = 0$ isocline the vector field must be vertical, either up or down. Since

$$\frac{dI}{dt} = \frac{I}{1000}(S - 250)$$

the vector field is upward if $S > 250$ and downward if $S < 250$. On Figure 11.84 we complete our sketch of the behavior of the solutions, including both the directions of the vector field on the zero isoclines and in the regions of the SI-plane between isoclines.

The SIRS model is very flexible and can be tailored to represent different models for disease transmission and recovery. In the next example, we will consider what happens when recovering from a disease confers lifetime immunity to the disease.

EXAMPLE 2 **Lifetime Immunity** We will model what occurs when recovering from a disease gives an individual lifelong immunity. Individuals who recover from the disease enter the recovered class, but they never return to the susceptible class (i.e., they remain immune to the disease for the rest of their lives). We can incorporate these biological facts in our model by setting a, the loss of immunity rate, to 0. So our model becomes:

$$\begin{aligned}\frac{dS}{dt} &= -\frac{kb}{N}SI\\ \frac{dI}{dt} &= \frac{kb}{N}SI - cI\\ \frac{dR}{dt} &= cI\end{aligned} \tag{11.106}$$

Since individuals progress from susceptible, to infected, to recovered, and then stay recovered, we call the equation system an **SIR model**. We will assume that the other coefficients in our model, k, b, N, and c, are all non-zero.
Assume that initially the population contains only susceptible and infected individuals. Show that whatever the initial number of infected individuals, the number of infected individuals will decrease to 0 as $t \to \infty$, and that some individuals will not catch the disease before it disappears.

Solution This problem is different from the ones that we have encountered previously because it has infinitely many equilibria. To find possible equilibria we need to solve:

$$-\frac{kb}{N}SI = 0 \tag{11.107}$$

and

$$\left(\frac{kb}{N}S - c\right)I = 0 \tag{11.108}$$

Equation (11.107) is satisfied if $S = 0$ or $I = 0$, and (11.108) is satisfied if $S = cN/kb$ or if $I = 0$. So both equations are satisfied if (and only if) $I = 0$. This condition makes sense biologically: If $I = 0$, then there is no one to transmit the disease, but we can still divide the population into susceptible and recovered classes, and any division of individuals between the two classes will produce an equilibrium [for example, we could have $(\hat{S}, \hat{R}) = (0, N)$, or $(N, 0)$, or $(N/2, N/2)$, and so on].

Since the equilibria are not isolated, we can't use the methods for analyzing equilibria from Section 11.3 to understand the behavior of solutions. However, it is still possible to analyze the equations: You should be able to understand the steps of the following argument, but don't worry about being able to make this kind of argument for general systems of differential equations.

First observe that $I \geq 0$ implies that $\frac{dR}{dt} = cI \geq 0$, so the number of recovered individuals increases monotonically as the infection spreads. We can therefore imagine using R as an independent variable, and solving for S and I in terms of R, rather than in terms of t. To do this, we must consider:

$$\frac{dS}{dR} = \frac{\frac{dS}{dt}}{\frac{dR}{dt}} \qquad \text{Chain rule: } \frac{dS}{dt} = \frac{dS}{dR}\frac{dR}{dt}$$

$$= -\frac{\frac{kb}{N}SI}{cI} = -\frac{kb}{Nc}S \qquad \text{Using Equations (11.106)}$$

We may solve this equation by separation of variables;

$$\int \frac{1}{S}\,dS = -\frac{kb}{Nc}\int dR \qquad k, b, N, c \text{ are all constants}$$

$$\ln S = -\frac{kb}{Nc}R + C \qquad S \geq 0 \text{ so } \ln|S| = \ln S \; C \text{ is a constant of integration}$$

$$S(R) = e^C \exp\left(-\frac{kb}{Nc}R\right)$$

$$= C_1 \exp\left(-\frac{kb}{Nc}R\right) \qquad \text{Define } C_1 = e^C$$

Here, C_1 is a constant of integration that we can obtain by applying the initial conditions. Note that although we assume $S \neq 0$ in our derivation, $S = 0$ is actually a constant solution and is attained if we let $C_1 = 0$. We are told that initially the population contains only susceptible and infected individuals, so when $t = 0$, $R = 0$. Let S_0 be the initial number of susceptible individuals; then

$$S_0 = C_1 \exp(0) = C_1$$

so

$$S = S_0 \exp\left(-\frac{kbR}{cN}\right) \tag{11.109}$$

We have now solved for S as a function of R. It is not obvious what to do with this expression; however, since we were asked to find how the population behaves as $t \to \infty$. Because R is an increasing function of t, Equation (11.109) implies that S must be a decreasing function of t. Can it decrease all the way to 0? No, because we always have $R(t) \leq N$. So:

$$S \geq S_0 \exp\left(-\frac{kb}{c}\right)$$

Hence, so long as $S_0 > 0$, S cannot decrease all of the way to 0–it is bounded below. Since $S(t)$ is a decreasing function of t, and it does not decrease below $S_0 e^{-kb/c}$, the graph of $S(t)$ against t must have a horizontal asymptote; that is, $\lim_{t\to\infty} S(t) = S_\infty$, for some limit $S_\infty > 0$. Hence, not every member of the population will become infected with the disease, even when $t \to \infty$.

Because S converges to a constant, dS/dt must converge to 0. Since $\frac{dS}{dt} = -\frac{kb}{N}SI$ and $S(t)$ has a non-zero limit as $t \to \infty$, for $dS/dt \to 0$ it is necessary that the disease die out, i.e., $I(t)$ must converge to 0.

So long as the number of susceptible exceeds $\frac{c}{kb}N$, $\frac{dI}{dt} > 0$; i.e., the infected population increases. However, the infection "uses up" susceptibles (meaning $S(t)$ decreases monotonically). Eventually the number of susceptibles will drop below $\frac{cN}{kb}$, whereupon $\frac{dI}{dt}$ becomes negative, and the infected population declines. Eventually the infection comes to an end. The infected population size decays to 0, before everyone has become infected. Figure 11.85 shows sample solutions for $S(t)$, $I(t)$, and $R(t)$. •

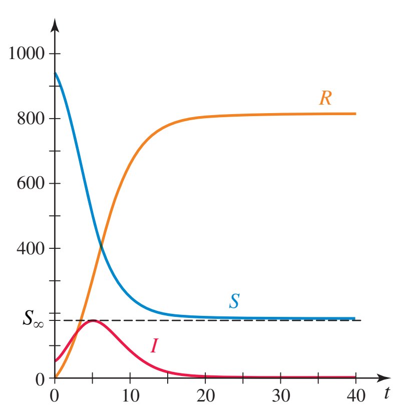


Figure 11.85 Solution of SIR model (Equations 11.106) with $N = 1000$, $k = 1$, $b = 1$, and $c = 1/2$. $\lim_{t\to\infty} S(t) = S_\infty \neq 0$.

Section 11.5 Problems

11.5.1

In Problems 1–8, classify each community matrix at equilibrium according to the five cases considered in Subsection 11.5.1 and determine whether the equilibrium is stable. (Assume in each case that the equilibrium exists.)

1. $\begin{bmatrix} -1 & -1.3 \\ 0.3 & -2 \end{bmatrix}$ **2.** $\begin{bmatrix} -3 & -1.2 \\ -1 & -2 \end{bmatrix}$

3. $\begin{bmatrix} -1.5 & 1.6 \\ 2.3 & -5.1 \end{bmatrix}$ **4.** $\begin{bmatrix} -0.3 & 0 \\ 0.4 & -0.7 \end{bmatrix}$

5. $\begin{bmatrix} -1 & 1.3 \\ 2 & -1.5 \end{bmatrix}$ **6.** $\begin{bmatrix} -2.7 & 0 \\ -1.3 & -0.6 \end{bmatrix}$

7. $\begin{bmatrix} -5 & -1.7 \\ -2.3 & -0.2 \end{bmatrix}$ **8.** $\begin{bmatrix} -2.3 & -4.7 \\ 1.2 & -3.2 \end{bmatrix}$

9. We assume that the diagonal elements a_{ii} of the community matrix of an ecosystem containing two species in equilibrium are negative. Explain why this assumption implies that species i exhibits self-regulation.

10. Consider a community composed of two species. Assume that both species inhibit themselves. Explain why mutualistic and competitive interactions lead to qualitatively similar predictions about the stability of the corresponding equilibria. That is, show that if $A = [a_{ij}]$ is the community matrix at equilibrium for the case of mutualism, and if $B = [b_{ij}]$ is the community matrix at equilibrium for the case of competition, then the following holds: If $|a_{ij}| = |b_{ij}|$ for $1 \le i, j \le 2$, then either both equilibria are locally stable or both are unstable.

11. The classical Lotka–Volterra model of predation is given by

$$\frac{dN}{dt} = rN - aNP$$

$$\frac{dP}{dt} = bNP - dP$$

where $N = N(t)$ is the prey density at time t and $P = P(t)$ is the predator density at time t. The constants a, b, d, and r are all positive.

(a) Find the nontrivial equilibrium $(\hat{N}, \hat{P})$ with $\hat{N} > 0$ and $\hat{P} > 0$.

(b) Find the community matrix corresponding to the nontrivial equilibrium.

(c) Explain each entry of the community matrix found in (b) in terms of how individuals in this community affect each other.

12. The modified Lotka–Volterra model of predation is given by

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

$$\frac{dP}{dt} = bNP - dP$$

where $N = N(t)$ is the prey density at time t and $P = P(t)$ is the predator density at time t. The constants a, b, d, r and K are positive. Assume that $d/b < K$.

(a) Find the nontrivial equilibrium $(\hat{N}, \hat{P})$ with $\hat{N} > 0$ and $\hat{P} > 0$.

(b) Find the community matrix corresponding to the nontrivial equilibrium.

(c) Explain each entry of the community matrix found in (b) in terms of how individuals in this community affect each other.

11.5.2

13. Use a graphing calculator to study the following example of the Fitzhugh–Nagumo model:

$$\frac{dV}{dt} = -V(V - 0.3)(V - 1) - w,$$

$$\frac{dw}{dt} = V - cw$$

Explain whether or not the model predicts multiple equilibria for the following values of c:
(a) $c = 2$, **(b)** $c = 16$.

14. Use a graphing calculator to study the following example of the Fitzhugh–Nagumo model:

$$\frac{dV}{dt} = -V(V - 0.6)(V - 1) - w,$$

$$\frac{dw}{dt} = V - cw$$

Explain whether or not the model predicts multiple equilibria for the following values of c:
(a) $c = 8$, **(b)** $c = 20$, **(c)** $c = 50$.

15. Assume the following example of the Fitzhugh–Nagumo model:

$$\frac{dV}{dt} = -V(V - 1/2)(V - 1) - w,$$

$$\frac{dw}{dt} = V - cw$$

Show that the model predicts multiple equilibria provided $c \ge 16$. What happens if $c < 16$?

16. Assume the following example of the Fitzhugh–Nagumo model:

$$\frac{dV}{dt} = -V(V - 3/5)(V - 1) - w,$$

$$\frac{dw}{dt} = V - cw$$

Find the smallest value of c for which the model predicts the existence of multiple equilibria.

11.5.3

In Problems 17–20, use the mass action law to translate each chemical reaction into a system of differential equations.

17. $\mathrm{A} + \mathrm{B} \xrightarrow{k} \mathrm{C}$ **18.** $\mathrm{A} + \mathrm{B} \underset{k_-}{\overset{k_+}{\rightleftarrows}} \mathrm{C}$

19. $\mathrm{E} + \mathrm{S} \xrightarrow{k_1} \mathrm{ES} \xrightarrow{k_2} \mathrm{E} + \mathrm{P}$ **20.** $\mathrm{A} + \mathrm{B} \xrightarrow{k} \mathrm{A} + \mathrm{C}$

21. Show that the following system of differential equations has a conserved quantity, and find it:

$$\frac{dx}{dt} = 2x - 3y$$

$$\frac{dy}{dt} = 3y - 2x$$

22. Show that the following system of differential equations has a conserved quantity, and find it:

$$\frac{dx}{dt} = -4x + 2y$$

$$\frac{dy}{dt} = -y + 2x$$

23. Show that the following system of differential equations has a conserved quantity, and find it:

$$\frac{dx}{dt} = -x + 2xy + z$$

$$\frac{dy}{dt} = -2xy$$

$$\frac{dz}{dt} = x - z$$

24. Suppose that $x(t) + y(t)$ is a conserved quantity. If

$$\frac{dx}{dt} = -3x + 2xy$$

find the differential equation for $y(t)$.

25. The Michaelis–Menten law [Equation (11.93)] states that

$$\frac{dp}{dt} = \frac{v_m s}{K_m + s}$$

where $p = p(t)$ is the concentration of the product of the enzymatic reaction at time t, $s = s(t)$ is the concentration of the substrate at time t, and v_m and K_m are positive constants. Set

$$f(s) = \frac{v_m s}{K_m + s}$$

where v_m and K_m are positive constants.

(a) Show that

$$\lim_{s\to\infty} f(s) = v_m$$

(b) Show that

$$f(K_m) = \frac{v_m}{2}$$

(c) Show that, for $s \geq 0$, $f(s)$ is **(i)** nonnegative, **(ii)** increasing, and **(iii)** concave down. Sketch a graph of $f(s)$. Label v_m and K_m on your graph.

(d) Explain why we said that the reaction rate dp/dt is limited by the availability of the substrate.

11.5.4

26. The growth of microbes in a chemostat was described by (11.94) and (11.95). Using the notation of that equation, together with the relationship

$$q(s) = \frac{v_m s}{K_m + s}$$

where v_m and K_m are positive constants, we will investigate how the substrate concentration $\hat{s}$ in equilibrium depends on the uptake rate Y.

(a) Assume that the microbes have a positive equilibrium density. Find the equilibrium concentration $\hat{s}$ algebraically, and investigate how the uptake rate Y affects $\hat{s}$.

(b) Assume that the microbes have a positive equilibrium density. Sketch a graph of $q(s)$, and explain how you would determine $\hat{s}$ graphically. Use your graph to explain how the uptake rate Y affects $\hat{s}$.

27. The growth of microbes in a chemostat was described by (11.94) and (11.95). Using the notation of that equation, together with the relationship

$$q(s) = \frac{v_m s}{K_m + s}$$

we will investigate how the substrate concentration $\hat{s}$ in equilibrium depends on D, the rate at which the medium enters the chemostat.

(a) Assume that the microbes have a positive equilibrium density. Find the equilibrium concentration $\hat{s}$ algebraically. Investigate how the rate D affects $\hat{s}$.

(b) Assume that the microbes have a positive equilibrium density. Sketch a graph of $q(s)$, and explain how you would determine $\hat{s}$ graphically. Use your graph to explain how the rate D affects $\hat{s}$.

In Problems 28 and 29, we investigate specific examples of microbial growth described by (11.94) and (11.95). We use the notation of Subsection 11.5.4. In each case, determine all equilibria and their stability.

28. $$\frac{ds}{dt} = 2(4-s) - \frac{3s}{2+s}x$$
$$\frac{dx}{dt} = \frac{sx}{2+s} - 2x$$

29. $$\frac{ds}{dt} = 2(4-s) - \frac{3s}{1+s}x$$
$$\frac{dx}{dt} = \frac{3sx}{1+s} - 2x$$

11.5.5

30. The spread of a disease through a population of 100 elephants is modeled by the following system of SIRS equations:

$$\frac{dS}{dt} = \frac{R}{10} - \frac{1}{200}SI$$

$$\frac{dI}{dt} = \frac{1}{200}SI - \frac{I}{10}I$$

$$\frac{dR}{dt} = \frac{1}{10}I - \frac{R}{10}$$

Find all of the equilibria for this model and classify them (e.g., as stable nodes, unstable nodes, or saddles) by analyzing the linearized system.

31. The spread of a disease through a herd of 50 cattle is modeled by the following system of SIRS equations:

$$\frac{dS}{dt} = R - \frac{1}{50}SI$$

$$\frac{dI}{dt} = \frac{1}{50}SI - 2I$$

$$\frac{dR}{dt} = 2I - R$$

Find all of the equilibria for this model and classify them (e.g., as stable nodes, unstable nodes, or saddles) by analyzing the linearized system.

32. The spread of influenza through a UCLA dormitory that houses 150 students is modeled by the following SIRS model:

$$\frac{dS}{dt} = \frac{1}{10}R - \frac{1}{50}SI$$

$$\frac{dI}{dt} = \frac{1}{50}SI - 4I$$

(a) Write down the missing differential equation for $\frac{dR}{dt}$.

(b) Find all of the equilibria for this model and classify them (e.g., as stable nodes, unstable nodes, or saddles) by analyzing the linearized system.

33. The spread of a fungal disease through an orchard containing 200 trees is modeled by the following SIRS model

$$\frac{dS}{dt} = \frac{1}{100}R - \frac{1}{800}SI$$
$$\frac{dR}{dt} = \frac{1}{8}I - \frac{1}{100}R$$

(a) Write down the missing differential equation for $\frac{dI}{dt}$.

(b) Find all of the equilibria for this model and classify them (e.g., as stable nodes, unstable nodes, or saddles) by analyzing the linearized system.

34. The spread of a disease through a population of 100 individuals is represented by the following SIRS model:

$$\frac{dS}{dt} = \frac{1}{10}R - \frac{1}{100}SI$$
$$\frac{dI}{dt} = \frac{1}{100}SI - \frac{1}{2}I$$
$$\frac{dR}{dt} = \frac{1}{2}I - \frac{1}{10}R$$

In this problem we will sketch the directions of the solution in the SI-plane.

(a) Eliminate R to rewrite the equation system as a system of differential equations in the dependent variables S and I.

(b) Draw the zero isoclines of your system from part (a).

(c) Find all of the equilibria for this model and classify them (e.g., as stable nodes, unstable nodes, or saddles) by analyzing the linearized system.

(d) Add to your plot from part (b) arrows showing the direction of the vector field on the isoclines, and in the regions between the isoclines.

35. The spread of a disease through a population of 250 individuals is represented by the following SIRS model;

$$\frac{dS}{dt} = R - \frac{1}{50}SI$$
$$\frac{dI}{dt} = \frac{1}{50}SI - \frac{1}{10}I$$
$$\frac{dR}{dt} = \frac{1}{10}I - R$$

In this problem we will sketch the directions of the solution in the SI-plane.

(a) Eliminate R to rewrite the equation system as a system of differential equations in the dependent variables $S(t)$ and $I(t)$.

(b) Draw the zero isoclines of your system from part (a).

(c) Find all of the equilibria for this model and classify them (e.g., as stable nodes, unstable nodes, or saddles) by analyzing the linearized system.

(d) Add to your plot from part (b) arrows showing the direction of the vector field on the isoclines, and in the regions between the isoclines.

36. Controlling the Spread of a Disease In our analysis of the SIRS system of equations (i.e., Equations 11.102) we determined that the model predicted very different behaviors depending on whether the basic reproductive number $R_0 > 1$ or $R_0 < 1$. In particular, the disease will die out if $R_0 < 1$. R_0 is defined by the equation

$$R_0 = \frac{kb}{c}.$$

Discuss the possible effect of the following public health measures on the constants k, b, and c, and whether they can prevent a disease from spreading.

(a) Primary care: treating people who have been infected by the disease so they recover more quickly from the disease.

(b) Quarantining: in a community affected by the disease, people are told to remain indoors and to minimize their contact with others.

(c) Handwashing: a campaign promotes frequent washing of hands to reduce transmission of the disease.

37. Lethal Diseases Some diseases are lethal; not every individual infected by the disease will recover; some will die. Assume that in one unit of time a fraction m of infected individuals will die (m is called the *mortality rate*).

(a) Explain how the SIRS model equations should be modified to incorporate deaths. In particular you should write down a new differential equation for $\frac{dI}{dt}$.

(b) Explain why it is no longer possible to eliminate $R(t)$ from the SIRS model equations.

Lethal Diseases

Some diseases are lethal; not every individual infected by the disease will recover; some will die. Assume that in one unit of time a fraction m of infected individuals will die (m is called the mortality rate). We will assume that the habitat this population lives in is at its carrying capacity. If no individuals die then no reproduction occurs. If individuals die, then resources are freed up and more individuals will be born: one birth for every death that occurs. That is, the number of individuals born in one unit of time is equal to the number of individuals who die in that unit of time. In Problems 38–41 you will analyze models for lethal diseases. In Problems 38–40 you should assume that infants are initially uninfected by the disease but are also not immune to it, so new individuals added to the population are all in the susceptible class.

38. (a) Explain how the SIRS model equations should be modified to incorporate deaths and births. In particular, you should write down new differential equations for $\frac{dS}{dt}$ and $\frac{dI}{dt}$.

(b) Show directly from your differential equations that

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0.$$

That is, $S + I + R = N$ is a conserved quantity.

(c) Use the result $S + I + R = N$ to eliminate R from the differential equation system.

(d) Find all of the equilibria of the differential equation. Show that if

$$kb > m + c$$

then there are two equilibria in the domain of the differential equation.

39. In this problem we will determine the stability of equilibria in an SIRS model that includes mortality. Consider a population of size $N = 100$. The SIRS model with mortality for this population is:

$$\frac{dS}{dt} = -\frac{1}{200}SI + \frac{1}{10}R + \frac{1}{3}I$$
$$\frac{dI}{dt} = \frac{1}{200}SI - \frac{2}{3}I$$
$$\frac{dR}{dt} = \frac{1}{3}I - \frac{1}{10}R$$

(a) What is the mortality rate for this disease?

(b) Rewrite the system of differential equations as a part of differential equations with S and I as dependent variables.

(c) Find all equilibria lying within the domain of the system.

(d) By linearizing the differential equations around the equilibria that you discovered in (c), classify each of the equilibria (e.g., as stable node, spiral, or saddle).

40. In this problem we will determine the stability of equilibria in an SIRS model that includes mortality. Consider a population of size $N = 250$. Assuming a mortality rate $m = 1/4$, our SIRS model becomes:

$$\frac{dS}{dt} = -\frac{1}{500}SI + \frac{1}{5}R + \frac{1}{4}I$$
$$\frac{dI}{dt} = \frac{1}{500}SI - \frac{1}{3}I$$
$$\frac{dR}{dt} = \frac{1}{12}I - \frac{1}{5}R$$

(a) Write the system of differential equations as a part of differential equations with S and I as dependent variables.

(b) Find all equilibria lying within the domain for this model.

(c) By linearizing the differential equations around the equilibria that you discovered in (c), classify each of the equilibria (e.g., as stable node, spiral, or saddle).

41. Assume that all individuals in the population are equally likely to be parents to the mI offspring added to the population in each unit of time. If an offspring is born to an infected parent it will be born infected (i.e., into the infectious class). Similarly, offspring born to susceptible parents are susceptible, and offspring born to recovered parents are recovered. Derive an SIRS model to describe the spread of this disease. There is no need to analyze your model.

42. Lifelong Immunity A particular infectious disease confers lifelong immunity to any individual who recovers from the disease. The population size is $N = 200$. Assume that the spread of the disease can be described by an SIR model:

$$\frac{dS}{dt} = -\frac{1}{100}SI$$
$$\frac{dI}{dt} = \frac{1}{100}SI - 6I$$
$$\frac{dR}{dt} = 6I$$

Assuming that $R(0) = 0$ initially and $I(0) = 5$, calculate a bound on the maximum number of individuals who will catch the disease.

43. Effect of Vaccination A particular infectious disease confers lifelong immunity to any individual who recovers from the disease. The population size is $N = 100$. Assume that the spread of the disease can be described by an SIR model:

$$\frac{dS}{dt} = -\frac{1}{300}SI$$
$$\frac{dI}{dt} = \frac{1}{300}SI - \frac{1}{9}I$$
$$\frac{dR}{dt} = \frac{1}{9}I$$

(a) Assuming that $R(0) = 0$ initially and $I(0) = 5$, calculate a bound on the maximum number of individuals who will catch the disease.

(b) Assume that a vaccination program means that half of the population start out immune to the disease, i.e., $R(0) = 50$. Assume also that there are initially 5 infected individuals (i.e., $I(0) = 5$). Recalculate the maximum bound on the number of individuals who will eventually catch the disease.

Relapsing Infections

In some diseases (such as herpes simplex), an individual may apparently recover from the disease, and in fact gain immunity to it, but the disease continues to be harbored in the person's body, breaking out some time after they recover from the initial infection. To model this process we will modify our SIRS model as follows: Since there is no loss of immunity, $a = 0$. However, in each unit of time a fraction r (r is a constant called the rate of relapse) of the individuals from the recovered class become infected with the disease. In Problems 44–47 you will analyze models for relapsing infections.

44. (a) Write down a system of differential equations to describe the spread of this disease through the population.

(b) Find all of the equilibria that lie in the domain of the system of differential equations you derived in part (a).

45. Assume that the following model can be used to represent the spread of a relapsing infection in a population of size $N = 100$ and with relapse rate $r = \frac{1}{100}$:

$$\frac{dS}{dt} = -\frac{1}{50}SI$$
$$\frac{dI}{dt} = \frac{1}{50}SI - \frac{1}{10}I + \frac{1}{100}R$$
$$\frac{dR}{dt} = \frac{1}{10}I - \frac{1}{100}R$$

(a) What is the domain for this differential equation system?

(b) Find all of the possible equilibria for this system of differential equations.

(c) Use the fact that $S + I + R = 100$ to eliminate S from the system, and to write it as a pair of differential equations with I and R as dependent variables.

(d) By linearizing the differential equation system near each of the equilibria that you discovered in part (b), classify these equilibria (e.g., as a stable node, spiral, or saddle).

46. Assume that the following model can be used to represent the spread of the infection in a population of size $N = 250$ and with relapse rate $r = \frac{1}{50}$:

$$\frac{dS}{dt} = -\frac{1}{100}SI$$
$$\frac{dI}{dt} = \frac{1}{100}SI - \frac{1}{5}I + \frac{1}{50}R$$
$$\frac{dR}{dt} = \frac{1}{5}I - \frac{1}{50}R$$

(a) What is the domain for this differential equation system?

(b) Find all of the possible equilibria for this system of differential equations.

(c) Use the fact that $S + I + R = 250$ to eliminate S from the system and to write it as a pair of differential equations with I and R as dependent variables.

(d) By linearizing the differential equation system near each of the equilibria that you discovered in part (b), classify these equilibria (e.g., as a stable node, spiral, or saddle).

(e) We will now sketch the direction of flow for the system in the IR-plane.

(i) First sketch the $\frac{dI}{dt} = 0$ and $\frac{dR}{dt} = 0$ isoclines.

(ii) Add arrows to show the directions of the vector field on the isoclines that you drew in part (i). Then add arrows showing the direction of the vector field in the regions between isoclines.

Chapter 11 Review

Key Terms

Discuss the following definitions and concepts:

1. Linear first-order equation
2. Homogeneous
3. Vector field, direction vector
4. Solution of a system of linear differential equations
5. Eigenvalue, eigenvector
6. Superposition principle
7. General solution
8. Stability
9. Sink, or stable node
10. Saddle point
11. Source, or unstable node
12. Spiral
13. Euler's formula
14. Compartment model
15. Conserved quantity
16. Harmonic oscillator
17. Models for love
18. Nonlinear autonomous system of differential equations
19. Critical point
20. Zero isoclines, or null clines
21. Graphical approach to stability
22. Lotka–Volterra model of competition
23. Monoculture
24. Founder control, coexistence
25. Lotka–Volterra predator–prey model
26. Community matrix
27. Fitzhugh–Nagumo model
28. Bistable
29. Excitable
30. Michaelis–Menten law
31. Chemostat model
32. SIRS model
33. Endemic disease
34. Basic reproductive number

Review Problems

In Problems 1–4 classify the equilibrium point at $(x, y) = (0, 0)$.

1. $\frac{dx}{dt} = 2x + 3y, \quad \frac{dy}{dt} = -2y$

2. $\frac{dx}{dt} = 2y, \quad \frac{dy}{dt} = 3x$

3. $\frac{dx}{dt} = 2y, \quad \frac{dy}{dt} = -3x$

4. $\frac{dx}{dt} = x + y, \quad \frac{dy}{dt} = y - x$

5. Two-Compartment Model Matter flows into and between two compartments, as shown in Figure 11.86.

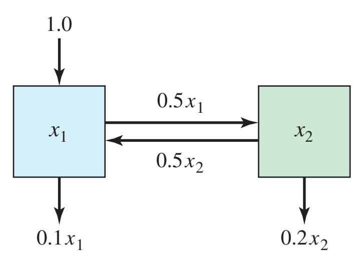


Figure 11.86 Problem 5

(a) Based on the flows shown in the figure, write down differential equations for the amounts of material $x_1(t)$ and $x_2(t)$ in the two compartments.

(b) Find the equilibrium values for $\hat{x}_1$ and $\hat{x}_2$.

(c) By linearizing your equations from part (a), determine whether the equilibrium you found in part (b) is stable or unstable.

6. Carbohydrate Flow in a Forest As we saw in Section 11.3, plants receive nitrogen from microorganism partners, fungi and bacteria, in their roots. In return, the plant shares the carbohydrates that it makes by photosynthesis with those partners. We will build a two-compartment model for this flow of carbohydrates, with one compartment representing the free carbohydrates stored in a plant, and the other representing the free carbohydrates in the plant's fungal partners.

Suppose that each day the plant produces 1 g of free carbohydrate. Now assume that each day 50% of the free carbohydrate in the plant is lost (actually used for growth), and 25% is shared with the fungal partner. For the fungal partner each day 10% of carbohydrates is returned to the plant (usually combined with other nutrients) and 75% is used for growth.

(a) Draw a two-compartment diagram (like Figure 11.40) showing the flows of carbohydrate in this system.

(b) Using your diagram from part (a), write down differential equations for the amounts of material $x_1(t)$ and $x_2(t)$ in the two compartments.

(c) Find the equilibrium values for $\hat{x}_1$ and $\hat{x}_2$.

(d) By linearizing your equations from part (a), determine whether the equilibrium you found in part (b) is stable or unstable.

7. **Population Growth** Let $N_1(t)$ and $N_2(t)$ denote the respective sizes of two populations at time t, and assume that their dynamics are respectively given by

$$\frac{dN_1}{dt} = r_1 N_1$$
$$\frac{dN_2}{dt} = r_2 N_2$$

where r_1 and r_2 are positive constants denoting the intrinsic rate of growth of the two populations. Set $Z(t) = N_1(t)/N_2(t)$, and show that $Z(t)$ satisfies

$$\frac{d}{dt} \ln Z(t) = r_1 - r_2 \tag{11.110}$$

Solve (11.110), and show that $\lim_{t\to\infty} Z(t) = \infty$ if $r_1 > r_2$. Conclude from this that population 1 becomes numerically dominant when $r_1 > r_2$.

8. **Population Growth** Let $N_1(t)$ and $N_2(t)$ denote the respective sizes of two populations at time t, and assume that their dynamics are respectively given by

$$\frac{dN_1}{dt} = r_1 N_1$$
$$\frac{dN_2}{dt} = r_2 N_2$$

where r_1 and r_2 are positive constants denoting the intrinsic rate of growth of the two populations. Denote the combined population size at time t by $N(t)$; that is, $N(t) = N_1(t) + N_2(t)$. Define the relative proportions

$$p_1 = \frac{N_1}{N} \quad \text{and} \quad p_2 = \frac{N_2}{N}$$

Use the fact that $p_1/p_2 = N_1/N_2$ to show that

$$\frac{dp_1}{dt} = p_1(1 - p_1)(r_1 - r_2)$$

Show that if $r_1 > r_2$ and $0 < p_1(0) < 1$, $p_1(t)$ will increase for $t > 0$ and population 1 will become numerically dominant.

9. **Predator–Prey Interactions** An unrealistic feature of the Lotka–Volterra model is that the prey exhibits unlimited growth in the absence of the predator. The model described by the following system remedies this shortcoming (in the model, we assume that the prey evolves according to logistic growth in the absence of the predator; the other features of the model are retained):

$$\begin{aligned} \frac{dN}{dt} &= 2N\left(1 - \frac{N}{10}\right) - 3PN \\ \frac{dP}{dt} &= PN - 3P \end{aligned} \tag{11.111}$$

(a) Draw the zero isoclines of (11.111).

(b) Determine whether the nontrivial equilibrium is locally stable.

10. **Resource Competition** ***Tilman (1982)*** developed a model for how plants compete for a single resource—for instance, nitrogen. If $B(t)$ denotes the total biomass at time t and $R(t)$ is the amount of the resource available at time t, then the dynamics are described by the following system of differential equations:

$$\frac{dB}{dt} = B\,[f(R) - m]$$
$$\frac{dR}{dt} = a(S - R) - cBf(R)$$

The first equation describes the rate of change of biomass, where the function $f(R)$ describes how the species growth rate depends on the resource, and m is the per capita plant mortality rate. The second equation describes the resource dynamics; the constant S is the maximal amount of the resource in a given habitat. The rate of resource supply (dR/dt) is assumed to be proportional to the difference between the current resource level and the maximal amount of the resource; the constant a is the constant of proportionality. The term $cBf(R)$ describes the resource uptake by the plants; the constant c can be considered a conversion factor.

In what follows, we assume that $f(R)$ follows the Monod growth function

$$f(R) = \frac{dR}{k + R}$$

where d and k are positive constants.

(a) Find all equilibria. Show that if $d > m$ and $S > mk/(d-m)$, then there exists a nontrivial equilibrium.

(b) Sketch the zero isoclines for the case in which the system admits a nontrivial equilibrium, and sketch the direction of the vector field on the isoclines and in the regions between them.

11. **Plant Competition** In this problem, we describe a simple competition model in which two species of plants compete for vacant space. Assume that the entire habitat is divided into a large number of patches. Each patch can be occupied by at most one species. We denote by $p_i(t)$ the fraction of patches occupied by species i. Note that $0 \le p_1(t) + p_2(t) \le 1$. The dynamics are described by

$$\frac{dp_1}{dt} = c_1 p_1(1 - p_1 - p_2) - m_1 p_1$$
$$\frac{dp_1}{dt} = c_2 p_2(1 - p_1 - p_2) - m_2 p_2$$

where c_1, c_2, m_1, and m_2 are positive constants. The first term on the right-hand side of each equation describes the colonization of vacant patches; the second term on the right-hand side of each equation describes how occupied patches become vacant.

(a) Show that the dynamics of species 1 in the absence of species 2 are given by

$$\frac{dp_1}{dt} = c_1 p_1(1 - p_1) - m_1 p_1 \tag{11.112}$$

and find conditions on c_1 and m_1 so that (11.112) admits a nontrivial equilibrium (an equilibrium in which $0 < p_1 \le 1$).

(b) Assume now that $c_1 > m_1$ and $c_2 > m_2$. Show that if

$$\frac{c_1}{m_1} > \frac{c_2}{m_2}$$

then species 1 will exclude species 2 if species 1 initially occupies a positive fraction of the patches.

T **12. Paradox of Enrichment** ***Rosenzweig (1971)*** analyzed a number of predator–prey models and concluded that enriching the system by increasing the nutrient supply destabilizes the nontrivial equilibrium. We will think of the predator–prey model as a plant–herbivore system in which plants represent prey and herbivores represent predators. The models analyzed were of the form

$$\frac{dN}{dt} = f(N, P) \tag{11.113}$$

$$\frac{dP}{dt} = g(N, P) \tag{11.114}$$

where $N = N(t)$ is the plant abundance at time t and $P = P(t)$ is the herbivore abundance at time t. The models all shared the property that the zero isocline for the herbivore was a vertical line and the zero isocline for the plants was a hump-shaped curve. We will look at one of the models, namely,

$$\begin{aligned} \frac{dN}{dt} &= N\left(1 - \frac{N}{K}\right) - P\left(1 - e^{-N}\right) \\ \frac{dP}{dt} &= P\left(1 - e^{-N}\right) - 0.9P \end{aligned} \tag{11.115}$$

K is the carrying capacity for plants.

(a) Find the zero isoclines for (11.115), and show that **(i)** the zero isocline of the herbivore ($dP/dt = 0$) is a vertical line in the N–P plane and **(ii)** the zero isocline for the plants ($dN/dt = 0$) intersects the N-axis at $N = K$.

(b) Plot the zero isoclines in the N–P plane for three levels of the carrying capacity: **(i)** $K = 1$, **(ii)** $K = 4$, and **(iii)** $K = 10$.

(c) For each of the three carrying capacities, determine whether a nontrivial equilibrium exists.

(d) Determine the stability of the existing nontrivial equilibria in (c).

(e) Enriching the community could mean increasing the carrying capacity of the plants. For instance, adding nitrogen or phosphorus to plant communities frequently results in an increase in biomass, which can be interpreted as an increase in the carrying capacity of the plants (the K-value). On the basis of your answers in (d), explain why enriching the community (increasing the carrying capacity of the plants) can result in a destabilization of the nontrivial equilibrium. The surprising consequence of this destabilization is that the herbivores can be driven to extinction.

13. Microbial Growth The growth of microbes in a chemostat was described by Equations (11.94) and (11.95). We will investigate how the microbial abundance in equilibrium depends on the characteristics of the system.

(a) Assume that $q(s)$ is a nonnegative function. Show that the equilibrium abundance of the microbes is given by

$$\hat{x} = Y(s_0 - \hat{s})$$

where $\hat{s}$ is the substrate equilibrium abundance. When is $\hat{x} > 0$?

(b) Assume now that

$$q(s) = \frac{v_m s}{K_m + s}$$

Investigate how the uptake rate Y and the rate D at which new medium enters the chemostat affect the equilibrium abundance of the microbes.

14. Antibiotics and Bacteria Bacteria compete among each other for resources like space and nutrients. To assist in this competition some bacteria produce antibiotics to kill off their competitors. Some bacteria are resistant (immune) to these antibiotics and therefore outcompete the antibiotic-producing bacteria. However, when an antibiotic-resistant bacteria competes with one that is neither antibiotic resistant nor capable of producing antibiotics (these bacteria are known as *wild type*), the wild-type bacteria wins because they are not burdened by the cost of making antibiotics or protecting themselves from them. We represent these dynamics in Figure 11.87.

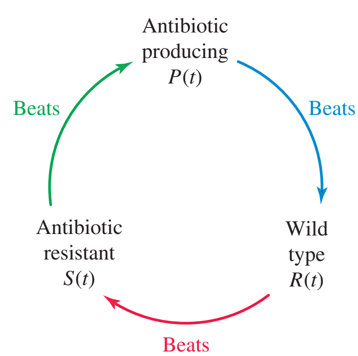


Figure 11.87 Dynamics of rock-paper-scissors game between bacteria.

Strogatz (2015) modeled these interactions as a **rock-paper-scissors game**. In this game, rock loses to paper, paper to scissors, but scissors lose to rock, just as the wild-type bacteria lose to the antibiotic-producing bacteria, which lose to the antibiotic-resistant bacteria, which in turn lose to the wild-type bacteria.

Assume that in a mixed community of bacteria, a fraction $P(t)$ are antibiotic producing, a fraction $R(t)$ are wild type (i.e., don't produce antibiotics and aren't resistant to it either), and a fraction $S(t)$ are antibiotic resistant. Since all bacteria are one of the three types, $P(t) + R(t) + S(t) = 1$. The growth rate of each population depends on their interactions with other bacteria. For example, if an antibiotic-resistant bacterium encounters mainly antibiotic-producing bacterium, then it will reproduce faster since it has a competitive advantage over those bacteria; if, on the other hand, it encounters wild-type bacteria, it will be outcompeted.

(a) Explain how the model

$$\frac{dS}{dt} = S(P - R)$$

is consistent with the biological model described above. Write down similar equations for $\frac{dR}{dt}$ and $\frac{dP}{dt}$.

(b) Show that $\frac{dR}{dt} + \frac{dP}{dt} + \frac{dS}{dt} = 0$. Why is this to be expected?

(c) We have a system of three differential equations, but our methods are designed for systems of two differential equations.

However, since we know that $R + P + S = 1$, we can eliminate R using $R = 1 - P - S$. Write down two differential equations for dP/dt and dS/dt that do not include R.

(d) Assume that there are no antibiotic-resistant bacteria present (i.e., $S(t) = 0$). Write down a differential equation for dP/dt. Find the equilibria for this differential equation and describe how $P(t)$ will change with time.

(e) Now consider the case where both P and S can be non-zero. Find all possible equilibria of the system, and classify them by linearization.

(f) One of the equilibria that you found in (d) is predicted to be a linear center. In such circumstances we cannot trust linearization. However, just as for the predator–prey system, we can write solutions as level curves of a function. Show using the chain rule that the function

$$F(P, S) = PS(1 - P - S)$$

is constant on solution curves.

T **(g)** Use your graphical calculator to draw the level curves of F. Show that, based on these level curves, we expect $P(t)$, $S(t)$, and $R(t)$ to all oscillate in time, if all three species of bacteria are present in a habitat.

15. Successional Niche ***Pacala and Rees (1998)*** modeled how plants compete to fill gaps created by forest fires. In this model, two species—an early successional and a late successional—occupy discrete patches. Each patch experiences disturbances (such as fire) at rate D. After a patch is disturbed, both species are present. Over time, however, the late successional species outcompetes the early successional species, causing the early successional species to become extinct. This change, from a patch that is occupied by both species to a patch that is occupied by the late successional species only, happens at rate a. We keep track of the number of patches occupied by both species at time t, denoted by $x(t)$, and the number of patches occupied by just the late successional species at time t, denoted by $y(t)$. The dynamics are given by the system of linear differential equations

$$\begin{aligned} \frac{dx}{dt} &= -ax + Dy \\ \frac{dy}{dt} &= ax - Dy \end{aligned} \tag{11.116}$$

where a and D are positive constants.

(a) Show that all equilibria are of the form $(\hat{x}, a\hat{x}/D)$.

(b) Find the eigenvalues and eigenvectors corresponding to each equilibrium.

(c) Show that

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = C_1 \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + C_2 e^{\lambda_2 t} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

where $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is the eigenvector corresponding to the zero eigenvalue and $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is the eigenvector corresponding to the nonzero eigenvalue λ_2, is a solution of (11.116).

(d) Show that $x(t)+y(t)$ does not depend on t. [*Hint*: Show that $\frac{d}{dt}(x(t) + y(t)) = 0$.] Show also that the line $x + y = A$ (where A is a constant) is parallel to the line in the direction of the eigenvector corresponding to the nonzero eigenvalue.

(e) Show that the zero isoclines of (11.116) are given by

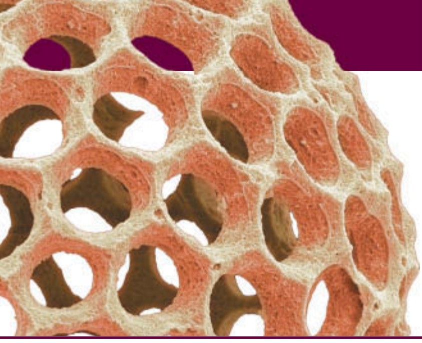
$$y = \frac{a}{D}x$$

and that this line is the line in the direction of the eigenvector corresponding to the zero eigenvalue.

(f) Suppose now that $x(t) + y(t) = c$, where c is a positive constant. Show that (11.116) can be reduced to just one equation, namely,

$$\frac{dx}{dt} = -(a + D)x + Dc$$

Show that $\hat{x} = c\frac{D}{D+a}$ is the only equilibrium, and determine its stability.



CHAPTER

12

Probability and Statistics

This chapter develops the principles of probability theory and statistics. Specifically, we will learn how to

- apply the principles of counting;
- define and calculate probabilities for discrete and continuous random variables;
- analyze the behavior of averages; and
- describe data, estimate parameters, and find statistical relationships.

We conclude this book with a chapter on probability and statistics. Although neither field is part of calculus, they both rely on calculus in providing indispensable tools for life scientists.

Many phenomena in nature are not deterministic. To give just a few examples, consider the number of eggs laid by a bird, the life span of an organism, the inheritance of genes, and the number of people infected during an outbreak of a disease. To deal with the inherent randomness (or stochasticity) of natural phenomena, we need to develop special tools; these tools are supplied by the disciplines of probability and statistics.

A short description of the role of probability and statistics in the life sciences might be as follows: Probability theory provides tools for modeling randomness and forms the foundation of statistics. Statistics provides tools for analyzing data from scientific experiments.

12.1 Counting

It is often necessary to count the ways in which a certain task can be performed. Sometimes we can make a list of all possible choices. Frequently, the total number of possible ways is very large, making it impractical to write down a list of all possible choices. There are three basic counting principles that will help us to count in a more systematic way. The first is the multiplication principle; the other two, which follow from it, are rules concerning permutations and combinations.

12.1.1 The Multiplication Principle

To illustrate the multiplication principle, consider the next example.

EXAMPLE 1 Imagine that we wish to experimentally manipulate growth conditions for plants to measure the effect of fertilizer and temperature. We want to grow plants in pots in a greenhouse at two different levels of fertilizer (low and high) and four different temperatures (10°C, 15°C, 20°C, and 25°C). If we want three replicates of each possible combination of fertilizer and temperature treatment, how many pots will we need?

Solution We can answer this question with the help of a tree diagram, as shown in Figure 12.1. We see from the tree that we will need

$$2 \cdot 4 \cdot 3 = 24$$

pots for our experiment. •

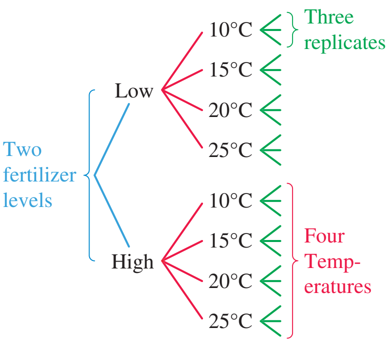


Figure 12.1 The tree diagram illustrates how many pots are needed in the experiment described in Example 1.

The counting principle that we just used is called the *multiplication principle*, which we can summarize as follows:

> **Multiplication Principle** Suppose that an experiment consists of m ordered tasks. Task 1 has n_1 possible outcomes, task 2 has n_2 possible outcomes, …, and task m has n_m possible outcomes. Then the total number of possible outcomes of the experiment is
> $$n_1 \cdot n_2 \cdot n_3 \cdots n_m$$

Looking back at Example 1, we see that the experiment consisted of three tasks: first, to select the fertilizer level; second, to select the temperature; and third, to replicate each combination of fertilizer and temperature three times. The successive tasks are illustrated in the tree diagram, and the total number of pots required for the experiment can be obtained by counting the number of tips of the tree.

We present one more example that illustrates this counting principle.

EXAMPLE 2 Suppose that after a long day in the greenhouse you decide to order pizza. You call a local pizza parlor and learn that there are three choices of crust and five choices of toppings and that you can order the pizza with or without cheese. If you want only one topping, how many different choices do you have for selecting a pizza?

Solution Your "experiment," which consists of ordering a pizza, involves three tasks. The first task is to choose a crust, the second is to choose the topping, and the third is to decide whether or not you want cheese. Using the multiplication principle, we find that there are

$$3 \cdot 5 \cdot 2 = 30$$

different pizzas that you could order. •

12.1.2 Permutations

EXAMPLE 3 Suppose that you grow plants in a greenhouse. To control for spatially varying environmental conditions, you rearrange the pots every other day. If you have six pots arranged in a row on a bench, in how many ways can you arrange the pots?

Solution To answer this question, imagine that you arrange the pots on the bench from left to right: You have six choices for the leftmost position on the bench, for the next position you can choose any of the remaining five pots, for the third position you can choose any of the remaining four pots, and so on, until there is one pot left that must go into the rightmost position. Using the multiplication principle, we find that there are

$$6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$$

ways to arrange the six pots on the bench. •

As shorthand notation for the type of descending products of positive integers we encountered in Example 3, we define

$$n! = n(n-1)(n-2) \cdots 3 \cdot 2 \cdot 1$$

and read $n!$ as "n factorial." We can now write 6! instead of $6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$.

We define

$$0! = 1$$

Then, for $n = 0, 1, 2, \ldots,$

$$(n+1)! = (n+1) \cdot n!$$

The quantity $n!$ grows very quickly. Suppose that instead of 6 pots you have 7; then there are $7! = 5040$ ways to arrange the seven pots. With 12 pots, there are $12! = 479{,}001{,}600$ possible ways to arrange them.

We will look at another example and then state a general principle.

EXAMPLE 4 Suppose that a track team has 10 sprinters, any 4 of whom can form a relay team. Assume that each person can run in any position on the team. How many teams can be formed if teams that consist of the same 4 people in different running orders are considered different teams?

Solution We select the members of the team in the order in which they run. There are 10 available sprinters for the first position. After having chosen a person for the first position, there are 9 left, and we can choose any of the 9 for the second position. For the third position, we can choose among the 8 remaining people and, finally, for the fourth position, we can select a person from the remaining 7. Using the multiplication principle, we find that there are

$$10 \cdot 9 \cdot 8 \cdot 7 = 5040$$

different relay teams. •

In Example 3, we selected $k = 6$ objects from a set of $n = 6$; the order of selection was important. In Example 4, we selected $k = 4$ objects from a set of $n = 10$, where again the order of selection was important. Such selections are called **permutations**. Using the multiplication principle, we can find the number of possible permutations of a given number of objects.

> **Permutations** A **permutation** of n different objects taken k at a time is an *ordered* subset of k out of the n objects. The number of ways that this can be done is denoted by $P(n, k)$ and is given by
>
> $$P(n, k) = n(n-1)(n-2)\cdots(n-k+1)$$

Note that the last term in the product defining $P(n, k)$ is of the form $(n-k+1)$, because there are k descending factors and the first factor is n.

Returning to Example 3, where we wanted to select six out of six objects in an ordered arrangement, we can now use our permutation rule to compute the number of ways that we can make the selection. Setting $n = 6$ and $k = 6$, we find that

$$P(6, 6) = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$$

for the number of different ways to arrange the pots. Since the product consists of six terms in descending order, starting with 6, the last term is $n - k + 1 = 6 - 6 + 1 = 1$.

Returning to Example 4, in which we wanted to select 4 out of 10 objects in an ordered arrangement, we can now use our permutation rule to compute the number of ways that we can make the selection. Setting $n = 10$ and $k = 4$, we obtain

$$P(10, 4) = 10 \cdot 9 \cdot 8 \cdot 7$$

for the number of different relay teams. Since the product consists of four terms in descending order, starting with 10, the last term is $n - k + 1 = 10 - 4 + 1 = 7$.

Another way to compute $P(n, k)$ follows from the calculation

$$P(n, k) = n(n-1)(n-2) \cdot (n-k+1)\underbrace{\frac{(n-k)(n-k-1)\cdots 3 \cdot 2 \cdot 1}{(n-k)(n-k-1)\cdots 3 \cdot 2 \cdot 1}}_{1}$$

which, after simplification, yields

$$P(n, k) = \frac{n!}{(n-k)!} \tag{12.1}$$

Let us consider one more example before we introduce the third counting principle.

EXAMPLE 5 How many 5-letter words with no repeated letters can you form out of the 26 letters of the alphabet? (Note that a "word" here need not be in the dictionary.)

Solution This task amounts to choosing 5 letters from 26, where the order is important. Hence, there are

$$P(26, 5) = \frac{26!}{21!} = 26 \cdot 25 \cdot 24 \cdot 23 \cdot 22 = 7{,}893{,}600$$

different words. •

12.1.3 Combinations

In choosing a permutation, the order of the objects is important. But what if the order is not important? How can we then compute the number of arrangements?

We return to Example 4, in which we chose a relay team. The order on the team is important when the members on the team actually run. But if we wanted to know only who was on the team, the order would no longer be important. We saw that there are $10 \cdot 9 \cdot 8 \cdot 7$ different relay teams. But since 4 people can be arranged in $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$ different ways, each choice of 4 people appears in 24 different teams. If we divide $10 \cdot 9 \cdot 8 \cdot 7$ by $4 \cdot 3 \cdot 2 \cdot 1$, we obtain the number of ways that we can choose 4 people from a group of 10 if the order does not matter. We find that this number is

$$\frac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1} = 210$$

Such unordered selections are called **combinations**. The approach we just used gives us a general formula for the number of combinations, summarized as follows:

> **Combinations** A **combination** of n different objects taken k at a time is an *unordered* subset of k out of n objects. The number of ways that this can be done is denoted by $C(n, k)$ and is given by
>
> $$C(n, k) = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$$

Note that it follows from the probabilistic meaning of $C(n, k)$ that $C(n, k)$ is always an integer. Instead of $C(n, k)$, we often write $\binom{n}{k}$, which we read "n choose k." The symbol $\binom{n}{k}$ is called a **binomial coefficient**. Using (12.1), we find that

$$C(n, k) = \binom{n}{k} = \frac{P(n, k)}{k!} = \frac{n!}{k!\,(n-k)!} \tag{12.2}$$

Looking at the rightmost expression, we see that it is also equal to $\binom{n}{n-k}$. We therefore find the identity

$$\binom{n}{k} = \binom{n}{n-k} \tag{12.3}$$

The following counting argument also explains this identity: The expression $\binom{n}{k}$ denotes the number of ways that we can select an unordered subset of size k from a set of size n. Instead of choosing the elements that go into the set, we could choose the elements that *do not* go into the set. There are $n - k$ such elements, and we can select them in $\binom{n}{n-k}$ different ways.

Another identity follows from setting $k = 0$ in (12.3) and using (12.2) and $0! = 1$:

$$\binom{n}{0} = \binom{n}{n} = \frac{n!}{0!\,n!} = 1$$

This identity can also be understood from the following counting argument: The expression $\binom{n}{0}$ means that we select a subset of size 0 from a set of size n, where the order is not important. But there is only one such set: the empty set. Similarly, $\binom{n}{n}$ means that we select a subset of n objects from a set of size n where the order is not important. There is only one way to do this: We must take the entire set of n objects.

We can use similar reasoning to argue that $\binom{n}{1} = n$; this represents the number of ways that we can choose subsets of size 1 where the order is not important. There are n such subsets: all the singletons. [Actually, the order does not play a role when we consider sets with one element, as is reflected in the fact that $P(n, 1) = n$ as well.]

In the next example, we also use the rule for counting combinations.

EXAMPLE 6 Suppose that you wish to plant 5 grass species in a plot. You can choose among 12 different species. How many choices do you have?

Solution Since the order is not important for this selection, there are

$$C(12, 5) = \frac{12!}{5!\,7!} = \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 792$$

different ways to make the selection. Alternatively, we could have written

$$C(12, 5) = \binom{12}{5} = \frac{P(12, 5)}{5!}$$

These expressions are equivalent and are evaluated as before. •

As a last example in this subsection, we prove the binomial theorem, which we encountered in Section 4.3 when we used the formal definition of derivatives to prove the rule for differentiating the function $f(x) = x^n$ for n a positive integer.

EXAMPLE 7 Show that if n is a positive integer, then

$$(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \qquad (12.4)$$

$\sum$-notation was introduced in Section 2.2

Solution The term $(x + y)^n$ consists of n factors, each of the form $x + y$. When we multiply out, each factor contributes either an x or a y. Thus, the product consists of sums that contain terms of the form $x^k y^{n-k}$ for $k = 0, 1, 2, \ldots, n$. A term of the form $x^k y^{n-k}$ occurs $\binom{n}{k}$ times, since there are $\binom{n}{k}$ ways of selecting the factor x exactly k times from among the n factors $(x + y)$. Equation (12.4) then follows. •

12.1.4 Combining the Counting Principles

The difficult part of counting is to decide which rule to use. To gain experience with this decision, we discuss several examples in which we combine the three counting principles.

EXAMPLE 8 How many different 11-letter "words" can be formed from the letters in the word MISSISSIPPI?

Solution There are four S's, four I's, two P's, and one M. There are 11! ways to arrange the letters, but some of the resulting words will be indistinguishable from one another, since letters that repeat in a word can be swapped without creating a new word.

Therefore, we need to divide by the order of the repeated letters. We then find that there are

$$\frac{11!}{4!\,4!\,2!\,1!} = 34,650$$

different words. •

EXAMPLE 9 Returning to Example 2, suppose that you now want *two* different toppings on your pizza. How does this change affect your answer?

Solution Since there are five toppings and the order in which we choose them is not important, we have $\binom{5}{2}$ choices for our two toppings. Everything else remains the same, and we find that there are

$$3 \cdot \binom{5}{2} \cdot 2 = 60 \qquad \binom{5}{2} = \frac{5!}{2!\,3!} = 10$$

different pizzas to choose from. •

EXAMPLE 10 Suppose that a license plate consists of three letters followed by three digits. How many license plates can there be if repetition of letters, but not of digits, is allowed?

Solution The order is important in this case. For each letter, there are 26 choices, since repetition is allowed. There are 10 choices for the first digit, 9 choices for the second digit, and 8 choices for the third digit, since repetition of digits is not allowed. Hence, there are

$$26 \cdot 26 \cdot 26 \cdot 10 \cdot 9 \cdot 8 = 12,654,720$$

different license plates with the aforementioned restriction. •

EXAMPLE 11 An urn contains six green and four blue balls. You take out three balls at random without replacement. How many different selections contain exactly two green balls and one blue ball? (Assume that the balls are distinguishable.)

Solution The order in this selection is not important. To obtain two green balls and one blue ball, we select two of the six green and one of the four blue balls and then combine our choices. That is, there are

$$\binom{6}{2}\binom{4}{1} = 60 \qquad \binom{6}{2} = \frac{6!}{2!\,4!} = 15, \quad \binom{4}{1} = 4$$

different selections. •

In the preceding example, we explicitly stated that all of the balls were distinguishable, and from now on we will always assume that the objects we select are distinguishable without explicitly stating this assumption. Our counting principles apply only to distinguishable objects, as assumed in the definitions ("*n* different objects"). There are cases in physics where objects are indistinguishable—for instance, electrons—but we will not deal with this possibility here.

EXAMPLE 12 A collection contains seeds for five different annual plants; two produce yellow flowers, and the other three produce blue flowers. You plan a garden bed with three different annual plants from this selection, but you do not want both of the plants with yellow flowers in the bed. How many different selections can you make?

Solution We choose three different plants out of the available five. Possible plant selections contain either three plants with blue flowers or one plant with yellow and two plants with blue flowers. There are $\binom{3}{3}\binom{2}{0}$ choices with only blue flowers, since we must first choose three blue flowers and then zero yellow flowers and $\binom{3}{2}\binom{2}{1}$ choices with exactly one yellow flower. Hence, there are

$$\binom{3}{3}\binom{2}{0} + \binom{3}{2}\binom{2}{1} = (1)(1) + (3)(2) = 7 \qquad \text{Add the two cases}$$

different selections.

Alternatively, we could have approached this problem in the following way: There are $\binom{5}{3}$ ways to select three plants from the available five. Choices with two plants with yellow flowers and one plant with blue flowers are not acceptable; there are $\binom{3}{1}\binom{2}{2}$ such choices. There are no choices of plants with three yellow and no blue flowers, since there are only two plants with yellow flowers. All other choices are acceptable. Hence, there are

$$\binom{5}{3} - \binom{3}{1}\binom{2}{2} = 10 - (3)(1) = 7$$

different selections. •

EXAMPLE 13 A standard deck of cards consists of 52 cards, arranged in 4 suits, each with 13 different values. In the game of poker, a hand consists of 5 cards drawn at random from the deck without replacement.

(a) How many hands are possible?

(b) How many hands consist of exactly one pair (i.e., two cards of equal value, with the three other cards having different values)?

Solution **(a)** There are

$$\binom{52}{5} = 2{,}598{,}960$$

different ways of choosing 5 cards from a deck of 52 cards.

(b) To pick exactly one pair, we first assign the value to the pair (13 ways) and then choose their suits [$\binom{4}{2}$ ways]. The 3 remaining cards all have different values that are different from the pair and from each other [$\binom{12}{3}$ ways]. There are four ways to assign a suit to each card and thus a total of 4^3 ways. Combining the different steps, we find that there are

$$13 \cdot \binom{4}{2} \cdot \binom{12}{3} \cdot 4^3 = 1{,}098{,}240$$

ways to pick exactly one pair. •

Section 12.1 Problems

12.1.1

1. Suppose that you want to investigate the influence of light and fertilizer levels on plant performance. You plan to use five fertilizer levels and two light levels. For each combination of fertilizer level and light level, you want four replicates. What is the total number of replicates?

2. Suppose that you want to investigate the effects of leaf damage on the performance of drought-stressed plants. You plan to use three levels of leaf damage and four different watering protocols. For each combination of leaf damage and watering protocol, you plan to have three replicates. What is the total number of replicates?

3. *Coleomegilla maculata*, a beetle, is an important predator of egg masses of *Ostrinia nubilalis*, the European corn borer. *C. maculata* also feeds on aphids and maize pollen. To study its food preferences, you choose two ages of *C. maculata* beetles and all combinations of two of the three food sources (i.e., either egg masses and aphids, egg masses and pollen, or aphids and pollen). For each experimental protocol, you want 20 replicates. What is the total number of replicates?

4. To test the effects of a new drug, you plan the following clinical trial: Each patient receives either the new drug or an established drug, or a placebo. You enroll 50 patients. In how many ways can you assign them to the three treatments?

5. The Muesli-Mix is a popular breakfast hangout near a campus. A typical breakfast there consists of one beverage, one bowl of cereal, and a piece of fruit. If you can choose among three different beverages, seven different cereals, and four different types of fruit, how many choices for breakfast do you have?

6. To study sex differences in food preferences in rats, you offer one of three choices of food to each rat. You plan to have 12 rats for each food-and-sex combination. How many rats will you need?

7. The genome of the HIV virus consists of 9749 nucleotides. There are four different types of nucleotides. Determine the total number of different genomes of size 9749 nucleotides.

8. Automated chemical synthesis of DNA has made it possible to custom-order moderate-length DNA sequences from commercial suppliers. Assume that a single nucleotide weighs about 5.6×10^{-22} gram and that there are four kinds of nucleotides. If you wish to order all possible DNA sequences of a fixed length, at what length will your order exceed **(a)** 100 kg and **(b)** the mass of the Earth (5.9736×10^{24} kg)?

12.1.2

9. You plan a trip to Europe during which you wish to visit London, Paris, Amsterdam, Rome, and Heidelberg. Because you want to buy a railway ticket before you leave, you must decide on the order in which you will visit these five cities. How many different routes are there?

10. Five people line up for a photograph. How many different lineups are possible?

11. You have just bought seven different books. In how many ways can they be arranged on your bookshelf?

12. Four cars arrive simultaneously at an intersection. Only one car can go through at a time. In how many different ways can they leave the intersection?

13. How many four-letter "words" with no repeated letters can you form from the 26 letters of the alphabet?

14. A committee of 3 people must be chosen from a group of 10. The committee consists of a president, a vice president, and a treasurer. How many committees can be selected?

15. Three different awards are to be given to a class of 15 students. Each student can receive at most one award. Count the number of ways these awards can be given out.

16. You have just enough time to play 4 songs out of 10 stored on your phone. In how many ways can you program your phone to play the 4 songs?

17. Six customers arrive at a bank at the same time. Only one customer at a time can be served. In how many ways can the six customers be served?

18. An amino acid is encoded by triplet nucleotides. How many different amino acids are possible if there are four different nucleotides that can be chosen for a triplet?

12.1.3

19. A bag contains 10 different candy bars. You are allowed to choose 3. How many choices do you have?

20. During International Movie Week, 60 movies are shown. You have time to see 5 movies. How many different plans can you make?

21. A committee of 3 people must be formed from a group of 10. How many committees can there be if no specific tasks are assigned to the members?

22. A standard deck contains 52 different cards. In how many ways can you select 5 cards from the deck?

23. An urn contains 15 different balls. In how many ways can you select 4 balls without replacement?

24. Twelve people wait in front of an elevator that has room for only 5. Count the number of ways that the first group of people to take the elevator can be chosen.

25. Four A's and five B's are to be arranged into a nine-letter word. How many different words can you form?

26. Suppose that you want to plant a flower bed with four different plants. You can choose from among eight plants. How may different choices do you have?

27. Amin owns a 4-GB music storage device that holds 1000 songs. How many different playlists of 20 songs are there if the order of the songs is important?

28. A bookstore has 300 science fiction books. Molly wants to buy 5 of the 300 science fiction books. How many selections are there?

12.1.4

29. A box contains five red and four blue balls. You choose two balls.

(a) How many possible selections contain exactly two red balls, how many exactly two blue balls, and how many exactly one of each color?

(b) Show that the sum of the number of choices for the three cases in (a) is equal to the number of ways that you can select two balls out of the nine balls in the box.

30. Twelve children are divided up into three groups, of five, four, and three children, respectively. In how many ways can this be done if the order within each group is not important?

31. Five A's, three B's, and six C's are to be arranged into a 14-letter "word". How many different words can you form?

32. A bag contains 45 beans of three different varieties. Each variety is represented 15 times in the bag. You grab 9 beans out of the bag.

(a) Count the number of ways that each variety can be represented exactly three times in your sample.

(b) Count the number of ways that only one variety appears in your sample.

33. Let $S = \{a, b, c\}$. List all possible subsets, and argue that the total number of subsets is $2^3 = 8$.

34. Suppose that a set contains n elements. Argue that the total number of subsets of this set is 2^n.

35. In how many ways can José, Hilary, Peter, and Jessica sit on a bench if Peter and Jessica want to be next to each other?

36. Paula, Crystal, Gloria, and Lan have dinner at a round table. In how many ways can they sit around the table if Crystal wants to sit to the left of Paula?

37. In how many ways can you form a committee of three people from a group of seven if two of the people do not want to serve together?

38. In how many ways can you form two committees of three people each from a group of nine if

(a) no person is allowed to serve on more than one committee?

(b) people can serve on both committees simultaneously?

39. A collection contains seeds for four different annual and three different perennial plants. You plan a garden bed with three different plants, and you want to include at least one perennial. How many different selections can you make?

40. In diploid organisms, chromosomes appear in pairs in the nuclei of all cells except gametes (sperm or ovum). Gametes are formed during meiosis, a process in which the number of chromosomes in the nucleus is halved; that is, only one member of each pair of chromosomes ends up in a gamete. Humans have 23 pairs of chromosomes. How many kinds of gametes can a human produce?

41. Sixty patients are enrolled in a small clinical trial to test the efficacy of a new drug against a placebo and the currently used drug. The patients are divided into 3 groups of 20 each. Each group is assigned one of the three treatments. In how many ways can all of the patients be assigned?

42. One hundred patients wish to enroll in a study in which patients are divided into four groups of 25 patients each. In how many ways can this be done if each patient is assigned to exactly one group?

43. Expand $(x + y)^4$.

44. Expand $(2x - 3y)^5$.

45. In how many ways can four red and five black cards be selected from a standard deck of cards if cards are drawn without replacement?

46. In how many ways can two aces and three kings be selected from a standard deck of cards if cards are drawn without replacement?

47. In the game of poker, determine the number of ways exactly two pairs can be picked.

48. In the game of poker, determine the number of ways a *flush* (five cards of the same suit) can be picked.

49. In the game of poker, determine the number of ways *four of a kind* (four cards of the same value, plus one other card) can be picked.

50. In the game of poker, determine the number of ways a *straight* (five cards with consecutive values, such as A 2 3 4 5 or 7 8 9 10 J, but not necessarily all of the same suit) can be picked.

51. **Counterpoint** *Counterpoint* is a musical term that means the combination of simultaneous voices; it is synonymous with *polyphony*. In *triple counterpoint*, three voices are arranged such that any voice can take any place of the three possible positions: highest, intermediate, and lowest voice. In how many ways can the three voices be arranged?

52. **Counterpoint** *Counterpoint* is a musical term that means the combination of simultaneous voices; it is synonymous with *polyphony*. In *quintuple counterpoint*, five voices are arranged such that any voice can take any place of the five possible positions: from highest to lowest voice. In how many ways can the five voices be arranged?

12.2 What Is Probability?

12.2.1 Basic Definitions

A **random experiment** is a repeatable experiment in which the outcome is uncertain. Tossing a coin and rolling a die are examples of random experiments. The set of all possible outcomes of a random experiment is called the **sample space** and is often denoted by Ω (uppercase Greek omega). We look at some examples in which we describe random experiments and give the associated sample space.

EXAMPLE 1 Suppose that we toss a coin labeled heads (H) on one side and tails (T) on the other. If we toss the coin once, the possible outcomes are H and T, and the sample space is therefore

$$\Omega = \{H, T\}$$

If we toss the coin twice in a row, then each outcome is an ordered pair describing the outcome of the first toss followed by the outcome of the second toss, such as HT, which means heads followed by tails. The sample space is

$$\Omega = \{HH, HT, TH, TT\}$$ •

EXAMPLE 2 Consider a population for which we keep track of one gene. We assume that the gene occurs in three different forms, called alleles and denoted by A_1, A_2, and A_3. Furthermore, the individuals in the population are diploid; that is, the chromosomes occur in pairs. This means that each individual has a pair of genes, such as A_1A_1 (we call this pair their **genotype**). The order of the chromosomes is not important, so the genotype A_1A_2 is the same as A_2A_1. If our random experiment consists of picking one individual and noting their genotype, then the sample space is given by

$$\Omega = \{A_1A_1, A_1A_2, A_1A_3, A_2A_2, A_2A_3, A_3A_3\}$$ •

EXAMPLE 3 An urn contains five balls, numbered 1–5, respectively. We draw two balls from the urn without replacement and note the numbers drawn.

There is some ambiguity in the formulation of this experiment. We can draw the two balls one after the other (without replacing the first in the urn after having noted its number), or we can draw the two balls simultaneously. In the first case, the sample space consists of ordered pairs (i, j), where the first entry is the number on the first ball and the second entry is the number on the second ball. Because the sampling is done without replacement, the two numbers are different. The sample space can then be written as

$$\begin{aligned}\Omega = \{&(1,2),\ (1,3),\ (1,4),\ (1,5),\\ &(2,1),\ (2,3),\ (2,4),\ (2,5),\\ &(3,1),\ (3,2),\ (3,4),\ (3,5),\\ &(4,1),\ (4,2),\ (4,3),\ (4,5),\\ &(5,1),\ (5,2),\ (5,3),\ (5,4)\}\end{aligned}$$

or, in short,

$$\Omega = \{(i, j) : 1 \le i \le 5, 1 \le j \le 5, i \ne j\}$$

In the second case, there is no first or second ball, because we draw the balls simultaneously. We can write this sample space as

$$\begin{aligned}\Omega = \{(1,2),\ (1,3),\ (1,4),\ &(1,5),\\ (2,3),\ (2,4),\ &(2,5),\\ (3,4),\ &(3,5),\\ &(4,5)\}\end{aligned}$$

or, in short,

$$\Omega = \{(i, j) : 1 \le i < j \le 5\}$$

where the first entry of (i, j) represents the smaller of the two numbers on the balls in our sample.

The specifics of the random experiment determine which description of the sample space we prefer. •

When we perform random experiments, we often consider a particular set of outcomes, or, more formally, a particular subset of the sample space. We call subsets of the sample space **events**. Since each outcome is an element of the sample space, an outcome is an event as well, namely, a subset that consists of just one element. We will use the basic set operations to deal with events.

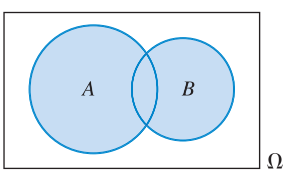


Figure 12.2 The union of A and B, $A \cup B$, illustrated in a Venn diagram. The rectangle represents the set Ω.

Basic Set Operations. Suppose that A and B are events of the sample space Ω. The **union** of A and B, denoted by $A \cup B$ (read "A union B"), is the set of all outcomes that belong to either A or B (or both). The **intersection** of A and B, denoted by $A \cap B$ (read "A intersected with B"), is the set of all outcomes that belong to both A and B. Figures 12.2 and 12.3 show these first two set operations. These figures, in which sets are visualized as "bubbles," are called **Venn diagrams**.

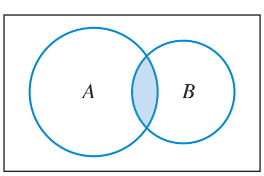


Figure 12.3 The intersection of A and B, $A \cap B$, illustrated in a Venn diagram.

We can generalize the union and intersection of two events to a finite number of events. Let $A_1, A_2, \ldots, A_n$ be a finite number of events. Then

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n = (A_1 \cup A_2 \cup \cdots \cup A_{n-1}) \cup A_n$$
$$= \begin{bmatrix}\text{the set of all outcomes that}\\ \text{belong to at least one set } A_i\end{bmatrix}$$

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n = (A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \cap A_n$$
$$= \begin{bmatrix}\text{the set of all outcomes that}\\ \text{belong to all sets } A_i,\ i = 1, 2, \ldots, n\end{bmatrix}$$

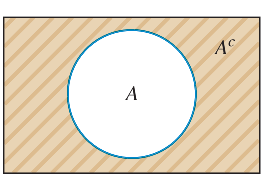


Figure 12.4 The complement of A is A^c.

The **complement** of A, denoted by A^c, is the set of all outcomes contained in Ω that are not in A. (See Figure 12.4.) It follows that

$$\Omega^c = \emptyset \qquad \text{and} \qquad \emptyset^c = \Omega$$

where $\emptyset$ denotes the empty set. Furthermore,

$$(A^c)^c = A$$

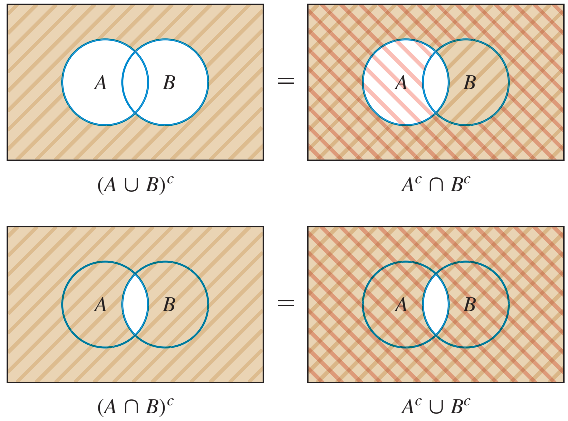


Figure 12.5 De Morgan's laws.

When we take complements of unions or intersections, the following two identities are useful (Figure 12.5):

De Morgan's Laws

$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$

EXAMPLE 4 Let us consider the experiment described in Example 1, in which we tossed a coin twice. The sample space is given by

$$\Omega = \{HH, HT, TH, TT\}$$

We denote by A the event that at least one head occurred:

$$A = \{HH, HT, TH\}$$

Let B denote the event that the first toss resulted in tails:

$$B = \{TH, TT\}$$

We see that

$$A \cup B = \{HH, HT, TH, TT\} \qquad \text{and} \qquad A \cap B = \{TH\}$$

Furthermore,

$$A^c = \{TT\} \qquad \text{and} \qquad B^c = \{HH, HT\}$$

To see how De Morgan's laws work, we compute both $(A \cup B)^c$ and $A^c \cap B^c$. We find that

$$(A \cup B)^c = \emptyset \qquad \text{and} \qquad A^c \cap B^c = \emptyset$$

which is consistent with De Morgan's first law. The second De Morgan's law claims that $(A \cap B)^c$ and $A^c \cup B^c$ are the same. We find that, indeed,

$$(A \cap B)^c = \{HH, HT, TT\} = A^c \cup B^c$$

•

We say that $A_1, A_2, \ldots, A_n$ are **pairwise disjoint** (or, simply, disjoint) if

$$A_i \cap A_j = \emptyset \qquad \text{whenever} \qquad i \neq j$$

This situation is illustrated for four sets by the Venn diagram in Figure 12.6.

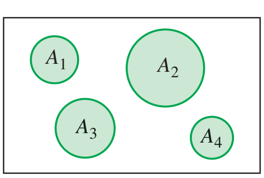


Figure 12.6 The four sets A_1, A_2, A_3, and A_4 are disjoint.

EXAMPLE 5 Is it true that if $A_1 \cap A_2 \cap A_3 = \emptyset$, then A_1, A_2, and A_3 are pairwise disjoint?

Solution No. Figure 12.7 shows a counterexample: $A_1 \cap A_2 \cap A_3 = \emptyset$, but $A_2 \cap A_3 \neq \emptyset$, implying that A_1, A_2, and A_3 are not pairwise disjoint.

•

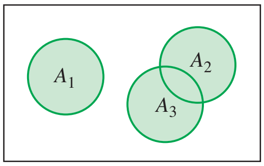


Figure 12.7 A counterexample to Example 5.

The Definition of Probability. In the following definition, we assume that the sample space Ω has finitely many elements:

> **Definition** Let Ω be a finite sample space and A and B be events in Ω. A **probability** is a function that assigns values between 0 and 1 to events. The probability of an event A, denoted by $P(A)$, satisfies the following properties:
>
> **1.** For any event A, $0 \leq P(A) \leq 1$.
>
> **2.** $P(\emptyset) = 0$ and $P(\Omega) = 1$.
>
> **3.** For two disjoint events A and B,
>
> $$P(A \cup B) = P(A) + P(B)$$

Note that a probability is a number that is *always* between 0 and 1. If you compute a probability and get either a negative number or a number greater than 1, you know immediately that your answer must be wrong.

EXAMPLE 6 Assume that $\Omega = \{1, 2, 3, 4, 5\}$ and that

$$P(1) = P(2) = 0.2, \quad P(3) = P(4) = 0.1, \quad \text{and} \quad P(5) = 0.4$$

where we wrote $P(i)$ for $P(\{i\})$. Set $A = \{1, 2\}$ and $B = \{4, 5\}$. Find $P(A \cup B)$, and show that $P(\Omega) = 1$.

Solution Since A and B are disjoint ($A \cap B = \emptyset$), it follows that

$$P(A \cup B) = P(A) + P(B) = P(\{1, 2\}) + P(\{4, 5\})$$

Also, since $\{1, 2\} = \{1\} \cup \{2\}$ and $\{4, 5\} = \{4\} \cup \{5\}$, and both are unions of disjoint sets,

$$P(\{1, 2\}) = P(1) + P(2)$$
$$P(\{3, 4\}) = P(3) + P(4)$$

Hence,

$$\begin{aligned} P(A \cup B) &= P(1) + P(2) + P(4) + P(5) \\ &= 0.2 + 0.2 + 0.1 + 0.4 = 0.9 \end{aligned}$$

To show that $P(\Omega) = 1$, we observe that

$$\Omega = \{1, 2, 3, 4, 5\} = \{1\} \cup \{2\} \cup \{3\} \cup \{4\} \cup \{5\}$$

Because this is a union of disjoint sets,

$$\begin{aligned} P(\Omega) &= P(1) + P(2) + P(3) + P(4) + P(5) \\ &= 0.2 + 0.2 + 0.1 + 0.1 + 0.4 = 1 \end{aligned}$$

We next derive two additional basic properties of probabilities. The first is

$$\boxed{P(A^c) = 1 - P(A)}$$

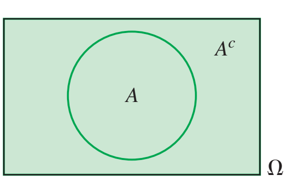


Figure 12.8 The set Ω, represented by the rectangle, can be written as a disjoint union of the sets A and A^c.

To see why this is true, we observe in Figure 12.8 that $\Omega = A \cup A^c$ and that A and A^c are disjoint. Therefore,

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c)$$

from which the claim follows after a rearrangement of terms.

Note that in Property 2 of the definition we wrote $P(\emptyset) = 0$ *and* $P(\Omega) = 1$. It would have been sufficient to require just one of these two identities. For instance, since $\Omega^c = \emptyset$, we can write $\Omega = \Omega \cup \emptyset$. Now, Ω and $\emptyset$ are disjoint and $P(\Omega) = 1$. It then follows that

$$1 = P(\Omega) = P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset)$$

and therefore,

$$P(\emptyset) = 1 - P(\Omega) = 0$$

The second property allows us to compute probabilities of unions of two sets (which are not necessarily disjoint, counter to Property 3 of the definition):

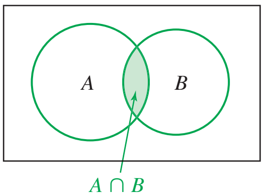


Figure 12.9 To compute $P(A \cup B)$, we add $P(A)$ and $P(B)$, but since we count $A \cap B$ twice, we need to subtract $P(A \cap B)$.

> For any sets A and B,
>
> $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

The preceding equation is illustrated in Figure 12.9 and follows from the fact that we count $A \cap B$ twice when we compute $P(A) + P(B)$. The proof of this property is left to the reader in Problem 19. Here, we give an example in which we use both of the two additional properties we just described.

EXAMPLE 7 Assume that $\Omega = \{1, 2, 3, 4, 5\}$ and that

$$P(1) = P(2) = 0.2, \quad P(3) = P(4) = 0.1, \quad \text{and} \quad P(5) = 0.4$$

Set $A = \{1, 3, 4\}$ and $B = \{4, 5\}$. Find $P(A \cup B)$.

Solution Observe that A and B are not disjoint. We find that

$$A \cap B = \{4\}$$

Using the second of the additional properties, we obtain

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= P(\{1, 3, 4\}) + P(\{4, 5\}) - P(\{4\}) \\ &= (0.2 + 0.1 + 0.1) + (0.1 + 0.4) - (0.1) = 0.8 \end{aligned}$$

We could have gotten the same result by observing that

$$A \cup B = \{1, 3, 4, 5\}, \qquad \text{and therefore,} \qquad A \cup B = \{2\}^c$$

which yields

$$P(A \cup B) = P(\{2\}^c) = 1 - P(\{2\}) = 1 - 0.2 = 0.8$$

•

12.2.2 Equally Likely Outcomes

An important class of random experiments with finite sample spaces is that in which all outcomes are equally likely. That is, if $\Omega = \{1, 2, \ldots, n\}$, then $P(1) = P(2) = \cdots = P(n)$, where we wrote $P(i)$ for $P(\{i\})$. Then

$$1 = P(\Omega) = \sum_{i=1}^{n} P(i) = nP(1)$$

$\sum$-notation was introduced in Section 2.2

which implies that

$$P(1) = P(2) = \cdots = P(n) = \frac{1}{n}$$

If we denote the number of elements in A by $|A|$, and if $A \subset \Omega$ with $|A| = k$, then

$$P(A) = \frac{|A|}{|\Omega|} = \frac{k}{n}$$

In the next two examples, we will discuss random experiments in which all outcomes are equally likely. You should pay particular attention to the sample space, to make sure that you understand that all of its elements are indeed equally likely.

EXAMPLE 8 Toss a fair coin three times and find the probability of the event $A = \{\text{at least two heads}\}$.

Solution The sample space in this case is

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

All outcomes are equally likely, since we assumed that the coin is fair. Because $|\Omega| = 8$, it follows that each possible outcome has probability $1/8$. Thus,

$$\begin{aligned} P(A) &= P(HHH, HHT, HTH, THH) \qquad \text{List all outcomes with two or more heads} \\ &= \frac{|A|}{|\Omega|} = \frac{4}{8} = \frac{1}{2} \end{aligned}$$

•

EXAMPLE 9 An urn contains five blue and six green balls. We draw two balls from the urn without replacement. The order in which the balls are drawn does not matter.

(a) Determine the sample space Ω and find $|\Omega|$.

(b) What is the probability that the two balls are of a different color?

(c) What is the probability that at least one of the two balls is green?

Solution **(a)** As physical objects, the balls are distinguishable and we can imagine them being numbered from 1 to 11, where we assign the first 5 numbers to the 5 blue balls and the remaining 6 numbers to the 6 green balls. The sample space for this random experiment then consists of all subsets of size 2 that can be drawn from the set of 11 balls. Each subset of size 2 is then equally likely. Using the counting techniques from Section 12.1, we find that the size of the sample space is

$$|\Omega| = \binom{11}{2} = 55.$$

since the order in which the balls are removed from the urn does not need to be accounted for, and each ball is drawn at most once.

(b) Let A denote the event that the two balls are of a different color. To obtain an outcome in which one ball is blue and the other is green, we must select one blue ball from the five blue balls, which can be done in $\binom{5}{1}$ different ways, and one green ball from the six green balls, which can be done in $\binom{6}{1}$ ways. Using the multiplication rule from Section 12.1, we find that

$$|A| = \binom{5}{1}\binom{6}{1}$$

Hence,

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\binom{5}{1}\binom{6}{1}}{\binom{11}{2}} = \frac{5 \cdot 6}{\frac{11 \cdot 10}{2}} = \frac{6}{11}$$

(c) Let B denote the event that at least one ball is green. This event can be written as a union of the following two disjoint sets:

$$B_1 = \{\text{exactly one ball is green}\}$$
$$B_2 = \{\text{both balls are green}\}$$

Since $B = B_1 \cup B_2$ with $B_1 \cap B_2 = \emptyset$, it follows that

$$P(B) = P(B_1) + P(B_2)$$

Using a similar argument as in (b), we find that

$$P(B) = \frac{\binom{5}{1}\binom{6}{1}}{\binom{11}{2}} + \frac{\binom{5}{0}\binom{6}{2}}{\binom{11}{2}} = \frac{5 \cdot 6}{55} + \frac{15}{55} = \frac{9}{11}$$

•

EXAMPLE 10 Four cards are drawn at random and without replacement from a standard deck of 52 cards. What is the probability of at least two kings?

Solution We find

$$P(\text{at least two kings}) = 1 - [P(\text{no kings}) + P(\text{one king})]$$

There are $\binom{52}{4}$ ways of selecting four cards. There are four kings in a standard deck of cards. There are $\binom{48}{4}$ ways of selecting a hand of four cards that does not contain any kings and there are $\binom{4}{1}\binom{48}{3}$ ways of selecting a hand of four cards that contains exactly one king. Hence,

$$\begin{aligned} P(\text{at least two kings}) &= 1 - \frac{\binom{48}{4}}{\binom{52}{4}} - \frac{\binom{4}{1}\binom{48}{3}}{\binom{52}{4}} \\ &= 1 - \frac{48 \cdot 47 \cdot 46 \cdot 45}{52 \cdot 51 \cdot 50 \cdot 49} - \frac{16 \cdot 48 \cdot 47 \cdot 46}{52 \cdot 51 \cdot 50 \cdot 49} \approx 0.0257 \end{aligned}$$

An Application from Genetics. Gregor Mendel, an Austrian monk, experimented with peas to study the laws of inheritance. He started his experiments in 1856. His work was fundamental in understanding the laws of inheritance. It took over 35 more years until Mendel's original work was publicized and his conclusions were confirmed in additional experiments.

We will use the current knowledge about inheritance to determine the likelihood of outcomes of certain crossings. We describe one of Mendel's experiments that studies the inheritance of flower color in peas. Mendel had seeds that produced plants with either red or white flowers. Flower color in Mendel's peas is determined by a single gene. This gene occurs in two forms, called alleles, which we denote by C and c, respectively. Since pea plants are diploid organisms, each plant has two genes that determine flower color, one from each parent plant. The following genotypes (i.e., pairs of genes) are thus possible: CC, Cc, and cc. The genotypes CC and Cc have red flowers, whereas the genotype cc has white flowers.

EXAMPLE 11 Suppose that you cross two pea plants, both of type Cc. Determine the probability of each genotype occurring in the next generation. What is the probability that a randomly chosen seed from this crossing results in a plant with red flowers?

Solution We denote the offspring of the crossing as a pair [e.g., (C, c)] whose first entry is the maternal contribution and second entry is the paternal contribution. For instance, an offspring of type (c, C) inherited a c from the mother and a C from the father. (See Figure 12.10.) We list all possible outcomes of this crossing in the sample space

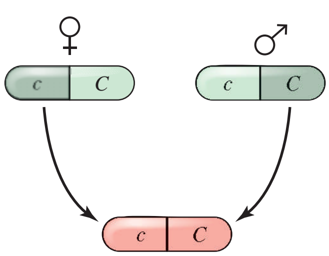


Figure 12.10 The crossing $Cc \times Cc$ resulted in an offspring of type Cc.

$$\Omega = \{(C, C), (C, c), (c, C), (c, c)\}$$

The laws of inheritance imply that gametes form at random and that, therefore, all outcomes in Ω are equally likely. (This fact is known as **Mendel's first law**.) Since $|\Omega| = 4$, it follows that each outcome has probability 1/4.

Although the sample space has four different outcomes, there are only three different genotypes, since (C, c) and (c, C) denote the same genotype, Cc. In what follows, we will often denote the event $\{(C, c), (c, C)\}$ simply by Cc, as is customary in genetics. We therefore find that

$$P(CC) = \frac{1}{4}, \quad P(Cc) = \frac{1}{2}, \quad \text{and} \quad P(cc) = \frac{1}{4}$$

Since the two genotypes CC and Cc result in red flowers, it follows that

$$P(\text{red}) = P(\{(C, C), (C, c), (c, C)\}) = \frac{3}{4}$$

The Mark–Recapture Method. The mark–recapture method is commonly used to estimate population sizes. We illustrate the method with a fish population. Suppose that N fish are in a lake, where N is unknown. To get an idea of how big N is, we capture K

fish, mark them, and subsequently release them. We wait until the marked fish in the lake have had sufficient time to mix with the other fish. We then capture n fish. Suppose that k of the n fish are marked. (Assume that $k > 0$.) Then if the fish are mixed well again, the ratio of the marked to unmarked fish in the sample of size n should approximately be equal to the ratio of marked to unmarked fish in the lake; that is,

$$\frac{k}{n} \approx \frac{K}{N}$$

We might therefore want to conclude that there are about

$$N \approx K\frac{n}{k}$$

fish in the lake. Can we really estimate N in this way? We will explain in the next two examples why this approach works.

EXAMPLE 12 Given the mark–recapture experiment, compute the probability of finding k marked fish in a sample of size n.

Solution There are N fish in the lake, K of which are marked. We choose a sample of size n. Each outcome is therefore a subset of size n, and all outcomes are equally likely. Using the counting techniques from Section 12.1, we find that

$$|\Omega| = \binom{N}{n}$$

since the order in the sample is not important.

We denote by A the event that the sample of size n contains exactly k marked fish. To determine how many outcomes contain exactly k marked fish, we argue as follows: We need to select k fish from the K marked ones and $n-k$ fish from the $N-K$ unmarked ones. Selecting the k marked fish can be done in $\binom{K}{k}$ ways; selecting the $n-k$ unmarked fish can be done in $\binom{N-K}{n-k}$ ways. Since each choice of k marked fish can be combined with any choice of the $n-k$ unmarked fish, we use the multiplication principle to find the total number of ways of obtaining a sample of size n with exactly k marked fish. We obtain

$$|A| = \binom{K}{k}\binom{N-K}{n-k}$$

Therefore,

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

(This example is of the same basic type as the urn problem in Example 9.) •

We will now give an argument that explains why the total number of fish in the lake can be estimated from the formula $N \approx Kn/k$.

EXAMPLE 13 Assume that there are K marked fish in the lake. We take a sample of size n and observe k marked fish. Show that the value of N which maximizes the probability of finding k marked fish in a sample of size n is the largest integer less than or equal to Kn/k. We use this value as our estimate for the population size N. Since this estimate of N maximizes the probability of what we observe, it is called a **maximum likelihood estimate**.

Solution We denote by A the event that the sample of size n contains exactly k marked fish. We showed in Example 12 that

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}.$$

We now consider $P(A)$ as a function of N and denote it by p_N. To find the value of N that maximizes p_N, we look at the ratio p_N/p_{N-1}. (The function p_N is not continuous, since it is defined only for integer values of N; therefore, we cannot differentiate p_N to find its maximum.) The ratio is given by

$$\frac{p_N}{p_{N-1}} = \frac{\dfrac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}}{\dfrac{\binom{K}{k}\binom{N-1-K}{n-k}}{\binom{N-1}{n}}} = \frac{\binom{N-K}{n-k}}{\binom{N-1-K}{n-k}} \frac{\binom{N-1}{n}}{\binom{N}{n}}$$

$$= \frac{(N-K)!\,(n-k)!\,(N-1-K-n+k)!}{(n-k)!\,(N-K-n+k)!\,(N-1-K)!} \frac{(N-1)!\,n!\,(N-n)!}{n!\,(N-1-n)!\,N!}$$

When we cancel terms, we find that

$$\frac{p_N}{p_{N-1}} = \frac{N-K}{N-K-n+k} \cdot \frac{N-n}{N}$$

We will now investigate when this ratio is greater than or equal to 1, since, when it is, we can find the values of N for which p_N exceeds p_{N-1}. Values of N for which p_N exceeds both p_{N-1} and p_{N+1} are local maxima. The ratio p_N/p_{N-1} is greater than or equal to 1 if

$$(N-K)(N-n) \geq N(N-K-n+k)$$

Multiplying out both sides of this inequality gives

$$N^2 - Nn - KN + Kn \geq N^2 - NK - Nn + Nk$$

Simplifying yields

$$Kn \geq kN$$

or

$$N \leq K\frac{n}{k}$$

Thus, $p_N \geq p_{N-1}$ as long as $N \leq Kn/k$. If Kn/k is an integer, then $p_N = p_{N-1}$ for $N = Kn/k$ and both Kn/k and $Kn/k - 1$ maximize the probability of observing k fish in the sample of size n. Either of the two values can then be chosen as estimates for the number of fish in the lake. If Kn/k is not an integer, then the largest integer less than Kn/k maximizes the probability p_N. To arrive at just one value, we will always use the largest integer less than or equal to Kn/k to estimate the total number of fish in the lake. •

EXAMPLE 14 Assume that there are 15 marked fish in a lake. We take a sample of size 10 and observe 4 marked fish. Find an estimate of the number of fish in the lake on the basis of Example 13.

Solution It follows from Example 13 that an estimate for the number of fish in the lake, denoted by N, is the largest integer less than or equal to Kn/k, where, in this example, $K = 15$, $n = 10$, and $k = 4$. Since

$$K\frac{n}{k} = 15 \cdot \frac{10}{4} = 37.5$$

we estimate that there are 37 fish in the lake.

To see that this value indeed maximizes

$$p_N = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

(defined in Example 13), we graph p_N as a function of N for $K = 15$, $n = 10$, and $k = 4$ (Figure 12.11). Since there are 15 marked fish in the lake and we sample 6 unmarked fish, the number of fish in the lake must be at least 21. Therefore, $p_N = 0$ for $N < 21$. •

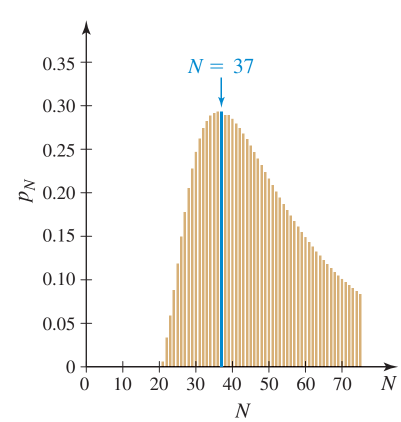


Figure 12.11 The function p_N for different values of N when $K = 15$, $n = 10$, and $k = 4$. The graph shows that p_N is maximal for $N = 37$.

Section 12.2 Problems

12.2.1

In Problems 1–4, determine the sample space for each random experiment.

1. The random experiment consisting of tossing a coin three times.

2. The random experiment consisting of rolling a six-sided die twice.

3. An urn contains five balls numbered 1–5, respectively. The random experiment consists of selecting two balls simultaneously without replacement.

4. An urn contains six balls numbered 1–6, respectively. The random experiment consists of selecting five balls simultaneously without replacement.

In Problems 5–8, assume that

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$A = \{1, 3, 5\}$, ***and*** $B = \{1, 2, 3\}$.

5. Find $A \cup B$ and $A \cap B$.

6. Find A^c and show that $(A^c)^c = A$.

7. Find $(A \cup B)^c$.

8. Are A and B disjoint?

In Problems 9–12, assume that

$$\Omega = \{1, 2, 3, 4, 5\}$$

$P(1) = 0.1$, $P(2) = 0.2$, ***and*** $P(3) = P(4) = 0.05$. ***Furthermore, assume that*** $A = \{1, 3, 5\}$ ***and*** $B = \{2, 3, 4\}$.

9. Find $P(5)$.

10. Find $P(A)$ and $P(B)$.

11. Find $P(A^c)$.

12. Find $P(A \cup B)$.

In Problems 13–15, assume that

$$\Omega = \{1, 2, 3, 4\}$$

and $P(1) = 0.1$. ***Furthermore, assume that*** $A = \{2, 3\}$ ***and*** $B = \{3, 4\}$, $P(A) = 0.7$, ***and*** $P(B) = 0.5$.

13. Find $P(3)$.

14. Set $C = \{1, 2\}$. Find $P(C)$.

15. Find $P((A \cap B)^c)$.

16. Assume that $P(A \cap B^c) = 0.1$, $P(B \cap A^c) = 0.5$, and $P((A \cup B)^c) = 0.2$. Find $P(A \cap B)$.

17. Assume that $P(A \cap B) = 0.1$, $P(A) = 0.4$, and $P(A^c \cap B^c) = 0.2$. Find $P(B)$.

18. Assume that $P(A) = 0.4$, $P(B) = 0.4$, and $P(A \cup B) = 0.7$. Find $P(A \cap B)$ and $P(A^c \cap B^c)$.

19. Prove the second of the additional properties, namely,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{12.5}$$

(a) Use a diagram to show that B can be written as a disjoint union of the sets $A \cap B$ and $B \cap A^c$.

(b) Use a diagram to show that $A \cup B$ can be written as a disjoint union of the sets A and $B \cap A^c$.

(c) Use your results in (a) and (b) to show that

$$P(A \cup B) = P(A) + P(B \cap A^c)$$

and

$$P(B \cap A^c) = P(B) - P(A \cap B)$$

Conclude from these two equations that (12.5) holds.

20. If $A \subset B$, we can define the difference between the two sets A and B, denoted by $B - A$ (read "B minus A"),

$$B - A = B \cap A^c$$

as illustrated in Figure 12.12.

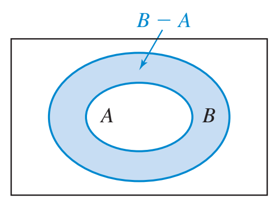


Figure 12.12 The set A is contained in B. The shaded area is the difference of A and B, $B - A$.

Go through the following steps to show that the difference rule

$$P(B - A) = P(B) - P(A) \tag{12.6}$$

holds:

(a) Use the diagram in Figure 12.12 to show that B can be written as a disjoint union of A and $B - A$.

(b) Use your result in (a) to conclude that

$$P(B) = P(A) + P(B - A)$$

and show that (12.6) follows from this equation.

(c) An immediate consequence of (12.6) is the result that if $A \subset B$, then

$$P(A) \leq P(B)$$

Use (12.6) to show this inequality.

12.2.2

21. Toss two fair coins and find the probability of at least one head.

22. Toss three fair coins and find the probability of no heads.

23. Toss four fair coins and find the probability of exactly two heads.

24. Toss four fair coins and find the probability of three or more heads.

25. Roll a fair die twice and find the probability of at least one 4.

26. Roll two fair dice and find the probability that the sum of the two numbers is even.

27. Roll two fair dice, one after the other, and find the probability that the first number is larger than the second number.

28. Roll two fair dice and find the probability that the minimum of the two numbers will be greater than 4.

29. In Example 11, we considered a cross between two pea plants, each of genotype Cc. Find the probability that a randomly chosen seed from this cross has white flowers.

30. In Example 11, we considered a cross between two pea plants, each of genotype Cc. Now we cross a pea plant of genotype cc with a pea plant of genotype Cc.

(a) What are the possible outcomes of this crossing?

(b) Find the probability that a randomly chosen seed from this crossing results in red flowers.

31. Suppose that two parents are of genotype Aa. What is the probability that their offspring is of genotype Aa? (Assume Mendel's first law.)

32. Suppose that one parent is of genotype AA and the other is of genotype Aa. What is the probability that their offspring is of genotype AA? (Assume Mendel's first law.)

33. A family has three children. Assuming a 1:1 sex ratio, what is the probability that all of the children are girls?

34. A family has three children. Assuming a 1:1 sex ratio, what is the probability that at least one child is a boy?

35. A family has four children. Assuming a 1:1 sex ratio, what is the probability that no more than two children are girls?

Color Blindness

In Problems 36–37, we discuss the inheritance of red–green color blindness. Color blindness is an X-linked inherited disease. A woman who carries the color blindness gene on one of her X chromosomes, but not on the other, has normal vision. A man who carries the gene on his only X chromosome is color blind.

36. If a woman with normal vision who carries the color blindness gene on one of her X chromosomes has a child with a man who has normal vision, what is the probability that their child will be color blind?

37. If a woman with normal vision who carries the color blindness gene on one of her X chromosomes has a child with a man who is red–green color blind, what is the probability that their child has normal vision?

38. **Cystic Fibrosis** Cystic fibrosis is an autosomal recessive disease, which means that two copies of the gene must be mutated for a person to be affected. Assume two unaffected parents who each carry a single copy of the mutated gene have a child. What is the probability that the child is affected?

39. An urn contains three red and two blue balls. You remove two balls without replacement. What is the probability that the two balls are of a different color?

40. An urn contains five blue and three green balls. You remove three balls from the urn without replacement. What is the probability that at least two out of the three balls are green?

41. You select 2 cards without replacement from a standard deck of 52 cards. What is the probability that both cards are spades?

42. You select 5 cards without replacement from a standard deck of 52 cards. What is the probability that you get four aces?

43. An urn contains four green, six blue, and two red balls. You take three balls out of the urn without replacement. What is the probability that all three balls are of different colors?

44. An urn contains three green, five blue, and four red balls. You take three balls out of the urn without replacement. What is the probability that all three balls are of the same color?

45. Four cards are drawn at random without replacement from a standard deck of 52 cards. What is the probability of at least one ace?

46. Four cards are drawn at random without replacement from a standard deck of 52 cards. What is the probability of exactly one pair?

47. Thirteen cards are drawn at random without replacement from a standard deck of 52 cards. What is the probability that all are red?

48. Four cards are drawn at random without replacement from a standard deck of 52 cards. What is the probability that all are of different suits?

49. Five cards are drawn at random without replacement from a standard deck of 52 cards. What is the probability of exactly two pairs?

50. Five cards are drawn at random without replacement from a standard deck of 52 cards. What is the probability of three of a kind and a pair (for instance, Q Q Q 3 3)? (This is called a *full house* in poker.)

51. A lake contains an unknown number of fish, denoted by N. You capture 100 fish, mark them, and subsequently release them. Later, you return and catch 10 fish, 3 of which are marked.

(a) Find the probability that exactly 3 out of 10 fish you just caught will be marked. This probability will be a function of N, the unknown number of fish in the lake.

(b) Find the value of N that maximizes the probability you computed in (a), and show that this value agrees with the value we computed in Example 13.

12.3 Conditional Probability and Independence

Before we define *conditional probability* and *independence*, we will illustrate these concepts by using Mendel's experiments on crossing of peas that we considered in the previous section.

Assume that two parent pea plants are of genotype Cc. Suppose you know that the offspring of the crossing $Cc \times Cc$ has red flowers. What is the probability that it is of genotype CC? We can find this probability by noting that one of the three equally likely possibilities that produce red flowers [namely, (C, C), (C, c), and (c, C) if we list the types according to maternal and paternal contributions as in Example 11 of the previous section] is of type CC. Hence, the probability that the offspring is of genotype CC is 1/3. Such a probability, conditioned on some prior knowledge (such as knowing flower color), is called a **conditional probability**.

Suppose now that the paternally transmitted gene in the offspring of the crossing $Cc \times Cc$ is of type C. What is the probability that the maternally transmitted gene in the offspring is of type c? To answer this question, we note that the paternal gene has no impact on the choice of the maternal gene in this case. The probability that the maternal gene is of type c is therefore 1/2. We say that the maternal gene is **independent** of the paternal gene: Knowing which of the paternal genes was chosen does not change the probability of the maternal gene.

12.3.1 Conditional Probability

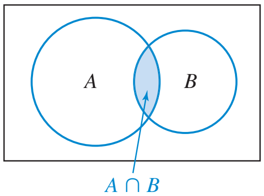


Figure 12.13 The conditional probability of A given B is the proportion of A in the set B, $A \cap B$, relative to the set B.

As illustrated in the introduction of this section, conditional probabilities have something to do with prior knowledge. Suppose we know that the event B has occurred and that $P(B) > 0$. Then the conditional probability of the event A given B, denoted by $P(A \mid B)$, is the probability that A will occur given the fact that B has occurred. This probability is defined as

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{12.7}$$

To explain this definition, we look at Figure 12.13. The probability of A given B is the proportion of A in the set B relative to B.

In the next example, we will use the definition (12.7) to repeat the introductory example and to find the probability that an offspring is of genotype CC given that its flower color is red.

EXAMPLE 1 Find the probability that the offspring of a $Cc \times Cc$ crossing of pea plants is of type CC given that its flowers are red.

Solution Let A denote the event that the offspring is of genotype CC and B represent the event that the flower color of the offspring is red. We want to find $P(A \mid B)$. Using (12.7), we have

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

The *unconditional* probabilities $P(B)$ and $P(A \cap B)$ are computed with the use of the sample space $\Omega = \{(C, C), (C, c), (c, C), (c, c)\}$, whose outcomes all have the same probability. The probability $P(B)$ is the probability that the genotype of the offspring is in the set $\{(C, C), (C, c), (c, C)\}$. Since the sample space has equally likely outcomes, $P(B) = 3/4$. To compute $P(A \cap B)$, we note that $A \cap B$ is the event that the offspring is of genotype CC. Using the sample space Ω, we find that $P(A \cap B) = 1/4$. Hence,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

which is the same answer that we obtained before. •

As we have seen in Example 1, (12.7) can be used to compute conditional probabilities. By rearranging terms, we can also use (12.7) to compute probabilities of the intersection of events. All we need do is multiply both sides of that equation by $P(B)$, to obtain

$$P(A \cap B) = P(A \mid B)P(B) \tag{12.8}$$

In Equation (12.7), we conditioned on the event B. If we condition on A instead, we have the following identity:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

Rearranging terms as in (12.8), we find that

$$P(A \cap B) = P(B \mid A)P(A) \tag{12.9}$$

Formulas (12.8) and (12.9) are particularly useful for computing probabilities in two-stage experiments. We illustrate the use of these identities in the next example. Often, there is a natural choice for which of the two events to condition on.

EXAMPLE 2 Suppose that we draw 2 cards at random without replacement from a standard deck of 52 cards. Compute the probability that both cards are diamonds.

Solution This example can be thought of as a two-stage experiment: We first draw one card and then, without replacing the first one, we draw a second card. We define the two events

$$A = \{\text{the first card is diamond}\}$$
$$B = \{\text{the second card is diamond}\}$$

Then

$$A \cap B = \{\text{both cards are diamonds}\}$$

Now, should we use (12.8) or (12.9)? Since the first card is drawn first, it will be easier to condition on the outcome of the first draw than on the second draw; that is, we will compute $P(B \mid A)$ rather than $P(A \mid B)$ and then use (12.9). We have

$$P(A) = \frac{13}{52}$$

since 13 out of the 52 cards are diamonds and each card has the same probability of being drawn. To compute $P(B \mid A)$, we note that if the first card is a diamond, then there are 12 diamonds left in the deck of the remaining 51 cards. Therefore,

$$P(B \mid A) = \frac{12}{51}$$

Using (12.9), we find that

$$P(A \cap B) = P(B \mid A)P(A) = \frac{12}{51} \cdot \frac{13}{52} = \frac{1}{17}$$

12.3.2 The Law of Total Probability

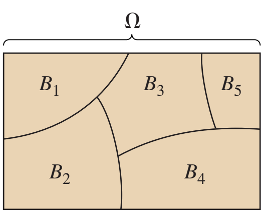


Figure 12.14 The sets $B_1, B_2, \ldots, B_5$ form a partition of Ω.

We begin this subsection by defining a partition of a sample space. Suppose the sample space Ω is written as a union of n disjoint sets $B_1, B_2, \ldots, B_n$. That is,

(i) $B_i \cap B_j = \emptyset$ whenever $i \neq j$
(ii) $\Omega = \bigcup_{i=1}^{n} B_i$

We then say that the sets $B_1, B_2, \ldots, B_n$ form a **partition** of the sample space Ω. (See Figure 12.14.)

Now, let A be an event. We can use our newly defined partition of Ω to write A as a union of disjoint sets:

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_n)$$

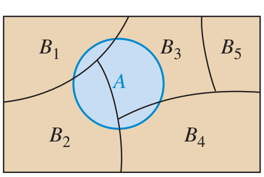


Figure 12.15 The set A is a disjoint union of sets of the form $A \cap B_i$.

We illustrate this union in Figure 12.15.

Since the sets $A \cap B_i$, $i = 1, 2, \ldots, n$, are disjoint, it follows that

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) \qquad \text{Use Property (3) in The Definition of Probability}$$

To evaluate $P(A \cap B_i)$, we might find it useful to condition on B_i; that is, $P(A \cap B_i) = P(A \mid B_i)P(B_i)$. Then

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i)P(B_i) \tag{12.10}$$

Equation (12.10) is known as the **law of total probability**.

EXAMPLE 3 A test for the HIV virus shows a positive result in 99% of all cases when the virus is actually present and in 5% of all cases when the virus is not present (a *false positive* result). If such a test is administered to a randomly chosen individual, what is the probability that the test result is positive? Assume that the prevalence of the virus in the population is 1/200.

Solution We set

$$A = \{\text{test result is positive}\}$$

Individuals in this population fall into two sets: those who are infected with the HIV virus and those who are not. These two sets form a partition of the population. If we pick an individual at random from the population, then the person belongs to one of the two sets. We define

$$B_1 = \{\text{person is infected}\}$$
$$B_2 = \{\text{person is not infected}\}$$

Using (12.10), we can write

$$P(A) = P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2)$$

Now, $P(B_1) = 1/200$ and $P(B_2) = 199/200$. Furthermore, $P(A \mid B_1) = 0.99$ and $P(A \mid B_2) = 0.05$. Hence,

$$P(A) = (0.99)\frac{1}{200} + (0.05)\frac{199}{200} = 0.0547$$

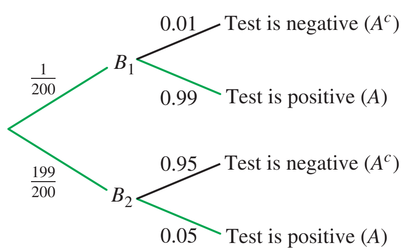


Figure 12.16 The green paths in this tree diagram lead to the event A. The numbers on the branches represent the respective probabilities.

We can illustrate Example 3 by a **tree diagram**, as shown in Figure 12.16. The numbers on the branches represent the respective probabilities. To find the probability of a positive test result, we multiply the probabilities along the paths that lead to tips labeled "test positive." (The two paths are marked in the tree.) Adding the results along the different paths then yields the desired probability—that is,

$$P(\text{test positive}) = \frac{1}{200}(0.99) + \frac{199}{200}(0.05) = 0.0547$$

as in Example 3. Tree diagrams are quite useful when used together with the law of total probability.

In the next example, we will return to our pea plants. Recall that red-flowering pea plants are of genotype CC or Cc and that white-flowering pea plants are of genotype cc.

EXAMPLE 4 Suppose that you have a batch of red-flowering pea plants, 20% of which are of genotype Cc and 80% of which are of genotype CC. You pick one of the red-flowering plants at random and cross it with a white-flowering plant. Find the probability that the offspring will produce red flowers.

Solution A white-flowering plant is of genotype cc. If the red-flowering parent plant is of genotype CC (probability 0.8) and is crossed with a white-flowering plant, then all offspring are of genotype Cc and therefore produce red flowers. If the red-flowering parent plant is of genotype Cc (probability 0.2) and is crossed with a white-flowering plant, then, with probability 0.5, an offspring is of genotype Cc (and therefore red flowering) and, with probability 0.5, an offspring is of genotype cc (and therefore white flowering). We use a tree diagram to illustrate the computation of the probability of a red-flowering offspring (Figure 12.17).

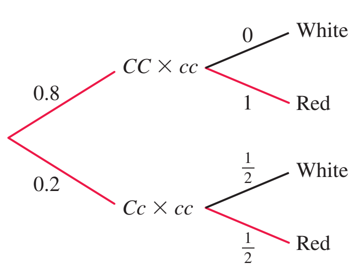


Figure 12.17 The tree diagram for Example 4: The red paths lead to red flowering offspring.

The paths that lead to a red-flowering offspring are marked. Using the tree diagram (or the law of total probability), we find that

$$P(\text{red-flowering offspring}) = (0.8)(1) + (0.2)(0.5) = 0.9$$

12.3.3 Independence

Suppose that you toss a fair coin twice. Let A be the event that the first toss results in heads and B the event that the second toss results in heads. Suppose that A occurs. Does this change the probability that B will occur? The answer is obviously no. The outcome of the first toss does not influence the outcome of the second toss. We can express this fact mathematically with conditional probabilities:

$$P(B|A) = P(B) \tag{12.11}$$

Although we say that A and B are *independent*, we will not use (12.11) as the definition of independence; rather, we will use the definition of conditional probabilities to rewrite (12.11). Since $P(B|A) = P(A \cap B)/P(A)$, (12.11) can be written as

$$\frac{P(A \cap B)}{P(A)} = P(B)$$

Multiplying both sides by $P(A)$, we obtain $P(A \cap B) = P(A)P(B)$. We use this formula as our definition:

> Two events A and B are **independent** if
>
> $$P(A \cap B) = P(A)P(B)$$

EXAMPLE 5 Suppose you draw 1 card from a standard deck of 52 cards. Let

$$A = \{\text{card is spade}\}$$

$$B = \{\text{card is king}\}$$

Show that A and B are independent.

Solution To show that A and B are independent, we compute

$$P(A) = \frac{13}{52} \qquad P(B) = \frac{4}{52}$$

and

$$P(A \cap B) = P(\text{card is king of spades}) = \frac{1}{52}.$$

Since

$$P(A)P(B) = \frac{13}{52} \cdot \frac{4}{52} = \frac{1}{52} = P(A \cap B),$$

it follows that A and B are independent. •

We now return to our pea plant example to illustrate how we can compute probabilities of intersections of events when we know that the events are independent.

EXAMPLE 6 What is the probability that the offspring of a $Cc \times Cc$ crossing is of genotype cc?

Solution Previously, we used a sample space with equally likely outcomes to compute the answer. But we can also use independence to compute this probability.

In order for the offspring to be of genotype cc, both parents must contribute a c gene. Let

$$A = \{\text{paternal gene is } c\}$$

$$B = \{\text{maternal gene is } c\}$$

Now, it follows from the laws of inheritance that A and B are independent and that $P(A) = P(B) = 1/2$. Hence,

$$P(cc) = P(A \cap B) = P(A)P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

This is the same result that we obtained when we looked at the sample space $\Omega = \{(C, C), (C, c), (c, C), (c, c)\}$ and observed that all four possible types were equally likely. •

We can extend independence to more than two events. We say that events A_1, $A_2, \ldots, A_n$ are independent if, for any $1 \leq i_1 < i_2 < \cdots < i_k \leq n$,

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}) \tag{12.12}$$

To see what (12.12) means when we have three events A, B, and C, we write the conditions explicitly: Three events A, B, and C are independent if

$$\left.\begin{aligned} P(A \cap B) &= P(A)P(B) \\ P(A \cap C) &= P(A)P(C) \\ P(B \cap C) &= P(B)P(C) \end{aligned}\right\} \tag{12.13}$$

and

$$P(A \cap B \cap C) = P(A)P(B)P(C) \tag{12.14}$$

That is, both (12.13) *and* (12.14) must hold.

The number of conditions we must verify increases quickly with the number of events. When there are just 2 sets, only one condition must be checked, namely, $P(A \cap B) = P(A)P(B)$. When there are 3 sets, as we just saw, there are four conditions: $\binom{3}{2}$ conditions involving pairs of sets and $\binom{3}{3}$ conditions involving all 3 sets. With 4 sets, there is a total of $\binom{4}{2} + \binom{4}{3} + \binom{4}{4} = 11$ conditions; with 5 sets, 26 conditions; and with 10 sets, 1013 conditions.

We emphasize that it is *not* enough to check independence between pairs of events to determine whether a collection of sets is independent. However, independence between pairs of sets is an important property itself, and we wish to define it. We say that events $A_1, A_2, \ldots, A_n$ are **pairwise independent** if

$$P(A_i \cap A_j) = P(A_i)P(A_j) \qquad \text{whenever } i \neq j$$

The next example presents a situation in which events are *pairwise independent* but not *independent*.

EXAMPLE 7 Roll two dice. Let

$$A = \{\text{the first die shows an even number}\}$$

$$B = \{\text{the second die shows an even number}\}$$

$$C = \{\text{the sum of the two dice is odd}\}$$

Show that A, B, and C are pairwise independent but not independent.

Solution To show pairwise independence, we compute

$$P(A) = \frac{18}{36} = \frac{1}{2} \qquad P(B) = \frac{18}{36} = \frac{1}{2} \qquad P(C) = \frac{18}{36} = \frac{1}{2}$$

$$P(A \cap B) = \frac{9}{36} = \frac{1}{4} = P(A)P(B)$$

$$P(A \cap C) = \frac{9}{36} = \frac{1}{4} = P(A)P(C) \qquad A \cap C = \{\text{first die is even and second die is odd}\}$$

$$P(B \cap C) = \frac{9}{36} = \frac{1}{4} = P(B)P(C)$$

However, if both A and B occur (i.e., both dice show even numbers), then their sum cannot be odd; i.e., C cannot also occur. So,

$$P(A \cap B \cap C) = P(\emptyset) = 0 \neq P(A)P(B)P(C)$$

which shows that A, B, and C are not independent. •

When events are independent, we can use (12.12) to compute the probability of their intersections, as illustrated in the next example.

EXAMPLE 8 Assume a 1:1 sex ratio. A family has five children. Find the probability that at least one of the children is a girl.

Solution Instead of computing the probability that at least one of the children is a girl, we will look at the complement of this event. This is a particularly useful trick when we look at events that ask for the probability of "at least one." We denote by

$$A_i = \{\text{the } i\text{th child is a boy}\}$$

Then the events $A_1, A_2, \ldots, A_5$ are independent. Let

$$B = \{\text{at least one of the children is a girl}\}$$

Instead of computing the probability of B directly, we compute the complement of B. Now, B^c is the event that all children are boys, which is expressed as

$$B^c = A_1 \cap A_2 \cap \cdots \cap A_5$$

It follows that

$$\begin{aligned} P(B) &= 1 - P(B^c) = 1 - P(A_1 \cap A_2 \cap \cdots \cap A_5) \\ &= 1 - P(A_1)P(A_2) \cdots P(A_5) \\ &= 1 - \left(\frac{1}{2}\right)^5 = \frac{31}{32} \end{aligned}$$

If we had tried to compute $P(B)$ directly, we would have needed to compute the probability of exactly one girl, exactly two girls, and so on, and then add all the probabilities. It is quicker and easier to compute the probability of the complement.

•

12.3.4 The Bayes Formula

In Example 3, we computed the probability that the result of an HIV test of a randomly chosen individual is positive. For the individual, however, it is much more important to know whether a positive test result actually means that he or she is infected. Recall that we defined

$$A = \{\text{test result is positive}\}$$

$$B_1 = \{\text{person is infected}\}$$

$$B_2 = \{\text{person is not infected}\}$$

We are interested in $P(B_1 \mid A)$—that is, the probability that a person is infected given that the result is positive. We saw in Example 3 that $P(A \mid B_1)$ and $P(A \mid B_2)$ followed immediately from the characteristics of the test. Now we wish to compute a conditional probability in which the roles of A and B_1 are reversed.

Before we compute the probability for this specific example, we look at the general case. We assume that the sets $B_1, B_2, \ldots, B_n$ form a partition of the sample space Ω, A is an event, and the probabilities $P(A \mid B_i)$, $i = 1, 2, \ldots, n$ are known. We are interested in computing $P(B_i \mid A)$. We can accomplish this as follows: Using the definition of conditional probabilities, we find that

$$P(B_i \mid A) = \frac{P(A \cap B_i)}{P(A)} \tag{12.15}$$

To compute $P(A \cap B_i)$, we now condition on B_i; that is,

$$P(A \cap B_i) = P(A \mid B_i)P(B_i) \tag{12.16}$$

To evaluate the denominator $P(A)$, we use the law of total probability:

$$P(A) = \sum_{j=1}^{n} P(A \mid B_j)P(B_j) \tag{12.17}$$

Combining (12.15), (12.16), and (12.17), we arrive at the following definition, known as the Bayes formula:

The Bayes Formula Let $B_1, B_2, \ldots, B_n$ form a partition of Ω, and let A be an event. Then

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{\sum_{j=1}^{n} P(A \mid B_j)P(B_j)}$$

We will now return to our example. We wish to find $P(B_1 \mid A)$—that is, the probability that a person is infected given a positive result. We partition the population into the two sets B_1 and B_2. Then, using the Bayes formula, we find that

$$P(B_1 \mid A) = \frac{P(A \mid B_1)P(B_1)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2)}$$

$$= \frac{(0.99) \cdot \frac{1}{200}}{(0.99) \cdot \frac{1}{200} + (0.05) \cdot \frac{199}{200}} \approx 0.090$$

The denominator is equal to $P(A)$, which we already computed in Example 3.

This example is worth discussing in more detail. The probability is quite small that a person is infected given a positive result. But you should compare this probability with the unconditional probability that a person is infected, namely, $P(B_1)$. The ratio of the conditional to the unconditional probability is

$$\frac{P(B_1 \mid A)}{P(B_1)} \approx 18.1$$

That is, if a test result is positive, the chance of actually being infected increases by a factor of 18 compared with the chance that a randomly chosen individual in the population who has not been tested is infected. (In practice, if a test result is positive, more than one test is performed to see whether a person is indeed infected or whether the first result was a false positive.)

If a test result is negative, we can also use the Bayes formula to compute the probability that the individual is not infected. We find that

$$P(B_2 \mid A^c) = \frac{P(B_2 \cap A^c)}{P(A^c)} = \frac{P(A^c \mid B_2)P(B_2)}{P(A^c)}$$

$$= \frac{(0.95)\frac{199}{200}}{1 - 0.0547} \approx 0.999947$$

where we used $P(A) = 0.0547$, which we computed in Example 3. This result is rather reassuring.

The reason that the probability of being infected given a positive result is so small comes from the fact that the prevalence of the disease is relatively low (1 in 200). To illustrate, we treat the prevalence of the disease as a variable and compute $P(B_1 \mid A)$ as a function of the prevalence of the disease. That is, we set

$$p = P(\text{a randomly chosen individual is infected})$$

Using the same test characteristics as before, we obtain

$$f(p) = P(B_1 \mid A) = \frac{p \cdot (0.99)}{p \cdot (0.99) + (1 - p)(0.05)} = \frac{0.99p}{0.05 + 0.94p}$$

(See Figure 12.18.) A graph of $f(p)$ is shown in Figure 12.19. We see that, for small p, $f(p)$ (the probability of being infected given a positive result) is quite small. The ratio

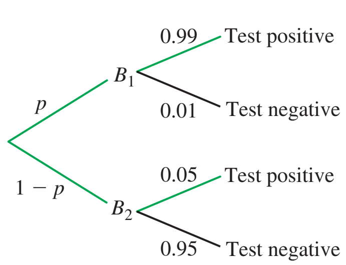


Figure 12.18 The tree diagram for computing the probability that a person is infected given that the test came back positive when the prevalence of the disease is p.

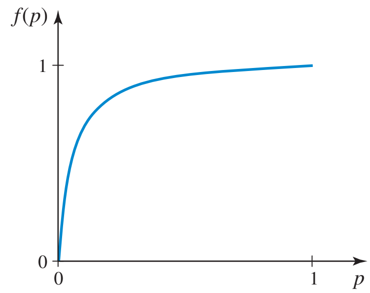


Figure 12.19 The probability of being infected given that the test came back positive as a function of the prevalence of the disease.

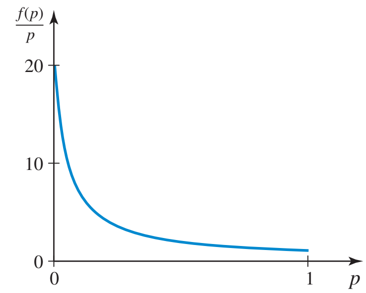


Figure 12.20 The ratio $f(p)/p$ illustrates by what factor the probability of being infected increases when the test result is positive compared with the prevalence of the disease in the population.

$f(p)/p$ is shown in Figure 12.20, from which we conclude that, although $f(p)$ is small when p is small, the ratio $f(p)/p$ is quite large for small p.

The Bayes formula is also important in genetic counseling. Hemophilia is a blood disorder that is characterized by a deficiency of a blood-clotting factor. Individuals afflicted with this disease suffer from excessive bleeding. The disease is caused by an abnormal gene that resides on the X chromosome. A female who carries the abnormal gene on one of her X chromosomes, but not on the other, is a carrier of the disease but will not develop symptoms. A male who carries the abnormal gene on his (only) X chromosome will develop symptoms of the disease. Almost all symptomatic individuals are males.

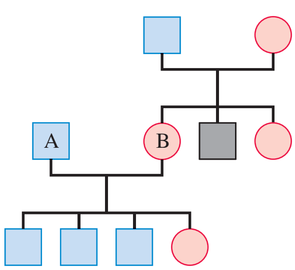


Figure 12.21 The pedigree of a family in which one member suffers from hemophilia. Squares indicate males, circles females. The black square shows an afflicted individual.

In what follows, we assume that only one parent carries the abnormal gene. If the father carries the abnormal gene (and thus suffers from hemophilia), all his daughters will be carriers, since they inherit their father's X chromosome; but all his sons will be disease free, since they inherit their father's Y chromosome. If the mother carries the abnormal gene, then her daughters have a 50% chance of being carriers and her sons have a 50% chance of suffering from the disease.

Pedigrees of families show family relationships among individuals and are indispensable tools for tracing diseases of genetic origin. In a pedigree, males are denoted by squares, females by circles; blackened symbols denote individuals who suffer from the disease that is tracked by the pedigree. Figure 12.21 shows the pedigree of a family in which one male (the black square) suffers from hemophilia. We will use this pedigree to determine the probability that individual B is a carrier of the disease given that all three sons of A and B are disease free.

We see from the pedigree that B has a hemophilic brother. Therefore, B's mother must be a carrier. There is a 50% chance that a sister of the affected individual is a carrier. We denote the event that B is a carrier by E. Then $P(E) = 1/2$. Now, assume that we are told that B has three sons with an unaffected male (A). If F denotes the event that all three sons are healthy, then

$$P(F \mid E) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

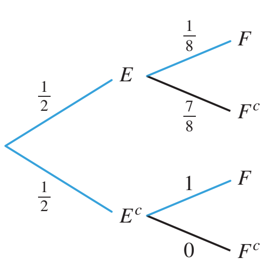


Figure 12.22 The sample space is partitioned into two sets—E and E^c—where E is the event that the individual B is a carrier of the hemophilia gene. Based on whether or not B is a carrier, the probability of the event that all three sons are healthy (F) can be computed as shown.

since if B is a carrier, each son has probability 1/2 of not inheriting the disease gene and thus being healthy.

We can use the Bayes formula to compute the probability that B is a carrier given that none of her three sons has the disease:

$$P(E \mid F) = \frac{P(E \cap F)}{P(F)} = \frac{P(F \mid E)P(E)}{P(F)}$$

To compute the denominator, we must use the law of total probability, as illustrated in the tree diagram in Figure 12.22.

We find that

$$P(F) = \frac{1}{2} \cdot \frac{1}{8} + \frac{1}{2} \cdot 1 = \frac{9}{16}$$

Therefore, using the Bayes formula, we obtain

$$P(E \mid F) = \frac{\frac{1}{8} \cdot \frac{1}{2}}{\frac{9}{16}} = \frac{1}{9}$$

or, in words, given that none of her three sons is symptomatic for the disease, the probability that B is a carrier of the gene causing hemophilia is $1/9$.

Section 12.3 Problems

12.3.1

1. Suppose you draw 2 cards from a standard deck of 52 cards. Find the probability that the second card is a spade given that the first card is a club.

2. Suppose you draw 2 cards from a standard deck of 52 cards. Find the probability that the second card is a spade given that the first card is a spade.

3. Suppose you draw 3 cards from a standard deck of 52 cards. Find the probability that the third card is a club given that the first two cards are spades.

4. Suppose you draw 3 cards from a standard deck of 52 cards. Find the probability that the third card is a club given that the first two cards are clubs.

5. An urn contains five blue and six green balls. You take two balls out of the urn, one after the other, without replacement. Find the probability that the second ball is green given that the first ball is blue.

6. An urn contains five green, six blue, and four red balls. You take three balls out of the urn, one after the other, without replacement. Find the probability that the third ball is green given that the first two balls were red.

7. A family has two children. The younger one is a girl. Find the probability that the other child is a girl as well.

8. A family has two children. One of their children is a girl. Find the probability that both children are girls.

9. You roll two fair dice. Find the probability that the first die is a 4 given that the sum is 7.

10. You roll two fair dice. Find the probability that the first die is a 5 given that the minimum of the two numbers is a 3.

11. You toss a fair coin three times. Find the probability that the first coin is heads given that at least one head occurred.

12. You toss a fair coin three times. Find the probability that at least two heads occurred given that the second toss resulted in heads.

13. You toss a fair coin four times. Find the probability that four heads occurred given that the first toss and the third toss resulted in heads.

14. You toss a fair coin four times. Find the probability of no more than three heads given that at least one toss resulted in heads.

12.3.2

15. A screening test for a disease shows a positive test result in 90% of all cases when the disease is actually present and in 15% of all cases when it is not. Assume that the prevalence of the disease is 1 in 100. If the test is administered to a randomly chosen individual, what is the probability that the result is negative?

16. A screening test for a disease shows a positive result in 92% of all cases when the disease is actually present and in 7% of all cases when it is not. Assume that the prevalence of the disease is 1 in 600. If the test is administered to a randomly chosen individual, what is the probability that the result is positive?

17. A patient underwent a diagnostic test for hypothyroidism. The diagnostic test correctly identifies patients who in fact have the disease in 93% of the cases and correctly identifies healthy patients in 81% of the cases. If 4 in 100 individuals have the disease, what is the probability that a test comes back negative?

18. A screening test for a disease shows a positive test result in 95% of all cases when the disease is actually present and in 20% of all cases when it is not. When the test was administered to a large number of people, 21.5% of the results were positive. What is the prevalence of the disease?

19. A drawer contains three bags numbered 1–3, respectively. Bag 1 contains three blue balls, bag 2 contains four green balls, and bag 3 contains two blue balls and one green ball. You choose one bag at random and take out one ball. Find the probability that the ball is blue.

20. A drawer contains six bags numbered 1–6, respectively. Bag i contains i blue balls and 2 green balls. You roll a fair die and then pick a ball out of the bag with the number shown on the die. What is the probability that the ball is blue?

21. You pick 2 cards from a standard deck of 52 cards. Find the probability that the second card is an ace. Compare this with the probability that the first card is an ace.

22. You pick 3 cards from a standard deck of 52 cards. Find the probability that the third card is an ace. Compare this with the probability that the first card is an ace.

23. You have a batch of red-flowering pea plants of which 40% are of genotype CC and 60% of genotype Cc. You pick one plant at random and cross it with a white-flowering pea plant. Find the probability that the offspring will have white flowers.

24. Suppose that you have a batch of red- and white-flowering pea plants, and suppose also that all three genotypes CC, Cc, and cc are equally represented in the batch. You pick one plant at random and cross it with a white-flowering pea plant. What is the probability that the offspring will have red flowers?

25. A bag contains two coins, one fair and the other with two heads. You pick one coin at random and flip it. Find the probability that the outcome is heads.

26. A drug company claims that a new headache drug will bring instant relief in 90% of all cases. If a person is treated with a placebo, there is a 20% chance that the person will feel instant relief. In a clinical trial, half the subjects are treated with the new drug and the other half receive the placebo. If an individual from this trial is chosen at random, what is the probability that the person will have experienced instant relief?

12.3.3

27. You are dealt 1 card from a standard deck of 52 cards. If A denotes the event that the card is a spade and if B denotes the event that the card is an ace, determine whether A and B are independent.

28. You are dealt 2 cards from a standard deck of 52 cards. If A denotes the event that the first card is an ace and B denotes the event that the second card is an ace, determine whether A and B are independent.

29. An urn contains five green and six blue balls. You take two balls out of the urn, one after the other, without replacement. If A denotes the event that the first ball is green and B denotes the event that the second ball is green, determine whether A and B are independent.

30. An urn contains four green and three blue balls. You take one ball out of the urn, note its color, and replace it. You then take a second ball out of the urn, note its color, and replace it. If A denotes the event that the first ball is green and B denotes the event that the second ball is green, determine whether A and B are independent.

31. Assume a 1:1 sex ratio. A family has three children. Find the probability of each event:

(a) $A = \{\text{all children are girls}\}$ **(b)** $B = \{\text{at least one boy}\}$

(c) $C = \{\text{at least two girls}\}$ **(d)** $D = \{\text{at most two boys}\}$

32. Assume that 20% of a very common insect species in your study area is parasitized. Assume that insects are parasitized independently of each other. If you collect 10 specimens of this species, what is the probability that no more than 2 specimens in your sample are parasitized?

33. A multiple-choice question has four choices, and a test has a total of 10 multiple-choice questions. A student passes the test only if he or she answers all questions correctly. If the student guesses the answers to all questions randomly, find the probability that he or she will pass.

34. Assume that A and B are disjoint and that both events have positive probability. Are they independent?

35. Assume that the probability that an insect species lives more than five days is 0.1. Find the probability that, in a sample of size 10 of this species, at least one insect will still be alive after five days.

36. **(a)** Use a Venn diagram to show that

$$(A \cup B)^c = A^c \cap B^c$$

(b) Use your result in (a) to show that if A and B are independent, then A^c and B^c are independent.

(c) Use your result in (b) to show that if A and B are independent, then

$$P(A \cup B) = 1 - P(A^c)P(B^c)$$

12.3.4

37. A screening test for a disease shows a positive result in 95% of all cases when the disease is actually present and in 10% of all cases when it is not. If the prevalence of the disease is 1 in 50 and an individual tests positive, what is the probability that the individual actually has the disease?

38. A screening test for a disease shows a positive result in 95% of all cases when the disease is actually present and in 10% of all cases when it is not. If a result is positive, the test is repeated. Assume that the second test is independent of the first test. If the prevalence of the disease is 1 in 50 and an individual tests positive twice, what is the probability that the individual actually has the disease?

39. A bag contains two coins, one fair and the other with two heads. You pick one coin at random and flip it. What is the probability that you picked the fair coin given that the outcome of the toss was heads?

40. You pick 2 cards from a standard deck of 52 cards. Find the probability that the first card was a spade given that the second card was a spade.

41. **Hemophilia** Suppose a woman has a hemophilic brother and one healthy son. Suppose furthermore that neither her mother nor her father were hemophilic but that her mother was a carrier for hemophilia. Find the probability that she is a carrier of the hemophilia gene.

The pedigree in Figure 12.23 shows a family in which one member (III-4) is hemophilic. In Problems 42 and 43, refer to this pedigree.

42. **(a)** Given the pedigree, find the probability that I-2 is a carrier of the hemophilia gene.

(b) Given the pedigree, find the probability that II-3 is a carrier of the hemophilia gene.

43. **(a)** Given the pedigree, find the probability that II-4 is a carrier of the hemophilia gene.

(b) Given the pedigree, find the probability that III-2 is a carrier of the hemophilia gene.

(c) Given the pedigree, find the probability that II-2 is a carrier of the hemophilia gene.

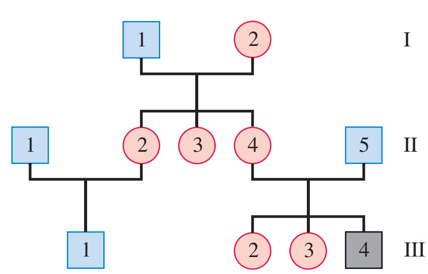


Figure 12.23 The pedigree for Problems 42 and 43. The solid black square (individual III-4) represents an afflicted male.

12.4 Discrete Random Variables and Discrete Distributions

Outcomes of random experiments frequently are real numbers, such as the number of heads in a coin-tossing experiment, the number of seeds produced in a crossing between two plants, or the life span of an insect. Such numerical outcomes can be described by **random variables**. A random variable is a function from the sample space Ω into the set of real numbers. Random variables are typically denoted by X, Y, or Z, or other capital letters chosen from the end of the alphabet. For instance,

$$X : \Omega \to \mathbf{R}$$

describes the random variable X as a map from the sample space Ω into the set of real numbers.

Random variables are classified according to their range. If X takes on a discrete set of values (finite or infinite), X is called a **discrete random variable**. If X takes on a continuous range of values—for instance, values that range over an interval—X is called a **continuous random variable**. Discrete random variables are the topic of this section; continuous random variables are the topic of the next section.

12.4.1 Discrete Distributions

In the first two examples in this section, we look at random variables that take on a discrete set of values. In the first example, this set is finite; in the second example, the set is infinite.

EXAMPLE 1 Toss a fair coin three times. Let X be a random variable that counts the number of heads in each outcome. The sample space is

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

and the random variable

$$X : \Omega \to \mathbf{R}$$

takes on values 0, 1, 2, or 3. For instance,

$$X(HHH) = 3 \qquad \text{or} \qquad X(TTH) = 1 \qquad \text{or} \qquad X(TTT) = 0$$

•

EXAMPLE 2 Toss a fair coin repeatedly until the first time heads appears. Let Y be a random variable that counts the number of trials until the first time heads shows up. The sample space is

$$\Omega = \{H, TH, TTH, TTTH, \dots\}$$

and the random variable

$$Y : \Omega \to \mathbf{R}$$

takes on values 1, 2, 3, ... For instance,

$$Y(H) = 1, \qquad Y(TH) = 2, \qquad Y(TTH) = 3 \qquad \dots$$

•

We will now turn to the problem of how to assign probabilities to the different values of a random variable X. For the moment, we will restrict the discussion to the case when the range of X is finite.

Let's go back to Example 1. The coin in Example 1 is fair. This means that each outcome in Ω has the same probability, namely, 1/8. We can translate this set of probabilities into probabilities for X. For instance,

x	$P(X = x)$
0	1/8
1	3/8
2	3/8
3	1/8

$$\begin{aligned} P(X = 1) &= P(\{HTT, THT, TTH\}) \\ &= P(HTT) + P(THT) + P(TTH) \\ &= \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8} \end{aligned}$$

We can perform similar computations for all other values of X. The table on the previous page summarizes the results.

The function $p(x) = P(X = x)$ is called a **probability mass function**. Note that $p(x) \geq 0$ and $\sum_x p(x) = 1$; these are defining properties of a probability mass function.

> **Definition** A random variable is called a **discrete random variable** if it takes on a finite or infinite set of discrete values. The probability distribution of X can be described by the probability mass function $p(x) = P(X = x)$, which has the following properties:
>
> 1. $p(x) \geq 0$ for all x
> 2. $\sum_x p(x) = 1$, where the sum is over all values of x with $P(X = x) > 0$

The probability mass function is one way to describe the probability distribution of a discrete random variable. Another important function that describes the probability distribution of a random variable X is the (cumulative) distribution function $F(x) = P(X \leq x)$. This function is defined for *any* random variable, not just discrete ones.

> **Definition** The **(cumulative) distribution function** $F(x)$ of a random variable X is defined as
>
> $$F(x) = P(X \leq x)$$

Instead of "cumulative distribution function," we will simply say "distribution function."

The probability mass function and the distribution function are equivalent ways of describing the probability distribution of a discrete random variable, and we can obtain one from the other, as illustrated in the next two examples.

EXAMPLE 3 Suppose that the probability mass function of a discrete random variable X is given by the table on the left below. Find and graph the corresponding distribution function $F(x)$.

Solution The function $F(x)$ is defined for all values of $x \in \mathbf{R}$. For instance, $F(-2.3) = P(X \leq -2.3) = P(\emptyset) = 0$, while $F(1) = P(X \leq 1) = P(X = -1 \text{ or } 0) = 0.3$. Since $F(x) = P(X \leq x)$, we must be particularly careful when x is in the range of X. To illustrate, we compute $F(1.4)$ and $F(1.5)$. We find that

x	$P(X = x)$
−1	0.1
0	0.2
1.5	0.05
3	0.15
5	0.5

$$F(1.4) = P(X \leq 1.4) = P(X = -1 \text{ or } 0) = 0.1 + 0.2 = 0.3$$

$$F(1.5) = P(X \leq 1.5) = P(X = -1, 0, \text{ or } 1.5) = 0.1 + 0.2 + 0.05 = 0.35$$

The distribution function $F(x)$ is a piecewise-defined function. We have

$$F(x) = \begin{cases} 0 & \text{for } x < -1 \\ 0.1 & \text{for } -1 \leq x < 0 \\ 0.3 & \text{for } 0 \leq x < 1.5 \\ 0.35 & \text{for } 1.5 \leq x < 3 \\ 0.5 & \text{for } 3 \leq x < 5 \\ 1 & \text{for } x \geq 5 \end{cases}$$

The graph of $F(x)$ is shown in Figure 12.24. •

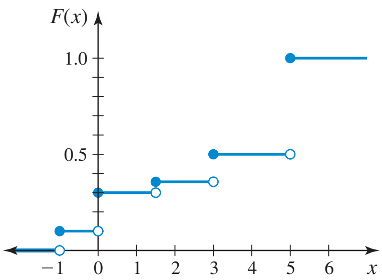


Figure 12.24 The distribution function $F(x)$ of Example 3. The solid circles indicate the values that the distribution function takes at the points where the function jumps.

Looking at Figure 12.24, we see that the graph of $F(x)$ is a nondecreasing and piecewise-constant function that takes jumps at those values x where $P(X = x) > 0$. The function $F(x)$ is right continuous; that is, for any $c \in \mathbf{R}$,

$$\lim_{x \to c^+} F(x) = F(c)$$

It is not left continuous everywhere, since, at values $c \in \mathbf{R}$ where $P(X = c) > 0$,

$$\lim_{x \to c^-} F(x) \neq F(c)$$

For instance, when $c = 3$,

$$\lim_{x \to 3^-} F(x) = 0.35 \neq F(3) = 0.5$$

Furthermore, a distribution function has the following additional characteristics:

$$\lim_{x \to -\infty} F(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} F(x) = 1$$

It is possible to obtain the probability mass function from the distribution function. Let's look at the distribution function of Example 3. The function jumps at $x = 3$ and the jump height is 0.15. Since $F(x) = P(X \leq x)$, it follows that

$$\begin{aligned} p(3) = P(X = 3) &= P(X \leq 3) - P(X < 3) \\ &= F(3) - \lim_{x \to 3^-} F(x) = 0.5 - 0.35 = 0.15 \end{aligned}$$

We see that the distribution function jumps at the values of X for which $P(X = x) > 0$. The jump height is then equal to the probability that X takes on this value.

EXAMPLE 4 Suppose the distribution function of a discrete random variable X is given by

$$F(x) = \begin{cases} 0 & \text{for } x < -5 \\ 0.2 & \text{for } -5 \leq x < 2 \\ 0.6 & \text{for } 2 \leq x < 3 \\ 0.7 & \text{for } 3 \leq x < 6.5 \\ 1.0 & \text{for } x \geq 6.5 \end{cases}$$

Find the corresponding probability mass function.

Solution We need to look at the points $x \in \mathbf{R}$ where $F(x)$ jumps. Those are the points where $p(x) = P(X = x) > 0$. The jump height is equal to the probability that X takes on this value. We find that

$$\begin{aligned} p(-5) = P(X = -5) &= P(X \leq -5) - P(X < -5) \\ &= F(-5) - \lim_{x \to 5^-} F(x) = 0.2 - 0.0 = 0.2 \end{aligned}$$

Likewise,

$$\begin{aligned} p(2) &= P(X = 2) = 0.6 - 0.2 = 0.4 \\ p(3) &= P(X = 3) = 0.7 - 0.6 = 0.1 \\ p(6.5) &= P(X = 6.5) = 1.0 - 0.7 = 0.3 \end{aligned}$$

There are no other values of x where $P(X = x) > 0$. We can check our result by adding up the probabilities we just found:

$$p(-5) + p(2) + p(3) + P(6.5) = 0.2 + 0.4 + 0.1 + 0.3 = 1.0$$

The sum adds to 1, which indicates that there cannot be any other values x where $P(X = x) > 0$. •

12.4.2 Mean and Variance

Knowing the distribution of a random variable tells us everything about the random variable. In practice, however, it is often impossible or unnecessary to know the full probability distribution of a random variable that describes a particular random experiment. Instead, it might suffice to determine a few characteristic quantities, which include the average value and a measure that describes the amount of spread.

The Average Value, or the Mean, of a Discrete Random Variable.

EXAMPLE 5 Clutch size can be thought of as a random variable. Let X denote the number of eggs per clutch laid by a certain species of bird, and assume that the distribution of X is described by the probability mass function on the left. The average number of eggs per clutch is computed as the weighted sum

x	$P(X = x)$
1	0.05
2	0.1
3	0.2
4	0.3
5	0.25
6	0.1

$$\begin{aligned}\text{average value} &= \sum_x xP(X = x)\\ &= (1)(0.05) + (2)(0.1) + (3)(0.2)\\ &\quad + (4)(0.3) + (5)(0.25) + (6)(0.1) = 3.9\end{aligned}$$

and we find that the average clutch size is 3.9. •

The average value of X is called the **expected value**, or **mean**, of X and is denoted by $E(X)$. The expected value is a very important quantity. Here is its definition:

> If X is a discrete random variable, then the expected value, or mean, of X is
>
> $$E(X) = \sum_x xP(X = x)$$
>
> where the sum is over all values of x with $P(X = x) > 0$.

When the range of X is finite, the sum in the definition is always defined. When the range of X is countably infinite, we must sum an infinite number of terms. Such sums can be finite or infinite, depending on the distribution of X. The expected value of X is defined only if both $\sum_{x<0} xP(X = x)$ and $\sum_{x\geq 0} xP(X = x)$ are finite. Determining whether such infinite sums are finite is beyond the scope of this book, and we will therefore restrict the discussion to cases in which these sums are finite.

The next example shows that the definition of the mean of a discrete random variable coincides with our everyday notion of average values.

EXAMPLE 6 On a winter day somewhere in southern Minnesota, the following temperature readings T_k (in Fahrenheit) at hour k were obtained:

k	0	1	2	3	4	5	6	7	8	9	10	11
T_k	6	6	6	5	5	5	5	5	8	10	12	12
k	12	13	14	15	16	17	18	19	20	21	22	23
T_k	12	12	12	12	10	8	8	8	5	5	3	3

On the basis of these hourly observations, the average temperature on that day, denoted by $\overline{T}$, is

$$\begin{aligned}\overline{T} &= \frac{1}{24}(6 + 6 + 6 + 5 + 5 + 5 + 5 + 5 + 8 + 10 + 12\\ &\quad + 12 + 12 + 12 + 12 + 12 + 10 + 8 + 8 + 8 + 5 + 5 + 3 + 3)\\ &= \frac{183}{24} = 7.625\end{aligned}$$

Rearranging these values according to size, we get

$$\begin{aligned}
\overline{T} &= \frac{1}{24}[(3+3)+(5+5+5+5+5+5+5)+(6+6+6) \\
&\quad +(8+8+8+8)+(10+10)+(12+12+12+12+12+12)] \\
&= \frac{1}{24}[(3)(2)+(5)(7)+(6)(3)+(8)(4)+(10)(2)+(12)(6)] \\
&= 3\cdot\frac{2}{24}+5\cdot\frac{7}{24}+6\cdot\frac{3}{24}+8\cdot\frac{4}{24}+10\cdot\frac{2}{24}+12\cdot\frac{6}{24} \\
&= \sum[\text{temperature}]\times[\text{relative frequency of that temperature}] \\
&= \frac{183}{24} = 7.625
\end{aligned}$$

In Example 6, we introduced the notion of a **relative frequency**, which tells us how often a value appears in a sample relative to the total sample size. For instance, 3°F appears twice in the sample of 24 measurements, so the relative frequency of 3°F is $2/24$.

If we interpret the relative frequencies as probabilities, we see that $\overline{T}$ in Example 6 is indeed the expected value of the temperature T on that day.

EXAMPLE 7 The following table contains the number of leaves per basil plant in a sample of 25 basil plants:

16	15	13	16	16
14	16	15	18	17
16	18	16	13	16
16	16	15	15	16
15	18	16	16	15

To find the relative frequency distribution, we must count how often each value occurs and then divide by the sample size, which is 25 in this case. The result is summarized in the following table:

No. of leaves	13	14	15	16	17	18
Relative frequency	$\frac{2}{25}$	$\frac{1}{25}$	$\frac{6}{25}$	$\frac{12}{25}$	$\frac{1}{25}$	$\frac{3}{25}$

We interpret relative frequencies as probabilities. If the random variable X denotes the number of leaves per plant with probability distribution given by the relative frequency distribution, then the expected value of the number of leaves per plant is

$$\begin{aligned}
E(X) &= 13\cdot\frac{2}{25}+14\cdot\frac{1}{25}+15\cdot\frac{6}{25}+16\cdot\frac{12}{25}+17\cdot\frac{1}{25}+18\cdot\frac{3}{25} \\
&= \frac{393}{25} = 15.72
\end{aligned}$$

Note that although the number of leaves per plant is an integer, the average number of leaves per plant is not. You would actually lose valuable information if you rounded the average number to the closest integer.

The expected value of an integer-valued random variable need not be an integer. To emphasize this point, consider how the average number of lifetime births expected by women depends on their educational attainment. (The data that follow are for women 45–50 years old in 2016, from the U.S. Census Bureau.) The number of lifetime births expected for a woman who is not a high school graduate is 2.48, whereas the corresponding number for a woman with a graduate or professional degree is 1.63.

If we rounded these numbers to the closest integer, they would be the same, namely, 2; we would no longer see the difference between the two groups of women.

We can extend the definition of the expected value of X to the expected value of a function of X. Let $g(x)$ be a function of x. Then

$$E[g(X)] = \sum_x g(x)P(X = x) \tag{12.18}$$

EXAMPLE 8 Compute $E(X^2)$ for the random variable X in Example 5.

Solution Using the probability mass function given in Example 5, we find that

$$\begin{aligned} E(X^2) &= \sum_x x^2 P(X = x) \\ &= (1)^2(0.05) + (2)^2(0.1) + (3)^2(0.2) + (4)^2(0.3) + (5)^2(0.25) + (6)^2(0.1) \\ &= 16.9 \end{aligned}$$

k	$P(X = k)$	$P(Y = k)$
−10	0	0.2
−1	0.2	0
0	0.6	0.6
1	0.2	0
10	0	0.2

The Variance of a Discrete Random Variable. Another important quantity that characterizes the distribution of a random variable, the **variance** describes how spread out the range of the random variable is. To motivate the definition, let's look at the two random variables X and Y, with probability mass functions as shown on the left. We illustrate these two distributions in Figure 12.25. Both random variables have mean 0, but the range of Y is much more spread out than the range of X.

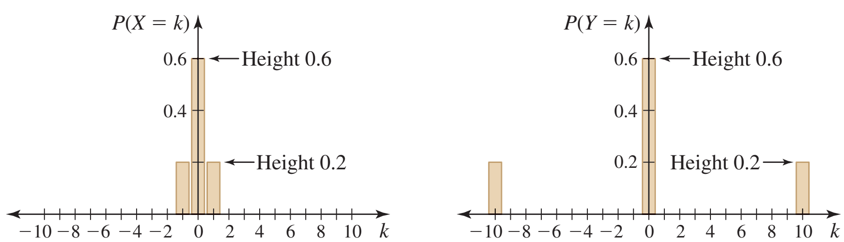


Figure 12.25 The probability mass functions of X and Y. The distribution of Y is more spread out than the distribution of X.

To capture this idea in a single quantity, we will compute the variance, which is defined as a weighted average of the squared distances to the mean:

> For any random variable X with mean μ, the **variance** of X is defined as
>
> $$\text{var}(X) = E[(X - \mu)^2]$$
>
> If X is a discrete random variable, then
>
> $$\text{var}(X) = \sum_x (x - \mu)^2 P(X = x)$$

Since the variance is an average value of a squared quantity, it is always nonnegative.

Let's return to the random variables X and Y. Their means are both equal to 0, so their variances are

$$\text{var}(X) = (-1 - 0)^2(0.2) + (0 - 0)^2(0.6) + (1 - 0)^2(0.2) = 0.4$$

$$\text{var}(Y) = (-10 - 0)^2(0.2) + (0 - 0)^2(0.6) + (10 - 0)^2(0.2) = 40$$

We see that the variance of Y is larger than the variance of X, reflecting the fact that the range of Y is more spread out than the range of X.

The variance of X is often denoted by σ^2 (read "sigma squared"). A quantity that is closely related to the variance is the **standard deviation**, denoted by s.d. or σ. The standard deviation is defined as the square root of the variance:

$$\text{s.d.} = \sigma = \sqrt{\operatorname{var}(X)}$$

The standard deviation has the advantage that it has the same units as the mean and so can be interpreted more easily than the variance.

EXAMPLE 9 Compute the variance and the standard deviation of the number of leaves per plant in Example 7.

Solution Denote by X the random variable that counts the number of leaves per plant. In Example 7, we found that $E(X) = 15.72$. Therefore, the variance of X is

$$\begin{aligned}\operatorname{var}(X) &= \underbrace{(13-15.72)^2}_{(x-\mu)^2}\underbrace{\frac{2}{25}}_{P(X=x)} + (14-15.72)^2\frac{1}{25} + (15-15.72)^2\frac{6}{25}\\ &\quad + (16-15.72)^2\frac{12}{25} + (17-15.72)^2\frac{1}{25} + (18-15.72)^2\frac{3}{25}\\ &= 1.5616\end{aligned}$$

We get $P(X = x)$ for $x = 13, 14, \ldots$ from Example 7.

$$\text{s.d.}(X) = \sqrt{\operatorname{var}(X)} = \sqrt{1.5616} \approx 1.2496$$

We will now collect some important rules regarding expected values and variances. The first rule tells us how to compute the expected value and the variance of a linear transformation of X. This rule holds for any random variable, not just discrete ones.

> Let a and b be constants. Then
>
> $$E(aX + b) = a[E(X)] + b$$
>
> $$\operatorname{var}(aX + b) = a^2 \operatorname{var}(X)$$

The first property says that the expected value of a linear function of X is the linear function evaluated at the expected value of X. The second property tells us what happens to the variance when we multiply a random variable by a constant factor; it is important to note that the constant factor is squared when we pull it out of the variance. Furthermore, we see that the variance is unchanged when we add a constant term to a random variable. The latter fact can be understood intuitively: Adding a constant term merely shifts the distribution without changing its shape. We will prove these two properties in Problems 25 and 26 for the case when X is a discrete random variable.

EXAMPLE 10 Suppose the average minimum temperature, measured in degrees Fahrenheit, in Minneapolis, Minnesota, in January is 2°F. Find the average minimum temperature in degrees Celsius.

Solution The linear transformation

$$C = \frac{5(F-32)}{9}$$

converts temperature measured in degrees Fahrenheit (F) into temperature measured in degrees Celsius (C). Hence,

$$E(C) = E\left[\frac{5(F-32)}{9}\right] = 5 \cdot \frac{E(F)-32}{9}$$
$$= 5 \cdot \frac{2-32}{9} = -\frac{150}{9} \approx -16.67$$

and we find that the average minimum temperature in January in Minneapolis is about $-16.67°$C. •

EXAMPLE 11 Find a formula that converts the variance of a temperature measured in degrees Celsius into the variance of the temperature measured in degrees Fahrenheit.

Solution We use the linear transformation of Example 10. This transformation relates a temperature measured in degrees Fahrenheit (F) to the temperature measured in degrees Celsius (C):

$$C = \frac{5(F-32)}{9}$$

Solving this equation for F, we obtain

$$F = \frac{9}{5}C + 32$$

Therefore,

$$\text{var}(F) = \text{var}\left(\frac{9}{5}C + 32\right) = \left(\frac{9}{5}\right)^2 \text{var}(C)$$
$$= \frac{81}{25}\text{var}(C) = 3.24 \cdot \text{var}(C)$$ •

It is often necessary to look at sums of random variables. We collect some rules without proof. Let X and Y be two random variables. Then $X + Y$ is also a random variable, and we have

$$\boxed{E(X+Y) = E(X) + E(Y)}$$

This formula holds for any random variables, not just discrete ones.

EXAMPLE 12 Suppose the average number of women who enter a coffee shop during lunch hour is 52.2 and the average number of men is 47.3. Find the average total number of people entering the coffee shop during lunch hour.

Solution If we denote the number of women by X and the number of men by Y, then we are interested in finding $E(X+Y)$. With $E(X) = 52.2$ and $E(Y) = 47.3$, we have

$$E(X+Y) = E(X) + E(Y) = 52.2 + 47.3 = 99.5$$ •

We can use our rules to find an alternative formula for the variance. We start with

$$(X-\mu)^2 = X^2 - 2X\mu + \mu^2$$

Taking expectations on both sides, we find that

$$E(X-\mu)^2 = E(X^2 - 2\mu X + \mu^2)$$

Since the expectation of a sum is the sum of the expectations, the right-hand side simplifies to

$$E(X^2) - E(2\mu X) + E(\mu^2) = E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - [E(X)]^2$$

because $\mu = E(X)$, $E(\mu^2) = \mu^2 = [E(X)]^2$, and $E(2\mu X) = 2\mu E(X) = 2[E(X)]^2$. With $E(X-\mu)^2 = \text{var}(X)$, we have

$$\boxed{\text{var}(X) = E(X^2) - [E(X)]^2}$$

This formula for the variance is often more convenient to use, since it leads to algebraically simpler expressions. Note that $E(X^2) \neq [E(X)]^2$, unless $\text{var}(X) = 0$, and that $E(X^2) \geq [E(X)]^2$, since $\text{var}(X) \geq 0$. In the next example, we apply this formula to the random variable X in Example 5.

EXAMPLE 13 Use the random variable X in Example 5, the result of Example 8, and the preceding formula to compute the variance of X.

Solution In Example 5, we found that $E(X) = 3.9$. In Example 8, we computed $E(X^2)$ and obtained

$$E(X^2) = 16.9$$

Hence,

$$\text{var}(X) = 16.9 - (3.9)^2 = 1.69$$

•

Joint Distributions. It is often important to investigate the relationship between random variables.

EXAMPLE 14 Gout is a type of arthritis in which uric acid is deposited in crystalline form within joints. A medical study might focus on whether the prevalence of the disease is dependent on gender. A survey in 1986 revealed that about 13.6 per 1000 men and 6.4 per 1000 women are affected. We can treat gender as one random variable and the presence of gout as another by defining

$$X = \begin{cases} 1 & \text{if male} \\ 0 & \text{if female} \end{cases}$$

and

$$Y = \begin{cases} 1 & \text{if gout is present} \\ 0 & \text{if gout is not present} \end{cases}$$

The following table lists the number of individuals in each of the four combinations of X and Y in a study of 10,000 men and 10,000 women:

	$X = 0$	$X = 1$	Total
$Y = 0$	9936	9864	19,800
$Y = 1$	64	136	200
Total	10,000	10,000	20,000

We see that the fraction of individuals in this study that are both male and affected by gout is $136/20,000 = 0.0068$. If we interpret this ratio as a probability, we could write

$$P(X = 1, Y = 1) = 0.0068$$

Converting all numbers in the table into relative frequencies and interpreting them as probabilities produces the **joint probability distribution** of X and Y:

	$X = 0$	$X = 1$	Total
$Y = 0$	0.4968	0.4932	0.99
$Y = 1$	0.0032	0.0068	0.01
Total	0.5	0.5	1.0

•

In general, when X and Y are discrete random variables, we define the joint probability distribution of X and Y by

$$p(x, y) = P(X = x, Y = y)$$

for all values of x in the range of X and all values of y in the range of Y. We can then obtain the distribution of X or of Y called the **marginal distribution** as follows:

$$p_X(x) = P(X = x) = \sum_y P(X = x, Y = y)$$

$$p_Y(y) = P(Y = y) = \sum_x P(X = x, Y = y)$$

EXAMPLE 15 Use the data from Example 14 to determine the probability that a randomly chosen person in the study described there has gout.

Solution We want to find the probability that $Y = 1$:

$$\begin{aligned} P(Y = 1) &= P(X = 0, Y = 1) + P(X = 1, Y = 1) \\ &= 0.0032 + 0.0068 = 0.01 \end{aligned}$$

•

We can define conditional probabilities as in Section 12.3, or

$$P(X = x \mid Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \tag{12.19}$$

provided that $P(Y = y) > 0$.

EXAMPLE 16 Use the data of Example 14 to determine the probability that a randomly chosen man in the study described there has gout.

Solution We want to find $P(Y = 1 \mid X = 1)$. Using (12.19), we find that

$$P(Y = 1 \mid X = 1) = \frac{P(X = 1, Y = 1)}{P(X = 1)} = \frac{0.0068}{0.5} = 0.0136$$

or 13.6 per 1000 men. •

Whether or not gender influences the prevalence of gout leads us to the concept of **independence** of random variables. The definition of independence of random variables follows from that of the independence of events. Events X and Y are **independent** if, for any two sets A and B,

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

When X and Y are discrete random variables, this equation simplifies to

$$P(X = x, Y = y) = P(X = x)P(Y = y) \tag{12.20}$$

for all values of x in the range of X and for all values of y in the range of Y.

EXAMPLE 17 Use the data of Example 14 to determine whether X and Y are independent.

Solution We check whether $P(X = 1, Y = 1)$ is equal to $P(X = 1)P(Y = 1)$. We find that $P(X = 1, Y = 1) = 0.0068$, $P(X = 1) = 0.5$, and $P(Y = 1) = 0.01$. Consequently, $P(X = 1)P(Y = 1) = (0.5)(0.01) = 0.005$, which is different from $P(X = 1, Y = 1)$. We conclude that X and Y are not independent; they are then said to be *dependent*. Note that to show that X and Y are *not* independent, it is enough to find one pair (x, y) for which (12.20) does not hold. •

EXAMPLE 18 Suppose that X and Y are two independent discrete random variables with probability mass functions as listed in the following table:

k	$P(X = k)$	$P(Y = k)$
-1	0.1	0.3
0	0.0	0.2
1	0.7	0.1
2	0.2	0.4

(a) Find the probability that X takes on the value -1 and Y takes on the value 2.

(b) Find the probability that X is negative and Y is positive.

Solution **(a)** We want to find $P(X = -1, Y = 2)$. Since X and Y are independent, we have

$$P(X = -1, Y = 2) = P(X = -1)P(Y = 2) = (0.1)(0.4) = 0.04$$

(b) We want to find the probability of the event that X is negative and Y is positive. This is the event $\{X = -1 \text{ and } Y = 1 \text{ or } 2\}$. Since X and Y are independent, we have

$$\begin{aligned} P(X = -1 \text{ and } Y = 1 \text{ or } 2) &= P(X = -1)P(Y = 1 \text{ or } 2) \\ &= (0.1)(0.1 + 0.4) = (0.1)(0.5) = 0.05 \end{aligned}$$

•

The definition of independence in (12.20) allows us to find the expected value of a product of independent discrete random variables. The following calculation shows the result when X and Y have finite ranges:

$$\begin{aligned} E(XY) &= \sum_{x,y} xyP(X = x, Y = y) = \sum_{x,y} xyP(X = x)P(Y = y) \\ &= \sum_{x} xP(X = x) \sum_{y} yP(Y = y) = E(X)E(Y) \end{aligned}$$

This relationship holds more generally:

> If X and Y are two independent random variables, then
> $$E(XY) = E(X)E(Y)$$

EXAMPLE 19 For the random variables X and Y in Example 18, find $E(XY)$.

Solution Since X and Y are independent, we have $E(XY) = E(X)E(Y)$. Now,

$$E(X) = (-1)(0.1) + (0)(0.0) + (1)(0.7) + (2)(0.2) = 1.0$$

and

$$E(Y) = (-1)(0.3) + (0)(0.2) + (1)(0.1) + (2)(0.4) = 0.6$$

Hence,

$$E(XY) = E(X)E(Y) = (1.0)(0.6) = 0.6$$

•

We can use the rule about the expected value of a product of independent random variables to compute the variance of the sum of two independent random variables. Suppose that X and Y are independent. Then

$$\begin{aligned} \operatorname{var}(X + Y) &= E(X + Y)^2 - [E(X + Y)]^2 \\ &= E(X^2 + 2XY + Y^2) - [E(X) + E(Y)]^2 \\ &= E(X^2) + 2E(XY) + E(Y^2) - [E(X)]^2 - 2E(X)E(Y) - [E(Y)]^2 \end{aligned}$$

But $E(XY) = E(X)E(Y)$, so the right-hand side simplifies to

$$E(X^2) - [E(X)]^2 + E(Y^2) - [E(Y)]^2$$

We recognize this quantity as $\text{var}(X) + \text{var}(Y)$. Hence,

If X and Y are independent random variables, then

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$$

Both formulas, $E(XY) = E(X)E(Y)$ and $\text{var}(X+Y) = \text{var}(X)+\text{var}(Y)$, hold for any independent random variables X and Y, not just for discrete ones. However, we will use these identities only in the context of discrete random variables. It is important to keep in mind that these two formulas only hold when X and Y are independent.

In the remaining subsections, we introduce four of the most important discrete distributions.

12.4.3 The Binomial Distribution

In this subsection, we will discuss a discrete random variable that models the number of successes among a fixed number of trials. A trial (also called a ***Bernoulli trial***) is a random experiment with two possible outcomes: success or failure. Suppose we perform a sequence of these trials. The trials are independent and the probability of success in each trial is p. We define the random variables X_k, $k = 1, 2, \dots, n$, as

$$X_k = \begin{cases} 1 & \text{if the } k\text{th trial is successful} \\ 0 & \text{otherwise} \end{cases}$$

Then $P(X_k = 1) = p$ and $P(X_k = 0) = 1 - p$ for $k = 1, 2, \dots, n$.

If we repeat these trials n times, we might want to know the total number of successes. We set

$$S_n = \text{number of successes in } n \text{ trials}$$

We can define S_n in terms of the random variables X_k as

$$S_n = \sum_{k=1}^{n} X_k \tag{12.21}$$

Since the trials are independent, Equation (12.21) shows that S_n can be written as a sum of independent random variables, all having the same distribution. We will use (12.21) subsequently.

The random variable S_n is discrete; its range is the set $\{0, 1, 2, \dots, n\}$. To find its probability mass function $p(k) = P(S_n = k)$, we argue as follows: All outcomes of the experiment can be represented as a string of zeros and ones, where 0 represents failure and 1 represents success. For instance, 01101 could be interpreted as the outcome of five trials, the first resulting in failure, followed by two successes, then a failure, and finally a success. The probability of this particular outcome is easy to compute, since the trials are independent. We obtain

$$P(01101) = (1 - p)pp(1 - p)p = p^3(1 - p)^2$$

The event $\{S_5 = 3\}$ consists of all outcomes that can be represented by a string of length 5 with exactly three ones. All of these strings have the same probability. To determine the number of different strings with exactly three ones, note that there are $\binom{5}{3}$ different ways of placing the three ones in the five possible positions and there is exactly one way to place the zeros in the remaining two positions. Hence, there are $\binom{5}{3} \cdot 1 = \binom{5}{3}$ different strings of length 5 with exactly three ones. Alternatively, notice, there are $5!$ ways of arranging the three ones and the two zeros if the zeros and the ones are distinguishable. The zeros and the ones can be rearranged among themselves without changing the outcome, so we must divide by the number of ways the zeros and ones can be reordered amongst themselves—that is, there are

$$\frac{5!}{3!2!} = \binom{5}{3} \qquad \text{Two 0s can be rearranged in 2! ways, three 1s in 3! ways}$$

different outcomes. As all outcomes are equally likely, we have

$$P(S_5 = 3) = \binom{5}{3} p^3 (1-p)^2 \qquad \text{Number of strings} \times \text{Prob. of each string}$$

We can use similar reasoning to derive the general formula, which we summarize as follows:

Binomial Distribution Let S_n be a random variable that counts the number of successes in n independent trials, each having probability p of success. Then S_n is said to be **binomially distributed** with parameters n and p, and

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n$$

The random variable S_n is called a **binomial** random variable, and its distribution is called the **binomial distribution**.

EXAMPLE 20 Toss a fair coin four times. Find the probability that there are exactly three heads.

Solution Let S_4 denote the number of heads. If heads denote success, then the probability of success is $p = 1/2$. S_4 is thus binomially distributed with parameters $n = 4$ and $p = 1/2$. Therefore,

$$P(S_4 = 3) = \binom{4}{3} \left(\frac{1}{2}\right)^3 \left(1 - \frac{1}{2}\right) = 4 \cdot \frac{1}{16} = \frac{1}{4}$$

EXAMPLE 21 In a shipment of 10 boxes, each box has probability 0.2 of being damaged. Find the probability of having two or more damaged boxes in the shipment.

Solution Let S_{10} denote the number of damaged boxes in the shipment. S_{10} is binomially distributed with parameters $n = 10$ and $p = 0.2$. The event of two or more damaged boxes can then be written as $S_{10} \geq 2$. To compute $P(S_{10} \geq 2)$, we use the formula

$$\begin{aligned} P(S_{10} \geq 2) &= 1 - P(S_{10} < 2) = 1 - [P(S_{10} = 0) + P(S_{10} = 1)] \\ &= 1 - \left[\binom{10}{0}(0.2)^0(0.8)^{10} + \binom{10}{1}(0.2)(0.8)^9\right] \\ &\approx 0.6242 \end{aligned}$$

EXAMPLE 22 **Down Syndrome** Down syndrome, or trisomy 21, is a genetic disorder in which three copies of chromosome 21 instead of two copies are present. In the United States, the prevalence is about 1 in 700 pregnancies. What is the probability that out of 100 pregnancies, at least one is affected?

Solution If S_{100} is the number of pregnancies and $p = 1/700$ is the probability of a pregnancy being affected, then S_{100} is binomially distributed with parameters $n = 100$ and $p = 1/700$. Thus,

$$\begin{aligned} P(S_{100} \geq 1) &= 1 - P(S_{100} = 0) \\ &= 1 - \left(1 - \frac{1}{700}\right)^{100} \approx 0.1332 \end{aligned}$$

If we use the representation $S_n = \sum_{k=1}^{n} X_k$ from (12.21) for the binomial random variable S_n, we can compute its mean and its variance. We find that

$$E(X_1) = (1)p + (0)(1-p) = p$$

and, with $E(X_1^2) = (1)^2 p + (0)^2(1-p) = p$, we have

$$\text{var}(X_1) = E(X_1^2) - [E(X_1)]^2 = p - p^2 = p(1-p)$$

Since all X_k, $k = 1, 2, \ldots, n$, have the same distribution, it follows that

$$E(S_n) = E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} E(X_k) = np \tag{12.22}$$

In addition, because the X_k are independent,

$$\text{var}(S_n) = \text{var}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \text{var}(X_k) = np(1-p) \tag{12.23}$$

We summarize these results as follows:

> If S_n is binomially distributed with parameters n and p, then
>
> $$E(S_n) = np \qquad \text{and} \qquad \text{var}(S_n) = np(1-p)$$

We present two more applications of the binomial distribution.

EXAMPLE 23 **Genetics** We consider the flowering pea plants again. Suppose that 20 independent offspring result from $Cc \times Cc$ crossings. Find the probability that at most two offspring have white flowers, and compute the expected value and the variance of the number of offspring that have white flowers.

Solution In a $Cc \times Cc$ crossing, the probability of a white-flowering offspring (genotype cc) is 1/4 and the probability of a red-flowering offspring (genotype CC or Cc) is 3/4. The flower colors of different offspring are independent. We can therefore think of this experiment as one with 20 trials, each having a probability of success of 1/4. We want to know the probability of at most two successes. With $n = 20$, $k \leq 2$, and $p = 1/4$, we find that

$$\begin{aligned} P(S_{20} \leq 2) &= P(S_{20} = 0) + P(S_{20} = 1) + P(S_{20} = 2) \\ &= \binom{20}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^{20} + \binom{20}{1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^{19} + \binom{20}{2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^{18} \\ &\approx 0.0913 \end{aligned}$$

The expected value of the number of white-flowering offspring is

$$E(S_{20}) = (20)\left(\frac{1}{4}\right) = 5$$

and the variance is

$$\text{var}(S_{20}) = (20)\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) = \frac{15}{4} = 3.75$$

EXAMPLE 24 **Hemophilia** Suppose that a woman who is a carrier for hemophilia has four daughters with a man who is not hemophilic. Find the probability that at least one daughter carries the hemophilia gene.

Solution Each daughter has probability 1/2 of carrying the disease gene, independently of all others. We can think of this experiment as one with four independent trials and a probability of success of 1/2. ("Success" in this case is being a carrier.) Therefore,

$$\begin{aligned} P(S_4 \geq 1) &= 1 - P(S_4 = 0) \\ &= 1 - \binom{4}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^4 \\ &= 1 - \frac{1}{16} = \frac{15}{16} = 0.9375 \end{aligned}$$

Sampling With and Without Replacement. Consider an urn with 10 green and 15 blue balls. Sampling balls from this urn can be done with or without replacement. If we sample with replacement, we take out a ball, note its color, and then place the ball back into the urn. The number of balls of a specific color is then binomially distributed.

EXAMPLE 25 If we sample five balls with replacement from the aforementioned urn, what is the probability of three blue balls in the sample?

Solution Denote the number of blue balls in the sample by S_5. Then S_5 is binomially distributed, with n, the number of trials, equal to 5 and probability of success $p = 15/(15+10) = 3/5$. We find that

$$P(S_5 = 3) = \binom{5}{3}\left(\frac{3}{5}\right)^3\left(\frac{2}{5}\right)^2 = 0.3456$$

If we sample without replacement, we take the balls out one after the other without putting them back into the urn and note their colors. If we take the balls out one after the other without replacing them, then the composition of the urn changes every time we remove a ball and the number of balls of a certain color is no longer binomially distributed.

EXAMPLE 26 If we sample five balls without replacement from the urn, what is the probability of three blue balls in the sample?

Solution (We encountered a similar problem in Example 9 of Section 12.2.) There are $\binom{25}{5}$ ways of sampling 5 balls from this urn, so $\binom{25}{5}$ is the size of the sample space. Each outcome in this sample space has the same probability. In order to have 3 blue balls, we need to select 3 out of the 15 blue balls, which can be done in $\binom{15}{3}$ ways. Since we want a total of 5 balls, we also need to select 2 out of the 10 green balls, which can be done in $\binom{10}{2}$ ways. Combining the blue and the green balls, we find that there are $\binom{15}{3}\binom{10}{2}$ ways of selecting 3 blue and 2 green balls from the urn. Therefore, the probability of obtaining 3 blue balls in a sample of size 5 when sampling is done without replacement is

$$\frac{\binom{15}{3}\binom{10}{2}}{\binom{25}{5}} = \frac{455 \cdot 45}{53130} \approx 0.3854$$

Note that the answer is different from that in Example 25.

The probability distribution in Example 26 is called the **hypergeometric distribution**. The hypergeometric distribution describes sampling without replacement if two types of objects are in the urn. Suppose the urn has M green and N blue balls, and a sample of size n is taken from the urn without replacement. If X denotes the number of blue balls in the sample, then

$$P(X = k) = \frac{\binom{N}{k}\binom{M}{n-k}}{\binom{M+N}{n}}, \quad k = 0, 1, 2, \ldots, n$$

12.4.4 The Multinomial Distribution

In the previous subsection, we considered experiments in which each trial resulted in exactly one of two possible outcomes. We will now extend this situation to more than two possible outcomes. The distribution is then called the **multinomial distribution**.

EXAMPLE 27 To study food preferences in the lady beetle *Coleomegilla maculata*, we present each beetle with three different food choices: maize pollen, egg masses of the European corn borer, and aphids. We suspect that 20% of the time the beetle prefers the aphids, 35% of the time egg masses, and 45% of the time pollen. We carry out this experiment with 30 beetles and find that 8 beetles prefer aphids, 10 egg masses, and 12 pollen. Compute the probability of this event, assuming that the trials are independent.

Solution We define the random variables

$$N_1 = \text{number of beetles that prefer aphids}$$

$$N_2 = \text{number of beetles that prefer egg masses}$$

$$N_3 = \text{number of beetles that prefer pollen}$$

and the probabilities

$$p_1 = P(\text{beetle prefers aphids}) = 0.2$$

$$p_2 = P(\text{beetle prefers egg masses}) = 0.35$$

$$p_3 = P(\text{beetle prefers pollen}) = 0.45$$

We claim that

$$P(N_1 = 8, N_2 = 10, N_3 = 12) = \frac{30!}{8!\,10!\,12!}(0.2)^8(0.35)^{10}(0.45)^{12}$$

The term $\frac{30!}{8!\,10!\,12!}$ counts the number of ways we can arrange 30 objects—8 of one type, 10 of another, and 12 of a third—that represent the beetles preferring aphids, egg masses, and pollen, respectively. The term $(0.2)^8(0.35)^{10}(0.45)^{12}$ comes from the probability of a particular arrangement, just as in the binomial case. •

A more involved example for the multinomial distribution is another of Mendel's experiments, one in which he crossed pea plants that had round, yellow seeds with plants that had green, wrinkled seeds. The round shape and yellow color are dominant traits, and wrinkled shape and green color are recessive traits. We denote the allele for round seeds by R, the allele for wrinkled seeds by r, the allele for yellow seeds by Y, and the allele for green seeds by y. Then a crossing between plants that are homozygous for round, yellow seeds (genotype RR/YY) and plants that are homozygous for wrinkled, green seeds (genotype rr/yy) is written as

$$RR/YY \times rr/yy$$

This crossing results in offspring of type Rr/Yy. That is, all offspring are heterozygous with round, yellow seeds. Crossing plants from this offspring generation then results in all possible combinations, as illustrated in the following table:

		Alleles from Parent 2			
		RY	***Ry***	***rY***	***ry***
Alleles from Parent 1	***RY***	RR/YY round, yellow	RR/Yy round, yellow	Rr/YY round, yellow	Rr/Yy round, yellow
	Ry	RR/Yy round, yellow	RR/yy round, green	Rr/yY round, yellow	Rr/yy round, green
	rY	rR/YY round, yellow	rR/Yy round, yellow	rr/YY wrinkled, yellow	rr/Yy wrinkled, yellow
	ry	rR/yY round, yellow	rR/yy round, green	rr/yY wrinkled, yellow	rr/yy wrinkled, green

	Yellow	Green
Round	9/16	3/16
Wrinkled	3/16	1/16

The laws of inheritance tell that each outcome of this crossing is equally likely. There are 16 different genotypes, some of which give rise to the same morphological type of seed (e.g., every genotype on the top row produces round yellow seeds). The appearance of the seed is called the **phenotype**. We summarize the phenotypes and their probability distribution in the table on the left.

EXAMPLE 28 Suppose that you obtain 50 independent offspring from the crossing

$$Rr/Yy \times Rr/Yy$$

where 25 seeds are round and yellow, 9 are round and green, 12 are wrinkled and yellow, and 4 are wrinkled and green. Find the probability of this outcome.

Solution This is another application of the multinomial distribution. Arguing as in the previous example, we find that the probability of this outcome is

$$\frac{50!}{25!\,9!\,12!\,4!}\left(\frac{9}{16}\right)^{25}\left(\frac{3}{16}\right)^{9}\left(\frac{3}{16}\right)^{12}\left(\frac{1}{16}\right)^{4}$$

•

12.4.5 Geometric Distribution

We again consider a sequence of independent Bernoulli trials, namely, a random experiment of repeated trials where each trial has two possible outcomes—success or failure—and the trials are independent. As in Subsection 12.4.3, we denote the probability of success by p. This time, however, we define a random variable X that counts the number of trials until the first success. The random variable X takes on values $1, 2, 3, \ldots$ and is therefore a discrete random variable. Its probability distribution is called the **geometric distribution** and is given by

$$P(X = k) = (1 - p)^{k-1}p, \quad k = 1, 2, 3, \ldots \tag{12.24}$$

since the event $\{X = k\}$ means that the first $k - 1$ trials resulted in failure (each one has probability $1 - p$ and the trials are independent) and were followed by a trial that resulted in a success (with probability p and independently of all other trials).

The range of X is the set of all positive integers; thus, X takes on infinitely (though still countably) many values. This is the first time we encounter a random variable of this kind. There are some issues we need to discuss that pertain to a countably infinite range. For instance, to show that $P(X = k)$ in (12.24) is indeed a probability mass function, we will need to sum up the probabilities from $k = 1$ to $k = \infty$. This means summing up an infinite number of terms, and we will need to explain what that means.

To show that (12.24) is a probability mass function, we need to check that $P(X = k) \geq 0$ and that it sums to 1. The first part is straightforward: Since p is a probability, it follows that $0 \leq p \leq 1$, making $(1 - p)^{k-1}p \geq 0$ for all $k = 1, 2, 3, \ldots$. For the second part, we need to show that

$$\sum_k P(X = k) = 1$$

where the sum ranges over all values of k in the range of X, namely, $k = 1, 2, 3, \ldots$. Mathematically, we write this as

$$\sum_{k=1}^{\infty} P(X = k)$$

and define this **infinite sum** as

$$\sum_{k=1}^{\infty} P(X = k) = \lim_{n\to\infty} \sum_{k=1}^{n} P(X = k)$$

To compute the infinite sum (and later the mean and the variance), we introduce the **geometric series**: the infinite sum

$$\sum_{k=0}^{\infty} q^k = 1 + q + q^2 + q^3 + \cdots$$

The finite sum

$$S_n = \sum_{k=0}^{n} q^k = 1 + q + q^2 + \cdots + q^n$$

can be computed with the use of the following computational "trick": Write

$$S_n = 1 + q + q^2 + \cdots + q^{n-1} + q^n$$

q, q^2, q^n appear in both sums

$$qS_n = q + q^2 + q^3 + \cdots + q^n + q^{n+1}$$

and then subtract qS_n from S_n. Most terms cancel, and we find that

$$S_n - qS_n = 1 - q^{n+1}$$

Factoring out S_n on the left-hand side and solving for S_n yields

$$(1-q)S_n = 1 - q^{n+1}$$

$$S_n = \frac{1-q^{n+1}}{1-q}$$

provided that $q \neq 1$.

If $|q| < 1$, then $\lim_{n\to\infty} q^{n+1} = 0$, and therefore,

$$\sum_{k=0}^{\infty} q^k = \lim_{n\to\infty} \sum_{k=0}^{\infty} q^k = \lim_{n\to\infty} \frac{1-q^{n+1}}{1-q} = \frac{1}{1-q} \quad \text{for } |q| < 1$$

These are important results, which we summarize as follows:

For $q \neq 1$,

$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q} \tag{12.25}$$

For $|q| < 1$,

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q} \tag{12.26}$$

We can use the preceding results to check that the probability mass function of the geometric distribution adds up to 1:

$$\sum_{k=1}^{\infty} P(X=k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p$$

$$= p \sum_{l=0}^{\infty} (1-p)^l = p \cdot \frac{1}{1-(1-p)} = p \cdot \frac{1}{p} = 1$$

In the summation, we made the substitution $l = k - 1$. Then the summation range $k = 1, 2, 3, \ldots$ changes to $l = 0, 1, 2, \ldots$, allowing us to apply the results for the geometric series we derived in (12.26).

EXAMPLE 29 A random experiment consists of rolling a fair die until the first time a six appears. Find the probability that the first six appears at the fifth trial.

Solution Denote by X the first time a six appears and by p the probability that the die shows a six in a single trial (the success in this experiment). Since the die is fair, all six numbers on the die are equally likely and we find that $p = 1/6$. Then

$$P(X=5) = \left(1 - \frac{1}{6}\right)^4 \frac{1}{6} \approx 0.0804$$

EXAMPLE 30 Consider a sequence of independent Bernoulli trials with probability of success p. Find the probability of no success in the first k trials.

Solution Denote by X the number of trials until the first success. We want to find the event $\{X > k\}$. Now, this event can be phrased in terms of a binomial random variable S_k that counts the number of successes in the first k trials. The event $\{X > k\}$ is equivalent to the event $\{S_k = 0\}$. Therefore,

$$P(X > k) = P(S_k = 0) = (1 - p)^k$$

•

EXAMPLE 31 Compare the probability of no success in the first k trials of independent Bernoulli trials with the probability of no success in k trials following n unsuccessful trials.

Solution If X denotes the number of trials with probability of success p, then we want to compare $P(X > k)$ with $P(X > n + k | X > n)$. From Example 30, we conclude that

$$P(X > k) = (1 - p)^k$$

To compute the conditional probability $P(X > n + k | X > n)$, we use

$$P(X > n + k | X > n) = \frac{P(X > n + k, X > n)}{P(X > n)}$$

Since the event $\{X > n + k\}$ is contained in the event $\{X > n\}$, it follows that

$$P(X > n + k, X > n) = P(\{X > n + k\} \cap \{X > n\}) = P(X > n + k)$$

and

$$\frac{P(X > n + k, X > n)}{P(X > n)} = \frac{P(X > n + k)}{P(X > n)} = \frac{(1 - p)^{n+k}}{(1 - p)^n} = (1 - p)^k$$

We then find that

$$P(X > k) = P(X > n + k | X > n)$$

That is, not having had a success in the first n trials does not change the probability of not having any successes in the following k trials. This result is a consequence of the independence of trials. •

EXAMPLE 32 **Genetic Disease** If both parents are carriers of a recessive autosomal disease, but are not symptomatic for the disease, then there is a 25% chance that a child of theirs will be symptomatic for the disease. Suppose the parents have three asymptomatic children and plan on having a fourth child. What is the probability that the fourth child will not be symptomatic for the disease?

Solution Denote by X the waiting time for the first symptomatic child in this family. With probability of "success" $p = 1/4$, we find, from Example 31, that

$$P(X > 4 | X > 3) = P(X > 1) = 1 - P(X = 1) = 1 - \frac{1}{4} = \frac{3}{4}$$

We can also argue as follows: The fact that the first three children are asymptomatic for the disease does not change the probability that their next child will be asymptomatic for the disease, since these events are independent. •

We will now compute the mean and the variance of the geometric distribution. If X is a geometrically distributed random variable with $P(X = k) = (1 - p)^{k-1}p$, then

$$E(X) = \sum_{k=1}^{\infty} k(1 - p)^{k-1}p = p\sum_{k=1}^{\infty} k(1 - p)^{k-1}$$

To compute this infinite sum, we need (12.26) and the following result, which we cannot prove here: If $|q| < 1$, then

$$\frac{d}{dq}\sum_{k=0}^{\infty} q^k = \sum_{k=0}^{\infty} \frac{d}{dq}(q^k)$$

In words, the derivative of this infinite sum can be obtained by differentiating each term separately and then taking the sum of all the derivatives. Interchanging differentiation and summation when the sum is an infinite sum cannot always be done, but can be justified in this case. Using this result, we find that

$$\sum_{k=0}^{\infty} \frac{d}{dq} q^k = \sum_{k=0}^{\infty} k q^{k-1} = \sum_{k=1}^{\infty} k q^{k-1}$$

where, in the last step, we used the fact that the term with $k = 0$ is equal to 0. Since $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$ when $|q| < 1$, it also follows that

$$\frac{d}{dq} \sum_{k=0}^{\infty} q^k = \frac{d}{dq}\left(\frac{1}{1-q}\right) = \frac{1}{(1-q)^2}$$

and hence,

$$\sum_{k=1}^{\infty} k q^{k-1} = \frac{1}{(1-q)^2}$$

With $q = 1 - p$,

$$E(X) = p \sum_{k=1}^{\infty} k(1-p)^{k-1} = p \frac{1}{(1-(1-p))^2} = \frac{p}{p^2} = \frac{1}{p}$$

To compute the variance of X, we employ a similar argument. We first derive the following result, which we will need to carry out the calculation of the variance: If $|q| < 1$, then

$$\frac{d^2}{dq^2} \sum_{k=0}^{\infty} q^k = \sum_{k=0}^{\infty} \frac{d^2}{dq^2} q^k = \sum_{k=2}^{\infty} k(k-1) q^{k-2}$$

Now, since

$$\frac{d^2}{dq^2}\left(\frac{1}{1-q}\right) = \frac{2}{(1-q)^3}$$

we have

$$\sum_{k=2}^{\infty} k(k-1) q^{k-2} = \frac{2}{(1-q)^3}$$

To compute the variance, it is useful to first compute

$$\begin{aligned} E[X(X-1)] &= \sum_{k=2}^{\infty} k(k-1) P(X = k) = \sum_{k=2}^{\infty} k(k-1)(1-p)^{k-1} p \\ &= (1-p)p \sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2} = (1-p)p \cdot \frac{2}{(1-(1-p))^3} \\ &= \frac{2p(1-p)}{p^3} = \frac{2(1-p)}{p^2} \end{aligned}$$

Since $E[X(X-1)] = E(X^2) - E(X)$, it follows that

$$\begin{aligned} \text{var}(X) &= E(X^2) - [E(X)]^2 = E[X(X-1)] + E(X) - [E(X)]^2 \\ &= \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2} \end{aligned}$$

We summarize our results as follows:

> If X is geometrically distributed with $P(X = k) = (1-p)^{k-1} p$, then
>
> $$E(X) = \frac{1}{p} \qquad \text{and} \qquad \text{var}(X) = \frac{1-p}{p^2}$$

EXAMPLE 33 You roll a fair die until the first time a 6 appears. How long do you have to wait, on average?

Solution To answer this question, we need to find $E(X)$, where X is geometrically distributed with probability of success $p = 1/6$. Therefore,

$$E(X) = \frac{1}{p} = 6$$

In words, you have to roll a die six times, on average, until the first time a 6 appears.

12.4.6 The Poisson Distribution

The Poisson distribution is one of the most important probability distributions. It is used to model, for instance, amino acid substitutions in proteins, the escape probability of hosts from parasitism, and the distributions of plants. It often models "rare events," as we will see.

We say that X is **Poisson** distributed with parameter $\lambda > 0$ if

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

Just as with the geometric distribution, the range of the random variable X is the set of all nonnegative integers. Thus, the range is infinite but countable.

To show that the probability distribution we defined sums to 1, or to find the mean and the variance of X, we need some additional results. Recall that the Taylor polynomial of order n of $f(x) = e^x$ is

$$P_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}$$ See Section 7.6

It turns out (but we cannot prove this here) that, in the limit as $n \to \infty$,

$$e^x = \lim_{n\to\infty} P_n(x) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!}$$

for all $x \in \mathbf{R}$. On the right-hand side, we have an infinite sum, and we will use the notation (as we did in the previous section)

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!}$$

to denote this infinite sum. Thus, for any $x \in \mathbf{R}$,

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{12.27}$$

We can use (12.27) to show that the probability mass function for the Poisson distribution indeed sums to 1:

$$\sum_{k=0}^{\infty} P(X = k) = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1$$

Furthermore, since $\lambda > 0$, $P(X = k) \geq 0$. The probability mass function for the Poisson distribution therefore satisfies the two necessary conditions for a probability mass function.

EXAMPLE 34 Suppose the number of plants per hectare of a certain species is Poisson distributed with parameter $\lambda = 3$ plants per hectare. Find the probability that there are **(a)** no plants in a given hectare and **(b)** at least two plants in a given hectare.

Solution Denote by X the number of plants in a given hectare. Then X is Poisson distributed with parameter $\lambda = 3$.

(a) The probability that there are no plants in a given hectare is

$$P(X = 0) = e^{-\lambda} = e^{-3} \approx 0.0498$$

(b) The probability that there are at least two plants in a given hectare is

$$\begin{aligned} P(X \geq 2) &= 1 - P(X \leq 1) = 1 - [P(X = 0) + P(X = 1)] \\ &= 1 - e^{-\lambda}(1 + \lambda) = 1 - e^{-3}(1 + 3) \approx 0.8009 \end{aligned}$$

•

To find the mean and the variance of a Poisson-distributed random variable, we need to use (12.27) repeatedly. Let X be Poisson distributed with parameter λ. Then, formally,

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} kP(X = k) = \sum_{k=0}^{\infty} ke^{-\lambda}\frac{\lambda^k}{k!} \\ &= e^{-\lambda}\lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda}\lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda e^{-\lambda}e^{\lambda} = \lambda \end{aligned}$$

To compute the variance of X, we begin by computing

$$\begin{aligned} E[X(X-1)] &= \sum_{k=0}^{\infty} k(k-1)P(X = k) = \sum_{k=2}^{\infty} k(k-1)P(X = k) \\ &= \sum_{k=2}^{\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} \\ &= e^{-\lambda}\lambda^2 \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda}\lambda^2 e^{\lambda} = \lambda^2 \end{aligned}$$

Therefore,

$$\begin{aligned} \text{var}(X) &= E(X^2) - [E(X)]^2 = E[X(X-1)] + E(X) - [E(X)]^2 \\ &= \lambda^2 + \lambda - \lambda^2 = \lambda \end{aligned}$$

We summarize our findings as follows:

> If X is Poisson distributed with parameter $\lambda > 0$, then
>
> $$E(X) = \lambda \qquad \text{and} \qquad \text{var}(X) = \lambda$$

EXAMPLE 35 **Amino Acid Evolution** The number of substitutions on a given amino acid sequence during a fixed period is modeled by a Poisson distribution. Suppose the number of substitutions on a sequence of 100 amino acids over a period of 1 million years is Poisson distributed with average number of substitutions equal to 1. What is the probability that at least one substitution occurred?

Solution If X denotes the number of substitutions, then X is Poisson distributed with mean 1. Since the mean of a Poisson distribution is equal to its parameter, we find that $\lambda = 1$. Hence,

$$P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-\lambda} = 1 - e^{-1} \approx 0.6321$$

•

We mentioned previously that the Poisson distribution frequently models rare events. The next result makes this precise. Consider a sequence of independent Bernoulli trials with probability of success p. The number of successes among n trials is binomially distributed. We denote the number of successes in n trials by S_n. We consider the case when the number of trials n is very large but the probability of success p is very small, so successes are rare. To make this concept mathematically precise, we will need to take the limit as n tends to infinity such that the product np, which denotes the expected number of successes among n trials, approaches a constant. To do so, we need to let p tend to 0 as n tends to infinity. To indicate that the probability of success depends on n, we will denote it by p_n. The following result says that the number of successes among a large number of trials is approximately Poisson distributed if the probability of success is small:

Poisson Approximation to the Binomial Distribution Suppose S_n is binomially distributed with parameters n and p_n. If $p_n \to 0$ as $n \to \infty$ such that $\lim_{n\to\infty} np_n = \lambda > 0$, then

$$\lim_{n\to\infty} P(S_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

In all the examples of the binomial distribution we have seen thus far, the probability of success p was fixed and did not depend on n. How, then, should we interpret the preceding result? It allows us to use a Poisson distribution as an approximation to the binomial distribution when the number of trials n is large and the probability of success p is small. Let's use the result to compare the Poisson approximation to the exact results of a binomial distribution. This will illustrate how to use the Poisson approximation result. We will prove the result afterward.

EXAMPLE 36 Suppose we toss a biased coin 100 times and denote the number of heads by S_{100}. If the probability of heads is $1/50$, compute $P(S_{100} = k)$ for $k = 0, 1$, and 2 exactly and compare your answer with the Poisson approximation.

Solution Since $n = 100$ and $p = 1/50$, we compare the distribution of S_{100} with a Poisson distribution with parameter$\lambda = np = 100/50 = 2$. We find that

$$P(S_{100} = k) = \binom{100}{k}\left(\frac{1}{50}\right)^k\left(\frac{49}{50}\right)^{100-k} \approx e^{-2}\frac{2^k}{k!} \qquad \text{By Poisson approximation}$$

For $k = 0$,

$$P(S_{100} = 0) = \left(\frac{49}{50}\right)^{100} \approx 0.1326$$

$$e^{-2} \approx 0.1353$$

For $k = 1$,

$$P(S_{100} = 1) = 100 \cdot \frac{1}{50}\left(\frac{49}{50}\right)^{99} \approx 0.2707$$

$$2e^{-2} \approx 0.2707$$

For $k = 2$,

$$P(S_{100} = 2) = \frac{100 \cdot 99}{2}\left(\frac{1}{50}\right)^2\left(\frac{49}{50}\right)^{98} \approx 0.2734$$

$$e^{-2}\frac{2^2}{2} \approx 0.2707$$

In each case, we see that the approximate value is close to the exact value. The advantage is that the Poisson distribution is much easier to calculate than the binomial distribution, since the binomial coefficients $\binom{n}{k}$ are computationally intensive. •

We will now prove the Poisson approximation:

$$\lim_{n\to\infty} P(S_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

We will need to use the following result: If $\lim_{n\to\infty} x_n = x$, then

$$\lim_{n\to\infty} \left(1 + \frac{x_n}{n}\right)^n = e^x$$

To prove this result, we show that

$$\lim_{n\to\infty} \ln\left(1 + \frac{x_n}{n}\right)^n = x$$

Now,

$$\lim_{n\to\infty} \ln\left(1 + \frac{x_n}{n}\right)^n = \lim_{n\to\infty} n\ln\left(1 + \frac{x_n}{n}\right)$$

This limit is of the form $\infty \cdot 0$, which suggests that we should rewrite the limit in the form $\frac{0}{0}$ or $\frac{\infty}{\infty}$ and use l'Hôpital's rule. We rewrite it in the form

$$\lim_{n\to\infty} n\ln\left(1 + \frac{x_n}{n}\right) = \lim_{n\to\infty} \frac{\ln\left(1 + \frac{x_n}{n}\right)}{\frac{1}{n}} = \lim_{n\to\infty} \frac{\ln\left(1 + \frac{x_n}{n}\right)}{\frac{x_n}{n}} x_n$$

To evaluate this limit, we use l'Hôpital's rule to compute

$$\lim_{y\to 0} \frac{\ln(1+y)}{y}$$

We find that

$$\lim_{y\to 0} \frac{\ln(1+y)}{y} = \lim_{y\to 0} \frac{\frac{1}{1+y}}{1} = 1$$

This result, together with $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} \frac{x_n}{n} = 0$, yields

$$\lim_{n\to\infty} \frac{\ln\left(1 + \frac{x_n}{n}\right)}{\frac{x_n}{n}} x_n = (1)(x) = x$$

We can now prove the Poisson approximation. Observe that

$$P(S_n = k) = \binom{n}{k} p_n^k (1 - p_n)^{n-k}$$

We define $\lambda_n = np_n$, so $p_n = \frac{\lambda_n}{n}$. Then

$$\begin{aligned} P(S_n = k) &= \binom{n}{k}\left(\frac{\lambda_n}{n}\right)^k \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \\ &= \frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k}\frac{\lambda_n^k}{k!}\left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-k} \end{aligned}$$

Since $\lim_{n\to\infty} np_n = \lim_{n\to\infty} \lambda_n = \lambda$,

$$\lim_{n\to\infty} \frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k} = \lim_{n\to\infty} \frac{n}{n}\frac{n-1}{n}\frac{n-2}{n}\cdots\frac{n-k+1}{n} = 1,$$

$$\lim_{n\to\infty} \frac{\lambda_n^k}{k!} = \frac{\lambda^k}{k!}, \qquad \lim_{n\to\infty} \left(1 - \frac{\lambda_n}{n}\right)^{-k} = 1$$

and

$$\lim_{n\to\infty} \left(1 - \frac{\lambda_n}{n}\right)^n = \lim_{n\to\infty} \left(1 + \frac{-\lambda_n}{n}\right)^n = e^{-\lambda},$$

it follows that

$$\lim_{n\to\infty} P(S_n = k) = (1)\frac{\lambda^k}{k!}e^{-\lambda}(1) = e^{-\lambda}\frac{\lambda^k}{k!}$$

which is the Poisson approximation to the binomial distribution.

We introduced parasitoids in Section 10.7. Parasitoids are insects that lay their eggs on, in, or near the (in most cases, immature) body of another arthropod, which serves as the host for the developing parasitoid. The hatching offspring then devour the host. Parasitoids make up about 14% of all insect species. A key component in modeling host–parasitoid interactions is the probability of a host escaping parasitism. The first models were developed by ***Nicholson and Bailey (1935)***, who assumed that the probability that a host escapes parasitism is given by e^{-aP}, where P is the parasitoid density and a is a positive constant called the **search efficiency**. The next example explains where this model for the escape probability comes from.

EXAMPLE 37 **Host–Parasitoid Models** Parasitoid encounters with their hosts are sometimes modeled with a Poisson distribution. Suppose a host is surrounded by P parasitoids. Each parasitoid, independently of all others, has a probability a of encountering the host. We can consider this probability as a sequence of P Bernoulli trials with probability of success a. If no parasitoid encounters the host, the host will escape parasitism. If P is large and a is small, we can use the Poisson approximation. The number of encounters is then approximately Poisson distributed with parameter aP, and we find that

$$P(0 \text{ encounters with parasitoids}) = e^{-aP}$$

This is the probability the host escapes parasitism used in the Nicholson–Bailey host–parasitoid model (see Section 10.7). It is the zeroth term of a Poisson distribution and comes about because parasitoids are assumed to search randomly. •

The Poisson approximation plays a crucial role in using amino acid sequence data to estimate the time of divergence of species. Sequences of the same protein across different species are compared, and the number of amino acid differences between the two proteins gives an indication of the evolutionary distance between each pair of species. The simplest mathematical model for estimating times of divergence based on amino acid sequences assumes that the probability of a substitution at a given site in the amino acid sequence is the same for all sites and depends only on the time since divergence. Furthermore, all sites are assumed to be independent. The number of amino acid substitutions along a sequence of length n is then binomially distributed with probability of success equal to the probability of a substitution at the given site, provided that multiple substitutions at the site can be ignored and the time since divergence is not too long. If the sequence is sufficiently long and the time since divergence is not too long, so that the probability of substitution is small, then the number of substitutions is equal to the number of differences between the two sequences and can be approximated by a Poisson distribution.

EXAMPLE 38 **Amino Acid Evolution** Suppose you compare the hemoglobin-α chain (length 140 amino acids) of two vertebrate species that diverged about 10 million years ago. A previous study found that, for any amino acid along this chain, the probability of an amino acid difference is about 0.014.

(a) How many amino acid differences would you expect when comparing the two sequences?

(b) What is the probability of finding at least three sites with amino acid differences?

Solution The number of amino acid differences is approximately Poisson distributed with parameters equal to the product of the length of the sequence and the probability of finding a difference at a given site. We find that $\lambda = 140 \cdot 0.014 = 1.96$.

(a) The expected number of amino acid differences is equal to the parameter of the distribution, 1.96 in this case.

(b) The probability of finding at least three differences is equal to

$$1 - e^{-\lambda}(1 + \lambda + \lambda^2/2) = 1 - e^{-1.96}(1 + 1.96 + 1.96^2/2) \approx 0.3125$$ •

Sums of Poisson Random Variables. Suppose X is a Poisson random variable with parameter λ and Y is a Poisson random variable with parameter μ. Then, if X and Y are independent, it follows that $X + Y$ is Poisson distributed with parameter $\lambda + \mu$. To prove this result, we calculate $P(X + Y = n)$ for $n = 0, 1, 2, \ldots$.

First, note that the event $\{X + Y = n\}$ can be decomposed into mutually exclusive events of the form $\{X = k, Y = n - k\}$ for $k = 0, 1, 2, \ldots, n$, so that

$$P(X + Y = n) = \sum_{k=0}^{n} P(X = k, Y = n - k)$$

Next, because of independence,

$$P(X = k, Y = n - k) = P(X = k)P(Y = n - k)$$

Therefore,

$$\begin{aligned} P(X + Y = n) &= \sum_{k=0}^{n} P(X = k)P(Y = n - k) \\ &= \sum_{k=0}^{n} e^{-\lambda}\frac{\lambda^k}{k!} \cdot e^{-\mu}\frac{\mu^{n-k}}{(n-k)!} \\ &= e^{-(\lambda+\mu)}\frac{1}{n!}\sum_{k=0}^{n}\frac{n!}{k!\,(n-k)!}\lambda^k\mu^{n-k} \end{aligned}$$

Recall from Example 7 in Section 12.1 that

$$(x + y)^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k} \qquad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Hence,

$$P(X + Y = n) = e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^n}{n!} \qquad k = 0 \text{ Poisson distribution with parameter } \lambda + \mu$$

EXAMPLE 39 **Cancer Deaths** The number of annual cancer deaths in a population is sometimes modeled with a Poisson distribution. In Ireland in the mid-1990s, the number of deaths from lung cancer was about 71 per 100,000 men and 34 per 100,000 women. Use these data to find the probability that no lung cancer deaths occurred during a given year in a group of 1000 men and 1000 women.

Solution We model the number of annual deaths with a Poisson distribution. We denote the number of deaths in men by X and use $\lambda = 0.71$ for the parameter of the Poisson distribution. We denote the number of deaths in women by Y and use $\mu = 0.34$ for the parameter of the Poisson distribution. The number of annual deaths in this group of 1000 men and 1000 women is then Poisson distributed with parameter

$$\lambda + \mu = 0.71 + 0.34 = 1.05 \qquad n = 1000,\ p = \tfrac{71}{105} \Rightarrow \lambda = np = 0.71$$

and it follows that

$$P(X + Y = 0) = e^{-(\lambda+\mu)} = e^{-1.05} \approx 0.3499$$

•

Section 12.4 Problems

12.4.1

1. Toss a fair coin twice. Let X be the random variable that counts the number of tails in each outcome. Find the probability mass function describing the distribution of X.

2. Toss a fair coin four times. Let X be the random variable that counts the number of heads. Find the probability mass function describing the distribution of X.

3. Roll a fair die twice. Let X be the random variable that gives the absolute value of the differences between the two numbers. Find the probability mass function describing the distribution of X.

4. Roll a fair die twice. Let X be the random variable that gives the maximum of the two numbers. Find the probability mass function describing the distribution of X.

5. An urn contains three green and two blue balls. You remove two balls at random without replacement. Let X denote the number of green balls in your sample. Find the probability mass function describing the distribution of X.

6. An urn contains five green balls, two blue balls, and three red balls. You remove three balls at random without replacement. Let X denote the number of red balls. Find the probability mass function describing the distribution of X.

7. You draw 3 cards from a standard deck of 52 cards without replacement. Let X denote the number of spades in your hand. Find the probability mass function describing the distribution of X.

8. You draw 5 cards from a standard deck of 52 cards without replacement. Let X denote the number of aces in your hand. Find the probability mass function describing the distribution of X.

9. Suppose that the probability mass function of a discrete random variable X is given by the following table:

x	$P(X = x)$
-3	0.2
-1	0.3
1.5	0.4
2	0.1

Find and graph the corresponding distribution function $F(x)$.

10. Suppose the probability mass function of a discrete random variable X is given by the following table:

x	$P(X = x)$
-1	0.2
-0.5	0.25
0.1	0.1
0.5	0.1
1	0.35

Find and graph the corresponding distribution function $F(x)$.

11. Let X be a random variable with distribution function

$$F(x) = \begin{cases} 0 & x < -2 \\ 0.2 & -2 \le x < 0 \\ 0.3 & 0 \le x < 1 \\ 0.7 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

Determine the probability mass function of X.

12. Let X be a random variable with distribution function

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.05 & 0 \le x < 1.3 \\ 0.30 & 1.3 \le x < 1.7 \\ 0.85 & 1.7 \le x < 1.9 \\ 0.90 & 1.9 \le x < 2 \\ 1.0 & x \ge 2 \end{cases}$$

Determine the probability mass function of X.

13. Let $S = \{1, 2, 3, \ldots, 10\}$, and assume that

$$p(k) = \frac{k}{N}, \ k \in S$$

where N is a constant.

(a) Determine N so that $p(k)$, $k \in S$, is a probability mass function.

(b) Let X be a discrete random variable with $P(X = k) = p(k)$. Find the probability that X is less than 8.

14. Geometric Distribution In Example 2, we tossed a coin repeatedly until the first heads showed up. Assume that the probability of heads is p, where $p \in (0, 1)$. Let Y be a random variable that counts the number of trials until the first heads shows up.

(a) Show that $P(Y = 1) = p$, $P(Y = 2) = (1 - p)p$, and $P(Y = 3) = (1 - p)^2 p$.

(b) Explain why

$$P(Y = j) = (1 - p)^{j-1} p$$

for $j = 1, 2, \ldots$. This equation is called the **geometric distribution**.

(c) Prove that

$$\sum_{j \ge 1} P(Y = j) = 1$$

as follows:

(i) For $0 \le q < 1$, define

$$S_n = 1 + q + q^2 + \cdots + q^n$$

Show that

$$S_n - qS_n = 1 - q^{n+1}$$

and conclude from this equation that

$$S_n = \frac{1 - q^{n+1}}{1 - q}$$

(ii) Show that

$$P(Y \le k) = \sum_{j=1}^{k} P(Y = j) = p \sum_{j=1}^{k} (1 - p)^{j-1}$$

Use your results in (i) to show that this formula simplifies to

$$1 - (1 - p)^k$$

and conclude from this equation that

$$\lim_{k \to \infty} P(Y \le k) = 1$$

which is equivalent to

$$\sum_{j \ge 1} P(Y = j) = 1$$

12.4.2

15. The following table contains the number of leaves per basil plant in a sample of size 25:

19	21	20	13	18
14	17	14	17	17
13	15	12	15	17
15	16	18	17	14
14	14	13	20	13

(a) Find the relative frequency distribution.

(b) Compute the average value by **(i)** averaging the values in the table directly and **(ii)** using the relative frequency distribution obtained in (a).

16. The following table contains the number of aphids per plant in a sample of size 30:

15	27	13	2	0	16
26	0	2	1	17	15
21	13	5	0	19	25
12	11	0	16	22	1
28	9	0	0	1	17

(a) Find the relative frequency distribution.

(b) Compute the average value by **(i)** averaging the values in the table directly and **(ii)** using the relative frequency distribution obtained in (a).

17. The following table contains the scores of 25 students on a certain exam:

7	8	8	3	2
5	6	9	10	6
8	8	7	6	9
10	4	4	8	6
9	10	5	5	8

(a) Find the relative frequency distribution.

(b) Compute the average value by **(i)** averaging the values in the table directly and **(ii)** using the relative frequency distribution obtained in (a).

18. The following table contains the number of flower heads per plant in a sample of size 20:

15	17	19	18	15
17	18	15	14	19
17	15	15	18	19
20	17	14	17	18

(a) Find the relative frequency distribution.

(b) Compute the average value by **(i)** averaging the values in the table directly and **(ii)** using the relative frequency distribution obtained in (a).

19. Suppose that the probability mass function of a discrete random variable X is given by the following table:

x	$P(X = x)$
-2	0.1
-1	0.4
0	0.3
1	0.2

(a) Find $E(X)$. **(b)** Find $E(X^2)$. **(c)** Find $E[X(X-1)]$.

20. Suppose that the probability mass function of a discrete random variable X is given by the following table:

x	$P(X = x)$
0	0.3
1	0.3
2	0.1
3	0.1
4	0.2

(a) Find $E(X)$. **(b)** Find $E(X^2)$. **(c)** Find $E(2X-1)$.

21. Suppose that the probability mass function of a discrete random variable X is given by the following table:

x	$P(X = x)$
-3	0.2
-1	0.3
1.5	0.4
2	0.1

Find the mean, the variance, and the standard deviation of X.

22. Suppose that the probability mass function of a discrete random variable X is given by the following table:

x	$P(X = x)$
-1	0.1
-0.5	0.2
0.1	0.1
0.5	0.25
1	0.35

Find the mean, the variance, and the standard deviation of X.

23. Let X be **uniformly distributed** on the set

$$S = \{1, 2, 3, \ldots, 10\}$$

That is,

$$P(X = k) = \frac{1}{10}, \quad k \in S$$

(a) Find $E(X)$. **(b)** Find $\text{var}(X)$.

24. Let X be uniformly distributed on the set

$$S = \{1, 2, 3, \ldots, n\}$$

where n is a positive integer; that is,

$$P(X = k) = \frac{1}{n}, \quad k \in S$$

(a) Find $E(X)$. **(b)** Find $\text{var}(X)$.

Hint: Recall that

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

and

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

25. Assume that X is a discrete random variable with finite range, and set

$$p(x) = P(X = x)$$

(a) Show that

$$E(aX + b) = \sum_x (ax + b)p(x)$$

(b) Use your result in (a) and the rules for finite sums to conclude that

$$E(aX + b) = aE(X) + b$$

26. Assume that X is a discrete random variable with finite range, and set

$$p(x) = P(X = x)$$

(a) Show that

$$\text{var}(aX + b) = a^2 \sum_x [x - E(X)]^2 p(x)$$

(b) Use your result in (a) and the rules for finite sums to conclude that

$$\operatorname{var}(aX + b) = a^2 \operatorname{var}(X)$$

27. Let X and Y be two random variables with the following joint distribution:

	$X = 0$	$X = 1$
$Y = 0$	0.3	0.1
$Y = 1$	0.2	0.4

(a) Find $P(X = 1, Y = 0)$. **(b)** Find $P(X = 1)$.
(c) Find $P(Y = 0)$. **(d)** Find $P(Y = 0 \mid X = 1)$.

28. Let X and Y be two random variables with the following joint distribution:

	$X = 0$	$X = 1$
$Y = 0$	0.2	0.0
$Y = 1$	0.3	0.5

(a) Find $P(X = 0, Y = 1)$. **(b)** Find $P(X = 0)$.
(c) Find $P(Y = 1)$. **(d)** Find $P(X = 0 \mid Y = 0)$.

29. Let X and Y be two independent random variables with probability mass function described by the following table:

k	$P(X = k)$	$P(Y = k)$
-2	0.1	0.2
-1	0	0.2
0	0.3	0.1
1	0.4	0.3
2	0.05	0
3	0.15	0.2

(a) Find $E(X)$ and $E(Y)$. **(b)** Find $E(X + Y)$.
(c) Find $\operatorname{var}(X)$ and $\operatorname{var}(Y)$. **(d)** Find $\operatorname{var}(X + Y)$.

30. Let X and Y be two independent random variables with probability mass function described by the following table:

k	$P(X = k)$	$P(Y = k)$
-3	0.1	0.1
-1	0.1	0.2
0	0.2	0.1
0.5	0.3	0.3
2	0.15	0.1
2.5	0.15	0.2

(a) Find $E(X)$ and $E(Y)$. **(b)** Find $E(X + Y)$.
(c) Find $\operatorname{var}(X)$ and $\operatorname{var}(Y)$. **(d)** Find $\operatorname{var}(X + Y)$.

31. We have two formulas for computing the variance of X, namely,

$$\operatorname{var}(X) = E[(X - E(X))^2],$$

and

$$\operatorname{var}(X) = E(X^2) - [E(X)]^2$$

(a) Explain why $\operatorname{var}(X) \ge 0$.

(b) Use your results in (a) to explain why

$$E(X^2) \ge [E(X)]^2$$

32. Assume that X is a discrete random variable with finite range. Show that if $\operatorname{var}(X) = 0$, then $P(X = E(X)) = 1$.

12.4.3

33. Toss a fair coin 10 times. Let X be the number of heads. Find

(a) $P(X = 5)$. **(b)** $P(X \ge 8)$. **(c)** $P(X \le 9)$.

34. Toss a coin with probability of heads 0.3 five times. Let X be the number of tails. Find

(a) $P(X = 2)$. **(b)** $P(X \ge 1)$.

35. Roll a fair die six times. Let X be the number of times you roll a 6. Find the probability mass function.

36. A loaded die has probability 0.5 of rolling a 6 and probability 0.1 of rolling each of the other five numbers. Find the probability of rolling a 6 three times in a row.

37. A loaded die is weighted so that rolling a 4 is three times as likely as rolling any of the other numbers. You roll the die twice and record the sum of the two numbers. What is the probability that the sum is equal to 7.

38. An urn contains four green and six blue balls. You draw a ball at random, note its color, and replace it. You repeat these steps four times. Let X denote the total number of green balls you obtain. Find the probability mass function of X.

39. An urn contains three blue and two white balls. You draw a ball at random, note its color, and replace it. You repeat these steps three times. Let X denote the total number of white balls. Find $P(X \le 1)$.

40. An urn contains four red, seven green, and two white balls. You draw a ball at random, note its color, and replace it. You repeat these steps four times. Let X denote the number of red balls and Y the number of green balls. Find $P(X + Y = 2)$.

41. Assume that 20% of all plants in a field are infested with aphids. Suppose that you pick 20 plants at random. What is the probability that none of them carried aphids?

42. **Blood Test** To test for a disease that has a prevalence of 1 in 100 in a population, blood samples of 10 individuals are mixed and the mixed blood is then tested. What is the probability that the test result is negative (i.e., the disease is not present in the pooled blood sample)?

43. Suppose that a box contains 10 apples. The probability that any one apple is spoiled is 0.1. (Assume that spoilage of the apples is an independent phenomenon.)

(a) Find the expected number of spoiled apples per box.

(b) A shipment contains 10 boxes of apples. Find the expected number of boxes that contain no spoiled apples.

44. Toss a fair coin 10 times. Let X denote the number of heads. What is the probability that X is within one standard deviation of its mean?

45. A multiple-choice exam contains 50 questions. Each question has four choices. Find the expected number of correct answers if a student guesses the answers at random.

46. A true–false exam has 20 questions. Find the expected number of correct answers if a student guesses the answers at random.

47. **Sampling With and Without Replacement** An urn contains 12 green and 24 blue balls.

(a) You take 10 balls out of the urn without replacing them. Find the probability that 6 of the 10 balls are blue.

(b) You take a ball out of the urn, note its color, and replace it. You withdraw a total of 10 balls this way. Find the probability that 6 of the 10 balls are blue.

48. Sampling With and Without Replacement An urn contains K green and $N - K$ blue balls.

(a) You take n balls out of the urn. Find the probability that k of the n balls are green.

(b) You take a ball out of the urn, note its color, and replace it. You repeat these steps n times. Find the probability that k of the n balls are green.

12.4.4

49. Repeat Example 27 when $N_1 = 10$, $N_2 = 14$, and $N_3 = 6$.

50. Repeat Example 27 when $N_1 = 5$, $N_2 = 15$, and $N_3 = 10$.

51. Repeat Example 28 when 20 seeds are round and yellow, 10 are round and green, 8 are wrinkled and yellow, and 2 are wrinkled and green.

52. Repeat Example 28 when 17 seeds are round and yellow, 22 are round and green, 13 are wrinkled and yellow, and 8 are wrinkled and green.

53. An urn contains six green, eight blue, and 10 red balls. You take one ball out of the urn, note its color, and replace it. You withdraw a total of six balls this way. What is the probability that you sampled two of each color?

54. An urn contains eight green, four blue, and six red balls. You take one ball out of the urn, note its color, and replace it. You repeat these steps four times. What is the probability that you sampled two green, one blue, and one red ball?

55. In a $Cc \times Cc$ crossing of peas, 5 offspring are of genotype CC, 12 are of genotype Cc, and 6 are of genotype cc. What is the probability of this event?

56. In a $Cc \times Cc$ crossing of peas, two offspring are of genotype CC, three are of genotype Cc, and one is of genotype cc. What is the probability of this event?

Recessive Genes

A number of traits are caused by recessive genes. The traits show up only in individuals who are homozygous (i.e., have two copies of the mutant gene). An individual with one normal and one mutant gene is a carrier, but does not exhibit the trait. In Problems 57–59, calculate each of the probabilites.

57. The inability to roll one's tongue is caused by a single pair of recessive genes (rr). For a couple consisting of a heterozygote individual (Rr) and an affected person (rr), what is the probability that, among their four children, at most one child is unable to roll his or her tongue?

58. An attached earlobe is caused by a single pair of recessive genes (aa). For a couple consisting of a heterozygous individual (Aa) and an affected person (aa), what is the probability that a child has an unattached earlobe?

59. Tay–Sachs disease is caused by a single pair of recessive genes. If both parents are carriers of the mutant gene, what is the likelihood that none of their four children will be affected?

60. Assume a 1:1 sex ratio. A woman who is a carrier of hemophilia has two daughters and two sons with a man who is not hemophilic. What is the probability that one daughter is not a carrier, one daughter is a carrier, one son is hemophilic, and one son is not hemophilic?

12.4.5

61. A random experiment consists of flipping a fair coin until the first time heads appears. Find the probability that the first heads appears on the kth trial for $k = 1, 2$, and 3.

62. A random experiment consists of flipping a biased coin with probability 0.3 of heads until the first time heads appears. Find the probability that heads appears for the first time on the fifth trial.

63. A random experiment consists of rolling a fair die until the first time an even number appears. Find the probability that the first even number appears on the third trial.

64. A random experiment consists of rolling a fair die until the first time a five or a six appears. Find the probability that the first five or six appears on the kth trial for $k = 1, 2, \ldots, 5$.

65. A random experiment consists of flipping a fair coin until the first time heads appears. Find the probability that the first heads appears after the third trial.

66. A random experiment consists of rolling a fair die until the first six appears. Find the probability that the first six appears after the seventh trial.

67. A random experiment consists of flipping a fair coin until the first time heads appears. Find the probability that the first heads appears within the first four trials.

68. A random experiment consists of rolling a fair die until the first time a 1 or a 2 appears. Find the probability that the first 1 or 2 appears within the first five trials.

69. An urn contains 1 black and 14 white balls. Balls are drawn at random, one at a time, until the black ball is selected. Each ball is replaced before the next ball is drawn. Find the probability that at least 20 draws are needed.

70. An urn contains 1 black and $n - 1$ white balls. Balls are drawn at random, one at a time, until the black ball is selected. Each ball is replaced before the next ball is drawn. Find the probability that at least n draws are needed. What happens as $n \to \infty$?

71. An urn contains 5 green and 25 blue balls. Balls are drawn at random, one at a time, until a green ball is selected. Each ball is replaced before the next ball is drawn. Let T denote the first time until a green ball is drawn. Find $E(T)$ and $\text{var}(T)$.

72. An urn contains 10 green and 20 blue balls. Balls are drawn at random, one at a time, until a green ball is selected. Each ball is replaced before the next ball is drawn. Let T denote the first time until a green ball is drawn. Find $E(T)$ and $\text{var}(T)$.

73. An urn contains 1 black and 9 white balls. Balls are drawn at random until the black ball is selected. Find the probability that exactly 6 white balls will be drawn before the black one is if **(a)** each ball is replaced before the next ball is drawn and **(b)** balls are not replaced.

74. An urn contains 1 black and $n-1$ white balls. Balls are drawn at random until the black ball is selected. Find the probability that exactly k white balls will be drawn before the black one is if **(a)** each ball is replaced before the next ball is drawn and **(b)** balls are not replaced.

75. Suppose the waiting time for the first success in an experiment is geometrically distributed with mean $1/p$.

(a) Find the probability that the first success occurs on the kth trial.

(b) The experiment is repeated after the first success. Assume that the waiting time for the second success has the same distribution as the waiting time for the first success. Find the probability mass function for the distribution of the second success.

76. A Bernoulli experiment with probability of success p is repeated until the nth success. Assume that each trial is independent of all others. Find the probability mass function of the distribution of the nth success. (This distribution is called the **negative binomial distribution**.)

12.4.6

77. Suppose X is Poisson distributed with parameter $\lambda = 2$. Find $P(X = k)$ for $k = 0, 1, 2$, and 3.

78. Suppose X is Poisson distributed with parameter $\lambda = 0.5$. Find $P(X = k)$ for $k = 0, 1, 2$, and 3.

79. Suppose X is Poisson distributed with parameter $\lambda = 1$.

(a) Find $P(X \geq 2)$. **(b)** Find $P(1 \leq X \leq 3)$.

80. Suppose X is Poisson distributed with parameter $\lambda = 0.2$.

(a) Find $P(X < 3)$. **(b)** Find $P(2 \leq X \leq 4)$.

81. Suppose X is Poisson distributed with parameter $\lambda = 1.5$. Find the probability that X exceeds 3.

82. Suppose X is Poisson distributed with parameter $\lambda = 1.2$. Find the probability that X is at most 3.

83. Suppose X is Poisson distributed with parameter $\lambda = 2$. Find the probability that X is at least 2.

84. Suppose X is Poisson distributed with parameter $\lambda = 0.6$. Find the probability that X is less than 3.

85. Suppose the number of phone calls arriving at a switchboard per hour is Poisson distributed with mean 7 calls per hour. Find the probability that no phone calls arrive during a certain hour.

86. Suppose the number of phone calls arriving at a switchboard per hour is Poisson distributed with mean 3 calls per hour.

(a) Find the probability that at least one phone call arrives between noon and 1 P.M.

(b) Assuming that phone calls in different hours are independent of each other, find the probability that no phone calls arrive between noon and 2 P.M.

87. Suppose the number of typos on a book page is Poisson distributed with mean 0.5. Find the probability that there is at least one typo on a given page.

88. Suppose the number of typos on a book page is Poisson distributed with mean 0.1.

(a) Find the probability that there are no typos on a page.

(b) How many pages with typos do you expect in a 200-page book?

89. Amino Acid Evolution The number of amino acid substitutions in a given sequence is Poisson distributed with mean 3. What is the probability of at least two substitutions?

90. Amino Acid Evolution The number of amino acid substitutions in a given sequence is Poisson distributed with mean 2. Given that there are substitutions on the sequence, what is the probability that there are at least two substitutions?

91. X and Y are independent and Poisson with mean 3.

(a) Find $P(X + Y = 2)$.

(b) Given that $X + Y = 2$, find the probability that $X = k$ for $k = 0, 1$, and 2.

92. X is Poisson distributed with mean 2, and Y is Poisson distributed with mean 3.

(a) Find $P(X + Y = 4)$

(b) Given that $X + Y = 1$, find the probability that $X = 1$.

93. Let X and Y be independent Poisson distributed variables with means 4 and 2 respectively. Calculate $P(X = 2|X + Y = 3)$.

94. Suppose X and Y are independent and Poisson with mean λ. Given that $X + Y = n$, find the probability that $X = k$ for $k = 0, 1, 2, \ldots, n$.

In Problems 95–99 use the Poisson approximation.

95. For a certain vaccine, 1 in 1000 individuals experiences some side effects. Find the probability that, in a group of 500 people, nobody experiences side effects.

96. For a certain vaccine, 1 in 500 individuals experiences some side effects. Find the probability that, in a group of 200 people, at least 1 person experiences side effects.

97. Down Syndrome About 1 in 700 births in the United States is affected by Down syndrome, a chromosomal disorder. Find the probability that there is at most 1 case of Down syndrome among 1000 births by **(a)** computing the exact probability and **(b)** using a Poisson approximation.

98. Fragile *X* Syndrome About 1 in 1000 boys is affected by fragile X syndrome, a genetic disorder that causes learning difficulties. Find the probability that, in a group of 500 boys, nobody is affected by this disorder by **(a)** computing the exact probability and **(b)** using a Poisson approximation.

99. (Refer to Example 37.) Suppose a parasitoid has a probability of 0.03 of detecting a given host. If 50 parasitoids are trying to find a particular host, what is the probability that the host will avoid detection?

12.5 Continuous Distributions

12.5.1 Density Functions

In the previous section, we discussed random variables that took on a discrete range of values. In this section, we will discuss random variables that take on a continuum of values (e.g., any number on an interval). Called **continuous random variables**, they arise, for instance, when we consider the length distribution of an organism: Within an appropriate interval, an individual's length can take on any value.

To illustrate, consider the following example adapted from ***de Roos (1996)***: The water flea *Daphnia pulex* feeds on the alga *Chlamydomonas rheinhardii*. An important component of the feeding behavior of *Daphnia* is the strong dependence of the amount of food consumed on the size of the individual. To model the feeding behavior of *Daphnia*, we would need to describe the distribution of *Daphnia* sizes. We view size as a random variable, denoted by X, and determine, for all possible values of x, the fraction of *Daphnia* whose size is less than or equal to x. If the population size is large, this fraction can be well approximated by a continuous function, which we denote by $F(x)$. This function serves the role of a distribution function, so $F(x) = P(X \leq x)$.

A distribution function completely characterizes the probability distribution of a random variable, as we saw in the previous section. This property is no different for continuous random variables. To describe the probability distribution of a continuous random variable X, we will therefore use its distribution function $F(x)$. The distribution function for a continuous random variable has the same definition as the one for a discrete random variable:

$$F(x) = P(X \leq x)$$

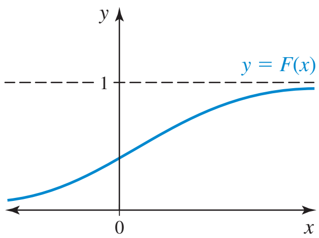


Figure 12.26 The distribution function of a continuous random variable.

$F(x)$ has the following properties (see Figure 12.26), some of which are absent in the case of a discrete random variable:

1. $0 \leq F(x) \leq 1$.
2. $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.
3. $F(x)$ is nondecreasing and continuous.

Note on the one hand that the distribution function of a continuous random variable is continuous, unlike the distribution function of a discrete random variable, which is piecewise constant and takes jumps at those values of x where $P(X = x) > 0$. On the other hand, both distribution functions are nondecreasing and take on values between 0 and 1.

EXAMPLE 1 Show that

$$F(x) = \begin{cases} 1 - e^{-2x} & \text{for } x > 0 \\ 0 & \text{for } x \leq 0 \end{cases}$$

is a distribution function of a continuous random variable.

Solution A graph of $F(x)$ is shown in Figure 12.27. We must check the three properties of distribution functions of continuous random variables:

1. Since $0 \leq 1 - e^{-2x} \leq 1$ for $x > 0$ and $F(x) = 0$ for $x \leq 0$, it follows that $0 \leq F(x) \leq 1$ for all $x \in \mathbf{R}$.
2. Since $F(x) = 0$ for $x \leq 0$ and $\lim_{x\to\infty} e^{-2x} = 0$, it also follows that
$$\lim_{x\to-\infty} F(x) = 0 \quad \text{and} \quad \lim_{x\to\infty} F(x) = 1$$
3. To show that $F(x)$ is continuous for all $x \in \mathbf{R}$, note that $F(x)$ is continuous for both $x > 0$ and $x < 0$. To check continuity at $x = 0$, we compute
$$\lim_{x\to 0^+} F(x) = \lim_{x\to 0^+} (1 - e^{-2x}) = 0 = \lim_{x\to 0^-} F(x)$$
which is equal to $F(0) = 0$. Hence, $F(x)$ is continuous at $x = 0$.

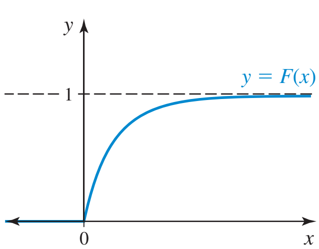


Figure 12.27 The distribution function in Example 1.

To show that $F(x)$ is nondecreasing, we compute $F'(x)$ for $x > 0$:

$$F'(x) = 2e^{-2x} > 0 \text{ for } x > 0$$

This equation implies that $F(x)$ is increasing for $x > 0$. Since $F(x)$ is continuous for all $x \in \mathbf{R}$ and equal to 0 for $x \leq 0$, it follows that $F(x)$ is nondecreasing for all $x \in \mathbf{R}$. •

If there is a nonnegative function $f(x)$ such that the distribution function $F(x)$ of a random variable X has the representation

$$F(x) = \int_{-\infty}^{x} f(u)\, du$$

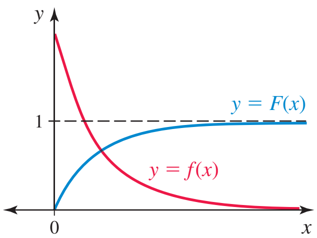


Figure 12.28 The distribution function $F(x)$ and corresponding density function $f(x)$ of a continuous random variable.

we say that X is a continuous random variable with **(probability) density function** $f(x)$. (See Figure 12.28.)

Since $F(x)$ is a distribution function, it follows that

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \tag{12.28}$$

Any nonnegative function that satisfies (12.28) defines a density function. The function $f(x)$ need not be continuous, although in all of our applications $f(x)$ will be continuous, except perhaps for a finite number of points. The function $f(x)$ will frequently be defined as a piecewise continuous function. Using part I of the fundamental theorem of calculus, we can obtain the density function $f(x)$ of a continuous random variable from the distribution function $F(x)$ by differentiating the distribution function. That is, $f(x) = F'(x)$ at all points x where $F(x)$ is differentiable. At points where $F(x)$ is not differentiable, we set $f(x) = 0$, a strategy that ensures that the density function is defined everywhere.

Furthermore, since

$$P(a < X \le b) = P(X \le b) - P(X \le a) = F(b) - F(a)$$

it follows that

$$P(a < X \le b) = \int_a^b f(x)\,dx \tag{12.29}$$

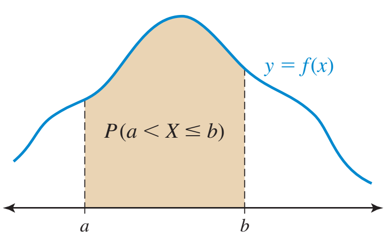


Figure 12.29 The area between the density function $f(x)$ and the x-axis between a and b represents the probability that the random variable X lies between a and b.

That is, the area under the curve $y = f(x)$ between a and b represents the probability that the random variable takes on values between a and b, as illustrated in Figure 12.29. It is important to realize that, in (12.29), probabilities are represented by areas; in particular, the *integral* of $f(x)$—not $f(x)$ itself—has this physical interpretation. The density function $f(x)$ does not have an immediate physical interpretation. Later in the section, we will explain how to find $f(x)$ empirically.

In contrast to discrete random variables, for which $P(X \le b)$ and $P(X < b)$ can differ, there is no difference for continuous random variables, since

$$P(X = b) = \int_b^b f(x)\,dx = 0$$

Therefore,

$$\begin{aligned}\int_a^b f(x)\,dx &= P(a < X \le b) = P(a \le X \le b)\\ &= P(a < X < b) = P(a \le X < b).\end{aligned}$$

EXAMPLE 2 The distribution function of a continuous random variable X is given by

$$F(x) = \begin{cases} 0 & \text{for } x \le 0 \\ x^2 & \text{for } 0 < x < 1 \\ 1 & \text{for } x \ge 1 \end{cases}$$

(a) Find the corresponding density function.

(b) Compute $P(-1 \le X \le 1/2)$.

Solution **(a)** To find the density function, we need to invoke part I of the fundamental theorem of calculus. Thus, we have

$$F(x) = \int_a^x f(u)\,du \qquad \text{implies} \qquad F'(x) = f(x)$$

The density function $f(x)$ can be found by differentiating the distribution function $F(x)$ in the intervals $(-\infty, 0)$, $(0, 1)$, and $(1, \infty)$:

$$f(x) = F'(x) = \begin{cases} 2x & \text{for } 0 < x < 1 \\ 0 & \text{for } x < 0 \text{ or } x > 1 \end{cases}$$

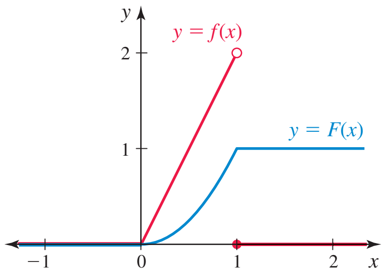


Figure 12.30 The distribution function $F(x)$ and corresponding density function $f(x)$ in Example 2.

To define the density function everywhere, we set $f(0) = f(1) = 0$. The distribution function, together with its density function, is shown in Figure 12.30.

(b) Using the distribution function, we immediately find that

$$P\left(-1 \le X \le \frac{1}{2}\right) = F\left(\frac{1}{2}\right) - F(-1) = \frac{1}{4} - 0 = \frac{1}{4}$$

If, instead, we use the density function, we must evaluate

$$P\left(-1 \le X \le \frac{1}{2}\right) = \int_{-1}^{1/2} f(x)\,dx = \int_0^{1/2} 2x\,dx$$
$$= x^2\Big]_0^{1/2} = \frac{1}{4}$$

The formulas for the mean and the variance of a continuous random variable are analogous to those for the discrete case. The expected value, or mean, $E(X)$ of a continuous random variable X with density function $f(x)$ is defined as

$$\boxed{E(X) = \int_{-\infty}^{\infty} x f(x)\,dx}$$

The expected value of a function of a random variable $g(X)$ is

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\,dx$$

where $f(x)$ is the density function of X.

The variance of a continuous random variable X with mean μ is defined as

$$\boxed{\operatorname{var}(X) = E(X - \mu)^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\,dx}$$

The alternative formula that we gave in the previous section holds as well:

$$\operatorname{var}(X) = E(X^2) - [E(X)]^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \left(\int_{-\infty}^{\infty} x f(x)\,dx\right)^2$$

Recall that these integrals are defined on unbounded intervals and $f(x)$ might be discontinuous. To evaluate such integrals, we must use the methods developed in Section 7.4.

EXAMPLE 3 The density function of a random variable X is given by

$$f(x) = \begin{cases} 3x^2 & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

(See Figure 12.31.) Compute the mean and the variance of X.

Solution To compute the mean, we must evaluate

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 3x^3\,dx = \frac{3}{4}x^4\Big]_0^1 = \frac{3}{4}$$

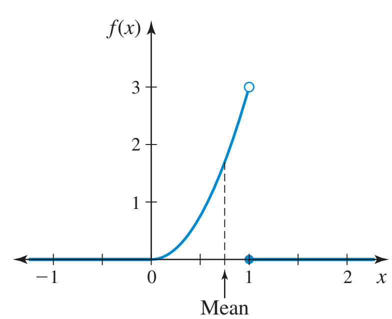


Figure 12.31 The density function $f(x)$ in Example 3, together with the location of the mean of X.

The mean is indicated in Figure 12.31. To compute the variance, we first evaluate

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^1 3x^4\,dx = \frac{3}{5}x^5\Big]_0^1 = \frac{3}{5}$$

The variance of X is then given by

$$\text{var}(X) = E(X^2) - [E(X)]^2 = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = \frac{3}{80}$$

•

EXAMPLE 4 **Seed Dispersal** The exponential function

$$f(r) = \begin{cases} \lambda e^{-\lambda r} & \text{for } r > 0 \\ 0 & \text{for } r \le 0 \end{cases}$$

where $\lambda > 0$ is a constant, is frequently used to model seed dispersal. The function $f(r)$ is a density function, and $\int_a^b f(r)\,dr$ describes the fraction of seeds dispersed between distances a and b from the source at 0. Find the average dispersal distance.

Solution We use the formula for the average value, namely,

$$\text{average dispersal distance} = \int_{-\infty}^{\infty} r f(r)\,dr = \int_0^{\infty} r\lambda e^{-\lambda r}\,dr$$

since $f(r) = 0$ for $r \le 0$. To evaluate this integral, we must integrate by parts:

$$\begin{aligned}\text{average dispersal distance} &= \int_0^{\infty} r\lambda e^{-\lambda r}\,dr = \lim_{z\to\infty} \int_0^z r\lambda e^{-\lambda r}\,dr \\ &= \lim_{z\to\infty} \left[r(-e^{-\lambda r})\right]_0^z + \lim_{z\to\infty} \int_0^z e^{-\lambda r} dr\end{aligned}$$

The first expression on the right is equal to

$$\lim_{z\to\infty} [-ze^{-\lambda z} + 0] = -\lim_{z\to\infty} \frac{z}{e^{\lambda z}}$$

This limit is of the form $\frac{\infty}{\infty}$. Using l'Hôpital's rule, we find that

$$\lim_{z\to\infty} \frac{z}{e^{\lambda z}} = \lim_{z\to\infty} \frac{1}{\lambda e^{\lambda z}} = 0$$

since $\lambda > 0$. The second expression is

$$\lim_{z\to\infty} \int_0^z e^{-\lambda r} dr = \lim_{z\to\infty} \left[-\frac{1}{\lambda} e^{-\lambda r}\right]_0^z = \frac{1}{\lambda}$$

Therefore,

$$\text{average dispersal distance} = \frac{1}{\lambda}$$

•

We will now discuss how to determine $f(x)$ empirically. We set

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

Then

$$F(x + \Delta x) - F(x) = \int_{-\infty}^{x+\Delta x} f(t)\,dt - \int_{-\infty}^{x} f(t)\,dt = \int_{x}^{x+\Delta x} f(t)\,dt$$

If Δx is sufficiently small, then $f(t)$ will not vary much over the interval $[x, x + \Delta x)$, and we approximate $f(t)$ by $f(x)$ over the interval $[x, x + \Delta x)$. Hence,

$$\int_{x}^{x+\Delta x} f(t)\,dt \approx f(x)\Delta x \tag{12.30}$$

for Δx sufficiently small, and we can think of $f(x)\Delta x$ as approximately representing the fraction that falls into the interval $[x, x + \Delta x)$. This is an important interpretation

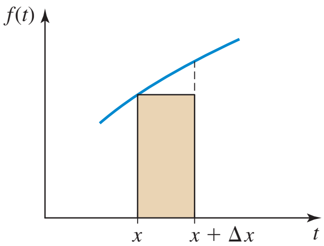


Figure 12.32 The area of the rectangle is $f(x)\Delta x$, a good approximation of $\int_x^{x+\Delta x} f(t)\,dt$, which is the area under the curve between x and $x + \Delta x$.

that will help us to determine density functions empirically. It should remind you of the Riemann sum approximation that we discussed in Section 6.1. For Δx small, just one rectangle gives a good approximation, as illustrated in Figure 12.32.

To determine the density $f(x)$ empirically, we will use the approximation (12.30). We take a sufficiently large sample from the population and measure the quantity of interest of each individual sampled. We partition the interval over which the quantity of interest varies into subintervals of length Δx_i. For each subinterval, we count the number of sample points that fall into the respective subintervals. To display the data graphically, we use a **histogram**, which consists of rectangles whose widths are equal to the lengths of the corresponding subintervals and whose *areas* are equal to the number of sample points that fall into the corresponding subintervals. This approach is analogous to approximating areas under curves by rectangles; we illustrate it in the example that follows.

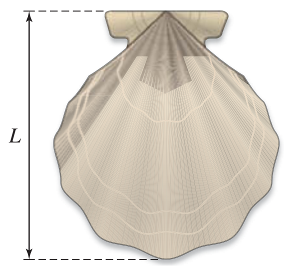


Figure 12.33 The sagittal length of a shell.

Brachiopods form a marine invertebrate phylum whose soft body parts are enclosed in shells. These organisms were the dominant seabed shelled animals in the Paleozoic era, but almost disappeared during the Permian–Triassic mass extinction.[1] They are still present today (with approximately 120 genera) and occupy a diverse range of habitats, but they are no longer the dominant seabed shelled animal, their place having been taken by bivalve mollusks. (The brachiopod story is described in ***Ward, 1992***.)

Brachiopod *Dielasma* fossils are common in Permian reef deposits in the north of England. The following table (adapted from ***Benton and Harper, 1997***) represents measurements of the sagittal length of the animal's shell, measured in mm (see Figure 12.33):

Length	5	7	9	11	13	15	17	19	21	23	25
Frequency	3	28	12	2	4	4	6	6	5	3	1

The length measurements are divided into classes $[0, 2), [2, 4), [4, 6), \ldots, [26, 28)$. The midpoint of each subinterval represents the size class. For instance, length 11 in the table corresponds to the size class of lengths between 10 mm and 12 mm. The number below each size class—the *frequency*—represents the number of brachiopods in the sample whose lengths fell into the corresponding size class. For instance, there were two brachiopods in the sample whose lengths fell into the size class $[10, 12)$.

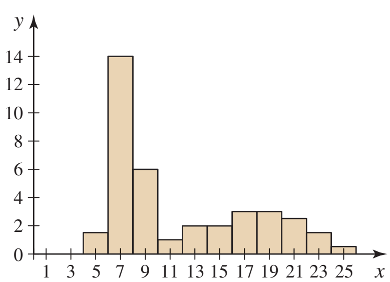


Figure 12.34 The histogram for the frequency distribution of the sagittal length.

To display this data set graphically, we use a histogram, as shown in Figure 12.34. The horizontal axis shows the midpoint of each size class. The graph consists of rectangles whose widths are equal to the length of the corresponding size class and whose *area* is proportional to the number of specimens in the corresponding class. It is very important to note that a histogram represents numbers by area, not by height. For instance, the number of specimens in size class $[8, 10)$ is equal to 12. This size class is represented by 9, the midpoint of the interval $[8, 10)$. Because the width of the size class is 2, the height of the rectangle must be equal to 6 units so that the area of the corresponding rectangle is equal to 12 units. In our example, all size classes are of the same length; of course, this need not be the case in general.

Displaying data graphically has certain advantages. For instance, we see immediately that the size distribution is biased toward smaller sizes, because the rectangles in the histogram that correspond to smaller shell lengths have larger areas. This bias toward smaller sizes might indicate, for instance, that brachiopods suffered a high juvenile mortality.

It is often convenient to scale the vertical axis of the histogram so that the total area of the histogram is equal to 1. One of the advantages of this approach is that the histogram then does not directly reflect the sample size, since only proportions are represented. Consequently, it is easier to compare histograms from different samples. For instance, if someone else had obtained a different sample of this type of

(1) The Permian geological period lasted from 286 million to 248 million years ago; the Triassic followed the Permian and lasted from 248 million to 213 million years ago. The Permian–Triassic mass extinction is believed to have been the most severe mass extinction of life that has ever occurred.

brachiopod in the same location, and if both samples are representative of its length distribution, then both histograms should look similar.

If we scale the total area of the histogram to 1, then the area of each rectangle in the histogram represents the fraction of the sample in the corresponding class. To obtain the fraction of sample points in a certain class, we divide the number of sample points in that class by the total sample size. In our example, the sample size was 74. The fraction of the sample in size class [8, 10), for instance, would then be $12/74 = 0.16$, or 16%.

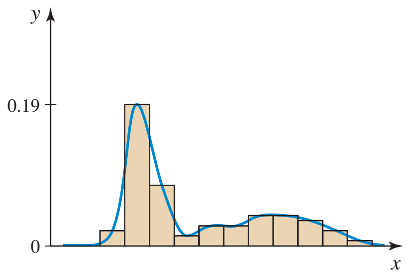


Figure 12.35 The scaled histogram of brachiopod lengths can be used as an approximation of the density function.

The choice of the widths of the classes in the histogram is somewhat arbitrary. The goal is to obtain an informative graph. Typically, the larger the sample size, the smaller the widths of the classes can be, and the better the approximation of $f(x)$. The outline of the normalized histogram (i.e., the total area of the histogram is equal to 1) can then be used as an approximation to the density function $f(x)$. (See Figure 12.35.)

In the subsections that follow, we will introduce some continuous distributions and their applications.

12.5.2 The Normal Distribution

The normal distribution was first introduced by Abraham De Moivre (1667–1754) in the context of computing probabilities in binomial experiments when the number of trials is large. Later, Gauss showed that this distribution was important in the error analysis of measurements. It is the most important continuous distribution, and we discuss an application first.

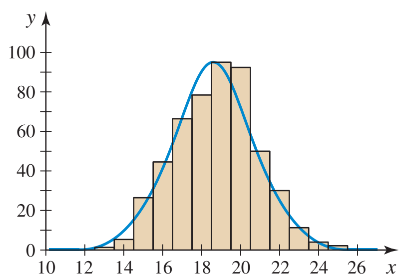


Figure 12.36 The number of abdominal bristles.

Quantitative genetics is concerned with measurable traits, such as plant height, litter size, body weight, and so on. Such traits are called **quantitative characters**. There are many quantitative characters whose frequency distributions follow a bell-shaped curve. For instance, counting the bristles on some particular part of the abdomen (fifth sternite) of a strain of *Drosophila melanogaster*, ***Mackay (1984)*** found that the number of bristles varied according to a bell-shaped curve. (This curve is shown in Figure 12.36, which is adapted from ***Hartl and Clark, 1989***.)

The smooth curve in Figure 12.36 that is fitted to the histogram is proportional to the density function of a **normal distribution**. (The curve is not scaled, so the area under the curve is not equal to 1.) The density function of the normal distribution is described by just two parameters, called μ and σ, which can be estimated from data. The parameter μ can be any real number; the parameter σ is a positive real number. The density function of a normal distribution is described as follows:

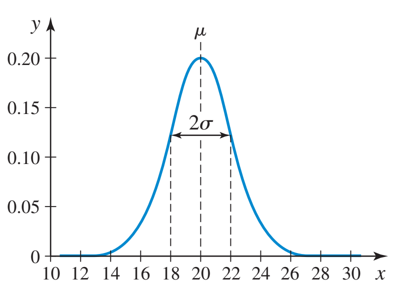


Figure 12.37 The graph of the normal density with $\mu = 20$ and $\sigma = 2$. The maximum is at $\mu = 20$; the two inflection points are at 18 and 22, respectively.

> A continuous random variable X is normally distributed with parameters μ and σ if it has density function
>
> $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \qquad -\infty < x < \infty$$

The parameter μ is the mean and the parameter σ is the standard deviation of the normal distribution. In Problem 11, we will investigate the shape of the density function of the normal distribution. (A graph is shown in Figure 12.37.) Following are the properties of $f(x)$:

1. $f(x)$ is symmetric about $x = \mu$.
2. The maximum of $f(x)$ is at $x = \mu$.
3. The inflection points of $f(x)$ are at $x = \mu - \sigma$ and $x = \mu + \sigma$.

Since $f(x)$ is a density function,

$$f(x) \geq 0 \qquad \text{and} \qquad \int_{-\infty}^{\infty} f(x)\,dx = 1$$

With the tools we have so far, we cannot show that the density function is normalized to 1. In Problem 12, we will show that the mean μ is indeed the expected value of X;

that is,

$$\mu = \int_{-\infty}^{\infty} x f(x)\, dx$$

Furthermore, if a quantity X is normally distributed with parameters μ and σ, then

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}\, dx$$

It is not possible to evaluate this integral by using elementary functions; it can be evaluated only numerically. There are tables for the normal distribution with parameters $\mu = 0$ and $\sigma = 1$ that list values for

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

A table for $F(x)$ is reproduced in the Appendix B and can be used to obtain probabilities for general μ and σ. Later, we will see how to do this.

k	$A(k)$
1	0.68
2	0.95
3	0.99

At the moment, we will only need a few values. The area $A(k)$ under the density function of the normal distribution with mean μ and standard deviation σ between $\mu - k\sigma$ and $\mu + k\sigma$, for $k = 1, 2,$ and 3, is shown in the table on the left. (See Figure 12.38.) That is, if a certain quantity X is normally distributed with mean μ and standard deviation σ, then 0.68 (or 68%) of the population falls within one standard deviation of the mean, 95% falls within two standard deviations of the mean, and 99% falls within three standard deviations of the mean.

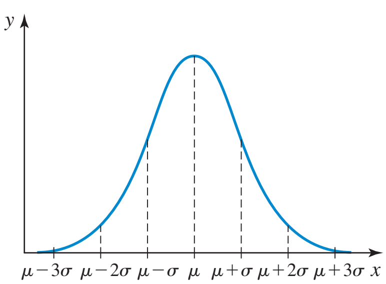


Figure 12.38 The density of the normal distribution with mean μ and standard deviation σ.

These percentages can also be interpreted in the following way: Suppose that a quantity X in a population is normally distributed with mean μ and standard deviation σ. If we sampled from this population—that is, if we picked one individual at random from the population—then there is a 68% chance that the observation would fall within one standard deviation of the mean. We can therefore say that the probability that X is in the interval $[\mu - \sigma, \mu + \sigma]$ is equal to 0.68, which we write as

$$P(X \in [\mu - \sigma, \mu + \sigma]) = 0.68$$

Likewise,

$$P(X \in [\mu - 2\sigma, \mu + 2\sigma]) = 0.95 \qquad \text{and} \qquad P(X \in [\mu - 3\sigma, \mu + 3\sigma]) = 0.99.$$

Finding Probabilities by Using the Mean and the Standard Deviation.

EXAMPLE 5 Assume that a certain quantitative character X is normally distributed with mean $\mu = 4$ and standard deviation $\sigma = 1.5$. Find an interval centered at the mean such that there is a 95% chance that an observation will fall into this interval. Then do the same for a 99% chance.

Solution Since 95% corresponds to a range within two standard deviations of the mean (see Figure 12.39), the resulting interval is

$$[4 - (2)(1.5), 4 + (2)(1.5)] = [1, 7]$$

We can therefore write

$$P(X \in [1, 7]) = 0.95.$$

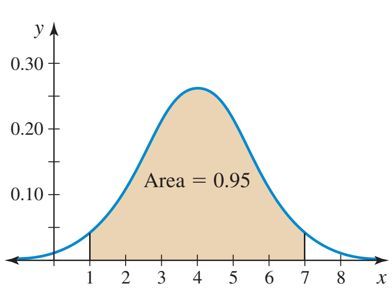


Figure 12.39 The normal density with mean 4 and standard deviation 1.5 in Example 5; 95% of the observations fall into the interval [1, 7].

Similarly, 99% corresponds to a range within three standard deviations of the mean, resulting in an interval of the form

$$[4 - (3)(1.5), 4 + (3)(1.5)] = [-0.5, 8.5].$$

We can therefore write

$$P(X \in [-0.5, 8.5]) = 0.99.$$

EXAMPLE 6 Assume that a certain quantitative character X is normally distributed with mean $\mu = 3$ and standard deviation $\sigma = 2$. We take a sample of size 1. What is the chance that we observe a value greater than 9?

Solution Since we know that $9 = 3 + (3)(2)$, we want to find the chance that the observation is three standard deviations above the mean. (See Figure 12.40.) Now, 99% of the population is within three standard deviations of the mean; therefore, 1% is outside of the interval $[\mu - 3\sigma, \mu + 3\sigma]$. Because the density function of the normal distribution is symmetric about the mean, the area to the left of $\mu - 3\sigma$ and the area to the right of $\mu + 3\sigma$ are the same. Hence, there is a $(1\%)/2 = 0.5\%$ chance that the observation is above 9. We can therefore write

$$P(X > 9) = 0.005$$

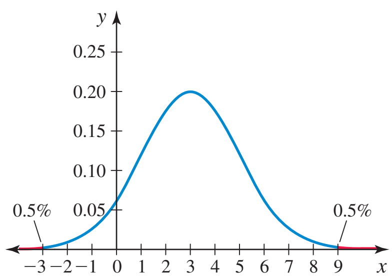


Figure 12.40 The normal density with mean 3 and standard deviation 2 in Example 6; 0.5% of the observations are greater than 9.

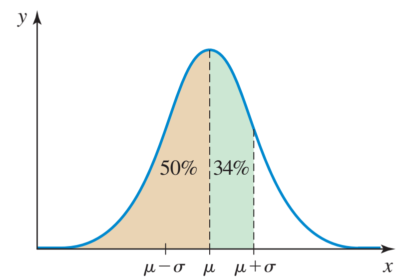


Figure 12.41 The area for Example 7.

EXAMPLE 7 Assume that a certain quantitative character X is normally distributed with parameters μ and σ. What is the probability that an observation lies below $\mu + \sigma$?

Solution Since 68% of the population falls within one standard deviation of the mean, and since the density curve is symmetric about the mean, it follows that 34% of the population falls into the interval $[\mu, \mu + \sigma]$. Furthermore, because of the symmetry of the density function, 50% of the population is below the mean. Hence, $50\% + 34\% = 84\%$ of the population lie below $\mu + \sigma$, as illustrated in Figure 12.41. We can therefore write

$$P(X < \mu + \sigma) = 0.84$$

Using the Table to Find Probabilities. The table for a normal distribution with mean 0 and standard deviation 1 (see Appendix B) can be used to compute probabilities when the distribution is normal with mean μ and standard deviation σ.

We begin by explaining how to use the table for the normal distribution with mean 0 and standard deviation 1, called the **standard normal distribution**, whose density is given by

$$f(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \qquad \text{for } -\infty < u < \infty$$

The table lists values for

$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$$

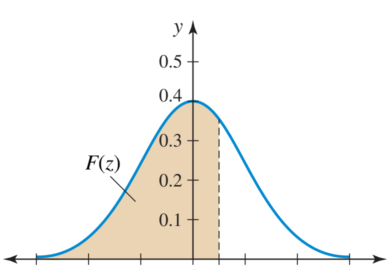


Figure 12.42 The area to the left of $x = z$ under the graph of the normal density $f(z)$, which is listed in the table for the normal distribution.

Geometrically, $F(z)$ is the area to the left of the line $x = z$ under the graph of the density function, as illustrated in Figure 12.42.

We interpret $F(z)$ as the probability that an observation is to the left of z. For instance, when $z = 1$, $F(1) = 0.8413$, and we say that the probability that an observation has a value less than or equal to 1 is 0.8413. In other words, 84.13% of the population has a value less than or equal to 1.

As you can see, the table does not provide entries for negative values of z. To compute such values, we take advantage of the symmetries of the density function.

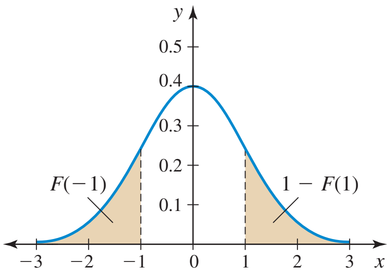


Figure 12.43 The area of the shaded region $F(-1)$ is the same as the area of the shaded region $1 - F(1)$.

For instance, we see from the graph of the function that the area to the left of -1 is the same as the area to the right of 1. (See Figure 12.43.) Thus, if we wish to compute $F(-1)$, we write

$$\begin{aligned} F(-1) &= \int_{-\infty}^{-1} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du = \int_{1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du - \int_{-\infty}^{1} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du \\ &= 1 - F(1) = 1 - 0.8413 = 0.1587 \end{aligned}$$

Here, we used the fact that the total area is equal to 1.

We can use the table to compute areas under the graph of a normal density with arbitrary mean μ and standard deviation σ. Thus, to find the value of

$$\int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}\, dx$$

where $-\infty < a < b < \infty$, we use the substitution

$$u = \frac{x-\mu}{\sigma} \quad \text{with} \quad \frac{du}{dx} = \frac{1}{\sigma}$$

which yields

$$\int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}\, dx = \int_{(a-\mu)/\sigma}^{(b-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du = F\left(\frac{b-\mu}{\sigma}\right) - F\left(\frac{a-\mu}{\sigma}\right).$$

We recognize the right-hand side as the area under the *standard normal density* between $(a-\mu)/\sigma$ and $(b-\mu)/\sigma$. Therefore, the area under the normal density with mean μ and standard deviation σ between a and b is the same as the area under the standard normal density between $(a-\mu)/\sigma$ and $(b-\mu)/\sigma$. (See Figure 12.44.) We illustrate this equality in the next example.

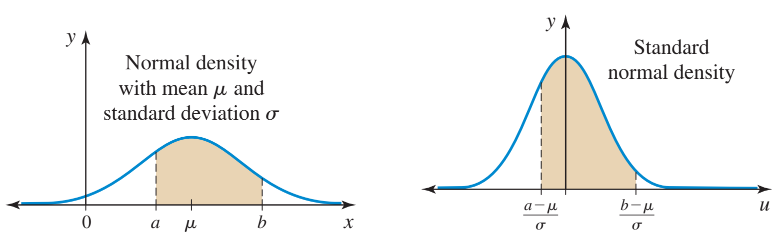


Figure 12.44 The area under the normal density with mean μ and standard deviation σ between a and b is the same as the area under the standard normal density between $(a-\mu)/\sigma$ and $(b-\mu)/\sigma$.

EXAMPLE 8 Suppose that a quantity X is normally distributed with mean 3 and standard deviation 2. Find the fraction of the population that falls into the interval $[2, 5]$; that is, find $P(X \in [2, 5])$.

Solution To solve this problem, we must compute

$$\int_2^5 \frac{1}{2\sqrt{2\pi}} e^{-(x-3)^2/8}\, dx \tag{12.31}$$

Using the transformation $u = (x-3)/2$, we find that when

$$x = 2, \quad u = \frac{2-3}{2} = -\frac{1}{2}$$

and when

$$x = 5, \quad u = \frac{5-3}{2} = 1$$

Therefore, the area under the normal density with mean 3 and standard deviation 2 between 2 and 5 is the same as the area under the standard normal density between $-1/2$ and 1. Hence, the integral in (12.31) is equal to

$$\begin{aligned}\int_{-1/2}^{1} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\,du &= \int_{-\infty}^{1} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\,du - \int_{-\infty}^{-1/2} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\,du \\ &= F(1) - F\left(-\frac{1}{2}\right) = F(1) - \left(1 - F\left(\frac{1}{2}\right)\right) \\ &= F(1) + F\left(\frac{1}{2}\right) - 1 = 0.8413 + 0.6915 - 1 = 0.5328\end{aligned}$$

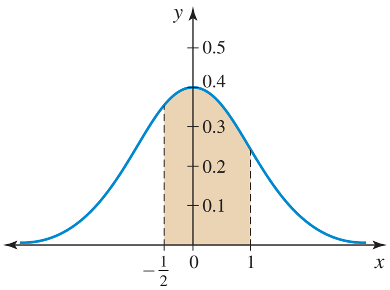


Figure 12.45 The area in Example 8.

and it follows that $P(X \in [2, 5]) = 0.5328$.

Instead of writing out these integrals, it is easier to determine what we need to compute when we sketch the relevant area under the standard normal curve. From Figure 12.45, we see that we need to compute $F(1) - F(-\frac{1}{2})$. Since $F(-\frac{1}{2}) = 1 - F(\frac{1}{2})$, we need to find $F(1) - 1 + F(\frac{1}{2})$, which we computed previously. •

EXAMPLE 9 Let X be normally distributed with mean 3 and variance 4. Find $P(1 \le X \le 6)$.

Solution We apply the transformation $z = (x - \mu)/\sigma$ with $\mu = 3$ and $\sigma = 2$ and denote a standard normally distributed random variable by Z:

$$\begin{aligned}P(1 \le X \le 6) &= P\left(\frac{1-3}{2} \le \frac{X - \mu}{\sigma} \le \frac{6-3}{2}\right) \\ &= P\left(-1 \le Z \le \frac{3}{2}\right) = F\left(\frac{3}{2}\right) - F(-1) \\ &= F(1.5) - (1 - F(1)) = 0.9332 - 1 + 0.8413 = 0.7745\end{aligned}$$ •

EXAMPLE 10 Suppose that a quantitative character X is normally distributed with mean 2 and standard deviation 1/2. Find x such that 30% of the population is above x.

Solution We need to find x such that $P(X > x) = 0.3$. Using the transformation $z = (x - \mu)/\sigma$ and letting Z denote a quantity that is normally distributed with mean 0 and standard deviation 1, we obtain

$$\begin{aligned}P(X > x) &= P\left(\frac{X - \mu}{\sigma} > \frac{x - \mu}{\sigma}\right) = P\left(Z > \frac{x-2}{1/2}\right) \\ &= P(Z > 2(x - 2)) = 0.3\end{aligned}$$

Now, $P(Z > z) = 1 - F(z) = 0.3$; hence, $F(z) = 0.7$. We then find from Appendix B that $F(0.52) = 0.6985 \approx 0.7$. Thus,

$$2(x - 2) = 0.52, \qquad \text{or} \qquad x = 2.26$$

That is, $P(X > 2.26) = 0.3$. Therefore, the value of x that we are seeking is 2.26. •

A Note on Samples. To obtain information about a quantity, such as size or bristle number, we cannot survey an entire population. Instead, we take a sample from the population and find the distribution of the quantity of interest in the sample. We must choose the sample so that it is representative of the population; this is a difficult problem, which we cannot discuss here. Even if we assume that each sample is representative of the population, two different samples will not be identical.

EXAMPLE 11 The numbers in the following table, representing two samples, each from the same population, show values of a quantity that is normally distributed with mean $\mu = 0$ and standard deviation $\sigma = 1$:

Sample 1				
−1.633	0.542	0.250	−0.166	0.032
1.114	0.882	1.265	−0.202	0.151
1.151	−1.210	−0.927	0.425	0.290
−1.939	0.891	−0.227	0.602	0.873
0.385	−0.649	−0.577	0.237	−0.289

Sample 2				
−0.157	0.693	1.710	0.800	−0.265
1.492	−0.713	0.821	−0.031	−0.780
−0.042	1.615	−1.440	−0.989	−0.580
0.289	−0.904	0.259	−0.600	−1.635
0.721	−1.117	0.635	0.592	−1.362

Both samples are obtained from a table of random numbers that are normally distributed with mean 0 and standard deviation 1 (***Beyer, 1991***).

(a) Count the number of observations in each sample that fall below the mean $\mu = 0$, and compare the number with what you would expect on the basis of properties of the normal distribution.

(b) Count the number of observations in each sample that fall within one standard deviation of the mean, and compare the number with what you would expect on the basis of properties of the normal distribution.

Solution **(a)** Since the mean is equal to 0, to find the number of observations that are below the mean, we simply count the number of observations that are negative. In the first sample, 10 observations are below the mean; in the second sample, 14 are. We expect that half of the sample points are below the mean. Because each sample is of size 25, we expect about 12 or 13 sample points to be below the mean.

(b) Since the standard deviation is 1, we count the number of observations that fall into the interval $[-1, 1]$. In the first sample, there are 19 such observations; in the second sample, there are 18. To compare these with the theoretical value, note that 68% of the population falls within one standard deviation of the mean. Since the sample size is equal to 25, and since $(0.68)(25) = 17$, we expect about 17 observations to fall into the interval $[-1, 1]$. •

The preceding example illustrates an important point: Even if random samples are taken from the same population, they are not identical. For instance, in the preceding example we expect that half of the observations will be below the mean. In the first sample fewer than half of the observations were below the mean, whereas in the second sample more than half of the observations were below the mean.

As the sample size increases, however, the sample will reflect the population increasingly more faithfully. That is, in order to determine the distribution of a quantitative character, such as the number of bristles in *D. melanogaster*, you would take a sample and find, for instance, the histogram associated with the quantity of interest. Then, if the sample is large enough, the histogram will reflect the population distribution quite well. But if you repeat the experiment, you should not expect the two histograms to be exactly the same. If the sample size is large enough, however, they will be close.

The importance of the normal distribution cannot be overstated. A large part of statistics is based on the assumption that observed quantities are normally distributed. You will probably ask why we can assume the normal distribution in the first place. The reason for this is quite deep, and we will discuss some of it in the next section. At this point, we wish to just give you the gist of it.

Many quantities can be thought of as a sum of a large number of small contributions. We can show that the distribution of any sum of independent random variables that all have the same distribution with a finite mean and a finite variance converges to a normally distributed random variable when the number of terms in the sum increases. This result is known as the **central limit theorem**. (See Section 12.6.)

The central limit theorem governs traits in quantitative genetics. Many quantitative traits (such as the height or birth weight of an organism) are thought of as resulting from numerous genetic and environmental factors that all act in an additive or multiplicative way. If these factors do in fact act in an additive way and are independent, the central limit theorem can be applied directly and the distribution of values of the trait will resemble a normal distribution. If the factors act in a multiplicative way, then a logarithmic transformation reduces this case to the additive one.

The same reasoning is used when we consider measurement errors. A measurement error is frequently thought of as a sum of a large number of independent contributions from different sources that act additively. This model for measurement errors is often supported in real experiments, so such errors are often assumed to be normally distributed.

12.5.3 The Uniform Distribution

The uniform distribution is in some ways the simplest continuous distribution. We say that a random variable U is **uniformly distributed** over the interval (a, b) if its density function is given by

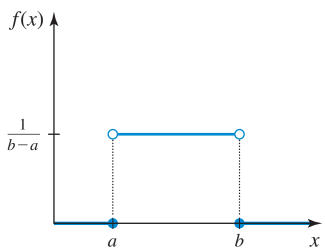


Figure 12.46 The density function of a uniformly distributed random variable over the interval (a, b).

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

as illustrated in Figure 12.46.

The reason for the term *uniform* can be seen when we compute the probability that the random variable U falls into the interval $(x_1, x_2) \subset (a, b)$. To compute probabilities of events for a uniformly distributed random variable, we compute the area of a rectangle. We therefore use the simple geometric formula asserting that the area is equal to the product of height and length, instead of formally integrating the density function. (See Figure 12.47.) When we do, we find that

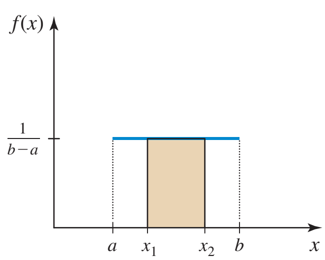


Figure 12.47 The probability $P(U \in (x_1, x_2))$ is equal to the area of the shaded region.

$$P(U \in (x_1, x_2)) = \begin{bmatrix} \text{area under } f(x) = \frac{1}{b-a} \\ \text{between } x_1 \text{ and } x_2 \end{bmatrix} = \frac{x_2 - x_1}{b - a}$$

We see that this probability depends only on the length of the interval (x_1, x_2) relative to the length of the interval (a, b), and *not* on the location of (x_1, x_2), provided that (x_1, x_2) is a subset of (a, b). Therefore, all intervals of equal lengths that are contained in (a, b) have equal chances of containing U.

To find the mean of a uniformly distributed random variable, we evaluate

$$\begin{aligned} E(U) &= \int_{-\infty}^{\infty} x f(x)\,dx = \int_a^b x \cdot \frac{1}{b-a}\,dx = \frac{1}{b-a}\left[\frac{1}{2}x^2\right]_a^b \\ &= \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{a+b}{2} \end{aligned}$$

The last identity comes about because $b^2 - a^2 = (b - a)(b + a)$. The mean of a uniformly distributed random variable over the interval (a, b) is therefore the midpoint of the interval (a, b). To find the variance, we first compute

$$\begin{aligned} E(U^2) &= \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_a^b x^2 \cdot \frac{1}{b-a}\,dx = \frac{1}{b-a}\left[\frac{1}{3}x^3\right]_a^b \\ &= \frac{1}{b-a} \cdot \frac{b^3 - a^3}{3} = \frac{b^2 + ab + a^2}{3} \end{aligned}$$

The last identity comes about because $b^3 - a^3$ factors into $(b-a)(b^2+ab+a^2)$. Then,

$$\begin{aligned}\text{var}(U) = E(U^2) - [E(U)]^2 &= \frac{b^2+ab+a^2}{3} - \frac{(a+b)^2}{4} \\ &= \frac{4b^2+4ab+4a^2-3a^2-6ab-3b^2}{12} = \frac{b^2-2ab+a^2}{12} \\ &= \frac{(b-a)^2}{12}\end{aligned}$$

Uniform distributions are frequently used in computer simulations of random experiments, as in the next example.

EXAMPLE 12 Suppose that you wish to simulate on the computer a random experiment that consists of tossing a coin five times with probability 0.3 of heads. The software that you want to use can generate uniformly distributed random variables in the interval (0, 1). How do you proceed?

Solution Each trial consists of flipping the coin and then recording the outcome. To simulate a coin flip with a probability of heads of 0.3, we draw a uniformly distributed random variable U from the interval (0, 1). The computer then returns a number u in the interval (0, 1). Since $P(U \leq 0.3) = 0.3$, if

$$u = \begin{cases} \leq 0.3, & \text{we record heads} \\ > 0.3, & \text{we record tails} \end{cases}$$

We repeat this experiment five times.

To be concrete, suppose that successive values of the uniform random variable are 0.2859, 0.9233, 0.5187, 0.8124, and 0.0913. These numbers are then translated into $HTTTH$, where H stands for heads and T for tails. •

In the next example, we compute the distribution function of a uniformly distributed random variable.

EXAMPLE 13 **(a)** Find the density and distribution functions of a uniformly distributed random variable on the interval (1, 5), and graph both in the same coordinate system.

(b) Suppose that we draw a uniformly distributed random variable from the interval (1, 5). Compute the probability that the first digit after the decimal point is a 2.

Solution **(a)** Since the length of the interval (1, 5) is 4, the density function is given by

$$f(x) = \begin{cases} \frac{1}{4} & \text{for } 1 < x < 5 \\ 0 & \text{otherwise} \end{cases}$$

The distribution function $F(x) = P(X \leq x)$ is given by

$$F(x) = \int_{-\infty}^{x} f(u)\,du = \begin{cases} 0 & \text{for } x \leq 1 \\ \int_1^x \frac{1}{4}\,du = \frac{1}{4}x - \frac{1}{4} & \text{for } 1 < x < 5 \\ 1 & \text{for } x \geq 5 \end{cases}$$

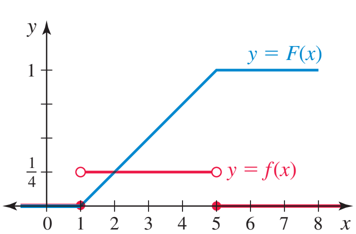


Figure 12.48 The density function $f(x)$ and the distribution function $F(x)$ of a uniformly distributed random variable over the interval (1, 5). (See Example 13.)

Graphs of $f(x)$ and $F(x)$ are shown in Figure 12.48.

(b) The event that the first digit after the decimal point is a 2 is the event that the random variable U falls into the set

$$A = [1.2, 1.3) \cup [2.2, 2.3) \cup [3.2, 3.3) \cup [4.2, 4.3)$$

Therefore,

$$P(U \in A) = (4)(0.1)\left(\frac{1}{4}\right) = 0.1 \qquad P(U \in [1 \cdot 2, 1 \cdot 3)) = (0.1)\left(\tfrac{1}{4}\right)$$

as illustrated in Figure 12.49. •

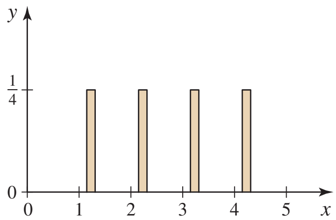


Figure 12.49 The probability $P(U \in A)$ in Example 13 is equal to the sum of the areas of the shaded regions.

12.5.4 The Exponential Distribution

We give the density function of the exponential distribution first and then explain where this distribution plays a role. We say that a random variable X is **exponentially distributed** with parameter $\lambda > 0$ if its density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x > 0 \\ 0 & \text{for } x \le 0 \end{cases}$$

Since

$$\int_0^x \lambda e^{-\lambda u}\,du = -e^{-\lambda u}\Big|_0^x = 1 - e^{-\lambda x} \qquad \frac{d}{dx}(-e^{-\lambda x}) = \lambda e^{-\lambda x}$$

its distribution function $F(x) = P(X \le x)$ is given by

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{for } x > 0 \\ 0 & \text{for } x \le 0 \end{cases}$$

The expected value of X is

$$E(X) = \int_0^\infty x\lambda e^{-\lambda x}\,dx = \frac{1}{\lambda}$$

and the variance of X is

$$\text{var}(X) = \int_0^\infty \left(x - \frac{1}{\lambda}\right)^2 \lambda e^{-\lambda x}\,dx = \frac{1}{\lambda^2}.$$

(The mean and the variance will be calculated in Problems 47 and 48.) The exponential distribution is frequently used to model lifetimes and waiting times (i.e., the time between events, such as a pollinator visiting a flower).

EXAMPLE 14 Suppose that the time between arrivals of insect pollinators to a flowering plant is exponentially distributed with parameter $\lambda = 0.3$/hr.

(a) Find the mean and the standard deviation of the waiting time between successive pollinator arrivals.

(b) If a pollinator just left the plant, what is the probability that you will have to wait for more than three hours before the next pollinator arrives?

Solution **(a)** Since $\lambda = 0.3$/hr, the mean is equal to $1/\lambda = 10/3$ hours and the standard deviation, which is the square root of the variance, is equal to $1/\lambda = 10/3$ hours.

(b) If we denote the waiting time by T, then

$$P(T > 3) = 1 - F(3) = e^{-(0.3)(3)} \approx 0.4066$$

•

EXAMPLE 15 Suppose that the lifetime of an organism is exponentially distributed with parameter $\lambda = (1/200)\,\text{yr}^{-1}$.

(a) Find the probability that the organism will live for more than 50 years.

(b) Given that the organism is 100 years old, find the probability that it will live for at least another 50 years.

Solution We denote the lifetime of the organism by T, measured in units of years. Then T is exponentially distributed with parameter $\lambda = (1/200)\,\text{yr}^{-1}$.

(a) We want to find the probability that T exceeds 50 years. We compute

$$P(T > 50) = 1 - P(T \le 50) = 1 - (1 - e^{-50/200})$$
$$= e^{-50/200} = e^{-1/4} \approx 0.7788$$

Note that the units in the exponent canceled out.

(b) We want to find $P(T > 150 \mid T > 100)$. This is a conditional probability. We evaluate it in the following way:

$$P(T > 150 \mid T > 100) = \frac{P(T > 150 \text{ and } T > 100)}{P(T > 100)}$$

Since $\{T > 150\} \subset \{T > 100\}$, it follows that

$$\{T > 150\} \cap \{T > 100\} = \{T > 150\}$$

Therefore,

$$P(T > 150 \text{ and } T > 100) = P(T > 150)$$

We can now continue to evaluate

$$\frac{P(T > 150 \text{ and } T > 100)}{P(T > 100)} = \frac{P(T > 150)}{P(T > 100)} = \frac{e^{-150/200}}{e^{-100/200}}$$
$$= e^{-3/4+1/2} = e^{-1/4} \approx 0.7788$$

This is the same answer as that in (a). The fact that the organism has lived for 100 years does not change its probability of living for another 50 years. We say that the organism does not age. This nonaging property is a characteristic feature of the exponential distribution. Of course, most organisms do age. Nevertheless, because of its mathematical simplicity the exponential distribution is still frequently used to model lifetimes—even if the organism ages, in which case the distribution should be considered as an approximation of the real situation. •

Let us examine the nonaging property in more detail. If T is an exponentially distributed lifetime, then we claim that

$$P(T > t + h \mid T > t) = P(T > h)$$

In words, if the organism is still alive after t units of time, then the probability that it will live for at least another h units of time is the same as the probability that the organism survived the first h units of time. This implies that death does not become more (or less) likely with age.

The nonaging property follows immediately from the calculation

$$P(T > t + h \mid T > t) = \frac{P(T > t + h \text{ and } T > t)}{P(T > t)}$$
$$= \frac{P(T > t + h)}{P(T > t)} = \frac{e^{-\lambda(t+h)}}{e^{-\lambda t}}$$
$$= e^{-\lambda h} = P(T > h)$$

which is the same as the one that we carried out in Example 15. Note that, for many organisms or objects, the exponential distribution is a poor lifetime model. However, as the next example illustrates, it is the correct distribution to model radioactive decay.

EXAMPLE 16 **Radioactive Decay** Assume that the lifetime of a radioactive atom is exponentially distributed with parameter $\lambda = 3/\text{days}$.

(a) Find the average lifetime of this atom.

(b) Find the time T_h such that the probability that the atom will not have decayed at time t is equal to $1/2$. (The time T_h is called the *half-life*.)

Solution **(a)** The average lifetime is $1/\lambda = (1/3)$ days.

(b) If the random variable T denotes the lifetime of the atom, then T_h satisfies

$$P(T > T_h) = \frac{1}{2}$$

That is,

$$e^{-\lambda T_h} = \frac{1}{2} \quad \text{or} \quad T_h = \frac{\ln 2}{\lambda}$$

With $\lambda = 3/\text{days}$, we find that

$$T_h = \frac{\ln 2}{3}\,\text{days} \approx 0.2310\,\text{days}$$

EXAMPLE 17 **Seed Dispersal** In Example 4, we used the exponential function

$$f(r) = \begin{cases} \lambda e^{-\lambda r} & \text{for } r > 0 \\ 0 & \text{for } r \le 0 \end{cases}$$

where $\lambda > 0$ is a constant, to model seed dispersal. The function $f(r)$ is a density function, and, for $0 < a < b$, $\int_a^b f(r)\,dr$ describes the fraction of seeds dispersed between distances a and b from the source at 0. We recognize this function as the density function of an exponentially distributed random variable with parameter λ.

(a) Show that $f(r)$ is a density function.

(b) Show that the fraction of seeds that are dispersed a distance R or more declines exponentially with R.

(c) Find R such that 60% of the seeds are dispersed within distance R of the source. How does R depend on λ?

Solution **(a)** To show that $f(r)$ is a density function, we need to show that $f(r) \ge 0$ for all $r \in \mathbf{R}$ and that $\int_{-\infty}^{\infty} f(r)\,dr = 1$. Since $\lambda > 0$ and $e^{-\lambda r} > 0$, it follows immediately that $f(r) \ge 0$ for $r > 0$. Combining this result with $f(r) = 0$ for $r \le 0$, we find that $f(r) \ge 0$ for all $r \in \mathbf{R}$. To check the second criterion, we need to carry out the integration. Since the function $f(r)$ is a piecewise-defined function, we need to split the integral into two parts:

$$\begin{aligned}\int_{-\infty}^{\infty} f(r)\,dr &= \int_{-\infty}^{0} f(r)\,dr + \int_0^{\infty} f(r)\,dr \\ &= \int_{-\infty}^{0} 0\,dr + \int_0^{\infty} \lambda e^{-\lambda r}\,dr\end{aligned}$$

The term $\int_{-\infty}^{0} 0\,dr$ is equal to 0. An antiderivative of $\lambda e^{-\lambda r}$ is $-e^{-\lambda r}$. Hence,

$$\begin{aligned}\int_0^{\infty} \lambda e^{-\lambda r} dr &= \lim_{z\to\infty} \int_0^{z} \lambda e^{-\lambda r}\,dr = \lim_{z\to\infty} \left[-e^{-\lambda r}\right]_0^z \\ &= \lim_{z\to\infty} \left[-e^{-\lambda z} - (-1)\right] = 1\end{aligned}$$

since $\lim_{z\to\infty} e^{-\lambda z} = 0$.

(b) For $R > 0$, let $G(R)$ denote the fraction of seeds that are dispersed a distance R or more. Then

$$\begin{aligned}G(R) &= \int_R^{\infty} f(r)\,dr = \int_R^{\infty} \lambda e^{-\lambda r}\,dr = \lim_{z\to\infty} \int_R^{z} \lambda e^{-\lambda r}\,dr \\ &= \lim_{z\to\infty} \left[-e^{-\lambda r}\right]_R^z = \lim_{z\to\infty} (-e^{-\lambda z} + e^{-\lambda R}) = e^{-\lambda R}\end{aligned}$$

This result shows that $G(R)$ declines exponentially with R.

(c) The number R satisfies

$$0.6 = \int_0^{R} \lambda e^{-\lambda r}\,dr$$

Carrying out the integration, we obtain

$$0.6 = \left[-e^{-\lambda r}\right]_0^R = 1 - e^{-\lambda R}$$

To find R, we need to solve

$$e^{-\lambda R} = 0.4$$
$$-\lambda R = \ln 0.4$$
$$R = -\frac{\ln 0.4}{\lambda} = \frac{1}{\lambda} \ln \frac{5}{2}$$

where, in the last step, we used the fact that $-\ln 0.4 = \ln \frac{1}{0.4} = \ln \frac{5}{2}$. We see that $R \propto 1/\lambda$ (i.e., the bigger λ, the smaller R is), which means that seeds tend to be dispersed more closely to the source for larger values of λ. •

EXAMPLE 18 Suppose that you wish to use a computer to generate exponentially distributed random variables, but the computer's software can generate only uniformly distributed random variables in the interval $(0, 1)$. How do you proceed?

Solution The key ingredient to solving this problem is the following result: If X is a continuous random variable with a strictly increasing distribution function $F(x)$, then $F(X)$ is uniformly distributed in the interval $(0, 1)$. To prove this result, we need to show that

$$P(F(X) \leq u) = u \quad \text{for } 0 < u < 1$$

Now, the event $\{F(X) \leq u\}$ is equivalent to the event $\{X \leq F^{-1}(u)\}$, where $F^{-1}(\cdot)$ is the inverse function of $F(\cdot)$. That the inverse function exists follows from the assumption that the distribution function is strictly increasing, which implies that $F(x)$ is one to one. Therefore,

$$P(F(X) \leq u) = P(X \leq F^{-1}(u))$$

Since $F(x) = P(X \leq x)$, we have

$$P(X \leq F^{-1}(u)) = F(F^{-1}(u)) = u$$

where the last identity follows from the properties of the inverse function.

How can we use the preceding result to generate exponentially distributed random variables from uniformly distributed random variables? The computer generates a uniformly distributed random variable U, which we interpret as $F(X)$, where X is distributed according to the distribution function $F(x)$. Since

$$U = F(X) \quad \text{is equivalent to} \quad X = F^{-1}(U)$$

we need to find the inverse function of $F(x)$ and compute $F^{-1}(U)$. The result of our computation is then the random variable X.

In the case of an exponential distribution with parameter λ, the distribution function $F(x)$ is

$$F(x) = 1 - e^{-\lambda x} \quad \text{for } x \geq 0$$

which is strictly increasing on $[0, \infty)$. Set $u = 1 - e^{-\lambda x}$ and solve for x:

$$1 - u = e^{-\lambda x}$$
$$-\frac{1}{\lambda} \ln(1 - u) = x$$

Therefore,

$$F^{-1}(u) = -\frac{1}{\lambda} \ln(1 - u)$$

To be concrete, assume that $\lambda = 2$ and that a computer generated the following three random variables, distributed uniformly in $(0, 1)$:

$$u_1 = 0.8890, \quad u_2 = 0.9394, \quad u_3 = 0.3586$$

Then

$$x_1 = F^{-1}(u_1) = -\frac{1}{2}\ln(1 - 0.8890) \approx 1.099$$

$$x_2 = F^{-1}(u_2) = -\frac{1}{2}\ln(1 - 0.9394) \approx 1.402$$

$$x_3 = F^{-1}(u_3) = -\frac{1}{2}\ln(1 - 0.3586) \approx 0.2221$$

are the corresponding realizations of the exponentially distributed random variable.

12.5.5 The Poisson Process

In Example 14, we modeled the time between arrivals of insect pollinators to a flowering plant. Our model was an exponential distribution with parameter λ. Suppose we start observing at time 0 and insects arrive at times $T_1, T_2, T_3, \ldots$. The interarrival times $T_1 - 0, T_2 - T_1, T_3 - T_2, \ldots$ are assumed to be independent and exponentially distributed with parameter λ. One can show that the time T_n of the nth arrival is a continous random variable whose distribution is given by the density function

$$f_{n,\lambda}(x) = \frac{\lambda^n x^{n-1}}{(n-1)!}e^{-\lambda x} \quad \text{for } x \geq 0$$

This density function is called the *gamma* (n, λ) *density function*. We will not derive it. Note, however, that if $n = 1$, then

$$f_{1,\lambda} = \lambda e^{-\lambda x} \quad \text{for } x > 0$$

which is indeed the density of an exponential distribution.

Instead of asking when the nth insect pollinator arrived, we can count the number of arrivals up to time t, which we denote by $N(t)$. Now, $N(t)$ is a discrete random variable that takes on values $0, 1, 2, \ldots$.

To calculate the probability mass function for $N(t)$, we begin with the event $\{N(t) = 0\}$. Note that $\{N(t) = 0\}$ is equivalent to $\{T_1 > t\}$. Since T_1 is exponentially distributed with parameter λ, it follows that

$$P(N(t) = 0) = P(T_1 > t) = e^{-\lambda t} \tag{12.32}$$

Generalizing this argument to arbitrary n, we find that the event $\{N(t) < n\}$ is equivalent to $\{T_n > t\}$. The event $\{T_n > t\}$ can be calculated with the use of the gamma (n, λ) density function

$$P(T_n > t) = \int_t^\infty f_{n,\lambda}(x)\,dx = \int_t^\infty \frac{\lambda^n x^{n-1}}{(n-1)!}e^{-\lambda x}\,dx \tag{12.33}$$

Using integration by parts, with $u = \frac{x^{n-1}}{(n-1)!}$ and $v' = \lambda^n e^{-\lambda x}$, on the right-hand side of (12.33), we obtain

$$\begin{aligned} P(T_n > t) &= -\frac{x^{n-1}}{(n-1)!}\,\lambda^{n-1}e^{-\lambda x}\Big]_t^\infty + \int_t^\infty \frac{x^{n-2}}{(n-2)!}\lambda^{n-1}e^{-\lambda x}\,dx \\ &= \frac{\lambda^{n-1}t^{n-1}}{(n-1)!}e^{-\lambda t} + P(T_{n-1} > t) \end{aligned}$$

Using (12.32), we can calculate

$$\begin{aligned} P(N(t) < 2) &= P(T_2 > t) = \lambda t e^{-\lambda t} + e^{-\lambda t} \\ P(N(t) < 3) &= P(T_3 > t) = \frac{(\lambda t)^2}{2!}e^{-\lambda t} + \lambda t e^{-\lambda t} + e^{-\lambda t} \\ &\vdots \\ P(N(t) < n) &= P(T_n > t) = \sum_{k=0}^{n-1} e^{-\lambda t}\frac{(\lambda t)^k}{k!} \end{aligned}$$

The last equation shows that $N(t)$ is Poisson distributed with parameter λt. We say that $N(t)$ is a **Poisson process** with rate λ and summarize as follows:

> If the interarrival times are independent and exponentially distributed with parameter λ, then the number of arrivals up to time t is a Poisson process with rate λ and
> $$P(N(t) = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \quad \text{for } n = 0, 1, 2, \ldots$$

EXAMPLE 19 **Continuation of Example 14** Suppose that the time between arrivals of insect pollinators to a flowering plant is exponentially distributed with parameter $\lambda = 0.3$/hour. Find the probability of fewer than two arrivals within four hours of observation.

Solution If $N(t)$ denotes the number of arrivals within t hours of observation, then $N(t)$ is a Poisson process with rate $\lambda = 0.3$/hour. Hence,

$$\begin{aligned} P(N(4) < 2) &= e^{-(0.3)(4)}[1 + (0.3)(4)] \\ &= 2.2e^{-1.2} \approx 0.6626 \end{aligned}$$

It follows from the properties of the Poisson distribution that

$$E[N(t)] = \lambda t \qquad \text{and} \qquad \operatorname{var}[N(t)] = \lambda t$$

The Poisson process has three important properties:

1. Nonoverlapping intervals are independent. That is, if $s < t$, then $N(s)$ and $N(t) - N(s)$ are independent.
2. For a small time interval Δt, the probability of an arrival occurring in the interval $[t + \Delta t)$ is approximately proportional to the length of the interval Δt:
$$\lim_{\Delta t \to 0} \frac{P(N([t, t + \Delta t)) = 1)}{\Delta t} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} e^{-\lambda \Delta t} \lambda \Delta t = \lambda$$
3. For a small time interval of length Δt, the probability of more than one arrival is negligible:
$$\begin{aligned} \lim_{\Delta t \to 0} \frac{P(N([t, t + \Delta t)) > 1)}{\Delta t} &= \lim_{\Delta t \to 0} \frac{1 - e^{-\lambda \Delta t} - \lambda \Delta t e^{-\lambda \Delta t}}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{\lambda e^{-\lambda \Delta t} - \lambda e^{-\lambda \Delta t} + \lambda^2 \Delta t e^{-\lambda \Delta t}}{1} = 0 \end{aligned}$$
Note that we used l'Hôpital's rule in the penultimate step.

These three properties characterize the Poisson process and are used to derive its distribution.

12.5.6 Aging

Aging is a universal feature of both living organisms and mechanical devices; it is described by a progressive loss of vitality or reliability. A number of mathematical models are used to describe the phenomenon. The starting point for these models is the **survival function** $S(x)$, which is defined as the probability that the individual or device is still alive or functioning at age x. If X is the lifetime, then

$$S(x) = P(X > x) = 1 - P(X \le x) = 1 - F(x)$$

where $F(x)$ is the distribution function of X. We assume now that X is a nonnegative, continuous random variable with density function $f(x) > 0$ for $x > 0$. This assumption implies that $F(0) = 0$ and that $F(x)$ is strictly increasing for $x > 0$. The **failure-rate**, or **hazard-rate**, function $\lambda(x)$ is defined as the relative rate of decline of the survival function:

$$\lambda(x) = -\frac{1}{S(x)}\frac{dS}{dx}$$

Note that $\lambda(x)$ is positive for $x > 0$, since $S(x)$ is strictly decreasing for $x > 0$. Also, because

$$S(x) = P(X > x) = \int_x^\infty f(u)\,du$$

for $x > 0$, it follows that

$$\lambda(x) = -\frac{1}{S(x)}\frac{dS}{dx} = \frac{1}{P(X > x)} f(x)$$

Consequently, $\lambda(x)\,dx$ can be interpreted as the conditional probability of dying within the age interval $[x, x + dx)$, given that the individual is still alive at age $x > 0$. Mathematically,

$$\lambda(x)\,dx = P(X \in [x, x + dx)\,|\,X > x) \quad \text{for } x > 0$$

Non-aging. Following the preceding interpretation of $\lambda(x)\,dx$, we say that a system does not age if the failure rate $\lambda(x)$ is constant. In this case, for $x > 0$,

$$-\frac{1}{S(x)}\frac{dS}{dx} = \lambda = \text{constant}$$

Separating variables and integrating yields

$$\int \frac{dS}{S} = -\int \lambda\,dx$$

or

$$\ln|S(x)| = -\lambda x + C_1$$

Thus,

$$S(x) = C_2 e^{-\lambda x}$$

with $C_2 = \pm e^{C_1}$. Since X is a nonnegative continuous random variable with $S(0) = P(X > 0) = 1$, it follows that $C_2 = 1$. Therefore,

$$S(x) = e^{-\lambda x} = 1 - F(x) \qquad \text{or} \qquad F(x) = 1 - e^{-\lambda x} \tag{12.34}$$

We conclude that X is exponentially distributed with parameter λ. We showed earlier that the exponential distribution has the non-aging property. Equation (12.34) shows that the reverse holds as well: If a device has the non-aging property (i.e., a constant failure rate), then its lifetime distribution is exponential.

EXAMPLE 20 Suppose the hazard-rate function $\lambda(x) = 3/\text{year}$ for $x \geq 0$. Find the probability that an individual will die before age one year.

Solution If $\lambda(x) = 3/\text{year}$, then $S(x) = e^{-3x}$ for $x \geq 0$, where x is measured in years. If X denotes the lifetime of the individual, then

$$P(X \leq 1) = 1 - S(1) = 1 - e^{-3} \approx 0.9502$$

Thus, there is about a 95% chance that the individual will die before age one year. •

Aging. When the hazard-rate function increases with age, then an older individual has a higher probability of dying than a younger individual; and we say that the system is an **aging system**. Figure 12.50 shows an empirical hazard-rate function based on 8926 males from an inbred line of *Drosophila melanogaster* obtained in the lab of Professor Jim Curtsinger at the University of Minnesota. The smoothed line is a curve fitted to the data, which are based on daily measurements of survival. The horizontal axis lists

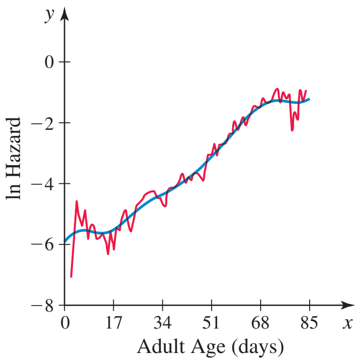


Figure 12.50 An empirical hazard-rate function (courtesy of Dr. Jim Curtsinger).

the age of individuals. The vertical axis lists $-\ln(N_{x+1}/N_x)$, where N_x is the number of adults alive at age x. This quantity can be interpreted as the hazard-rate function, averaged over the interval $[x, x+1]$. (See Problem 71.) We see from the graph that the hazard-rate function increases with age, but seems to level off at very old ages. This is the typical pattern one observes in these kinds of studies.

We again assume that the lifetime X is a nonnegative, continuous random variable with hazard-rate function $\lambda(x) > 0$ for $x > 0$ and survival function $S(x)$ for $x \geq 0$, with $S(0) = 1$.

Given the hazard-rate function $\lambda(x)$, we can find the survival function $S(x)$ by integration. For $x > 0$,

$$-\frac{1}{S(x)}\frac{dS}{dx} = \lambda(x)$$

$$\frac{dS}{S} = -\lambda(x)\,dx$$

$$\ln|S(x)| = -\int_0^x \lambda(u)\,du + C_1$$

and therefore,

$$S(x) = C\exp\left[-\int_0^x \lambda(u)\,du\right]$$

with $C = \pm e^{C_1}$. Since $S(0) = 1$, it follows that $C = 1$. Therefore,

$$\boxed{S(x) = \exp\left[-\int_0^x \lambda(u)\,du\right]}$$

The two most prominent hazard-rate functions that model aging are the **Gompertz law**, in which the function increases exponentially with age, and the **Weibull law**, in which the function increases according to a power law.

Gompertz Law:

$$\lambda(x) = A + Be^{\alpha x}, \quad x \geq 0$$

where A, B, and α are positive constants.

Weibull Law:

$$\lambda(x) = Cx^{\beta}, \quad x \geq 0$$

where C and β are positive parameters.

The Weibull law is often used to calculate the lifetimes of devices, whereas the Gompertz law is used to model the lifetimes of organisms. (See the review article by ***Gavrilov and Gavrilova, 2001***.)

EXAMPLE 21 Suppose the lifetime of an organism follows the Gompertz law with hazard-rate function

$$\lambda(x) = 1.5 + 0.3e^{0.1x}, \quad x \geq 0$$

where x is measured in years. Find the probability that the organism will live for more than one year.

Solution The survival function is given by

$$S(x) = \exp\left[-\int_0^x (1.5 + 0.3e^{0.1u})\,du\right], \quad x \geq 0$$

We evaluate the integral first. For $x \geq 0$,

$$\int_0^x (1.5 + 0.3e^{0.1u})\, du = 1.5u + \frac{0.3}{0.1}e^{0.1u}\Big]_0^x$$

$$= (1.5x + 3e^{0.1x}) - (0 + 3) = 1.5x + 3e^{0.1x} - 3$$

Therefore,

$$S(x) = \exp[-(1.5x + 3e^{0.1x} - 3)], \ x \geq 0$$

If X denotes the lifetime of the organism, then the probability that the organism will live for more than one year is

$$P(X > 1) = S(1) = \exp\left[-(1.5 + 3e^{0.1} - 3)\right] \approx 0.1628$$

The organism has about a 16% chance of living for more than one year. •

EXAMPLE 22 **Fruit Fly Lifetimes** Mortality data from *Drosophila melanogaster* were fitted to a Weibull law. It was found that the hazard-rate function

$$\lambda(x) = (3 \times 10^{-6})x^{2.5}, \quad x \geq 0$$

where x is measured in days, provided the best fit.

(a) Find the probability that an individual will die within the first 20 days.

(b) Find the age at which the probability of still being alive is 0.5.

Solution **(a)** The survival function is

$$S(x) = \exp\left[-\int_0^x (3 \times 10^{-6})u^{2.5}\, du\right], \ x \geq 0$$

We evaluate the integral first. For $x \geq 0$,

$$\int_0^x (3 \times 10^{-6})u^{2.5}\, du = (3 \times 10^{-6})\frac{1}{3.5}u^{3.5}\Big|_0^x$$

$$= (3 \times 10^{-6})\frac{1}{3.5}x^{3.5}$$

Thus, the probability that an individual will die within the first 20 days is

$$1 - S(20) = 1 - \exp\left[-(3 \times 10^{-6})\frac{1}{3.5}(20)^{3.5}\right] \approx 0.0302$$

(b) We need to find x such that

$$S(x) = 0.5$$

We solve for x:

$$\exp\left[-(3 \times 10^{-6})\frac{1}{3.5}x^{3.5}\right] = \frac{1}{2}$$

$$-(3 \times 10^{-6})\frac{1}{3.5}x^{3.5} = \ln\frac{1}{2}$$

$$x^{3.5} = \frac{(3.5)(\ln 2)}{3 \times 10^{-6}}$$

$$x = \left(\frac{(3.5)(\ln 2)}{3 \times 10^{-6}}\right)^{1/3.5}$$

$$x \approx 48.746$$

The age at which the probability of survival is 0.5 is about 48.7 days. •

Section 12.5 Problems

12.5.1

1. Show that

$$f(x) = \begin{cases} 3e^{-3x} & \text{for } x > 0 \\ 0 & \text{for } x \le 0 \end{cases}$$

is a density function. Find the corresponding distribution function.

2. Show that

$$f(x) = \begin{cases} \frac{1}{2} & \text{for } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

is a density function. Find the corresponding distribution function.

3. Determine c such that

$$f(x) = \frac{c}{1+x^2}, \quad x \in \mathbf{R}$$

is a density function.

4. Determine c such that

$$f(x) = \begin{cases} \frac{c}{x^2} & \text{for } x > 1 \\ 0 & \text{for } x \le 1 \end{cases}$$

is a density function.

5. Let X be a continuous random variable with density function

$$f(x) = \begin{cases} 2e^{-2x} & \text{for } x > 0 \\ 0 & \text{for } x \le 0 \end{cases}$$

Find $E(X)$ and $\text{var}(X)$.

6. Let X be a continuous random variable with density function

$$f(x) = \frac{1}{2}e^{-|x|}$$

for $x \in \mathbf{R}$. Find $E(X)$ and $\text{var}(X)$.

7. Let X be a continuous random variable with distribution function

$$F(x) = \begin{cases} 1 - \frac{1}{x^3} & \text{for } x > 1 \\ 0 & \text{for } x \le 1 \end{cases}$$

Find $E(X)$ and $\text{var}(X)$.

8. Let X be a continuous random variable with

$$P(X > x) = e^{-ax}, \quad x \ge 0$$

where a is a positive constant. Find $E(X)$ and $\text{var}(X)$.

9. Let X be a continuous random variable with density function

$$f(x) = \begin{cases} (a-1)x^{-a} & \text{for } x > 1 \\ 0 & \text{for } x \le 1 \end{cases}$$

(a) Show that $E(X) = \infty$ when $a \le 2$.

(b) Compute $E(X)$ when $a > 2$.

10. Suppose that X is a continuous random variable that takes on only nonnegative values. Set

$$G(x) = P(X > x)$$

(a) Show that

$$G'(x) = -f(x)$$

where $f(x)$ is the corresponding density function.

(b) Assume that

$$\lim_{x\to\infty} xG(x) = 0$$

and use integration by parts and (a) to show that

$$E(X) = \int_0^\infty G(x)\,dx \tag{12.35}$$

(c) Let X be a continuous random variable with

$$P(X > x) = e^{-ax}, \quad x > 0$$

where a is a positive constant. Use (12.35) to find $E(X)$. (If you did Problem 8, compare your answers.)

12.5.2

11. Denote by the density of a normal distribution with mean μ and standard deviation σ

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$$

for $-\infty < x < \infty$.

(a) Show that $f(x)$ is symmetric about $x = \mu$.

(b) Show that the maximum of $f(x)$ is at $x = \mu$.

(c) Show that the inflection points of $f(x)$ are at $x = \mu - \sigma$ and $x = \mu + \sigma$.

(d) Graph $f(x)$ for $\mu = 2$ and $\sigma = 1$.

12. Suppose that $f(x)$ is the density function of a normal distribution with mean μ and standard deviation σ. Show that

$$\mu = \int_{-\infty}^{\infty} x f(x)\,dx$$

is the mean of this distribution. (*Hint*: Use substitution.)

13. Suppose that a quantitative character is normally distributed with mean $\mu = 12.8$ and standard deviation $\sigma = 2.7$. Find an interval centered at the mean such that 95% of the population falls into the interval. Do the same for 99% of the population.

14. Suppose a quantitative character is normally distributed with mean $\mu = 15.4$ and standard deviation $\sigma = 3.1$. Find an interval centered at the mean such that 95% of the population falls into this interval. Do the same for 99% of the population.

In Problems 15–20, assume that a quantitative character is normally distributed with mean μ and standard deviation σ. Determine what fraction of the population falls into the given interval.

15. $[\mu, \infty)$

16. $[\mu - 2\sigma, \mu + \sigma]$

17. $(-\infty, \mu + 3\sigma]$

18. $[\mu + \sigma, \mu + 2\sigma]$

19. $(-\infty, \mu - 2\sigma]$

20. $[\mu - 3\sigma, \mu]$

21. Suppose that X is normally distributed with mean $\mu = 3$ and standard deviation $\sigma = 2$. Use the table in Appendix B to find the following:

(a) $P(X \le 4)$

(b) $P(2 \le X \le 4)$

(c) $P(X > 5)$

(d) $P(X \le 0)$

22. Suppose that X is normally distributed with mean $\mu = -1$ and standard deviation $\sigma = 1$. Use the table in Appendix B to find the following:

(a) $P(X > 0)$

(b) $P(0 < X < 1)$

(c) $P(-1.5 < X < 2.5)$

(d) $P(X > 1.5)$

23. Suppose that X is normally distributed with mean $\mu = 1$ and standard deviation $\sigma = 2$. Use the table in Appendix B to find x such that the following hold:

(a) $P(X \le x) = 0.9$ **(b)** $P(X > x) = 0.4$

(c) $P(X \le x) = 0.4$ **(d)** $P(|X - 1| < x) = 0.5$

24. Suppose that X is normally distributed with mean $\mu = -2$ and standard deviation $\sigma = 1$. Use the table in Appendix B to find x such that the following hold.

(a) $P(X \ge x) = 0.8$ **(b)** $P(X < 2x + 1) = 0.5$

(c) $P(X \le x) = 0.1$ **(d)** $P(|X - 2| > x) = 0.4$

25. Assume that the mathematics score X on the Scholastic Aptitude Test (SAT) is normally distributed with mean 500 and standard deviation 100.

(a) Find the probability that an individual's score exceeds 700.

(b) Find the math SAT score so that 10% of the students who took the test have that score or greater.

26. **Fruit Fly Bristles** In a study of *Drosophila melanogaster* by ***Mackey (1984)***, the number of bristles on the fifth abdominal sternite in males was shown to follow a normal distribution with mean 18.7 and standard deviation 2.1.

(a) What percentage of the male population has fewer than 17 abdominal bristles?

(b) Find an interval centered at the mean so that 90% of the male population have bristle numbers that fall into this interval.

27. Suppose the weight of an animal is normally distributed with mean 3720 g and standard deviation 527 g. What percentage of the population has a weight that exceeds 5000 g?

28. Suppose the height of an adult animal is normally distributed with mean 17.2 in. Find the standard deviation if 10% of the animals have a height that exceeds 19 in.

29. Suppose that X is normally distributed with mean 2 and standard deviation 1. Find $P(0 \le X \le 3)$.

30. Suppose that X is normally distributed with mean -1 and standard deviation 2. Find $P(-3.5 \le X \le 0.5)$.

31. Suppose that X is normally distributed with mean μ and standard deviation σ. Show that $E(X) = \mu$. [You may use the fact that if Z is standard normally distributed, then $E(Z) = 0$ and $\text{var}(X) = 1$.]

32. Suppose that X is normally distributed with mean μ and standard deviation σ. Show that $\text{var}(X) = \sigma^2$. [You may use the fact that if Z is standard normally distributed, then $E(Z) = 0$ and $\text{var}(X) = 1$.]

33. Suppose that X is standard normally distributed. Find $E(|X|)$.

34. **Seed Production** Suppose that the number of seeds a plant produces is normally distributed, with mean 142 and standard deviation 31. Find the probability that in a sample of five plants, at least one produces more than 200 seeds. Assume that the plants are independent.

35. The total maximum score on a calculus exam was 100 points. The mean score was 74 and the standard deviation was 11. Assume that the scores are normally distributed.

(a) Determine the percentage of students scoring 90 or above.

(b) Determine the percentage of students scoring between 60 and 80 (inclusive).

(c) Determine the minimum score of the highest 10% of the class.

(d) Determine the maximum score of the lowest 5% of the class.

36. The mean weight of goats on a farm is 123 lb, and the standard deviation is 9 lb. If the weights are normally distributed, determine what percentage of goats weigh **(a)** between 110 and 130 lb, **(b)** less than 100 lb, and **(c)** more than 150 lb.

12.5.3

37. Suppose that you pick a number at random from the interval $(0, 4)$. What is the probability that the first digit after the decimal point is a 3?

38. Suppose that you pick a number X at random from the interval $(0, a)$. If $P(X \ge 1) = 0.2$, find a.

39. Suppose that you pick a number X at random from the interval (a, b). If $E(X) = 4$ and $\text{var}(X) = 3$, find a and b.

40. Suppose that you pick five numbers at random from the interval $(0, 1)$. Assume that the numbers are independent. What is the probability that all numbers are greater than 0.7?

41. Suppose that X_1, X_2, and X_3 are independent and uniformly distributed over $(0, 1)$. Define

$$Y = \max(X_1, X_2, X_3)$$

Find $E(Y)$. [*Hint*: Compute $P(Y \le y)$, and use it to deduce the density of Y.]

42. Suppose that X_1, X_2, and X_3 are independent and uniformly distributed over $(0, 1)$. Define

$$Y = \min(X_1, X_2, X_3)$$

Find $E(Y)$. [*Hint*: Compute $P(Y > y)$, and use it to deduce the density of Y.]

43. Suppose that you wish to simulate a random experiment that consists of tossing a coin with probability 0.6 of heads 10 times. The computer generates the following 10 random variables: 0.1905, 0.4285, 0.9963, 0.1666, 0.2223, 0.6885, 0.0489, 0.3567, 0.0719, 0.8661. Find the corresponding sequence of heads and tails.

44. Suppose that you wish to simulate a random experiment that consists of rolling a fair die. The computer generates the following 10 random variables: 0.7198, 0.2759, 0.4108, 0.7780, 0.2149, 0.0348, 0.5673, 0.0014, 0.3249, 0.6630. Describe how you would find the corresponding sequence of numbers on the die, and find them.

45. Suppose $X_1, X_2, \ldots, X_n$ are independent random variables with uniform distribution on $(0, 1)$. Define $X = \min(X_1, X_2, \ldots, X_n)$.

(a) Compute $P(X > x)$.

(b) Show that $P(X > x/n) \to e^{-x}$ as $n \to \infty$.

46. Suppose $X_1, X_2, \ldots, X_n$ are independent random variables with uniform distribution on $(0, 1)$. Define $X = \max(X_1, X_2, \ldots, X_n)$.

(a) Find the distribution function of X.

(b) Use Problem 10 to compute $E(X)$.

12.5.4

47. Let X be exponentially distributed with parameter λ. Find $E(X)$.

48. Let X be exponentially distributed with parameter λ. Find $\text{var}(X)$.

49. Suppose that the lifetime of a battery is exponentially distributed with an average life span of three months. What is the probability that the battery will last for more than four months?

50. Suppose that the lifetime of a battery is exponentially distributed with an average life span of two months. You buy six batteries. What is the probability that none of them will last more than two months? (Assume that the batteries are independent.)

51. Suppose that the lifetime of a radioactive atom is exponentially distributed with an average life span of 27 days.

(a) Find the probability that the atom will not decay during the first 20 days after you start to observe it.

(b) Suppose that the atom does not decay during the first 20 days that you observe it. What is the probability that it will not decay during the next 20 days?

52. If X has distribution function $F(x)$, we can show that $F(X)$ is uniformly distributed over the interval $(0, 1)$. Use this fact to generate exponentially distributed random variables with mean 1. [Assume that a computer generated the following four uniformly distributed random variables on the interval $(0, 1)$: 0.0371, 0.5123, 0.1370, 0.9865.]

12.5.5

53. Suppose the number of customers per hour arriving at the post office is a Poisson process with an average of four customers per hour.

(a) Find the probability that no customer arrives between 2 and 3 P.M.

(b) Find the probability that exactly two customers arrive between 3 and 4 P.M.

(c) Assuming that the number of customers arriving between 2 and 3 P.M. is independent of the number of customers arriving between 3 and 4 P.M., find the probability that exactly two customers arrive between 2 and 4 P.M.

(d) Assume that the number of customers arriving between 2 and 3 P.M. is independent of the number of customers arriving between 3 and 4 P.M. Given that exactly two customers arrive between 2 and 4 P.M., what is the probability that both arrive between 3 and 4 P.M.?

54. Suppose the number of customers per hour arriving at the post office is a Poisson process with an average of five customers per hour.

(a) Find the probability that exactly one customer arrives between 2 and 3 P.M.

(b) Find the probability that exactly two customers arrive between 3 and 4 P.M.

(c) Assuming that the number of customers arriving between 2 and 3 P.M. is independent of the number of customers arriving between 3 and 4 P.M., find the probability that exactly three customers arrive between 2 and 4 P.M.

(d) Assume that the number of customers arriving between 2 and 3 P.M. is independent of the number of customers arriving between 3 and 4 P.M. Given that exactly three customers arrive between 2 and 4 P.M., what is the probability that one arrives between 2 and 3 P.M. and two between 3 and 4 P.M.?

55. You arrive at a bus stop at a random time. Assuming that busses arrive according to a Poisson process with rate 4/hr, what is the expected time to the next arrival?

56. Assume that $N(t)$ is a Poisson process with rate λ and T_1 is the time of the first arrival. Show that, for $s < t$,

$$P(T < s \mid N(t) = 1) = \frac{s}{t}$$

That is, show that, given that an arrival occurred in the interval $[0, t)$, the time of occurrence is uniform over the interval.

57. Suppose the lifetime of a lightbulb is exponentially distributed with mean 3 years. The lightbulb is instantly replaced upon failure.

(a) Find the probability that the lightbulb will have failed after two years.

(b) What is the probability that, over a period of five years, the lightbulb was replaced only once?

58. Suppose the lifetime of a light bulb is exponentially distributed with mean 1 year. The light bulb is instantly replaced upon failure. What is the probability that, over a period of five years, at most five light bulbs are needed?

12.5.6

59. Suppose the lifetime of a laptop computer is exponentially distributed with mean five years.

(a) Find the probability that the computer will have failed after three years.

(b) Given that the computer has worked for six years, find the probability that it will work for another year.

60. Suppose the lifetime of an organism is exponentially distributed with hazard rate function $\lambda(x) = 2/\text{day}$.

(a) Find the probability that an individual of this species lives for more than three days.

(b) What is the expected lifetime?

61. Suppose the lifetime of a printer is exponentially distributed with parameter $\lambda = 0.2/\text{year}$.

(a) What is the expected lifetime?

(b) The **median lifetime** is defined as the age x_m at which the probability of not having died by age x_m is 0.5. Find x_m.

62. The median lifetime is defined as the age x_m at which the probability of not having died by age x_m is 0.5. If the life span of an organism is exponentially distributed, and if $x_m = 4$ years, what is the hazard-rate function?

63. The hazard-rate function of an organism is given by

$$\lambda(x) = 0.3 + 0.1e^{0.01x}, \quad x \geq 0$$

where x is measured in days.

(a) What is the probability that the organism will live for more than five days?

(b) What is the probability that the organism will live between 7 and 10 days?

64. The hazard-rate function of an organism is given by

$$\lambda(x) = 0.1 + 0.5e^{0.02x}, \quad x \geq 0$$

where x is measured in days.

(a) What is the probability that the organism will live less than 10 days?

(b) What is the probability that the organism will live for another five days, given that it survived the first five days?

T **65.** The median lifetime is defined as the age x_m at which the probability of not having died by age x_m is 0.5. Use a graphing calculator to numerically approximate the median lifetime if the hazard-rate function is

$$\lambda(x) = 1.2 + 0.3e^{0.5x}, \quad x \geq 0$$

T **66.** The median lifetime is defined as the age x_m at which the probability of not having died by age x_m is 0.5. Use a graphing calculator to numerically approximate the median lifetime if the hazard-rate function is

$$\lambda(x) = 0.5 + 0.1e^{0.2x}, \quad x \geq 0$$

67. The hazard-rate function of an organism is given by

$$\lambda(x) = (2 \times 10^{-5})x^{1.5}, \quad x \geq 0$$

where x is measured in days.

(a) What is the probability that the organism will live for more than 50 days?

(b) What is the probability that the organism will live between 50 and 70 days?

68. The hazard-rate function of an organism is given by

$$\lambda(x) = 0.04x^{3.1}, \quad x \geq 0$$

where x is measured in years.

(a) What is the probability that the organism will live for more than three years?

(b) What is the probability that the organism will live for another three years, given that it survived the first three years?

69. The median lifetime is defined as the age x_m at which the probability of not having died by age x_m is 0.5. Find the median lifetime if the hazard-rate function is

$$\lambda(x) = (4 \times 10^{-5})x^{2.2}, \quad x \geq 0$$

70. The median lifetime is defined as the age x_m at which the probability of not having died by age x_m is 0.5. Find the median lifetime if the hazard-rate function is

$$\lambda(x) = (3.7 \times 10^{-6})x^{2.7}, \quad x \geq 0$$

71. Let N_x be the number of individuals that are still alive at age x. Show that

$$-\ln \frac{N_{x+1}}{N_x}$$

can be estimated by

$$\int_x^{x+1} \lambda(u)\, du$$

where $\lambda(x)$ is the hazard-rate function at age x.

12.6 Limit Theorems

12.6.1 The Law of Large Numbers

Prostate-specific antigen (PSA) levels in a man's blood are used to detect prostate cancer. They are also used to screen for "biochemical failure" after removal of the prostate. Biochemical failure is defined as a PSA level that exceeds 0.5 ng/ml; that level of PSA may mean that prostate cancer cells are still in the body of the patient. In a study by ***Iselin et al. (1999)***, the PSA levels of 429 men with prostate cancer were followed after removal of the prostate. (The disease was initially confined to the prostate in these men). After five years, 8% of the men whose cancer was initially confined to the prostate had biochemical failure.

Suppose now that you heard of a small study of 30 men whose prostate was surgically removed and in whom the disease was initially confined to the prostate. After five years, 3 out of the 30 men, or 10%, experienced biochemical failure. Which of the two figures, 8% or 10%, would you deem more reliable? You probably answered 8%. We tend to trust larger studies more than smaller ones. The reason is found in a result known as the **law of large numbers**, which implies that estimates of proportions become more reliable as the sample size increases.

To state the law of large numbers, we consider a sequence of independent random variables $X_1, X_2, \ldots, X_n$, all with the same distribution. We say that $X_1, X_2, \ldots, X_n$ are **independent and identically distributed** (i.i.d., for short). We assume that $E(X_i) = \mu$ and $\text{var}(X_i) = \sigma^2$. We can think of the sequence as coming from some random experiment that we repeat n times and X_i is the outcome on the ith trial. For instance, the X_i's could denote the successive outcomes of tossing a coin n times, with $X_i = 1$ if the ith toss results in heads and $X_i = 0$ otherwise.

We are interested in the (arithmetic) **average** of $X_1, X_2, \ldots, X_n$, denoted by

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

In the case of the coin-tossing example, $\overline{X}_n$ would be the fraction of heads in n trials.

To state the law of large numbers, we need a type of convergence called **convergence in probability**. We say that a random variable Z_n converges to a constant γ in probability as $n \to \infty$ if, for any $\epsilon > 0$,

$$\lim_{n\to\infty} P(|Z_n - \gamma| \geq \epsilon) = 0$$

Now we state the law:

Weak Law of Large Numbers[2] If $X_1, X_2, \ldots, X_n$ are i.i.d. with $E(|X_i|) < \infty$, then, as $n \to \infty$, $\overline{X}_n$ converges to $E(X_1)$ in probability.

The weak law of large numbers explains why taking a larger sample improves the reliability of estimates of proportions. In the prostate cancer example described at the beginning of this subsection, we were interested in the likelihood of biochemical failure after five years in men who underwent prostate surgery when the cancer was confined to the prostate. We can think of this likelihood as a (to us unknown) probability that we estimate by taking a large sample in which all individuals are considered independent. (Such a sample is called a **random sample**.) If we set

$$X_i = \begin{cases} 1 & \text{if biochemical failure appears after five years in individual } i \\ 0 & \text{otherwise} \end{cases}$$

then $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ is the fraction of men in the sample who experience biochemical failure after five years. Now, if we set

$$\mu = E(X_i) = P(X_i = 1)$$

then the law of large numbers tells us that $\overline{X}_n \to \mu$ in probability as $n \to \infty$. Thus, if the sample size is sufficiently large, $\overline{X}_n$ provides a good estimate of the probability of biochemical failure, in the sense that, with high probability, $\overline{X}_n$ will be close to μ.

Two inequalities help us to prove the weak law of large numbers.

Markov's Inequality If X is a nonnegative random variable with $E(X) < \infty$, then, for any $a > 0$,

$$P(X \geq a) \leq \frac{E(X)}{a}$$

X can be discrete or continuous

Proof We prove Markov's inequality when X is a nonnegative, continuous random variable with density function $f(x)$. Let $a > 0$. Then

$$E(X) = \int_0^\infty x f(x)\,dx = \int_0^a x f(x)\,dx + \int_a^\infty x f(x)\,dx \tag{12.36}$$

Since $\int_0^a x f(x)\,dx \geq 0$, it follows that

$$E(X) \geq \int_a^\infty x f(x)\,dx \tag{12.37}$$

Using $x \geq a$, we find that the right-hand side of (12.37) is bounded as follows:

$$\int_a^\infty x f(x)\,dx \geq \int_a^\infty a f(x)\,dx = a\int_a^\infty f(x)\,dx = aP(X \geq a)$$

Therefore,

$$E(X) \geq aP(X \geq a) \qquad \text{or} \qquad P(X \geq a) \leq \frac{E(X)}{a}$$

•

(2) In addition to the weak law of large numbers, there is also a strong law of large numbers. The weak law does not exclude the possibility that the averages $\overline{X}_n$ may occasionally be quite different from $E(X_1)$, even for large n. The strong law of large numbers excludes this possibility. Because it requires additional theoretical background to state, the strong law of large numbers will not be discussed in this book.

The next inequality is a consequence of Markov's inequality.

Chebyshev's Inequality If X is a random variable with finite mean μ and finite variance σ^2, then, for any $c > 0$,

$$P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}$$

Proof The events $\{|X - \mu| \geq c\}$ and $\{(X - \mu)^2 \geq c^2\}$ are the same. Therefore,

$$P(|X - \mu| \geq c) = P((X - \mu)^2 \geq c^2)$$

The random variable $(X - \mu)^2$ is nonnegative, and $E(X - \mu)^2 = \sigma^2 < \infty$ by assumption. We can thus apply Markov's inequality and obtain

$$P((X - \mu)^2 \geq c^2) \leq \frac{E(X - \mu)^2}{c^2} = \frac{\sigma^2}{c^2}$$

•

To prove the weak law of large numbers, we need to find the mean and the variance of $\overline{X}_n$. First, the mean is

$$E(\overline{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n} \cdot n\mu = \mu$$

Then, using independence, we calculate the variance:

$$\text{var}(\overline{X}_n) = \text{var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \text{var}(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}$$

Applying Chebyshev's inequality with $c > 0$ to $\overline{X}_n$ gives

$$P(|\overline{X}_n - \mu| \geq c) \leq \frac{\text{var}(\overline{X}_n)}{c^2} = \frac{\sigma^2}{nc^2}$$

If we let $n \to \infty$, the right-hand side tends to 0. Since probabilities are (always) nonnegative, we can conclude that

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| \geq c) = 0 \tag{12.38}$$

The result about the limit in (12.38) can thus be expressed as "$\overline{X}_n$ converges to μ in probability." This is the weak law of large numbers.

We proved the weak law under the additional assumption that $\text{var}(X_i) < \infty$. The weak law of large numbers still holds if $\text{var}(X_i) = \infty$, provided that $E(|X_i|) < \infty$, but the proof of this result would be quite a bit more complicated (and we won't do it).

EXAMPLE 1 Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. with

$$X_i = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1 - p \end{cases}$$

Set $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, and show that $\overline{X}_n$ converges to p in probability as $n \to \infty$.

Solution Since

$$E(|X_i|) = |1|(p) + |0|(1 - p) = p < \infty \quad \text{and} \quad E(X_i) = (1)(p) + (0)(1 - p) = p$$

we can invoke the law of large numbers and conclude that

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \to E(X_i) = p$$

in probability as $n \to \infty$. •

EXAMPLE 2 **Monte Carlo Integration** Suppose that $f(x)$ is an integrable function on $[0, 1]$ with $f(x) \geq 0$ for $x \in [0, 1]$. Let $U_1, U_2, \ldots, U_n$ be i.i.d. with U_i uniformly distributed on $(0, 1)$. Show that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(U_i) = \int_0^1 f(x)\,dx \qquad \text{in probability.}$$

Solution Define $X_i = f(U_i)$. Then, using the fact that the density function of a uniform distribution on $(0, 1)$ is equal to 1 on that same interval, we obtain

$$E(|X_i|) = E(|f(U_i)|) = \int_0^1 |f(x)|\,dx = \int_0^1 f(x)\,dx < \infty$$

and

$$E(X_i) = E[f(U_i)] = \int_0^1 f(x)\,dx$$

The X_i's are i.i.d. We can therefore apply the law of large numbers and conclude that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(U_i) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = E(X_i) = \int_0^1 f(x)\,dx$$

in probability. •

We can use Chebyshev's inequality to get an estimate of sample size.

EXAMPLE 3 Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. with

$$X_i = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

Use Chebyshev's inequality to find n such that $\overline{X}_n$ will differ from p by less than 0.01 with probability at least 0.95.

Solution We know from Example 1 that $\overline{X}_n$ converges to p in probability as $n \to \infty$. We want to investigate how fast the convergence is. More precisely, we want to find n such that

$$P(|\overline{X}_n - p| < 0.01) \geq 0.95$$

or, taking complements,

$$P(|\overline{X}_n - p| \geq 0.01) \leq 0.05$$

Using Chebyshev's inequality, we find that

$$P(|\overline{X}_n - p| \geq 0.01) \leq \frac{\operatorname{var}(\overline{X}_n)}{(0.01)^2}$$

Since the X_i's are independent,

$$\begin{aligned}\operatorname{var}(\overline{X}_n) &= \operatorname{var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{var}(X_i) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} p(1-p) = \frac{p(1-p)}{n}\end{aligned}$$

Therefore,

$$\frac{\operatorname{var}(\overline{X}_n)}{(0.01)^2} = 10{,}000\frac{p(1-p)}{n}$$

We want this expression to be less than or equal to 0.05. However, we don't know the value of p. Fortunately, the following reasoning allows us to find a bound on $p(1-p)$:

The function $f(p) = p(1-p)$ is an upside-down parabola with roots at $p = 0$ and $p = 1$. Its maximum is at $p = 1/2$ and is $1/4$; that is, $f(1/2) = 1/4$. Therefore,

$$p(1-p) \le \frac{1}{4} \quad \text{for all } 0 \le p \le 1.$$

We thus obtain

$$10{,}000\frac{p(1-p)}{n} \le 10{,}000\frac{1}{4n} \le 0.05, \qquad \text{or} \qquad n \ge 50{,}000.$$

We conclude that a sample size of 50,000 would suffice to estimate p within an error of 0.01 with 95% probability. However, it turns out that Chebyshev's inequality does not give very good estimates and this lower bound on the sample size is much larger than what we would really need. In the next subsection, we will learn a better way to estimate sample sizes. •

12.6.2 The Central Limit Theorem

We now come to a result that is central to probability theory. It says that if we add up a large number of independent and identically distributed random variables with finite mean and variance, then, after suitable scaling, the distribution of the resulting quantity is approximately normally distributed. We will not be able to prove the theorem here, but we will be able to explore some of its implications.

Theorem **Central Limit Theorem** Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. with mean $E(X_i) = \mu$ and variance $\text{var}(X_i) = \sigma^2 < \infty$. Define $S_n = \sum_{i=1}^n X_i$. Then, as $n \to \infty$,

$$P\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \le x\right) \to F(x) \qquad n\mu = E\,S_n\ , \quad \sqrt{n\sigma^2} = \text{s.d.}(S_n)$$

where $F(x)$ is the distribution function of the standard normal distribution.

EXAMPLE 4 Toss a fair coin 500 times. Use the central limit theorem to find an approximation for the probability of at least 265 heads.

Solution We define

$$X_i = \begin{cases} 1 & \text{if } i\text{th toss results in heads} \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\mu = E(X_i) = \frac{1}{2} \qquad \text{and} \qquad \sigma^2 = \text{var}(X_i) = \frac{1}{2}\left(1 - \frac{1}{2}\right) = \frac{1}{4}.$$

Denote the number of heads among the first 500 tosses by $S_{500} = \sum_{i=1}^{500} X_i$. Then

$$\begin{aligned} P(S_{500} \ge 265) &= P\left(\frac{S_{500} - 500\mu}{\sqrt{500\sigma^2}} \ge \frac{265 - 250}{\sqrt{125}}\right) \\ &\approx 1 - F\left(\frac{15}{\sqrt{125}}\right) = 1 - F(1.34) \\ &= 1 - 0.9099 = 0.0901 \end{aligned}$$

where $F(x)$ is the distribution function of a standard normally distributed random variable and the values of $F(x)$ are obtained from the table in Appendix B. •

When the central limit theorem is applied to integer-valued random variables, a correction is used to get a better approximation. The correction, called the **histogram correction**, is explained in the next example.

EXAMPLE 5 For S_{500} defined in Example 4, use the central limit theorem to find an approximation for $P(S_{500} = 250)$.

Solution If we applied the central limit theorem without any corrections, we would find that

$$P(S_{500} = 250) = P\left(\frac{S_{500} - 500\mu}{\sqrt{500\sigma^2}} = \frac{250 - 250}{\sqrt{125}}\right) \approx P(Z = 0) = 0$$

where Z is a standard normally distributed random variable. We can compare this result with the exact value. In that case, S_{500} is binomially distributed with parameters $n = 500$ and $p = 1/2$. We find that

$$P(S_{500} = 250) = \binom{500}{250}\left(\frac{1}{2}\right)^{500} \approx 0.036$$

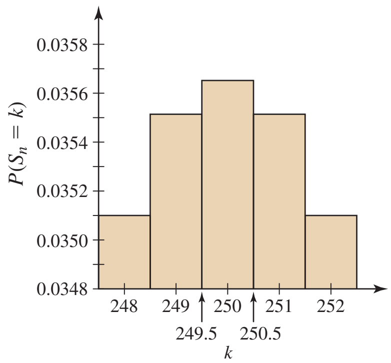


Figure 12.51 The histogram correction for Example 5.

The central limit theorem does not give a good approximation. We can do better by writing the event $\{S_{500} = 250\}$ as $\{249.5 \le S_{500} \le 250.5\}$. (See Figure 12.51.) Then

$$\begin{aligned} P(S_{500} = 250) &= P(249.5 \le S_{500} \le 250.5) \\ &= P\left(\frac{249.5 - 250}{\sqrt{125}} \le \frac{S_{500} - 500\mu}{\sqrt{500\sigma^2}} \le \frac{250.5 - 250}{\sqrt{125}}\right) \\ &= P(-0.04 \le Z \le 0.04) \end{aligned}$$

where Z is standard normally distributed. It then follows that

$$P(-0.04 \le Z \le 0.04) = 2F(0.04) - 1 = (2)(0.5160) - 1 = 0.032$$

where $F(x)$ is the distribution function of a standard normal distribution. •

EXAMPLE 6 Redo Example 4 with the histogram correction.

Solution We write the event $\{S_{500} \ge 265\}$ as $\{S_{500} \ge 264.5\}$. (See Figure 12.52.) Then

$$\begin{aligned} P(S_{500} \ge 265) = P(S_{500} \ge 264.5) &= P\left(\frac{S_{500} - 500\mu}{\sqrt{500\sigma^2}} \ge \frac{264.5 - 250}{\sqrt{125}}\right) \\ &\approx P(Z \ge 1.30) = 1 - F(1.30) = 1 - 0.9032 = 0.0968 \end{aligned}$$

where Z is a standard normally distributed random variable with distribution function $F(x)$. •

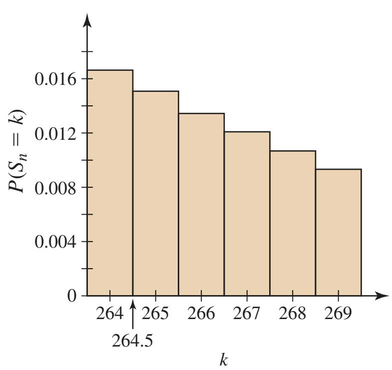


Figure 12.52 The histogram correction for Example 6.

Quantitative genetics is a field in biology that attempts to explain quantitative differences between individuals that are either of genotypic or environmental origin, such as differences in height, litter size, number of abdominal bristles in *Drosophila*, and so on. Estimates on the number of loci involved in a quantitative trait range from very few, such as five loci for skull length in rabbits (***Wright, 1968***), to very many, such as 98 loci for abdominal bristles in *Drosophila* (***Falconer, 1989***).

When many loci are involved in a quantitative trait, the **infinitesimal model** is used to model the genotypic value of that trait. The genotypic value G of a trait is considered to be a sum of the contributions of each of the loci involved:

$$G = X_1 + X_2 + \cdots + X_n$$

In the simplest case, the X_i's are assumed to be independent and identically distributed and represent the contribution of locus i to the genotypic value. If the X_i's have finite mean and variance, and if n is large, the distribution of G can be approximated by a normal distribution.

EXAMPLE 7 Suppose a trait is controlled by 100 loci. Each locus, independently of all others, contributes to the genotypic value of the trait either $+1$ with probability 0.6 or -0.7 with probability 0.4.

(a) Find the mean value of the trait.

(b) What proportion of the population has a trait value greater than 40?

Solution **(a)** The genotypic value of the trait can be written as

$$S_{100} = \sum_{i=1}^{100} X_i$$

with

$$X_i = \begin{cases} 1 & \text{with probability } 0.6 \\ -0.7 & \text{with probability } 0.4 \end{cases}$$

Hence,

$$E(S_{100}) = \sum_{i=1}^{100} E(X_i) = \sum_{i=1}^{100} [(1)(0.6) + (-0.7)(0.4)] = \sum_{i=1}^{100} 0.32 = 32$$

Hence, the mean value of the trait is 32.

(b) To find the proportion of the population that has a trait value greater than 40, we employ the central limit theorem. We compute the variance of X_i first:

$$E(X_i^2) = (1)^2(0.6) + (-0.7)^2(0.4) = 0.796$$

Thus,

$$\text{var}(X_i) = E(X_i^2) - [E(X_i)]^2 = 0.796 - (0.32)^2 = 0.6936$$

Now,

$$P(S_{100} > 40) = P\left(\frac{S_{100} - 32}{\sqrt{(100)(0.6936)}} > \frac{40.5 - 32}{\sqrt{69.36}}\right)$$
$$\approx 1 - F(1.02) = 1 - 0.8461 = 0.1539$$

where $F(x)$ is the distribution function of a standard normal distribution. Consequently, about 15% of the population has trait value greater than 40. •

EXAMPLE 8 **Estimating Sample Sizes** Suppose you wish to conduct a medical study to determine the fraction of people in the general population whose total cholesterol level is above 220 g/dl. How large a sample size would you need to estimate the proportion within 0.01 of the true value with probability at least 0.95?

Solution Define

$$X_i = \begin{cases} 1 & \text{if } i\text{th individual has cholesterol} \geq 220 \text{ mg/dl} \\ 0 & \text{otherwise} \end{cases}$$

and assume that the individuals are selected so that the X_i's are i.i.d. Then

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

is an estimate of the proportion of individuals whose cholesterol level exceeds 220 mg/dl.

If we set $S_n = \sum_{i=1}^{n} X_i$, then, with $p = E(X_i)$ and $\sigma^2 = \text{var}(X_i)$,

$$\frac{S_n - np}{\sqrt{n\sigma^2}} \quad \text{is approximately standard normally distributed}$$

Dividing both numerator and denominator by n, we find that, with $\sigma = \sqrt{\text{var}(X_i)} = \sqrt{p(1-p)}$,

$$\sqrt{n}\frac{\overline{X}_n - p}{\sqrt{p(1-p)}} \quad \text{is approximately standard normally distributed}$$

We are interested in finding n in order to estimate the proportion within 0.01 of the true value p with probability at least 0.95. That is,

$$P(|\overline{X}_n - p| \leq 0.01) \geq 0.95$$

We rewrite this inequality as

$$P(-0.01 \le \overline{X}_n - p \le 0.01) \ge 0.95$$

or

$$P\left(\sqrt{n}\frac{-0.01}{\sqrt{p(1-p)}} \le \sqrt{n}\frac{\overline{X}_n - p}{\sqrt{p(1-p)}} \le \sqrt{n}\frac{0.01}{\sqrt{p(1-p)}}\right) \ge 0.95$$

Since $\sqrt{n}\frac{\overline{X}_n - p}{\sqrt{p(1-p)}}$ is approximately standard normally distributed, we find that the left-hand side is approximately

$$2F\left(\sqrt{n}\frac{0.01}{\sqrt{p(1-p)}}\right) - 1$$

This is ≥ 0.95 if

$$F\left(\sqrt{n}\frac{0.01}{\sqrt{p(1-p)}}\right) \ge 0.975$$

or

$$\sqrt{n}\frac{0.01}{\sqrt{p(1-p)}} \ge 1.96.$$

Solving for n, we obtain

$$n \ge (196)^2 p(1-p)$$

Since we do not know p, we take the worst possible case that maximizes $p(1-p)$. As we saw in Example 3, this occurs for $p = 1/2$. Therefore,

$$n \ge (196)^2 \frac{1}{2}\left(1 - \frac{1}{2}\right) = 9604.$$

Thus, about 9604 individuals would suffice for this study. You should compare this result with that in Example 3, where we solved the same problem (in a different application) by using Chebyshev's inequality instead of the central limit theorem. •

Remark Both the normal and the Poisson distribution serve as approximations to the binomial distribution. As a rule of thumb, the approximations are reasonably good when $n \ge 40$. When $np \le 5$, the Poisson approximation should be used; when $np \ge 5$, the normal approximation should be used.

Section 12.6 Problems

12.6.1

1. Let X be exponentially distributed with parameter $\lambda = 1/2$. Use Markov's inequality to estimate $P(X \ge 3)$, and compare your estimate with the exact answer.

2. Let X be uniformly distributed over $(1, 4)$.

(a) Use Markov's inequality to estimate $P(X \ge a)$, $1 \le a \le 4$, and compare your estimate with the exact answer.

(b) Find the value of $a \in (1, 4)$ that minimizes the difference between the bound and the exact probability computed in (a).

3. Prove Markov's inequality when X is a nonnegative discrete random variable with $E(X) < \infty$.

4. Let X be a continuous random variable with density $f(x)$, and assume that $X \ge 2$. Why is $E(X) \ge 2$?

5. Let X be uniformly distributed over $(-2, 2)$. Use Chebyshev's inequality to estimate $P(|X| \ge 1)$, and compare your estimate with the exact answer.

6. Let X be standard normally distributed. Use Chebyshev's inequality to estimate **(a)** $P(|X| \ge 1)$, **(b)** $P(|X| \ge 2)$, and **(c)** $P(|X| \ge 3)$. Compare each estimate with the exact answer.

7. Suppose X is a random variable with mean 10 and variance 9. What can you say about $P(|X - 10| \ge 5)$?

8. Suppose X is a random variable with mean -5 and variance 2. What can you say about the probability that X deviates from its mean by at least 4?

9. Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. with

$$X_i = \begin{cases} -1 & \text{with probability } 0.2 \\ 1 & \text{with probability } 0.5 \\ 2 & \text{with probability } 0.3 \end{cases}$$

What can you say about $\frac{1}{n}\sum_{i=1}^n X_i$ as $n \to \infty$?

10. Suppose $X_1, X_2, \ldots, X_n$ are independent random variables with $P(X_i > x) = e^{-2x}$. What can you say about $\frac{1}{n}\sum_{i=1}^n X_i$ as $n \to \infty$?

11. Suppose $X_1, X_2, \ldots, X_n$ are independent random variables with density function

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbf{R}$$

Can you apply the law of large numbers to $\frac{1}{n}\sum_{i=1}^{n} X_i$? If so, what can you say about $\frac{1}{n}\sum_{i=1}^{n} X_i$ as $n \to \infty$?

12. How often do you have to toss a coin to determine $P(\text{heads})$ within 0.1 of its true value with probability at least 0.9?

13. A certain study showed that less than 5% of the population suffers from a certain disorder. To get a more accurate estimate of this proportion, you plan to conduct another study. What sample size should you choose if you want to be at least 95% sure that your estimate is within 0.05 of the true value?

14. Assume that $E(e^{cX}) < \infty$ for $c > 0$. Use Markov's inequality to prove **Bernstein's inequality**,

$$P(X \geq x) \leq e^{-cx} E(e^{cX})$$

for $c > 0$.

12.6.2

15. Toss a fair coin 400 times. Use the central limit theorem and the histogram correction to find an approximation for the probability of getting at most 190 heads.

16. Toss a fair coin 150 times. Use the central limit theorem and the histogram correction to find an approximation for the probability that the number of heads is at least 70.

17. Toss a fair coin 200 times.

(a) Use the central limit theorem and the histogram correction to find an approximation for the probability that the number of heads is at least 120.

(b) Use Markov's inequality to find an estimate for the event in (a), and compare your estimate with that in (a).

18. Toss a fair coin 300 times.

(a) Use the central limit theorem and the histogram correction to find an approximation for the probability that the number of heads is between 140 and 160.

(b) Use Chebyshev's inequality to find an estimate for the event in (a), and compare your estimate with that in (a).

19. Suppose S_n is binomially distributed with parameters $n = 200$ and $p = 0.3$. Use the central limit theorem to find an approximation for $P(99 \leq S_n \leq 101)$ **(a)** without the histogram correction and **(b)** with the histogram correction. **(c)** Use a graphing calculator to compute the exact probabilities, and compare your answers with those in (a) and (b).

20. Suppose S_n is binomially distributed with parameters $n = 150$ and $p = 0.4$. Use the central limit theorem to find an approximation for $P(S_n = 60)$ **(a)** without the histogram correction and **(b)** with the histogram correction. **(c)** Use a graphing calculator to compute the exact probabilities and compare your answers with those in (a) and (b).

21. Suppose a genotypic trait is controlled by 80 loci. Each locus, independently of all others, contributes to the genotypic value of the trait either $+0.3$ with probability 0.2, -0.1 with probability 0.5, or -0.5 with probability 0.3.

(a) Find the mean value of the trait.

(b) What proportion of the population has a trait value between -12 and -7?

22. Suppose a genotypic trait is controlled by 90 loci. Each locus, independently of all others, contributes to the genotypic value of the trait either 1.1 with probability 0.7, 0.9 with probability 0.1, or 0.1 with probability 0.2.

(a) Find the mean value of the trait.

(b) What proportion of the population has a trait value less than 72?

23. How often should you toss a coin to be at least 90% certain that your estimate of $P(\text{heads})$ is within 0.1 of its true value?

24. How often should you toss a coin to be at least 90% certain that your estimate of $P(\text{heads})$ is within 0.01 of its true value?

25. To forecast the outcome of a presidential election in which two candidates run for office, a telephone poll is conducted. How many people should be surveyed to be at least 95% sure that the estimate is within 0.05 of the true value? (Assume that there are no undecided people in the survey.)

26. A medical study is conducted to estimate the proportion of people suffering from seasonal affective disorder. How many people should be surveyed to be at least 99% sure that the estimate is within 0.02 of the true value?

In Problems 27–30, S_n is binomially distributed with parameters n and p.

27. For $n = 100$ and $p = 0.01$, compute $P(S_n = 0)$ **(a)** exactly, **(b)** by using a Poisson approximation, and **(c)** by using a normal approximation.

28. For $n = 100$ and $p = 0.1$, compute $P(S_n = 10)$ **(a)** exactly, **(b)** by using a Poisson approximation, and **(c)** by using a normal approximation.

29. For $n = 50$ and $p = 0.1$, compute $P(S_n = 5)$ **(a)** exactly, **(b)** by using a Poisson approximation, and **(c)** by using a normal approximation.

30. For $n = 50$ and $p = 0.5$, compute $P(S_n = 25)$ **(a)** exactly, **(b)** by using a Poisson approximation, and **(c)** by using a normal approximation.

31. Suppose you want to estimate the proportion of people in the United States who do not believe in evolution. You happen to take a class on evolutionary theory at a U.S. college that is attended by 200 students, all of whom are biology majors. Do you think you would get an accurate estimate if you asked all 200 students in your class? Discuss.

32. A soft-drink company introduces a new beverage. One month later, the company wants to know whether its marketing strategies have reached young adults of ages 18–20. You happen to work part time for the marketing company that is conducting the survey. At the same time, you are taking a calculus class that is attended by 250 students. It would be easy for you to hand out a survey in class. Would you suggest this to your supervisor in the marketing company? Discuss.

33. Clementine oranges are sold in boxes. Each box contains 50 clementines. The probability that a clementine in a box is spoiled is 0.01.

(a) Use an appropriate approximation to determine the probability that a box contains 0, 1, or at least 2 spoiled clementines.

(b) A shipment of clementines (said to be hybrid crossings between oranges and tangerines) with 100 boxes is considered unacceptable if 35% or more of the boxes contain spoiled clementines. What is the probability that a shipment is unacceptable?

34. Turner's Syndrome Turner's syndrome is a chromosomal disorder in which girls have only one X chromosome. It affects about 1 in 2000 girls. About 1 in 10 girls with Turner's syndrome suffers from narrowing of the aorta.

(a) In a group of 4000 girls, what is the probability that no girls are affected with Turner's syndrome? That one girl is affected? Two? At least three?

(b) In a group of 170 girls affected with Turner's syndrome, what is the probability that at least 20 of them suffer from an abnormal narrowing of the aorta?

Cystic Fibrosis

In Problems 35–37, use the following facts: Cystic fibrosis is an inherited disorder that causes abnormally thick body secretions. About 1 in 2500 white babies in the United States has this disorder. About 3 in 100 children with cystic fibrosis develop diabetes mellitus, and about 1 in 5 females with cystic fibrosis is infertile.

35. Find the probability that, in a group of 5000 newborn white babies in the United States, at least 4 babies suffer from cystic fibrosis.

36. Find the probability that, in a group of 1000 children with cystic fibrosis, at least 25 will develop diabetes mellitus.

37. Find the probability that, in a group of 250 women with cystic fibrosis, no more than 60 are infertile.

12.7 Statistical Tools

In the preceding sections, we learned how to model various random experiments. Using the underlying probability distribution of the model, we were able to compute the probabilities of events, such as the probability of obtaining white-flowering pea plants in Mendel's pea experiment.

To understand observations or outcomes of experiments in the biological or medical sciences, however, we often take the reverse approach: We infer the underlying probability distribution from events we have observed. On the basis of a collection of observations, called **data**, we estimate characteristics of the underlying probability distribution. For instance, in the case of the normal distribution, our goal might be to estimate the mean and the variance, which are parameters that describe a normal distribution.

This section provides an introduction to data, parameters estimation, confidence intervals, and finally linear regression, which we already met in Chapter 10. Throughout, we assume that all observations are expressed numerically.

12.7.1 Describing Univariate Data

Observations that are expressed numerically can be described by **variables**. We can measure one or more variables in an experiment or an observational study. **Univariate** data refers to data obtained from measuring a single variable. **Bivariate** data consists of pairs of observations for each sample point. If more than two variables are measured, we call the data **multivariate**. Data in medical studies are typically multivariate. For instance, ***Dyck et al. (1999)*** measured 26 variables in a cohort of patients with diabetes in Rochester, Minnesota, to understand the effects of chronic hyperglycemia (high blood sugar levels) on diabetic neuropathy (a neurological disorder that often accompanies long-term diabetes). Examples of variables included in the study are age, height, weight, duration of diabetes, cholesterol levels, and percent glycosylated hemoglobin. In this section, we will encounter only univariate data.

To learn something about the distribution of a character in a population (severity of diabetic neuropathy, clutch size, plant height, life span, effect of a new drug, and so forth), we cannot measure the occurrence of that character on every individual in the population. Instead, we take a subset of the population, called a **sample**, obtain individual observations on the character of interest, which are assumed to be independent of each other, and then infer the distribution of the character in the population from its distribution in the sample. A sample that is representative of the population and whose individual observations are independent of each other and have the same distribution is called a **random sample**.

To illustrate the importance of choosing a random sample, let's look at the recent debate regarding the benefit of hormone replacement therapy for postmenopausal women. For several decades, physicians recommended hormone replacement therapy for these women. The recommendation was based on observational studies that indicated health benefits, such as a decrease in coronary heart disease (CHD). Three well-designed and large drug trials—the Heart and Estrogen/Progestin Replacement

Study (HERS), the Women's Health Initiative (WHI) estrogen-plus-progestin trial, and the WHI estrogen-alone trial—found an *increase* in CHD in addition to other harmful effects. In 2002, after the conclusion of the first two studies, the U.S. Food and Drug Administration issued a warning about the potential harm of this treatment. An editorial by ***Hulley and Grady (2004)*** that accompanied the research article of the third study (the WHI estrogen-alone trial) stated, "Given the absence of evidence in all three trials that these hormone regimens prevent CHD in these populations. . . , it is now clear that previously available evidence was misleading. Observational studies were probably confounded by the tendency of healthier women to seek and comply with hormone treatment."

In what follows, we will always assume that our sample is a random sample and thus representative of the population. By definition, all observations in a random sample are independent and come from the same distribution (the distribution of the quantity in the entire population). A typical scheme to obtain a random sample of size n is to pick an individual at random from the population, record the quantity of interest, replace the individual in the population, and then select the next individual. This procedure is repeated until a sample of size n is obtained. Replacing the sampled individual after recording the quantity of interest ensures that the population always has the same composition and, hence, that all observations come from the same distribution. It also means that an individual may be chosen more than once, unless the population size is much larger than the sample size. In practice, individuals are often chosen without replacement; for instance, in medical studies, when a representative group of participants is chosen, each individual is represented only once. When the sample size is much smaller than the population size, the difference between sampling with replacement and without replacement is negligible. We denote the sample by the vector $(X_1, X_2, \dots, X_n)$, where X_k is the kth observation. The X_k are independent random variables that all have the same distribution. We say that the X_k, $k = 1, 2, \dots, n$, are **independent and identically distributed**.

Location	Number of Lions
Kenyangaga	0
Lemai	1
Kogatende	6
Bolagonja	4
Klein's Camp	7
Tabora B	0
Lobo	6
Tagoro	6
Kirawira A	3
Kirawira B	0
Ikimoa	6
Ndabaka	0
Handajega	4
Mereo	4
Seronera	4
Maswa HQ	0
Mamarehe	4

We use the data set on the left, which describes the number of lions at 17 locations in the Serengeti over a three-day period in late October 1990, to introduce some important definitions. These data were collected by Dr. Craig Packer of the University of Minnesota and a group of collaborators working with him.

One way to summarize data is to give the number of times a certain category (in this case, number of lions) occurs. These numbers are called **frequencies**. If we divide the frequencies by the total number of observations, we obtain **relative frequencies**. The list of (relative) frequencies is called a (relative) frequency distribution.

To obtain the frequency distribution of the number of lions, we count how often each of the values in the category "number of lions" appears. For instance, the value 4 appears five times, so its frequency is 5; the value 7 appears once, so its frequency is 1, and so on.

The frequencies appear in the following table, which shows that the range of values is between 0 and 7:

Number of Lions	0	1	2	3	4	5	6	7
Number of Locations	5	1	0	1	5	0	4	1

To obtain relative frequencies, we divide each frequency by 17, because there are 17 locations. The relative frequencies appear in the following table.

Number of Lions	0	1	2	3	4	5	6	7
Relative Frequency of Locations	5/17	1/17	0	1/17	5/17	0	4/17	1/17

A quantity that is computed from observations in a sample is called a **statistic**.

The first statistic that we define is the **sample median**: the middle of the observations when we order the data according to size. If the number of observations is odd, there is one data point in the middle of the ordered data. If the number of observations is even, we take the average of the two observations in the middle. The number of data points in our example is odd. The list of the ordered data is as follows:

Ordered Lion Data: 0, 0, 0, 0, 0, 1, 3, 4, **4**, 4, 4, 4, 6, 6, 6, 6, 7

After ordering the data points, we find that the middle of the observations is the ninth data point, which is the median. Thus, the median is 4.

The two most frequently used statistics are the **sample mean** and the **sample variance**. In addition, the sample standard deviation is used. To define these quantities, we recall that we denoted a sample of size n by the vector $(X_1, X_2, \ldots, X_n)$, where X_k is the kth observation. The sample mean, the sample variance, and the sample standard deviation are defined as follows:

Sample mean: $\overline{X}_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$

Sample variance: $S_n^2 = \dfrac{1}{n-1}\displaystyle\sum_{k=1}^{n}(X_k - \overline{X}_n)^2$

Sample standard deviation: $S_n = \sqrt{S_n^2}$

The sample mean is thus the arithmetic average of the observations. The sample variance is the sum of the squared deviations from the sample mean, divided by $n-1$. (We will explain in the next subsection why we divide by $n-1$ rather than n.) The sample standard deviation is the square root of the sample variance.

The preceding definition of the sample variance is not very convenient for computation. An alternative form that is typically easier to use follows from an algebraic manipulation of the definition of the sample variance:

$$\begin{aligned}
S_n^2 &= \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X}_n)^2 \\
&= \frac{1}{n-1}\sum_{k=1}^{n}(X_k^2 - 2X_k\overline{X}_n + \overline{X}_n^2) \\
&= \frac{1}{n-1}\left[\sum_{k=1}^{n} X_k^2 - 2\overline{X}_n\sum_{k=1}^{n} X_k + n\overline{X}_n^2\right]
\end{aligned}$$

Using the fact that $\sum_{k=1}^{n} X_k = n\overline{X}_n$, we simplify the equation for S_n^2 to

$$\begin{aligned}
S_n^2 &= \frac{1}{n-1}\left[\sum_{k=1}^{n} X_k^2 - n\overline{X}_n^2\right] \\
&= \frac{1}{n-1}\left[\sum_{k=1}^{n} X_k^2 - \frac{1}{n}\left(\sum_{k=1}^{n} X_k\right)^2\right]
\end{aligned}$$

EXAMPLE 1 Compute the sample mean, the sample variance, and the sample standard deviation for the data on lions.

Solution We find that

$$\sum_{k=1}^{17} X_k = 55 \qquad \text{and} \qquad \sum_{k=1}^{17} X_k^2 = 283$$

Hence,

$$\overline{X}_{17} = \frac{55}{17} \approx 3.2,$$

$$\begin{aligned}
S_{17}^2 &= \frac{1}{16}\left[\sum_{k=1}^{17} X_k^2 - \frac{1}{17}\left(\sum_{k=1}^{17} X_k\right)^2\right] \\
&= \frac{1}{16}\left(283 - \frac{1}{17}(55)^2\right) \approx 6.57
\end{aligned}$$

and

$$S_{17} = \sqrt{S_{17}^2} \approx 2.56.$$

The sample standard deviation (S.D.) is an estimate of the variance of the population. When we report the sample mean and wish to give an indication of the variance of the population, we report the sample mean and the sample standard deviation as

$$\text{Mean} \pm \text{S.D.}$$

In Example 1, we would thus report 3.2 ± 2.6.

Reporting the sample mean and the variance of the population is common practice in scientific publications. These statistics frequently are reported under a heading of the form "Mean ± S.D." The expression "Mean" stands for the sample mean $\overline{X}_n$. The abbreviation "S.D." stands for the sample standard deviation, an estimate of the variance of the population. For instance, in the study by ***Dyck et al. (1999)***, mentioned earlier, the baseline characteristics of 149 Type 2 diabetes patients are listed in a table as "Mean ± S.D." The age of this group is listed as 69.7 ± 9.7, the height (in cm) as 166.2 ± 4.4, and the weight (in kg) as 84.8 ± 16.9.

If the sample distribution is summarized in a frequency distribution, the sample mean and the sample variance take on the following form: Assume that a sample of size n has l distinct values $x_1, x_2, \ldots, x_l$, where x_k occurs f_k times in the sample. Then the sample mean is given by the formula

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{l} x_k f_k$$

and the sample variance has the form

$$S_n^2 = \frac{1}{n-1}\left[\sum_{k=1}^{l} x_k^2 f_k - \frac{1}{n}\left(\sum_{k=1}^{l} x_k f_k\right)^2\right]$$

EXAMPLE 2 Use the frequency distribution of the data on lions to calculate the sample mean and the sample variance.

Solution We expect to find the same answers as in Example 1. We have

$$\begin{aligned}\overline{X}_n &= \frac{1}{n}\sum_{k=1}^{l} x_k f_k \\ &= \frac{1}{17}[(0)(5) + (1)(1) + (2)(0) + (3)(1) + (4)(5) + (5)(0) + (6)(4) + (7)(1)] \\ &= \frac{55}{17}\end{aligned}$$

and

$$\begin{aligned}S_n^2 &= \frac{1}{n-1}\left[\sum_{k=1}^{l} x_k^2 f_k - \frac{1}{n}\left(\sum_{k=1}^{l} x_k f_k\right)^2\right] \\ &= \frac{1}{16}[(0)(5) + (1)(1) + (4)(0) + (9)(1) + (16)(5) + (25)(0) + (36)(4) + (49)(1) \\ &\quad - \frac{1}{17}(55)^2] \\ &= \frac{1}{16}\left[283 - \frac{1}{17}(55^2)\right] \approx 6.57\end{aligned}$$

It is important to realize that any statistic computed from a sample will vary from sample to sample, since the samples are random subsets of the population. Statistics are therefore random variables with their own probability distributions.

EXAMPLE 3 Assume that a population consists of the three numbers 2, 4, and 7. List all samples of size 2 that can be drawn from this population with replacement, and find the sample mean of each sample.

Solution There are nine equally likely samples. We list them in the table on the left, together with the sample mean of each sample. We see from the table that the sample mean 3.0 occurs twice. This means that if we drew a sample of size 2 from this population, we would obtain a sample mean 3.0 with probability 2/9. •

Sample	Sample Mean
(2, 2)	2.0
(2, 4)	3.0
(2, 7)	4.5
(4, 2)	3.0
(4, 4)	4.0
(4, 7)	5.5
(7, 2)	4.5
(7, 4)	5.5
(7, 7)	7.0

To illustrate the point further, we look at simulated data from a normal distribution with mean 5 and variance 2. We take 5000 samples from this population, each of size 20, and record the sample mean and the sample median for each sample. Let's denote the vector of sample means by $(y_1, y_2, \ldots, y_{5000})$ and the vector sample medians by $(z_1, z_2, \ldots, z_{5000})$.

This simulation generated a table filled with random numbers from a normal distribution with mean 5 and variance 2. Although we cannot reproduce the full matrix, the following table lists the first four and the last two samples, together with the sample means and sample medians of each of the samples:

	Sample 1	Sample 2	Sample 3	Sample 4	...	Sample 4999	Sample 5000
1	5.4920	3.6955	4.8933	3.4264	...	2.4151	6.1035
2	5.4436	6.0239	6.7905	4.8145	...	3.2596	3.7615
3	6.3139	4.2353	5.8876	4.9372	...	3.6499	8.5641
4	6.5158	4.7598	3.5983	3.9115	...	2.0195	2.3705
5	3.5321	5.1131	4.9673	4.5547	...	5.0145	6.6010
6	4.2789	6.0736	6.3267	8.3319	...	3.5154	4.8617
7	4.0582	4.4552	8.2660	5.1155	...	5.2439	6.9706
8	5.0427	7.4530	4.7665	4.3044	...	5.6628	4.5673
9	4.6626	9.5493	6.6253	3.6156	...	4.0330	4.0917
10	4.1105	5.4313	4.1620	5.8632	...	6.0880	4.5681
11	6.0461	5.7836	5.0604	3.6048	...	5.8100	5.0355
12	7.4127	5.9725	3.8046	5.4154	...	6.0803	7.7348
13	3.2608	5.2567	5.9747	5.3955	...	6.3623	5.0953
14	4.3100	5.6717	4.2149	7.2308	...	6.3486	7.3613
15	4.4968	5.3302	7.0006	6.7701	...	2.1907	4.7502
16	2.9105	4.3761	6.8677	6.9040	...	4.6863	3.3855
17	5.2193	5.9124	5.2160	5.4927	...	5.1773	4.1588
18	6.5259	0.8479	6.3996	5.4561	...	4.3043	3.6232
19	3.0414	4.8115	6.5505	4.7408	...	7.5806	4.6489
20	3.5006	3.2167	5.0989	5.8786	...	4.9495	5.9608
Mean	4.8087	5.1985	5.6236	5.2882	...	4.7196	5.2107
Median	4.5797	5.2935	5.5518	5.2555	...	4.9820	4.8060

Histograms of the values of the sample means and sample medians are shown in Figure 12.53.

We can treat the sample means $(y_1, y_2, \ldots, y_{5000})$ and the sample medians $(z_1, z_2, \ldots, z_{5000})$ as data and calculate the sample means and sample variances of these two statistics.

For the simulated data that generated the histograms in Figure 12.53, we find, for the sample mean and the sample variance of the sample mean,

$$\bar{y} = \frac{1}{5000}\sum_{k=1}^{5000} y_k = 4.9966 \qquad S_{\bar{y}}^2 = \frac{1}{4999}\sum_{k=1}^{5000}(y_k - \bar{y})^2 = 0.0979$$

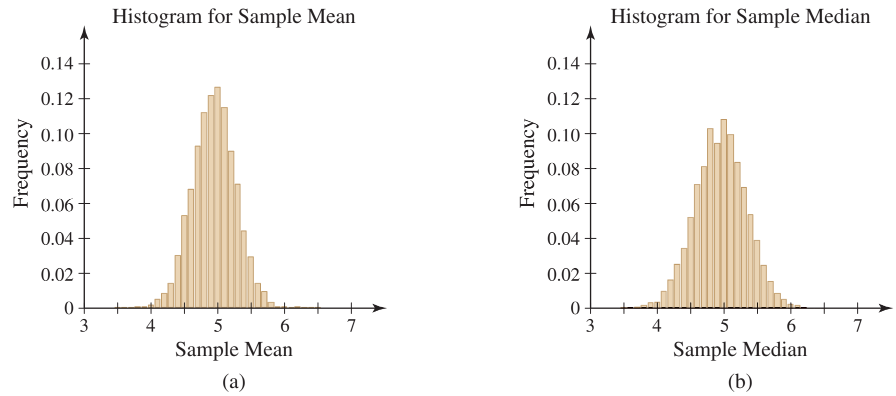


Figure 12.53 (a) The histogram of the sample means. (b) The histogram of the sample medians.

and, for the sample mean and the sample variance of the sample median,

$$\overline{z} = \frac{1}{5000}\sum_{k=1}^{5000} z_k = 5.0012 \qquad S_{\overline{z}}^2 = \frac{1}{4999}\sum_{k=1}^{5000}(z_k - \overline{z})^2 = 0.1442.$$

As we said earlier, statistics are random variables with their own probability distributions. We know that the distribution of the population in our example is a normal distribution with mean 5 and variance 2. It can be shown that the distribution of the sample mean of normally distributed random variables is also normally distributed. The mean of the distribution of the sample mean is equal to the mean of the population distribution, and the variance of the sample mean is equal to the variance of the population divided by the sample size. In our example with samples of size 20, the mean is therefore 5 and the variance is $2/20 = 0.1$. The simulated data agree with the theoretical predictions.

The histogram for the sample median (Figure 12.53b) suggests that the sampling distribution for the median of a sample from a normal distribution is approximately normal, and this is correct even for relatively small sample sizes. The mean of the distribution of the sample median is the same as the population distribution, namely, 5. However, it appears that the sample variance of the sample mean is smaller than that of the sample median (cf. Figures 12.53a and b). In fact the population is normally distributed with mean μ and variance σ^2, then the variance of the sample median of a sample of size n, denoted by σ_m^2, has the following relationship with σ^2 for large values of n:

$$\sigma_m^2 \approx \sigma^2 \frac{\pi}{2n}$$

We will not prove this result

When the population distribution is not normal, the sample mean is still approximately normally distributed for a large enough sample size, provided that the population distribution has a finite mean and a finite variance. This statement follows from the central limit theorem. The distribution of the sample *median* is often more complicated and may not even be approximately normally distributed.

It is important to note that a new simulation would generate different samples, and thus different values, for the various quantities we have computed. Because of the large number of samples, however, the sample means and sample variances of the two statistics would not differ much from simulation run to simulation run.

12.7.2 Estimating Parameters

We take random samples to learn something about the distribution of a variable in a population. For instance, we might be interested in knowing what the average cholesterol level in 50-year-old white males is. To estimate this level, we would take a random sample of 50-year-old white males, measure their cholesterol levels, and then compute the average of the measurements. This is an example of a *point estimate*. More generally, when estimating a parameter of a distribution, we can either give a single number, called a **point estimate**, or give a range, called an **interval estimate**.

In 1994, the National Institute of Standards and Technology (NIST)[3] published *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*. These guidelines are valuable for measurements in general. They state, "The result of a measurement is only an approximation or estimate of the value of the specific quantity subject to measurement, that is, the *measurand*, and thus the result is complete only when accompanied by a quantitative statement of uncertainty." The mean of a population is a parameter of particular interest. We will see in what follows how to provide point and interval estimates for this parameter and how to assess uncertainty in its measurement.

Estimates for parameters rely on outcomes of measurements. We expect measurements to be **accurate** and **precise**. *Accuracy* refers to how close measurements are to the true value, and *precision* refers to how close repeated measurements are to each other.

Point Estimates of Means. We assume that the population distribution has a finite mean μ and that the value of this parameter is unknown to us. For reasons that will become clear shortly, we will also assume that the distribution has a finite variance σ^2. We wish to estimate the mean by taking a random sample $(X_1, X_2, \ldots, X_n)$ of size n from the population. The X_k are independent and identically distributed according to the distribution of the population, with

$$E(X_k) = \mu \qquad \text{and} \qquad \operatorname{var}(X_k) = \sigma^2 \qquad \text{for } k = 1, 2, \ldots, n \tag{12.39}$$

To estimate the mean of the population distribution, we will use the sample mean defined in the previous section. The sample mean is the arithmetic average

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$$

We stated that statistics are random variables. Hence, we can compute their mean and variance. Using (12.39), we find that the mean of the sample mean is

$$E(\overline{X}_n) = E\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n}\sum_{k=1}^{n} E(X_k) = \frac{1}{n}(n\mu) = \mu$$

From the independence of the observations, in addition to (12.39), the variance of the sample mean is

$$\operatorname{var}(\overline{X}_n) = \operatorname{var}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} \operatorname{var}(X_k) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}$$

We see from these equations that the expected value of the sample mean is equal to the population mean. The spread of the distribution of $\overline{X}_n$ is described by the variance of $\overline{X}_n$. Since that variance becomes smaller as the sample size increases ($\sigma^2/n \to 0$ as $n \to \infty$), we conclude that the sample mean of large samples shows less variation about its mean than does the sample mean of small samples. This conclusion implies that the larger the sample size, the more accurately the mean of the population can be estimated. In fact, invoking the weak law of large numbers from the previous section, we find that

$$\overline{X}_n \to \mu \quad \text{in probability as } n \to \infty$$

This relationship justifies the use of $\overline{X}_n$ as an estimate for the mean of the distribution. Since $E(\overline{X}_n) = \mu$, we say that $\overline{X}_n$ is an **unbiased estimator** for μ.

(3) NIST, founded in 1901, is a nonregulatory federal agency within the U.S. Commerce Department's Technology Administration. NIST's mission is "to promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve quality of life."

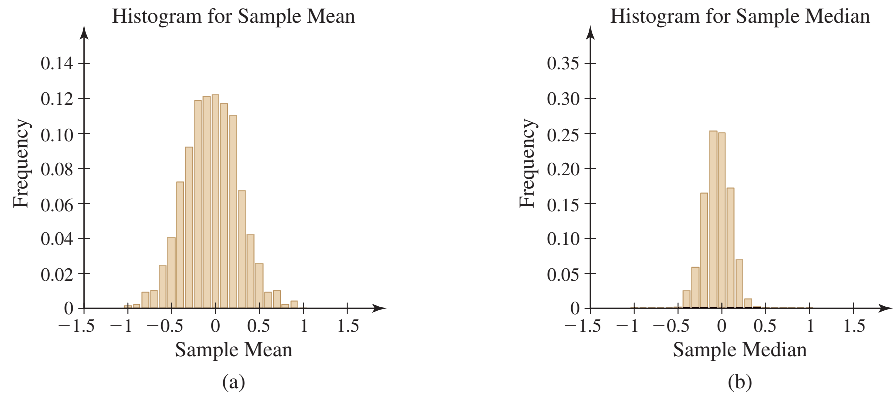


Figure 12.54 (a) The histogram of sample means when the sample size is 10. (b) The histogram of sample means when the sample size is 50.

We illustrate the behavior of the sample mean as a function of sample size in the following simulation, where we draw random samples from a standard normal distribution (i.e., $\mu = 0$ and $\sigma^2 = 1$): Figure 12.54(a) shows a histogram for the sample means of 1000 random samples, each of size 10; Figure 12.54(b) shows a histogram for the sample means of 1000 random samples, each of size 50. The simulations confirm that as the sample size increases, the variance of the sample mean decreases, resulting in a narrower histogram. If the sample is drawn from a normal distribution with mean 0 and variance 1, then the sample mean is normally distributed with mean 0 and variance $1/n$, where n is the sample size. When the distribution of the population is not normal, but has a finite mean and variance, we can invoke the central limit theorem to conclude that the sample mean is approximately normal. We can view each histogram as an approximation to the theoretical distribution of the sample mean.

EXAMPLE 4 Assume that we draw a random sample from a population that consists of individuals whose lifetimes are exponentially distributed with a mean of five years.

(a) Denote by X the random variable that is exponentially distributed with a mean of five years. What is the probability density function of the random variable that describes the lifetime of an individual in this population? What are the mean and the variance of this random variable?

(b) If you took a large number of random samples, each of size 50, from this population and graphed the histogram for the sample mean of these random samples, what do you expect the histogram to look like? Compute the mean and the variance of the sample mean.

Solution **(a)** The probability density of an exponential distribution with mean 5 has parameter $\lambda = 0.2$ and is given by

$$f(x) = 0.2 \exp(-0.2x)$$

The mean of an exponentially distributed random variable is $1/\lambda$ and the variance is $1/\lambda^2$. Hence, $E(X) = 5$ and $\text{var}(X) = 25$. The graph of the probability density is shown in Figure 12.55(a).

(b) Since the mean of the population distribution are finite, the sample mean of a large sample is approximately normally distributed with mean equal to the population mean and variance equal to the population variance divided by the sample size. The sample size is 50, so the mean of the sample mean is 5 and the variance of the sample mean is $25/50 = 0.5$. To verify these ideas, we simulated 10,000 samples, each of size 50, and displayed the sample means in the histogram in Figure 12.55(b). Note that the histogram resembles a normal distribution and that it is centered around 5, the mean of the sample mean. •

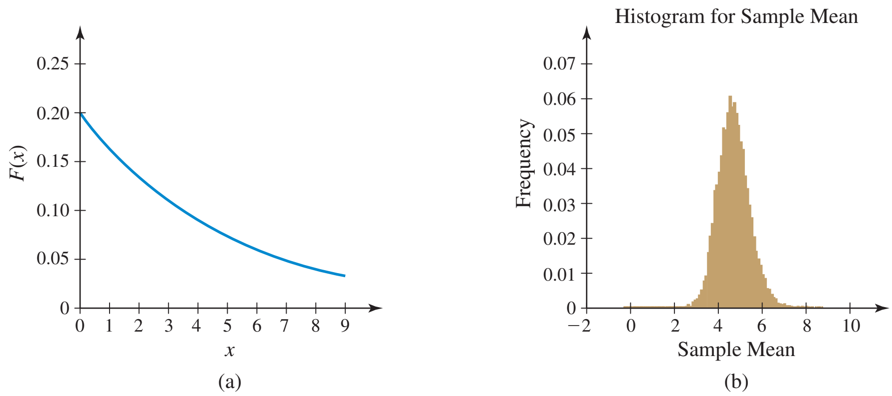


Figure 12.55 (a) The probability density of an exponentially distributed random variable with mean 5 years. (b) The histogram of sample means when the sample size is 50.

A Remark on Using the Sample Mean to Estimate the Mean. You might wonder why we used the sample mean, and not some other quantity, to estimate the mean of the population distribution. Statisticians have established criteria to assess the quality of an estimator. We mentioned one such criterion already: It is desirable to use unbiased estimators. The sample mean is unbiased, since its expected value is equal to the parameter it is supposed to estimate, the population mean. A biased estimator introduces a systematic error. Later we will see an example in which we will define an estimator that will turn out to be biased, but then we will find a way to remove the bias.

Another criterion is as follows: Choose an estimator with as small a variance as possible. We illustrate the application of this criterion in a population distribution that is normal with mean μ and variance σ^2. Suppose we wish to estimate the mean of the distribution. We already know that the sample mean is an unbiased estimator of the mean. In Subsection 12.7.1, we mentioned that the sample median is also an unbiased estimator of the mean (although not necessarily if the population distribution is not normal). If we know that the population distribution is normal, why don't we choose the sample median instead of the sample mean? After all, the sample median seems easier to compute. We saw in Subsection 12.7.1 that the variance of the sample mean is smaller than the variance of the sample median. If the sample size is denoted by n, then the former is equal to σ^2/n, and the latter is equal to $\pi\sigma^2/(2n) \approx 1.508\sigma^2/n$, if n is sufficiently large. A smaller variance increases the precision of our estimate, so an estimator with a smaller variance is preferable.

Point Estimates of Proportions. Estimating proportions is a special case of estimating means. Examples of estimating proportions are, for instance, estimating the proportion of white-flowering pea plants in a $Cc \times Cc$ crossing or estimating the proportion of patients in a clinical study who had a recurrence of a disease after a fixed length of time. If we take a random sample of size n from a population of which a proportion p has a certain characteristic, then the number of observations in the sample with this characteristic is binomially distributed with parameters n and p. If we denote this quantity by B_n, then we have

$$P(B_n = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for } k = 0, 1, 2, \ldots, n$$

To derive an estimator for the parameter p, we consider each single observation in the sample as a success or a failure, depending on whether the observation has the characteristic under investigation. We set

$$X_k = \begin{cases} 1 & \text{if the } k\text{th observation is a success} \\ 0 & \text{if the } k\text{th observation is a failure} \end{cases}$$

Then $B_n = \sum_{k=1}^{n} X_k$ is the total number of successes in the sample and p is the probability of success. The sample mean $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ is then the fraction of successes in the sample. We find that

$$E(\overline{X}_n) = \frac{1}{n}\sum_{k=1}^{n} E(X_k) = \frac{1}{n}(np) = p$$

and

$$\text{var}(\overline{X}_n) = \text{var}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}n(\text{var}(X_1))$$

$$= \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

The sample mean $\overline{X}_n$ will serve as an estimator for the probability of success p. It is customary to use $\hat{p}$ instead of $\overline{X}_n$; that is, we denote the estimate for p by $\hat{p}$ (read "p hat"). If we observe k successes in a sample of size n, then

$$\hat{p} = \frac{k}{n}$$

with

$$E(\hat{p}) = p \qquad \text{and} \qquad \text{var}(\hat{p}) = \frac{p(1-p)}{n}$$

The next example considers Mendel's experiment of crossing pea plants.

EXAMPLE 5 To estimate the probability of white-flowering pea plants in a $Cc \times Cc$ crossing, Mendel randomly crossed red-flowering pea plants of genotype Cc. He obtained 705 plants with red flowers and 224 plants with white flowers. Estimate the probability of a white-flowering pea plant in a $Cc \times Cc$ crossing.

Solution The sample mean is

$$\overline{X}_n = \frac{224}{224 + 705} \approx 0.24$$

The estimate for the probability of a white-flowering pea plant in a $Cc \times Cc$ crossing is therefore $\hat{p} = 0.24$. We know from the laws of inheritance that the expected value of $\overline{X}_n$ is $p = 0.25$. •

The next example illustrates the use of estimating proportions in a clinical trial.

EXAMPLE 6 The Women's Health Initiative Steering Committee (2004) conducted a clinical study on the effects of estrogen-alone therapy in postmenopausal women who had undergone hysterectomies. The study was halted in July 2002 because the risks exceeded the benefits. Of 5310 patients who had been randomly assigned to receive the hormone therapy, 158 suffered strokes during the 81.6 months of follow-up, whereas 118 of 5429 patients in the placebo group suffered strokes during the 81.9 months of follow-up. Estimate the number of strokes per 10,000 person-years in each group. Which group had the higher incidence of stroke?

Solution We first need to convert the number of patients in each group into person-years. The hormone therapy group has $(81.6)(5310)/12 = 36{,}108$ person-years; the placebo group has $(81.9)(5429)/12 = 37{,}053$ person-years. The incidence in the hormone therapy group is therefore $158/36{,}108 \approx 0.0044$, or 44 per 10,000 person-years. The incidence in the placebo group is $118/37{,}053 \approx 0.0032$, or 32 per 10,000 person-years. The group that received hormone therapy had a higher risk of stroke. •

Point Estimates of Variances. Recall the sample variance

$$S_n^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X}_n)^2$$

We will show that dividing by $n-1$ (instead of by n) yields an unbiased estimator for the variance. To compute the mean of the sample variance, we need the following identity: For any $c \in \mathbf{R}$,

$$\sum_{k=1}^{n}(X_k - c)^2 = \sum_{k=1}^{n}(X_k - \overline{X}_n)^2 + n(\overline{X}_n - c)^2 \tag{12.40}$$

To justify Equation (12.40), we expand its left-hand side after adding $0 = \overline{X}_n - \overline{X}_n$ inside the parentheses:

$$\begin{aligned}\sum_{k=1}^{n}(X_k - c)^2 &= \sum_{k=1}^{n}(X_k - \overline{X}_n + \overline{X}_n - c)^2 \\ &= \sum_{k=1}^{n}[(X_k - \overline{X}_n)^2 + 2(X_k - \overline{X}_n)(\overline{X}_n - c) + (\overline{X}_n - c)^2] \\ &= \sum_{k=1}^{n}(X_k - \overline{X}_n)^2 + 2(\overline{X}_n - c)\sum_{k=1}^{n}(X_k - \overline{X}_n) + n(\overline{X}_n - c)^2\end{aligned}$$

Since $\sum_{k=1}^{n}(X_k - \overline{X}_n) = 0$, the middle term is equal to 0, and (12.40) follows.

If we set $c = \mu$ in (12.40) and rearrange the equation, we obtain

$$\sum_{k=1}^{n}(X_k - \overline{X}_n)^2 = \sum_{k=1}^{n}(X_k - \mu)^2 - n(\overline{X}_n - \mu)^2$$

Taking expectations on both sides and using, on the right-hand side, the fact that the expectation of a sum is the sum of the expectations, we find that

$$E\left(\sum_{k=1}^{n}(X_k - \overline{X}_n)^2\right) = \sum_{k=1}^{n}E(X_k - \mu)^2 - nE(\overline{X}_n - \mu)^2$$

Now, $\sum_{k=1}^{n}(X_k - \overline{X}_n)^2 = (n-1)S_n^2$, $E(X_k - \mu)^2 = \sigma^2$, and $E(\overline{X}_n - \mu)^2 = \mathrm{var}(\overline{X}_n) = \frac{1}{n}\sigma^2$. Hence,

$$(n-1)E(S_n^2) = n\sigma^2 - \sigma^2 = (n-1)\sigma^2$$

and, therefore,

$$E(S_n^2) = \sigma^2$$

which shows that S_n^2 is an unbiased estimate of the variance of the population. This is the reason that we divided by $n-1$ instead of n when we computed the sample variance. We will not compute the variance of the sample variance; it is given by a complicated formula. One can show, however, that the variance of the sample variance goes to 0 as the sample size becomes infinite. Hence, invoking the weak law of large numbers, we find that

$$S_n^2 \to \sigma^2 \quad \text{in probability as } n \to \infty$$

The next example illustrates how we would estimate the mean and the variance of a characteristic of a population.

EXAMPLE 7 Suppose that a computer generates the following sample of independent observations from a population:

$$0.0201, 0.8918, 0.9619, 0.1713, 0.0357,$$
$$0.6325, 0.4276, 0.2517, 0.2330, 0.6754$$

Estimate the mean and the variance of these observations.

Solution To estimate the mean, we compute the sample mean $\overline{X}_n$. We sum the 10 numbers in the sample and divide the result by 10, which yields

$$\overline{X}_n = 0.4301$$

Thus, our estimate of the mean is 0.4301.

To estimate the variance, we compute the sample variance S_n^2. We square the difference between each sample point and the sample mean and add the results. We then divide the resulting number by 9 to obtain

$$S_n^2 = 0.1176$$

Thus, our estimate of the variance is 0.1176. •

Confidence Intervals. Earlier, we learned that the sample mean varies from sample to sample. The variation of the distribution of the sample mean is described by the **standard error** (S.E.), also denoted by $S_{\overline{x}}$ to indicate that it is the standard deviation of the sample mean.

The definition of the standard error is motivated by the following considerations (again, we assume that the mean and the variance of the population distribution are finite): To estimate the population mean μ, we use the sample mean $\overline{X}_n$. The variance of $\overline{X}_n$ gives us an idea as to how much the distribution of $\overline{X}_n$ varies. Now,

$$\operatorname{var}(\overline{X}_n) = \operatorname{var}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\operatorname{var}\left(\sum_{k=1}^{n} X_k\right)$$

Since the X_k are independent, the variance of the sum is the sum of the variances; moreover, all X_k have the same distribution. Hence,

$$\operatorname{var}(\overline{X}_n) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n} \tag{12.41}$$

The variance of $\overline{X}_n$ thus depends on another population parameter: the variance of X_n (i.e., σ^2). In problems where we wish to estimate the mean, we typically don't know the variance either. If the sample size is large, however, the sample variance will be close to the population variance. We can therefore approximate the variance of $\overline{X}_n$ by replacing σ^2 in (12.41) by S_n^2, giving S_n^2/n. The standard error is then the square root of that expression:

$$\text{S.E.} = S_{\overline{x}} = \frac{S_n}{\sqrt{n}}$$

The next example illustrates how to determine a sample mean and its standard error.

EXAMPLE 8 In a sample of six leaves from a morning glory plant that is infested with aphids, the following numbers of aphids per leaf are found: 12, 27, 17, 35, 14, and 18. Find the sample mean, the sample variance, and the standard error.

Solution The sample size is $n = 6$. We use the formulas

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k \qquad \text{and} \qquad S_n^2 = \frac{1}{n-1}\left[\sum_{k=1}^{n} X_k^2 - \frac{1}{n}\left(\sum_{k=1}^{n} X_k\right)^2\right]$$

with $n = 6$ to construct the following table:

$\sum_{k=1}^{n} X_k$	$\sum_{k=1}^{n} X_k^2$	$\overline{X}_n$	S_n^2	$\frac{S_n}{\sqrt{n}}$
123	2907	20.5	77.1	3.58

We thus find that S.E. = 3.58. •

To find the standard error in the case of k successes in a sample of size n, note that

$$S_n^2 = \frac{1}{n-1}\left[\sum_{j=1}^{n} X_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n} X_j\right)^2\right]$$

$$= \frac{1}{n-1}\left(k - \frac{k^2}{n}\right)$$

$$= \frac{n}{n-1}\frac{k}{n}\left(1 - \frac{k}{n}\right) = \frac{n\hat{p}(1-\hat{p})}{n-1}$$

The standard error of the sample mean p is therefore

$$S_{\hat{p}} = \frac{S_n}{\sqrt{n}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}} \tag{12.42}$$

In some textbooks, you will find the standard error for proportions stated as

$$S_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \tag{12.43}$$

For n large, the two numbers are very close, so that it does not matter which one you use.

EXAMPLE 5 **(continued)** In Example 5, we estimated the probability of white-flowering pea plants in Mendel's $Cc \times Cc$ crossing that produced 705 plants with red flowers and 224 plants with white flowers. We obtained $\hat{p} = 0.24$ as an estimate for the probability of a white-flowering pea plant in this crossing. Find the standard error of the sample mean p.

Solution The standard error of the sample mean is

$$\text{S.E.} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.24)(0.76)}{705+224}} \approx 0.014$$

We can thus report the result as 0.24 ± 0.014. •

In the scientific literature, variability is often reported as either "Mean ± S.D." or "Mean ± S.E." Since S.E. = S.D./$\sqrt{n}$, the two methods of describing variability are not the same. You must explain which method you are using.

EXAMPLE 9 During 2000–2003, the temperature in Medicine Lake, Minnesota, was measured repeatedly at a 3-m depth. The following table lists the mean temperature (in degree Celsius) and standard error for every other month from June through October, together with number of sample points (n) for each month:

Month	**Mean ± S.E.**	***n***
June	21.5 ± 1.52	6
August	25.1 ± 1.45	9
October	13.5 ± 1.42	7

Convert "Mean ± S.E." to "Mean ± S.D."

Solution Since S.E. = S.D./$\sqrt{n}$, we multiply the S.E. given in the table by $\sqrt{n}$. For instance, to obtain the S.D. for June, we calculate S.D. $= 1.52\sqrt{6} = 3.72$. We get the following results:

Month	**Mean ± S.D.**	***n***
June	21.5 ± 3.72	6
August	25.1 ± 4.35	9
October	13.5 ± 3.76	7

•

Interpreting Mean ± S.E.. What does "Mean ± S.E." mean? When we write "Mean ± S.E.," we specify an interval, namely, [Mean − S.E., Mean + S.E.]. Since we use "Mean" ($= \overline{X}_n$) as an estimate for the population mean μ, we would like this interval to contain μ. Surely, since $\overline{X}_n$ is a random variable, if we took repeated samples and computed such intervals for each sample, not *all* the intervals would contain μ. But maybe we can at least find out what fraction of these intervals are likely to contain the population mean μ. In other words, before taking the sample, we might wish to know what the probability is that the interval [Mean − S.E., Mean + S.E.], or, more generally, [Mean − a S.E., Mean + a S.E.], where a is a positive constant, will contain the actual value of μ. To be concrete, we will try to determine a so that this probability is equal to 0.95.

If the sample size n is large, it follows from the central limit theorem that

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$$

is approximately standard normally distributed. If Z is standard normally distributed, then

$$P(-1.96 \le Z \le 1.96) = 0.95$$

Hence, the event

$$-1.96 \le \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \le 1.96 \tag{12.44}$$

has a probability of approximately 0.95 for n sufficiently large.

Rearranging terms in (12.44), we find that

$$-1.96\frac{\sigma}{\sqrt{n}} \le \overline{X}_n - \mu \le 1.96\frac{\sigma}{\sqrt{n}}$$

or

$$\overline{X}_n - 1.96\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X}_n + 1.96\frac{\sigma}{\sqrt{n}}$$

We can thus write

$$P\left(\overline{X}_n - 1.96\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X}_n + 1.96\frac{\sigma}{\sqrt{n}}\right) \approx 0.95 \quad \text{for large } n \tag{12.45}$$

This equation tells us that if we repeatedly draw random samples of size n from a population with mean μ and standard deviation σ, then, in about 95% of the samples, the interval $\left[\overline{X}_n - 1.96\frac{\sigma}{\sqrt{n}}, \overline{X}_n + 1.96\frac{\sigma}{\sqrt{n}}\right]$ will contain the true mean μ. Such an interval is referred to as a **95% confidence interval**.

Notice that this interval contains the parameter σ. If we do not know μ, we probably do not know σ either. We might then wish to replace σ by the square root of the sample variance, denoted by S_n. Fortunately, when n is large, S_n will be very close to σ and (12.45) holds approximately when σ is replaced by S_n.

We thus find that, for large n,

$$P\left(\overline{X}_n - 1.96\frac{S_n}{\sqrt{n}} \le \mu \le \overline{X}_n + 1.96\frac{S_n}{\sqrt{n}}\right) \approx 0.95$$

or, with the event rewritten in interval notation,

$$P\left(\mu \in \left[\overline{X}_n - 1.96\frac{S_n}{\sqrt{n}}, \overline{X}_n + 1.96\frac{S_n}{\sqrt{n}}\right]\right) \approx 0.95$$

The interval in this expression is of the form

$$[\text{Mean} - (1.96)\text{S.E.}, \text{Mean} + (1.96)\text{S.E.}] \tag{12.46}$$

We have succeeded in determining what fraction of intervals of the form (12.46) contain the mean. That is, if n is large and we take repeated samples each of size n, then approximately 95% of the intervals [Mean − (1.96)S.E., Mean + (1.96)S.E.] would contain the true mean. [Or, equivalently, before we take the sample, the probability that the interval in (12.46) will contain the actual value of μ is 0.95.] If we wanted 99% of such intervals to contain the true mean, we would need to replace the factor 1.96 by

2.58, since if Z is standard normally distributed, then $P(-2.58 \le Z \le 2.58) = 0.99$ (95% and 99% are the most frequently used percentages for confidence intervals). Because $P(-1 \le Z \le 1) = 0.68$, we conclude that if the sample size is large, approximately 68% of intervals of the form [Mean − S.E., Mean + S.E.] contain the population mean μ.

EXAMPLE 10 In the study by ***Dyck et al. (1999)*** mentioned at the beginning of Section 12.7, the weight (in kg) of the group of 149 Type 2 diabetes patients was reported as Mean ± S.D. = 84.4 ± 16.9. Find a 99% confidence interval for the mean.

Solution If Z is a standard normally distributed random variable, we need to find z such that

$$P(-z \le Z \le z) = 0.99$$

Now,

$$P(-z \le Z \le z) = 2\Phi(z) - 1 = 0.99$$

where $\Phi(z)$ denotes the distribution function of a standard normal distribution. Hence,

$$\Phi(z) = \frac{1.99}{2} = 0.995$$

and it follows that $z = 2.58$. With S.D. = 16.9 and $n = 149$, we find that

$$\text{S.E.} = \frac{16.9}{\sqrt{149}} \approx 1.38$$

Therefore, the 99% confidence interval is of the form

$$[84.8 - (2.58)(1.38), 84.8 + (2.58)(1.38)] = [81.2, 88.4]$$

12.7.3 Linear Regression

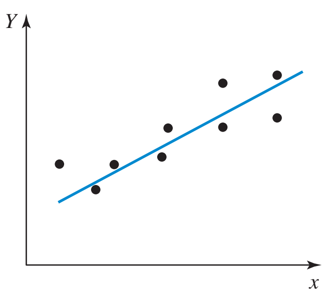


Figure 12.56 A straight line fitted to data.

We previously met least squares fitting (or regression) in Chapter 10. Now we will revisit this topic from the viewpoint of probability. We frequently meet plots that fit a straight line to data (as shown in Figure 12.56). A linear model is used to describe the relationship between the quantities on the horizontal and the vertical axes. We denote the quantity on the horizontal axis by x and the quantity on the vertical axis by Y. We think of x as a particular variable that is under the control of the experimenter and of Y as the response. In measurements of Y, errors are typically present, so that the data points will not lie exactly on the straight line (even if the linear model is correct) but will be scattered around it; that is, Y is not completely determined by x. The degree of scatter is an indication of how much random variation there is. In what follows, we will see how to separate the random variation from the actual relationship between the two quantities. We will discuss one particular model.

We assume that x is under the control of the experimenter to the extent that it can be measured without error. The response Y, however, shows random variation. We assume the linear model

$$Y = a + bx + \epsilon$$

where ϵ is a normal random variable representing the **error**, which has mean 0 and standard deviation σ. The standard deviation of the error does not depend on x and is thus the same for all values of x.

Our goal is to estimate a and b from data that consist of the points (x_i, y_i), $i = 1, 2, \ldots, n$. The approach will be to choose a and b such that the sum of the squared deviations

$$h(a, b) = \sum_{k=1}^{n} [y_k - (a + bx_k)]^2$$

is minimized. The deviations $y_k - (a + bx_k)$ are called **residuals**. The procedure of finding a and b is called the **method of least squares** and is illustrated in Figure 12.57. The resulting straight line is called the **least square line** (or **linear regression line**).

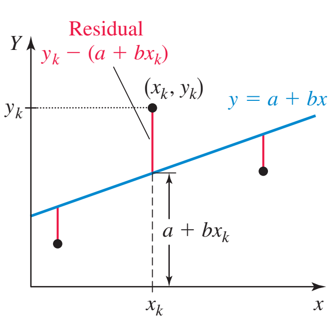


Figure 12.57 The line $Y = a + bx$ is chosen so that the sum of the squared residuals is minimized.

EXAMPLE 11 Given the three points (0, 2), (1, 0), and (2, 1), use the method of least squares to find the least square line.

Solution We wish to find a straight line of the form $y = a + bx$. For given values of a and b, the residuals are

$$2 - (a + 0b) \qquad 0 - (a + b) \qquad 1 - (a + 2b)$$

and the sum of their squares is

$$\begin{aligned}
&(2-a)^2 + (a+b)^2 + (1-a-2b)^2\\
&= (4 - 4a + a^2) + (a^2 + 2ab + b^2) + (1 + a^2 + 4b^2 - 2a - 4b + 4ab)\\
&= 5 - 6a + 3a^2 + 6ab + 5b^2 - 4b\\
&= (2b^2 + 2b) + (3 + 3b^2 + 3a^2 + 6ab - 6a - 6b) + 2\\
&= 2(b^2 + b) + 3(1 + b^2 + a^2 + 2ab - 2a - 2b) + 2
\end{aligned}$$

Grouping the terms in this way allows us to complete the squares; we find that the sum of the squares is then equal to

$$2\left(b + \frac{1}{2}\right)^2 + 3(1 - a - b)^2 + \frac{3}{2} \tag{12.47}$$

(If this looks like magic, don't worry; we will derive a general formula for a and b shortly.) Since (12.47) consists of two squares (plus a constant term), the expression is minimized when the two squares are both equal to 0. Accordingly, we solve

$$\begin{aligned}
b + \frac{1}{2} &= 0\\
1 - a - b &= 0
\end{aligned}$$

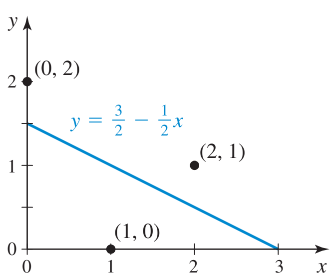


Figure 12.58 The least square line together with the three data points of Example 11.

which yields $b = -1/2$ and $a = 3/2$. Therefore, the least square line is of the form

$$y = \frac{3}{2} - \frac{1}{2}x$$

This line, togetherwith the given three points, is shown in Figure 12.58. •

We will now derive the general formula for finding a and b. The basic steps will be similar to those in Example 11. We will first rewrite the residuals. We set

$$\overline{x} = \frac{1}{n}\sum_{k=1}^{n} x_k \qquad \text{and} \qquad \overline{y} = \frac{1}{n}\sum_{k=1}^{n} y_k$$

and

$$y_k - (a + bx_k) = (y_k - \overline{y}) + (\overline{y} - a - b\overline{x}) - b(x_k - \overline{x})$$

In what follows, we will simply write $\sum$ instead of $\sum_{k=1}^{n}$. If we square the expression and sum over k, we find that

$$\begin{aligned}
\sum[y_k - (a + bx_k)]^2 = &\sum(y_k - \overline{y})^2 + n(\overline{y} - a - b\overline{x})^2\\
&+ b^2\sum(x_k - \overline{x})^2 - 2b\sum(x_k - \overline{x})(y_k - \overline{y})\\
&+ 2(\overline{y} - a - b\overline{x})\sum(y_k - \overline{y})\\
&- 2b(\overline{y} - a - b\overline{x})\sum(x_k - \overline{x})
\end{aligned} \tag{12.48}$$

The last two terms are equal to 0.

Next, we introduce notation to simplify our derivation. Let

$$\mathrm{SS}_{xx} = \sum (x_k - \overline{x})^2 = \sum x_k^2 - \frac{(\sum x_k)^2}{n}$$

$$\mathrm{SS}_{yy} = \sum (y_k - \overline{y})^2 = \sum y_k^2 - \frac{(\sum y_k)^2}{n}$$

$$\mathrm{SS}_{xy} = \sum (x_k - \overline{x})(y_k - \overline{y}) = \sum x_k y_k - \frac{(\sum x_k)(\sum y_k)}{n}$$

Using this notation, we can write the right-hand side of (12.48) as

$$\mathrm{SS}_{yy} + n(\overline{y} - a - b\overline{x})^2 + b^2\mathrm{SS}_{xx} - 2b\mathrm{SS}_{xy}$$

The last two terms suggest that we should complete the square:

$$\mathrm{SS}_{yy} + n(\overline{y} - a - b\overline{x})^2 + \mathrm{SS}_{xx}\left(b^2 - 2b\frac{\mathrm{SS}_{xy}}{\mathrm{SS}_{xx}} + \left(\frac{\mathrm{SS}_{xy}}{\mathrm{SS}_{xx}}\right)^2\right) - \frac{(\mathrm{SS}_{xy})^2}{\mathrm{SS}_{xx}}$$

$$= n(\overline{y} - a - b\overline{x})^2 + \mathrm{SS}_{xx}\left(b - \frac{\mathrm{SS}_{xy}}{\mathrm{SS}_{xx}}\right)^2 + \mathrm{SS}_{yy} - \frac{(\mathrm{SS}_{xy})^2}{\mathrm{SS}_{xx}}$$

As in Example 11, we succeeded in writing the sum of the squared deviations as a sum of two squares plus an additional term. We can minimize this expression by setting each squared expression equal to 0:

$$\overline{y} - a - b\overline{x} = 0$$

$$b - \frac{\mathrm{SS}_{xy}}{\mathrm{SS}_{xx}} = 0$$

Solving for a and b yields

$$b = \frac{\mathrm{SS}_{xy}}{\mathrm{SS}_{xx}}$$

$$a = \overline{y} - b\overline{x}$$

The right-hand sides serve as estimates of a and b, respectively denoted by $\hat{a}$ and $\hat{b}$. Summarizing, we have the following result:

The least square line (or linear regression line) is given by

$$y = \hat{a} + \hat{b}x$$

with

$$\hat{b} = \frac{\sum_{k=1}^{n}(x_k - \overline{x})(y_k - \overline{y})}{\sum_{k=1}^{n}(x_k - \overline{x})^2} \tag{12.49}$$

$$\hat{a} = \overline{y} - \hat{b}\overline{x} \tag{12.50}$$

We illustrate finding $\hat{a}$ and $\hat{b}$ in the next example.

EXAMPLE 12 Fit a linear regression line through the points

$$(1, 1.62), (2, 3.31), (3, 4.57), (4, 5.42), (5, 6.71)$$

Solution To facilitate the computation, we construct the following table:

x_k	y_k	$x_k - \overline{x}$	$y_k - \overline{y}$	$(x_k - \overline{x})(y_k - \overline{y})$
1	1.62	−2	−2.706	5.412
2	3.31	−1	−1.016	1.016
3	4.57	0	0.244	0
4	5.42	1	1.094	1.094
5	6.71	2	2.384	4.768

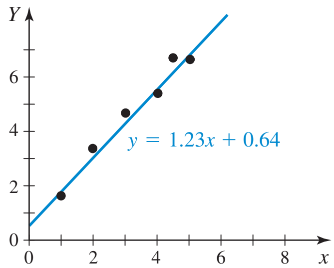


Figure 12.59 The linear regression line and the data points of Example 12.

Now,

$$\hat{b} = \frac{12.29}{10} = 1.229$$

$$\hat{a} = 4.326 - (1.229)(3) = 0.639$$

Hence, the linear regression line is given by

$$y = 1.23x + 0.64$$

This line and the given datapoints are shown in Figure 12.59. •

So far, we have only reproduced the results from Chapter 10; but in addition to fitting the straight line, we might want to measure how good the fit is. To this end, we will define a quantity known as the **coefficient of determination**. We motivate its definition as follows: We start with a set of observations (x_k, y_k), $k = 1, 2, \ldots, n$, and assume the linear model $Y = a + bx + \epsilon$. We set

$$\hat{y}_k = \hat{a} + \hat{b}x_k \qquad \text{and} \qquad \overline{y} = \frac{1}{n}\sum_{k=1}^{n} y_k$$

We think of $\hat{y}_k$ as the expected response under the linear model if $x = x_k$. Now, $y_k - \overline{y}$ is the deviation of the observation from the sample mean, $y_k - \hat{y}_k$ is the deviation of the observation from the expected response under the linear model, and $\hat{y}_k - \overline{y}$ is the deviation of the expected response under the linear model from the sample mean. The deviation $\hat{y}_k - \overline{y}$ can be thought of as being explained by the model, and the deviation $y_k - \hat{y}_k$ can be thought of as the unexplained part due to random variation (the stochastic error). We can write

$$y_k - \overline{y} = (\hat{y}_k - \overline{y}) + (y_k - \hat{y}_k) \tag{12.51}$$

If we look at $\sum(y_k - \overline{y})^2$, which is the total sum of the squared deviations, and use (12.51), we find

$$\begin{aligned}\sum(y_k - \overline{y})^2 &= \sum[(\hat{y}_k - \overline{y}) + (y_k - \hat{y}_k)]^2 \\ &= \sum(\hat{y}_k - \overline{y})^2 + 2\sum(\hat{y}_k - \overline{y})(y_k - \hat{y}) \\ &\quad + \sum(y_k - \hat{y}_k)^2\end{aligned} \tag{12.52}$$

We want to show that $\sum(\hat{y}_k - \overline{y})(y_k - \hat{y}) = 0$. To do so, we observe that

$$\hat{y}_k - \overline{y} = (\hat{a} + \hat{b}x_k) - (\hat{a} + \hat{b}\overline{x}) = \hat{b}(x_k - \overline{x})$$

$$y_k - \hat{y}_k = (y_k - \overline{y}) - (\hat{y}_k - \overline{y}) = (y_k - \overline{y}) - \hat{b}(x_k - \overline{x})$$

Therefore,

$$\begin{aligned}\sum(\hat{y}_k - \overline{y})(y_k - \hat{y}_k) &= \sum\hat{b}(x_k - \overline{x})[(y_k - \overline{y}) - \hat{b}(x_k - \overline{x})] \\ &= \hat{b}\sum(x_k - \overline{x})(y_k - \overline{y}) - \hat{b}^2\sum(x_k - \overline{x})^2\end{aligned}$$

Using (12.49) to substitute for one of the $\hat{b}$'s in the last term, we obtain

$$\begin{aligned}\sum(\hat{y}_k - \overline{y})(y_k - \hat{y}_k) &= \hat{b}\sum(x_k - \overline{x})(y_k - \overline{y}) \\ &\quad - \hat{b}\frac{\sum(x_k - \overline{x})(y_k - \overline{y})}{\sum(x_k - \overline{x})^2}\sum(x_k - \overline{x})^2 \\ &= 0\end{aligned} \tag{12.53}$$

This equation allows us to partition the total sum of squares into the explained and the unexplained sums of squares. Accordingly, continuing with (12.52) and using (12.53), we find that

$$\underbrace{\sum(y_k - \overline{y})^2}_{\text{total}} = \underbrace{\sum(\hat{y}_k - \overline{y})^2}_{\text{explained}} + \underbrace{\sum(y_k - \hat{y}_k)^2}_{\text{unexplained}}$$

The ratio

$$\frac{\text{explained}}{\text{total}} = \frac{\sum(\hat{y}_k - \overline{y})^2}{\sum(y_k - \overline{y})^2}$$

is therefore the proportion of variation that is explained by the model. It is denoted by r^2 and is called the *coefficient of determination*. With $\hat{y}_k - \overline{y} = \hat{b}(x_k - \overline{x})$ and $\hat{b}$ given in (12.49), the coefficient of determination can be written as

$$r^2 = (\hat{b})^2 \frac{\sum(x_k - \overline{x})^2}{\sum(y_k - \overline{y})^2} = \left(\frac{\sum(x_k - \overline{x})(y_k - \overline{y})}{\sum(x_k - \overline{x})^2}\right)^2 \frac{\sum(x_k - \overline{x})^2}{\sum(y_k - \overline{y})^2}$$

$$= \frac{\left[\sum(x_k - \overline{x})(y_k - \overline{y})\right]^2}{\sum(x_k - \overline{x})^2 \sum(y_k - \overline{y})^2}$$

We summarize this result as follows:

> The coefficient of determination is given by
>
> $$r^2 = \frac{\left[\sum(x_k - \overline{x})(y_k - \overline{y})\right]^2}{\sum(x_k - \overline{x})^2 \sum(y_k - \overline{y})^2}$$
>
> and represents the proportion of variation that is explained by the model.

Returning to Example 12, we find that

$$r^2 = \frac{(12.29)^2}{(10)(15.29)} = 0.988$$

That is, 98.8% of the variation is explained by the model.

Since r^2 is the ratio of explained to total variation, it follows that $r^2 \leq 1$. Furthermore, since r^2 is the square of an expression, it is always nonnegative. That is, we have

$$0 \leq r^2 \leq 1$$

The closer r^2 is to 1, the more closely the data points follow the straight line resulting from the linear model. In the extreme case, when $r^2 = 1$, all points lie on the line; there is no random variation.

Section 12.7 Problems

12.7.1

1. The following data represent the number of aphids per plant found in a sample of 10 plants:

17, 13, 21, 47, 3, 6, 12, 25, 0, 18

Find the median, the sample mean, and the sample variance.

2. The following data represent the number of seeds per flower head in a sample of nine flowering plants:

27, 39, 42, 18, 21, 33, 45, 37, 21

Find the median, the sample mean, and the sample variance.

3. The following data represent the age of patients in a clinical trial:

28, 45, 34, 36, 30, 42, 35, 45, 38, 27

Find the median, the sample mean, and the sample variance.

4. The following data represent blood cholesterol levels, in mg/dL, of patients in a clinical trial:

174, 138, 212, 203, 194, 245, 146, 149, 164, 209, 158

Find the median, the sample mean, and the sample variance.

5. The following data represent the frequency distribution of seed numbers per flower head in a flowering plant:

Seed Number	Frequency
9	37
10	48
11	53
12	49
13	61
14	42
15	31

Calculate the sample mean and the sample variance.

6. The following data represent the frequency distribution of the numbers of days that it took a certain ointment to clear up a skin rash:

Number of Days	Frequency
1	2
2	7
3	9
4	27
5	11
6	5

Calculate the sample mean and the sample variance.

7. The following data represent the relative frequency distribution of clutch size in a sample of 300 laboratory guinea pigs:

Clutch Size	Relative Frequency
2	0.05
3	0.09
4	0.12
5	0.19
6	0.23
7	0.12
8	0.13
9	0.07

Calculate the sample mean and the sample variance.

8. The following data represent the relative frequency distribution of clutch size in a sample of 42 mallards:

Clutch Size	Relative Frequency
6	0.10
7	0.24
8	0.29
9	0.21
10	0.16

Calculate the sample mean and the sample variance.

9. Assume that a population consists of the three numbers 1, 6, and 8. List all samples of size 2 that can be drawn from this population with replacement, and find the sample mean of each sample.

10. Use a graphing calculator to generate five samples, each of size 6, from a uniform distribution over the interval (0,1). Compute the sample means of each sample.

11. Let $(X_1, X_2, \ldots, X_n)$ denote a sample of size n. Show that

$$\sum_{k=1}^{n}(X_k - \overline{X}) = 0$$

where $\overline{X}$ is the sample mean.

12. Let $(X_1, X_2, \ldots, X_n)$ denote a sample of size n. Show that

$$n\overline{X}^2 = \frac{1}{n}\left(\sum_{k=1}^{n} X_k\right)^2$$

where $\overline{X}$ is the sample mean.

13. Assume that a sample of size n has l distinct values $x_1, x_2, \ldots, x_l$, where x_k occurs f_k times in the sample. Explain why the sample mean is given by the formula

$$\overline{X} = \frac{1}{n}\sum_{k=1}^{l} x_k f_k$$

14. Assume that a sample of size n has l distinct values $x_1, x_2, \ldots, x_l$, where x_k occurs f_k times in the sample. Explain why the sample variance is given by the formula

$$S^2 = \frac{1}{n-1}\left[\sum_{k=1}^{l} x_k^2 f_k - \frac{1}{n}\left(\sum_{k=1}^{l} x_k f_k\right)^2\right]$$

15. Assume that X is exponentially distributed with parameter $\lambda = 3.0$.

(a) Assume that a sample of size 50 is taken from this population. What is the approximate distribution of the sample mean?

(b) Assume now that 1000 samples, each of size 50, are taken from this population and a histogram of the sample means of each of the samples is produced. What shape will the histogram be approximately?

16. Assume that X is exponentially distributed with parameter $\lambda = 3.0$. Assume that a sample of size 50 is taken from this population and that the sample mean of this sample is calculated. How likely is it that the sample mean will exceed 0.43?

12.7.2

T **17.** Use the random-number generator on a graphing calculator to generate three samples, each of size 10, from a uniform distribution over the interval (0, 1).

(a) Compute the sample mean and the sample variance of each sample.

(b) Combine all three samples, and compute the mean and the sample variance of the combined sample.

(c) Compare your answers in (a) and (b) with the true values of the mean and the variance.

18. Suppose that X is exponentially distributed with mean 1. A computer generates the following sample of independent observations from the population X:

$$0.3169, 0.5531, 2.376, 1.150, 0.6174,$$
$$0.1563, 2.936, 1.778, 0.7357, 0.1024$$

Find the sample mean and the sample variance, and compare them with the corresponding population parameters.

19. Compute the sample mean and the standard error for the sample in Problem 1.

20. Compute the sample mean and the standard error for the sample in Problem 2.

21. The following data represent a sample from a certain population:

$$-0.68, 1.22, 1.33, -0.84, -0.06,$$
$$0.50, 0.03, -0.13, -0.29, -0.47$$

Construct a 95% confidence interval for the mean.

22. The following data represent a sample from a certain population:

$$-1.18, 0.52, 0.36, -0.16, 0.92,$$

$$0.68, -0.61, -0.54, 0.15, 1.04$$

Construct a 95% confidence interval for the mean.

T **23.** Use a graphing calculator to construct a 95% confidence interval for a sample of size 30 from a uniform distribution over the interval $(0, 1)$. Take a class poll to determine the percentage of confidence intervals that contain the true mean. Discuss the result in class.

T **24. (a)** If X has distribution function $F(x)$, we can show that $F(X)$ is uniformly distributed over the interval $(0, 1)$. Use this fact, a graphing calculator, and the table for the standard normal distribution to generate 15 standard normally distributed random variables.

(b) Use your data from (a) to construct a 95% percent confidence interval. Take a class poll to determine the percentage of confidence intervals that contain the true mean. Discuss the result in class.

25. To determine the germination success of seeds of a certain plant, you plant 162 seeds. You find that 117 of the seeds germinate. Estimate the probability of germination and give a 95% confidence interval.

26. To test a new drug for lowering cholesterol, 72 people with elevated cholesterol receive the drug; 51 of them show reduced cholesterol levels. Estimate the probability that the drug lowers cholesterol, and construct a 95% confidence interval.

12.7.3

In Problems 27 and 28, fit a linear regression line through the given points and compute the coefficient of determination.

27. $(-3, -6.3), (-2, -5.6), (-1, -3.3), (0, 0.1), (1, 1.7), (2, 2.1)$

28. $(0, 0.1), (1, -1.3), (2, -3.5), (3, -5.7), (4, -5.8)$

29. Show that the sum of the residuals about any linear regression line is equal to 0.

30. Show that the last two terms in (12.48), namely

$$2(\overline{y} - a - b\overline{x})\sum(y_k - \overline{y})$$

and

$$2(\overline{y} - a - b\overline{x})\sum(x_k - \overline{x})$$

are equal to 0.

31. To determine whether the frequency of chirping crickets depends on temperature, the following data were obtained ***(Pierce, 1949)***:

Temperature (°F)	Chirps/s
69	15
70	15
72	16
75	16
81	17
82	17
83	16
84	18
89	20
93	20

Fit a linear regression line to the data, and compute the coefficient of determination.

32. The initial velocity v of an enzymatic reaction that follows Michaelis–Menten kinetics is given by

$$v = \frac{v_{\max} s}{K_m + s} \qquad (12.54)$$

where s is the substrate concentration and $v_{\max}$ and K_m are two parameters that characterize the reaction. The following computer-generated table contains values of the initial velocity v when the substrate concentration s was varied:

s	v
1	4.1
2.5	6.1
5	9.3
10	12.9
20	17.1

(a) Invert (12.54) and show that

$$\frac{1}{v} = \frac{K_m}{v_{\max}}\frac{1}{s} + \frac{1}{v_{\max}} \qquad (12.55)$$

This is the Lineweaver–Burk equation. If we plot $1/v$ as a function of $1/s$, a straight line with slope $K_m/v_{\max}$ and intercept $1/v_{\max}$ results. Use (12.55) to transform the data, and fit a linear regression line to the transformed data. Find the slope and the intercept of the linear regression line, and determine K_m and $v_{\max}$.

(b) ***Dowd and Riggs (1965)*** proposed to use the transformation

$$v = v_{\max} - K_m \frac{v}{s} \qquad (12.56)$$

and then plot v against v/s. The resulting straight line has slope $-K_m$ and intercept $v_{\max}$. Use (12.56) to transform the data, and fit a linear regression line to the transformed data. Find the slope and the intercept of the linear regression line, and determine K_m and $v_{\max}$.

Chapter 12 Review

Key Terms

Discuss the following definitions and concepts:

1. Multiplication principle
2. Permutation
3. Combination
4. Random experiment
5. Sample space
6. Basic set operations, Venn diagram, De Morgan's laws
7. Definition of probability
8. Equally likely outcomes
9. Mendel's pea experiments
10. Mark–recapture method
11. Maximum likelihood estimate

12. Conditional probability
13. Partition of sample space
14. Law of total probability
15. Independence
16. Bayes formula
17. Random variable
18. Discrete distribution
19. Probability mass function
20. Distribution function of a discrete random variable
21. Mean and variance
22. Joint distributions
23. Binomial distribution
24. Multinomial distribution
25. Geometric distribution
26. Poisson distribution
27. Poisson approximation to the binomial distribution
28. Continuous random variable
29. Density function
30. Distribution function of a continuous random variable
31. Mean and variance of a continuous random variable
32. Histogram
33. Normal distribution
34. Uniform distribution
35. Exponential distribution
36. Aging
37. Gompertz law
38. Weibull law
39. Law of large numbers
40. Markov's inequality
41. Chebyshev's inequality
42. Central limit theorem
43. Histogram correction
44. Sample
45. Statistic
46. Sample median, sample mean, sample variance, standard error
47. Confidence interval
48. Estimating proportions
49. Linear regression line
50. Coefficient of determination

Review Problems

1. (a) There are 25 students in a calculus class. What is the probability that no two students have the same birthday?

(b) Let p_n denote the probability that, in a group of n people, no two people have the same birthday. Show that

$$p_1 = 1 \quad \text{and} \quad p_{n+1} = p_n \frac{365 - n}{365}$$

Use this formula to generate a table of p_n for $1 \leq n \leq 25$.

2. Thirty patients are to be randomly assigned to two different treatment groups. How many ways can this be done?

3. Fifteen different plants are to be equally divided among five plots. How many ways can this be done?

4. Assume that a certain disease either is caused by a genetic mutation or appears spontaneously. The disease will appear in 67% of all people with the mutation and in 23% of all people without the mutation. Assume that 3% of the population carries the disease gene.

(a) What is the probability that a randomly chosen individual will develop the disease?

(b) Given an individual who suffers from the disease, what is the probability that he or she has the genetic mutation?

5. Suppose that 42% of the seeds of a certain plant germinate.

(a) What is the expected number of germinating seeds in a sample of 10 seeds?

(b) You plant 10 seeds in one pot. What is the probability that none of the seeds will germinate?

(c) You plant five pots with 10 seeds each. What is the expected number of pots with no germinating seeds?

(d) You plant five pots with 10 seeds each. What is the probability that at least one pot has no germinating seeds?

6. Suppose that the amount of yearly rainfall in a certain area is normally distributed with mean 27 and standard deviation 5.7 (measured in inches).

(a) What is the probability that, in a given year, the rainfall will exceed 35 inches?

(b) What is the probability that, in five consecutive years, the rainfall will exceed 35 inches in each year?

(c) What is the probability that, in at least 1 out of 10 years, the rainfall will exceed 35 inches per year?

7. Suppose that, each time a student takes a particular test, he or she has a 20% chance of passing. (Assume that consecutive trials are independent.)

(a) What are the chances of passing the test on the second trial?

(b) Given that a student failed the test the first time, what are the chances that he or she will pass the test on the second trial?

8. Explain why

$$2^n = \sum_{k=0}^{n} \binom{n}{k}$$

9. A bag contains 170 chocolate-covered raisins, on average. Production standards require that, in 95% of all bags, the number of raisins does not deviate from 170 by more than 10. Assume that the number of raisins is normally distributed with mean μ and variance σ^2.

(a) Determine μ and σ.

(b) A shipment contains 100 bags. What is the probability that no bag contains fewer than 160 raisins?

10. Genetics Suppose that two parents are carriers of a recessive gene causing a metabolic disorder. Neither parent has the disease. If they have three children, what is the probability that none of the children will be afflicted with the disease? (Note that a recessive gene causes a disorder only if an individual has two copies of the gene.)

11. Suppose that you cross a white-flowering pea plant with each plant from a large batch of red-flowering pea plants. What percentage of the red-flowering parent plants are of genotype *Cc* if 90% of the offspring have red flowers?

12. Suppose that a random variable is normally distributed with mean μ and variance σ^2. How would you estimate μ and σ?

13. Let $(X_1, X_2, \ldots, X_n)$ be a sample of size n from a population with mean μ and variance σ^2. Define

$$V = \frac{1}{n}\sum_{k=1}^{n}(X_k - \overline{X})^2$$

where $\overline{X}$ is the sample mean.

(a) If S^2 is the sample variance, show that

$$V = \frac{n-1}{n}S^2$$

(b) Compute $E(V)$.

14. Assume that the weight of a certain species is normally distributed with mean μ and variance σ^2. The following data represent the weight (measured in grams) of 10 individuals:

171, 168, 151, 192, 175, 163, 182, 157, 177, 169

(a) Find the median, the sample mean, and the sample variance.

(b) Construct a 95% confidence interval for the population mean.

T **15.** **(a)** Generate five observations (x, y) from a random experiment, where

$$y = 2x + 1 + \epsilon$$

$x = 1, 2, 3, 4, 5$, and ϵ is normally distributed with mean 0 and variance 1.

(b) Use your data from (a) to find the least square line, and compare your results with the linear model that describes this experiment.

(c) What proportion of your data is explained by the model?

16. Suppose X is a continuous random variable with density function

$$f(x) = \begin{cases} 0 & \text{for } x < 1 \\ (r-1)x^{-r} & \text{for } x \geq 1 \end{cases}$$

where r is a constant greater than 1.

(a) For which values of r is $E(X) = \infty$?

(b) Compute $E(X)$ for those values of r for which $E(X) < \infty$.

Appendices

A Frequently Used Symbols

TABLE A.1 Greek Letters

Lowercase Letters

α	alpha	η	eta	ν	nu	τ	tau
β	beta	θ	theta	ξ	xi	υ	upsilon
γ	gamma	ι	iota	o	omicron	ϕ	phi
δ	delta	κ	kappa	π	pi	χ	chi
ϵ	epsilon	λ	lambda	ρ	rho	ψ	psi
ζ	zeta	μ	mu	σ	sigma	ω	omega

Uppercase Letters

Γ	Gamma	Λ	Lambda	Σ	Sigma
Δ	Delta	Π	Pi	Ω	Omega

TABLE A.2 Mathematical Symbols

$<$	less than	$\subset$	subset	$=$	equal to
$\leq$	less than or equal to	$\in$	element of	$\neq$	not equal to
$>$	greater than	$\perp$	perpendicular	$\approx$	approximately
$\geq$	greater than or equal to	$\parallel$	parallel	$\propto$	proportional to
$\cup$	union	$\cap$	intersection	Σ	sum

B Table of the Standard Normal Distribution

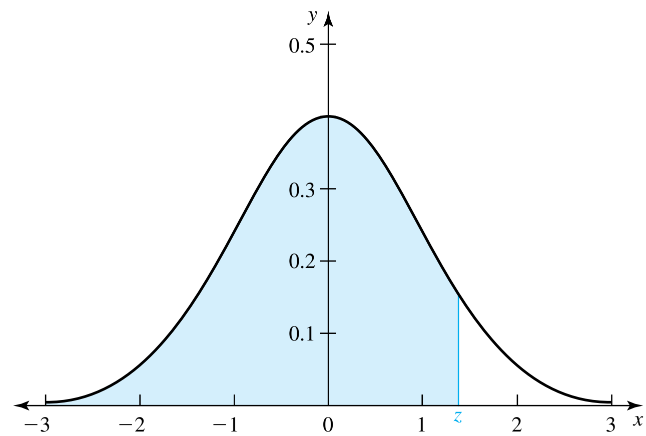


Figure B.1 Areas under the standard normal curve from $-\infty$ to z.

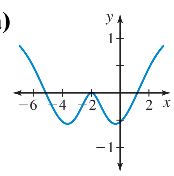
z	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5754
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7258	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7518	.7549
0.7	.7580	.7612	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7996	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986

Answers to Odd-Numbered Problems

Section 1.1 (Includes Even-Numbered Problems)

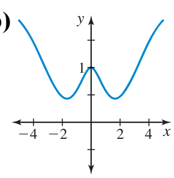
1. **(a)** $y = x + 5$; **(b)** $y = 2x - 1$; **(c)** $y = 3x - 5$ (Review 1.2.2)
2. **(a)** 26.7°C; **(b)** 14°F; **(c)** −40°F or −40°C (Review 1.2.2)
3. A circle with center $(x, y) = (-1, 5)$ and with radius 3 (Review 1.2.3) **4.** **(a)** $\frac{180}{7} \approx 25.7°$; **(b)** $\frac{-\pi}{3} + 2n\pi$ and $\frac{-2\pi}{3} + 2n\pi$ for any $n \in \mathbf{Z}$; **(c)** $\sin^2\theta + \cos^2\theta = 1 \Rightarrow \frac{\sin^2\theta + \cos^2\theta}{\cos^2\theta} = \frac{1}{\cos^2\theta}$, **(d)** $\frac{\pi}{12}, \frac{7\pi}{12}, \frac{3\pi}{4}$. (Review 1.2.4) **5.** **(a)** (i) $2^{7/3}$, (ii) $2^{8/3}$; **(b)** $x = -\frac{1}{2}$; **(c)** 4; **(d)** $\log_{10}(15x^2)$; **(e)** $x = e^{2/3}$; **(f)** $3\log_{10} e$ (Review 1.2.5)
6. **(a)** $-\frac{1}{2} \pm \frac{\sqrt{3}}{2}$, **(b)** $3 + i$, **(c)** $z + \overline{z} = 2a \in \mathbf{R}$ (Review 1.2.6)
7. **(a)** (i) $f(x) \in [0, 1]$, (ii) $f(x) \in [0, \infty)$; **(b)** (i) $\sqrt{3}$, (ii) $|x + 1|$, (iii) 9; **(c)** Since $f(x) = |x| \geq 0$, then $f \circ f(x) = |f(x)| = f(x)$ (Review 1.3.1) **8.** **(a)** (iii); **(b)** (i) degree 2, (ii) $0 \leq r \leq a$, (iii) $0 \leq u(r) \leq u_0$, (iv) $r = \frac{a}{\sqrt{2}}$ (Review 1.3.2) **9.** **(a)** (iii); **(b)** (iii); **(c)** $c = 20$; **(d)** No, doubling rate of metabolization is impossible starting at $c = 10$ (Review 1.3.3)
10. **(a)** $A = \frac{A_0}{32}$; **(b)** $A = \frac{A_0}{243}$; **(c)** $A = 32A_0$ (Review 1.3.4)
11. **(a)** $N_0 = 1000$, $r = \frac{\ln 2}{2} \approx 0.347$; **(b)** $t = \frac{\ln 3}{r} \approx 3.17$; **(c)** $t = 4$ (Review 1.3.5) **12.** **(a)** $f^{-1}(x) = \sqrt{x - 1}$; **(b)** $f^{-1}(x) = -1 + e^{x/2}$; **(c)** $f^{-1}(x) = x^{1/5}$ (Review 1.3.6)
13. **(a)** (i) $\ln(x(x^2 + 1))$, (ii) $\frac{1}{3}\log\left(\frac{x}{x+1}\right)$, (iii) $\log_2(4x)$; **(b)** (i) $\frac{\ln 7}{\ln 2}$, (ii) $\frac{\ln 6}{\ln 10}$, (iii) $\frac{\ln 2}{\ln x}$ (Review 1.3.7) **14.** **(a)** (i) amplitude = 2, period = 2π, (ii) amplitude = 2, period = $2\pi/3$, (iii) amplitude = 3, period = 4; **(b)** (i) $f(x) \in \mathbb{R}$ and $x \in \mathbb{R}$ except when $x \neq \pm\frac{\pi}{2}$, $\pm\frac{3\pi}{2}$, etc. (odd integer multiples of $\frac{\pi}{2}$), (ii) $-1 \leq f(x) \leq 1$ and $x \in \mathbb{R}$; **(c)** Compress curve by a factor of $\frac{1}{2}$ in x-direction and stretch by a factor of 3 in y-direction. (Review 1.3.8)

15. **(a)**



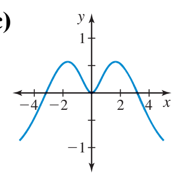
$y = f(x + 2)$

(b)



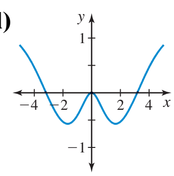
$y = f(x) + 1$

(c)



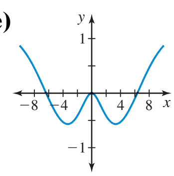
$y = -f(x)$

(d)



$y = f(-x)$

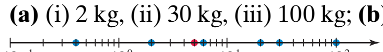
(e)



$y = f(x/2)$

(Review 1.4.1)

16. **(a)** (i) 2 kg, (ii) 30 kg, (iii) 100 kg; **(b)** $\frac{\text{Adult weight}}{\text{Puppy weight}} = \frac{20\text{ kg}}{0.4\text{ kg}} = 50$. **(c)**



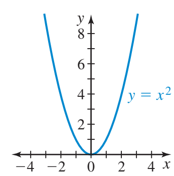
(Review 1.4.2)
17. **(a)** $D = kA^{0.41}$, **(b)** $N = k\,10^{-m}$, **(c)** $N = k\,7.24^t$, **(d)** $N = k\,0.81^t$ (Review 1.4.3)
18. **(a)** Curve (4); **(b)** Curve (3), **(c)** Curve (2), (Review 1.4.4)

Section 1.2

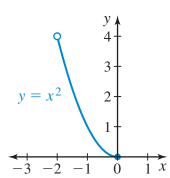
1. **(a)** $\{-5, 3\}$ **(b)** $\{-5, 3\}$ **3.** **(a)** $\{-5, 1\}$ **(b)** $\{1, 5\}$ **(c)** $\{-1, 4\}$ **(d)** $\left\{-1, \frac{7}{5}\right\}$ **5.** **(a)** $[-\frac{2}{5}, \frac{6}{5}]$, **(b)** $\left(-\infty, -\frac{5}{4}\right) \cup \left(\frac{11}{4}, \infty\right)$, **(c)** $(-\infty, -1] \cup [-\frac{1}{7}, \infty)$, **(d)** $(-5, 2)$ **7.** $2x + y - 8 = 0$
9. $3x + y + 2 = 0$ **11.** $7x - 3y + 5 = 0$ **13.** $x + y - 3 = 0$
15. $4y - 1 = 0$ **17.** $x + 2 = 0$ **19.** $3x - y + 2 = 0$
21. $x - 2y + 4 = 0$ **23.** $2x + y - 2 = 0$ **25.** $2x + 4y + 1 = 0$
27. $x + 2y + 4 = 0$ **29.** $2x + 3y + 5 = 0$ **31.** $2x + 5y - 22 = 0$
33. $x - y - 6 = 0$ **35.** $y - 3 = 0$ **37.** $x + 2 = 0$ **39.** $x - 1 = 0$
41. $y - 3 = 0$ **43.** **(a)** (i) 183 cm, (ii) ≈ 10.2 cm, (iii) ≈ 48.3 cm, (iv) ≈ 52.1 cm, **(b)** $x = \frac{y}{30.5}$, **(c)** (i) ≈ 6.39 ft, (ii) ≈ 0.39 ft, (iii) ≈ 1.57 ft **45.** $s(t) = (40\text{ mi/hr})t$, $k = 40\text{ mi/hr}$
47. $\frac{1}{(0.305)^2}\text{ ft}^2$ **49.** **(a)** $y = \frac{x}{33.81}$ **(b)** $\frac{12}{33.81} \approx 0.35$ liters
51. **(a)** 300 g, **(b)** $\frac{15}{8} = 1.875$ cups, **(c)** If c is amount measured in cups and g is amount measured in grams $c = \frac{g}{120}$.
53. **(a)** $y = \frac{1}{7}(3 + 4x)$, **(b)** (i) 0.714, (ii) 0.943, (iii) 0.429, **(c)** $x = \frac{1}{4}(7y - 3)$. **55.** $(x - 1)^2 + (y + 2)^2 = 4$
57. **(a)** $(x - 2)^2 + (y - 5)^2 = 16$, **(b)** $y = 5 - 2\sqrt{3}$ and $y = 5 + 2\sqrt{3}$, **(c)** No **59.** center: $(-2, 0)$; radius: 5 **61.** center: $(-3, -1)$, radius: $\sqrt{22}$ **63.** **(a)** $\frac{13}{36}\pi$, **(b)** 165° **65.** **(a)** $-\frac{1}{2}\sqrt{2}$, **(b)** $-\frac{1}{2}\sqrt{3}$, **(c)** $\frac{\sqrt{3}}{3}$ **67.** **(a)** $\alpha = \frac{2\pi}{3}$ or $\alpha = \frac{4\pi}{3}$ **(b)** $\alpha = \frac{\pi}{6}$ or $\alpha = \frac{7\pi}{6}$
69. Divide both left and right side by $\cos^2\theta$. **71.** $\frac{3\pi}{2}$ **73.** **(a)** 4 **(b)** 9 **(c)** 5^{4k-4} **75.** **(a)** $x = \frac{1}{16}$ **(b)** $x = 27$ **(c)** $x = \frac{1}{100}$
77. **(a)** $x = -5$ **(b)** $x = -4$ **(c)** $x = -3$ **79.** **(a)** $\ln 3$ **(b)** $\log_4(x - 2) + \log_4(x + 2)$ **(c)** $6x - 2$ **81.** **(a)** $x = \frac{1}{3}(\ln 2 + 1)$ **(b)** $x = -\frac{1}{2}\ln 10$ **(c)** $x = \pm\sqrt{1 + \ln 10}$ **83.** **(a)** $x = 3 + e^5$ **(b)** $x = \sqrt{4 + e}$ **(c)** $x = 18$ **85.** $8 - 4i$ **87.** $13 + 2i$ **89.** $20 + 12i$
91. 37 **93.** $1 - 2i$ **95.** $2 + 3i$ **97.** $6 + 4i$ **99.** $x_1 = \frac{3}{4} + i\frac{\sqrt{7}}{4}$, $x_2 = \frac{3}{4} - i\frac{\sqrt{7}}{4}$ **101.** $x_1 = -1, x_2 = 2$ **103.** $x_1 = -\frac{1}{2} + i\frac{\sqrt{23}}{2}$, $x_2 = -\frac{1}{2} - i\frac{\sqrt{23}}{2}$ **105.** $x_1 = \frac{7}{3}, x_2 = -1$ **107.** $x_1 = -1, x_2 = \frac{4}{3}$
109. $x_1 = \frac{5 + i\sqrt{47}}{6}, x_1 = \frac{5 - i\sqrt{47}}{6}$ **111.** $z + \overline{z} = 2a$, $z - \overline{z} = 2bi$
113. $z = \overline{(\overline{z})} = a + bi$. **115.** $\overline{zw} = \overline{z} \cdot \overline{w} = (a - bi)(c - di)$.

Section 1.3

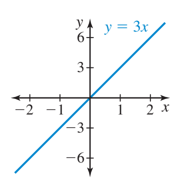
1. range: $y \geq 0$



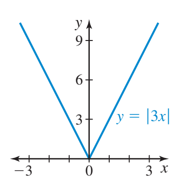
3. range: $y \in [0, 4)$



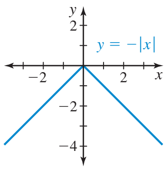
5. **(b)** No, their domains are different.
7. $f(x)$ is odd. $f(-x) = -3x = -f(x)$



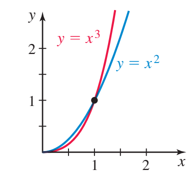
9. $f(x)$ is even. $f(-x) = |3(-x)| = |3x| = f(x)$



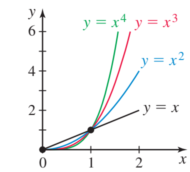
11. $f(x)$ is even. $f(-x) = -|-x| = -|x| = f(x)$



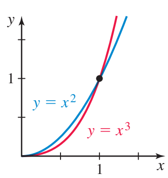
13. **(a)** $(f \circ g)(x) = (3+x)^2$ **(b)** $(g \circ f)(x) = 3 + x^2$
15. **(a)** $(f \circ g)(x) = 1 - \sqrt{x}, x \geq 0$ **(b)** $(g \circ f)(x) = \sqrt{1-x}, x \leq 1$
17. **(a)** $(f \circ g)(x) = \frac{1}{\sqrt{x}}, x > 0$ **(b)** $(g \circ f)(x) = \sqrt{\frac{1}{x}}, x > 0$
19. $(f \circ g)(x) = x, x \geq 0$; $(g \circ f)(x) = x, x \geq 0$
21. $x^2 > x^3$ for $0 < x < 1$; $x^2 < x^3$ for $x > 1$



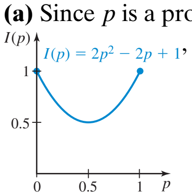
23. They intersect at $x = 0$ or 1.



25. **(a)**

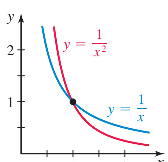


(b) $x \leq 1 \Rightarrow x \cdot x \leq x \cdot 1$ provided $x \geq 0$ (since we are multiplying both sides of the inequality by a quantity). Hence $x^2 \leq x$. **(c)** $x \geq 1 \Rightarrow x \cdot x \geq x \cdot 1$ (again $x \geq 0$). Hence $x^2 \geq x$.
27. **(a)** $f(-x) = f(x)$ **(b)** $f(-x) = -f(x)$
29. **(a)** Since p is a proportion, it is in the interval $[0, 1]$
(b)

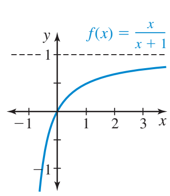


(c) range: $[0.5, 1]$

31. $s(t) = t$, polynomial of degree 1 **33.** domain: $x \neq 1$; range: $y \neq 0$ **35.** domain: $x \neq -3, 3$; range: **R**
37. $\frac{1}{x} < \frac{1}{x^2}$ for $0 < x < 1$; $\frac{1}{x} > \frac{1}{x^2}$ for $x > 1$; they intersect at $x = 1$.

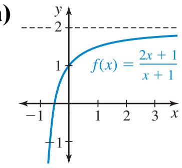


39. **(a)**



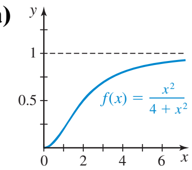
(b) range of $f(x)$ is $(-\infty, 1)$ **(c)** $x = 3$
(d) exactly one.

41. **(a)**



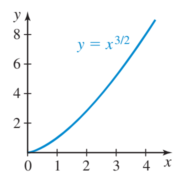
(b) range of $f(x)$ is $[1, 2)$, **(c)** $x = \frac{1}{3}$,
(d) $x = \frac{1-a}{a-2}$. **43.** 83.3 2.44

45. **(a)**

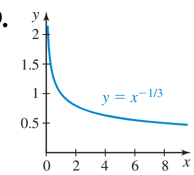


(b) range of $f(x)$ is $[0, 1)$,
(c) $f(x)$ approaches 1.

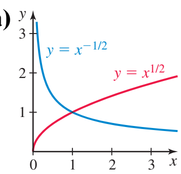
47.



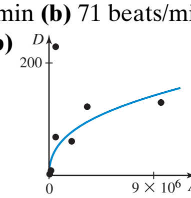
49.



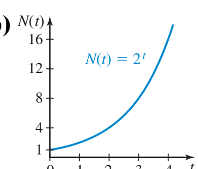
51. **(a)**



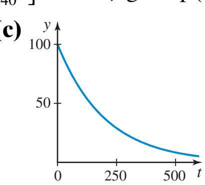
53. **(a)** 93 beats/min **(b)** 71 beats/min **(c)** Molly: 1.3, Prof. R: 0.18
55. **(a)** 147.39, **(b)**



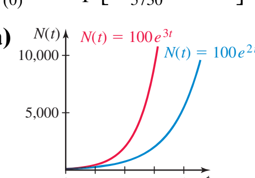
(i) 6, (ii) yes, Cameroon, (iii) 0.249, no.
57. **(a)** 1, 2, 4, 8, 16 **(b)**



59. $20 \exp\left[-\frac{\ln 2}{5730} 2000\right]$
61. $\lambda = \frac{\ln 2}{7 \text{ days}} \approx 0.099$ days
63. **(a)** $W(t) = (100\,\mu\text{g}) \exp\left[-\frac{\ln 2}{140} t\right] \approx 100\,\mu\text{g} \cdot \exp(-0.005t)$
(b) $t = \frac{\ln 10}{\ln 2} 140$ days ≈ 465 days **(c)**

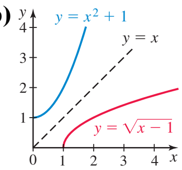


65. $\frac{W(t)}{W(0)} = \exp\left[-\frac{\ln 2}{5730} 10{,}000\right] \approx 29.8\%$
67. **(a)**

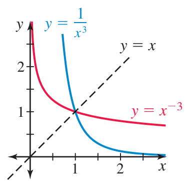


The population with $r = 3$ grows faster.
(b) (i) $N_0 = 100$, $r = \frac{\ln 3}{2} \approx 0.549$, (ii) At $t \approx 4.19$, (iii) At $t \approx 8.38$
69. **(a)** yes **(b)** no **(c)** yes **(d)** yes **(e)** no **(f)** yes

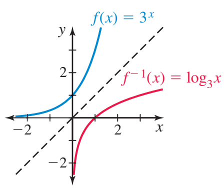
71. **(a)** $f^{-1}(x) = \sqrt{x-1}, x \geq 1$ **(b)**



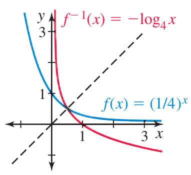
73. **(a)** $f^{-1}(x) = \sqrt[3]{\frac{1}{x}}, x > 0$, **(b)**



75. $f^{-1}(x) = \log_3 x, x > 0$

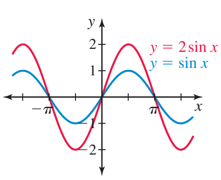


77. $f^{-1}(x) = \log_{1/4} x, x > 0$

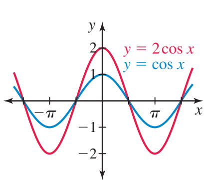


79. **(a)** x^5 **(b)** x^3 **(c)** x^{-5} **(d)** x^6 **(e)** x^{-3} **(f)** x^{-3} **81.** **(a)** $\ln x$ **(b)** $\frac{14}{3} \ln x$ **(c)** $\ln(x-1)$ **(d)** $-\frac{7}{2} \ln x$ **83.** **(a)** $e^{x \ln 3}$ **(b)** $e^{(x^2-1)\ln 4}$ **(c)** $e^{-(x+1)\ln 2}$ **(d)** $e^{(-4x+1)\ln 3}$ **85.** $\mu = \ln 2$ **87.** **(a)** $m = 5.04$ **(b)** $m = 3.28$ **(c)** (i) $8.318 \times 10^6 \mu$m (ii) 14454μm **(d)** 10

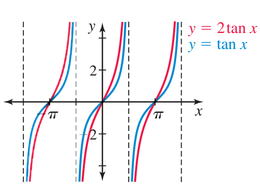
89. Same period; $2 \sin x$ has twice the amplitude of $\sin x$.



91. Same period; $2 \cos x$ has twice the amplitude of $\cos x$.



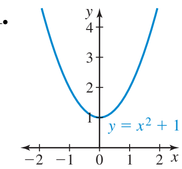
93. Same period; $y = 2 \tan x$ is stretched by a factor of 2.



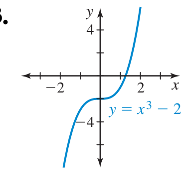
95. amplitude: 2; period: 4π **97.** amplitude: 4; period: 2 **99.** $p(t) = 3.2 + 1.6 \cos\left(2\pi \cdot \frac{t}{12}\right)$ **101.** $\cos x = 0$ when x is an odd integer multiple of $\pi/2$.

Section 1.4

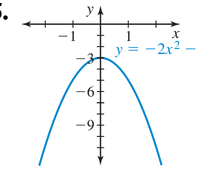
1.



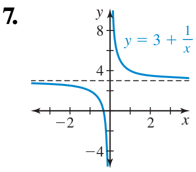
3.



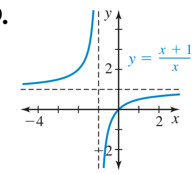
5.



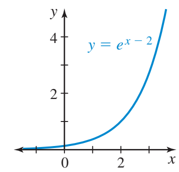
7.



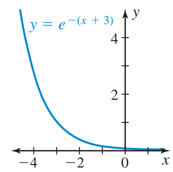
9.



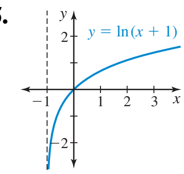
11.



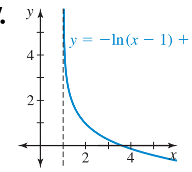
13.



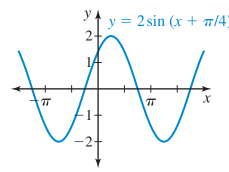
15.



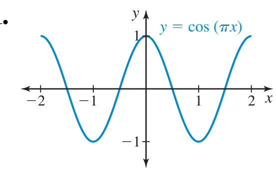
17.



19.



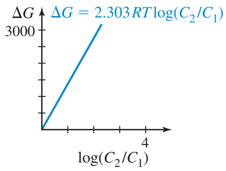
21.

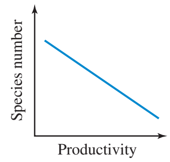


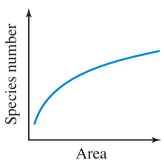
23. **(a)** Shift two units down. **(b)** Shift $y = x^2$ one unit to the right and then one unit up. **(c)** Shift $y = x^2$ two units to the left, stretch by a factor of 2, and then reflect about the x-axis. **25.** **(a)** Reflect $\frac{1}{x}$ about the x-axis, and then shift up one unit. **(b)** Shift $\frac{1}{x}$ one unit to the right, and then reflect about the x-axis. **(c)** $\frac{x}{x+1} = 1 - \frac{1}{x+1}$, so shift $y = \frac{1}{x}$ one unit to the left, reflect about x-axis, and then shift up one unit. **27.** **(a)** Shift $y = e^x$ three units up. **(b)** Reflect $y = e^x$ about the y-axis. **(c)** Shift $y = e^x$ two units to the right, stretch vertically by a factor of 2, and then shift three units up. **29.** **(a)** Shift $y = \ln x$ one unit to the right. **(b)** Reflect $y = \ln x$ about the x-axis, and then shift up one unit. **(c)** Shift $y = \ln x$ three units to the left, then down one unit. **31.** **(a)** Compress $y = \sin x$ horizontally by a factor of π. **(b)** Shift $y = \sin x$ by $\pi/4$ units to the left. **(c)** Shift $y = \sin x$ to the left by $1/\pi$ units, then compress horizontally by a factor of π, then stretch vertically by a factor of 2. Finally, reflect about the x-axis. **33.** Calculate log of each number, for instance, $\log 3 \approx 0.477$ and $\log 1000 = 3$. **35.** **(a)** $2, -3, -4, -7, -10$ **(b)** No **(c)** No **37.** four **39.** eight **41.** six to seven **43.** $y = 5 \times (0.58)^x$ **45.** $y = 3^{1/2} \times (3^{1/2})^x$ **47.** $\log y = \log 3 - 2x$ **49.** $\log y = \log 2 - (1.2)(\log e)x$ **51.** $\log y = \log 5 + (4 \log 2)x$ **53.** $\log y = \log 4 + (2 \log 3)x$ **55.** $y = (2)x^{-(\log 2)/\log 5}$ **57.** $y = \frac{1}{8}x^2$ **59.** $\log y = \log 2 + 5 \log x$ **61.** $\log y = 6 \log x$ **63.** $\log y = \log 2 - 2 \log x$ **65.** $\log y = \log 4 - 3 \log x$ **67.** $\log y = \log 3 + 1.7 \log x$, log-log transformation **69.** $\log N(t) = \log 130 + (1.2t) \log 2$, log-linear transformation **71.** $\log R(t) = \log 3.6 + 1.2 \log t$, log-log transformation **73.** $y = 1.8x^{0.2}$ **75.** $y = 4 \times 10^x$ **77.** $y = (5.7)x^{2.1}$ **79.** $\log_2 y = x$ **81.** $\log_2 y = -x$ **83.** **(a)** $\log N = \log 2 + 3t \log e$ **(b)** slope: $3 \log e \approx 1.303$ **85.** Yes; $\log S = \log C + z \log A$, $z =$ slope of straight line **87.** **(b)** $-1/K$ is the slope of the line, and v_{max}/K is the vertical axis intercept. **(c)** $v_{max} \approx 2$, $K \approx 10$. **89.** **(a)** $\log S = \log 1.162 + 0.93 \log B$ **91.** $b \approx 72.44$, $a \approx -0.5$. **93.** **(c)** $d \approx 0.751$.

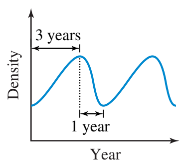
95. **(a)** $\alpha = -\ln 0.9/\text{m}$ **(b)** 10% **(c)** 1 m: 90%, 2 m: 81%, 3 m: 72.9% **(e)** slope $= \log 0.9 = -\alpha/\ln 10$ **(f)** $z = -\frac{1}{\alpha}\ln(0.01) = \frac{\ln(0.01)}{\ln(0.9)}$ **(g)** Clear lake: small α; milky lake: large α

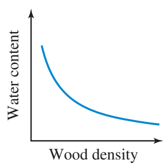
97. $y = (100)(10^{1/3})^x$ **99.** $y = (2^{1/3})(2^{2/3})^x$ **101.** $y = \log x$

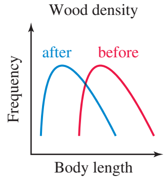
103.


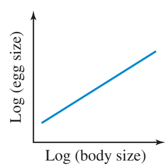
105.


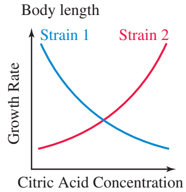
107.


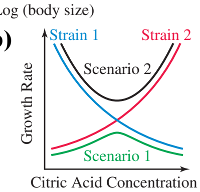
109.


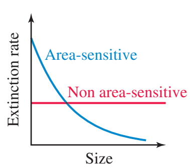
111.


113.


115.


117. **(a)**


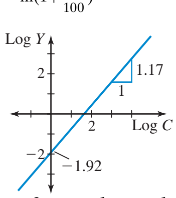
(b)


119.


Chapter 1 Review Problems

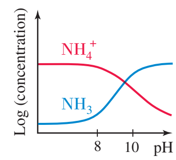
1. **(a)** 10^3, 1.1×10^3, 1.22×10^3, 1.35×10^3, 1.49×10^3 **(b)** $t = 10 \ln 100 \approx 46.1$ **3.** **(a)** $R(y) = ky(y+1)$, $0 \le y \le 1.5$, $0 \le R(y) \le 3.75k$ **(b)** $R(z) = kz(z-1)$, $1 \le z \le 2.5$, $0 \le R(z) \le 3.75k$ **5.** **(a)** $L(1) = \ln(1) + 1 = 1$, $E(t) = e^{1-1} = e^0 = 1$. **(b)** $L(2) = 1.693$, $L(10) = 3.303$, $L(100) = 5.605$, $E(2) = 2.718$, $E(10) = 8103$, $E(100) = 9.889 \times 10^{42}$ feet. **(c)** Species A: 1.718 years; Species B: 0.693 years (counting from $t = 1$). **(d)** Species A: 20.09 years; Species B: 2.386 years (counting from $t = 1$). **(e)** Species A: 7.389 years; Species B: 2.099 years (counting from $t = 1$).

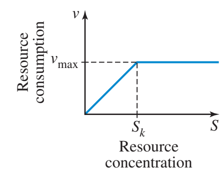
7. $T = \frac{\ln 2}{\ln(1+\frac{q}{100})}$, T goes to infinity as q gets closer to 0.

9. **(a)**


(b) $Y = C^{1.17}10^{-1.92}$. When log-transformed we obtain a straight line plot. The exponent can be calculated directly from the slope of the straight line.

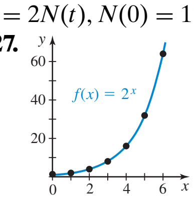
(c) $Y_p = 2.25Y_c$ **(d)** 8.5% **11.** **(a)** $21.\overline{8}$ hours per day, 400 days per year
(b) line through (0, 24) and (380, $21.\overline{8}$): $y = 24 - \frac{x}{180}$, $x = 4320 - 180y$ **(c)** 376×10^6 to 563×10^6 years ago
13. **(a)** males: $S(t) = \exp[-(0.019t)^{3.41}]$; females: $S(t) = \exp[-(0.022t)^{3.24}]$ **(b)** males: 47.27 days; females: 40.59 days **(c)** males should live longer **15.** **(a)** $x = k$, $v = \frac{a}{2}$ **(b)** $x_{0.9} = 81x_{0.1}$

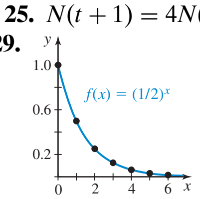
17.


19. $g(S) = \frac{v_{\max}}{S_k} S$


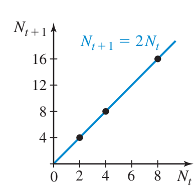
21. **(a)** $2000e^{t\ln 2}$; $N_0 = 2000$ and $r = \ln 2$, **(b)** 2.32 hours, **(c)** $0.492 \le r \le 0.894$.

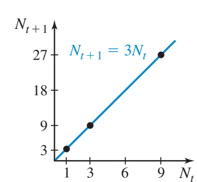
Section 2.1

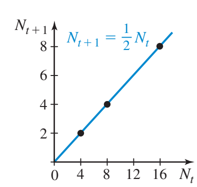
1. 1, 3, 9, 27, 81, 243 **3.** 1, $\frac{1}{3}$, $\frac{1}{9}$, $\frac{1}{27}$, $\frac{1}{81}$, $\frac{1}{243}$ **5.** $N_t = 2 \cdot 2^t = 2^{t+1}$, $t = 0, 1, 2, \ldots$ **7.** $N_t = 4^t$, $t = 0, 1, 2, \ldots$ **9.** About 3.17 hours, or about 190 minutes. **11.** 2, 4, 8, 16, 32 **13.** 420 minutes **15.** 60 minutes **17.** $N(t) = 40 \cdot 2^t$, $t = 0, 1, 2, \ldots$ **19.** $N(t) = 20 \cdot 5^t$, $t = 0, 1, 2, \ldots$ **21.** $N(t) = 5 \cdot 4^t$, $t = 0, 1, 2, \ldots$
23. $N(t+1) = 2N(t)$, $N(0) = 11$ **25.** $N(t+1) = 4N(t)$, $N(0) = 30$ **27.**


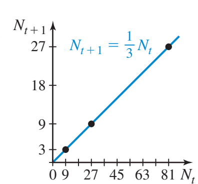
29.


31. 3, 6, 12, 24, 48, 96 **33.** 2, 6, 18, 54, 162, 486 **35.** 1, 5, 25, 125, 625, 3125 **37.** 640, 320, 160, 80, 40, 20 **39.** 1215, 405, 135, 45, 15, 5 **41.** 31250, 6250, 1250, 250, 50, 10

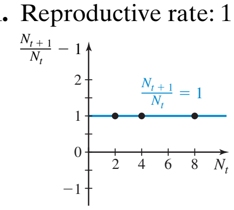
43.


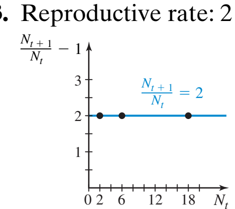
45.


47.


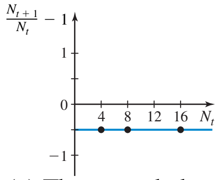
49.


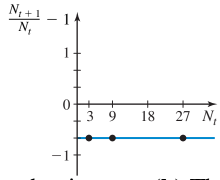
In Problems 51–58 graph the reproductive rate $(N_{t+1}/N_t - 1)$ against N_t for the indicated value of R and locate the points $(N_0, N_1/N_0 - 1)$ on the plot for the given value of N_0.

51. Reproductive rate: 1


53. Reproductive rate: 2


55. Reproductive rate: $-\frac{1}{2}$ **57.** Reproductive rate: $-\frac{2}{3}$





59. **(a)** The second; decreasing reproductive rate **(b)** The first; constant reproductive rate **(c)** The third; increasing reproductive rate **61.** **(a)** The third; increasing reproductive rate **(b)** The first; constant reproductive rate **(c)** The second; decreasing reproductive rate

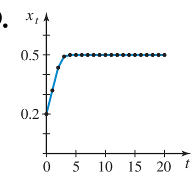
Section 2.2

1. 1, 2, 3, 4, 5, 6 **3.** undefined, 3, 2, $\frac{5}{3}$, $\frac{3}{2}$, $\frac{7}{5}$ **5.** 1, $\frac{1}{4}$, $\frac{1}{9}$, $\frac{1}{16}$, $\frac{1}{25}$, $\frac{1}{36}$ **7.** 1, 4, 9, 16, 25, 36 **9.** 0, 0, 0, 0, 0, 0 **11.** 0, $\frac{1}{2}$, $\frac{4}{3}$, $\frac{9}{4}$, $\frac{16}{5}$, $\frac{25}{6}$ **13.** 1, $e^{1/2} \approx 1.649$, $e \approx 2.718$, $e^{3/2} \approx 4.482$, $e^2 \approx 7.389$, $e^{5/2} \approx 12.183$ **15.** 1, $\frac{1}{3}$, $\frac{1}{9}$, $\frac{1}{27}$, $\frac{1}{81}$, $\frac{1}{243}$ **17.** 36, 49, 64, 81 **19.** $\frac{1}{6}$, $\frac{1}{7}$, $\frac{1}{8}$, $\frac{1}{9}$ **21.** $\frac{6}{7}$, $\frac{7}{8}$, $\frac{8}{9}$, $\frac{9}{10}$ **23.** $\sqrt{6+e^6}$, $\sqrt{7+e^7}$, $\sqrt{8+e^8}$, $\sqrt{9+e^9}$ **25.** $a_n = n$, $n = 0, 1, 2, \ldots$ **27.** $a_n = 2^n$, $n = 0, 1, 2, \ldots$ **29.** $a_n = \frac{1}{3^n}$, $n = 0, 1, 2, \ldots$ **31.** $a_n = (-1)^{n+1}(n+1)$, $n = 0, 1, 2, \ldots$ **33.** $a_n = 2n + 5$, $n = 0, 1, 2, \ldots$ **35.** $a_n = 1 - (-1)^{n+1}$, $n = 0, 1, 2, \ldots$ **37.** ≈ 1.618 **39.** -1.095 **41.** ≈ 1.490 **43.** ≈ 1.333 **45.** 1, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$; 0 **47.** 0, $\frac{1}{2}$, $\frac{2}{3}$, $\frac{3}{4}$, $\frac{4}{5}$; 1 **49.** 5, 3, $\frac{9}{5}$, $\frac{7}{5}$, $\frac{21}{17}$; 1 **51.** -1, $\frac{1}{2}$, $-\frac{1}{3}$, $\frac{1}{4}$, $-\frac{1}{5}$; 0 **53.** 0, $\frac{1}{2}$, $\frac{4}{3}$, $\frac{9}{4}$, $\frac{16}{5}$; limit does not exist. **55.** 0, 1, $\sqrt{2}$, $\sqrt{3}$, 2; limit does not exist. **57.** 1, 2, 4, 8, 16; limit does not exist. **59.** 1, 3, 9, 27, 81; limit does not exist. **61.** $a = 0, N = 100$ **63.** $a = 0, N = 10$ **65.** $a = 0, N = 100$ **67.** $a = 0, N = 100$ **69.** $a = 0, N = 2$ **71.** $a = 0, N = 6$ **73.** $N = 3/\epsilon$. **75.** $N = \sqrt{1/\epsilon}$. **77.** $N = -\frac{\log \epsilon}{3 \log 2}$ assuming $\epsilon < 1$. **79.** 0 **81.** 1 **83.** 1 **85.** 0 **87.** ∞ **89.** 1 **91.** 2, 4, 8, 16, 32 **93.** $-2, 4, -8, 16, -32$ **95.** 1, 3, 7, 15, 31 **97.** $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$, $\frac{1}{6}$ **99.** 2, $\frac{5}{2}$, $\frac{29}{10}$, $\frac{941}{290}$, $\frac{969{,}581}{272{,}890}$ **101.** 4 **103.** $\frac{5}{3}$ **105.** 2, -2 **107.** $-1+\sqrt{3}$, $-1-\sqrt{3}$ **109.** 0, 5 **111.** 2; 2 **113.** 0, 2; 2 **115.** 0, $\frac{1}{2}$; $\frac{1}{2}$ **117.** 1, -1; 1 **119.** $1+\sqrt{2}+\sqrt{3}+2$ **121.** $9+27+81+243+729$ **123.** $1+2+4+8$ **125.** $\sum_{i=1}^{n} 2i$ **127.** $\sum_{n=2}^{5} \ln n$ **129.** $\sum_{n=4}^{8} \left(1 - \frac{5}{n}\right)$ **131.** $\sum_{i=1}^{n} q^{i-1}$

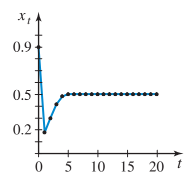
Section 2.3

1. The number of cod next year is equal to the number this year minus those that die from old age, are killed by predators, or are caught by fishermen during the year, plus the number that are born during the year. **3.** The size of the wild population of kakapo next year is equal to the number this year minus those that are removed for captive breeding, killed by predators, or die from disease during the year, plus the number that are born or are reintroduced from captive breeding during the year. **5.** The area of living coral next year is equal to the area of living coral this year minus the area killed by ocean acidification or fishing during the year, plus the area of reef restored or rebuilt during the year. **7.** The population of amoeba cells one hour from now is equal to the population now minus the number that die in one hour, plus the number that divide in two in one hour. **9.** **(a)** (i) The number born is computed from the current population by multiplying it by the fraction of females (0.5), the fraction of those that lay in a particular year (0.25), and the fraction that survive their first year (0.29). (ii) $\frac{1}{50}N_i = 0.02N_i$. (iii) $N_1 = 50.8125 \approx 51$, $N_2 \approx 51.64 \approx 52$, $N_3 \approx 52.48 \approx 52$, $N_4 \approx 53.33 \approx 53$, $N_5 \approx 54.20 \approx 54$. (iv) $N_t = N_0 \cdot 1.01625^t$; $N_{44} > 100 > N_{43}$, $N_{87} > 200 > N_{86}$. **(b)** (i) $N_{t+1} = N_t + 0.5 \cdot 0.5 \cdot 0.29N_t - 0.02N_t = 1.0525N_t$; $N_1 = 52.625 \approx 53$, $N_2 \approx 55.3878 \approx 55$, $N_3 \approx 58.2957 \approx 58$, $N_4 \approx 61.3562 \approx 61$, $N_5 \approx 64.5774 \approx 65$; (ii) $N_{t+1} = N_t + 0.5 \cdot 0.25 \cdot 0.75N_t - 0.02N_t = 1.07375N_t$; $N_1 = 53.6875 \approx 54$, $N_2 \approx 57.647 \approx 58$, $N_3 \approx 61.8984 \approx 62$, $N_4 \approx 66.4634 \approx 66$, $N_5 \approx 71.3651 \approx 71$; (iii) strategy 2. **11.** 0, 90 **13.** 0, 90 **15.** 0, 60 **17.** 2, 3.922, 7.547, 14.035, 24.615, 39.506; 100 **19.** 7, 15.556, 26.25, 34.054, 37.8, 39.239; 40 **21.** 2, 7.619, 25.6, 62.439, 97.524, 113.463; 120 **23.** $\frac{1}{50}$ **25.** $\frac{1}{60}$ **27.** $x_t = \frac{1}{10}N_t$ **29.** $x_t = \frac{1}{30}N_t$ **31.** $x_t = \frac{1}{75}N_t$ **33.** 0, $\frac{R_0-1}{b}$ **35.** $\frac{3}{80}$ **37.** $\frac{1}{15}$

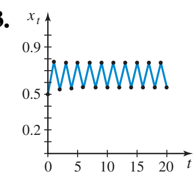
39.



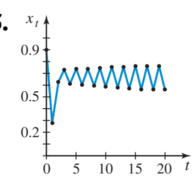
41.



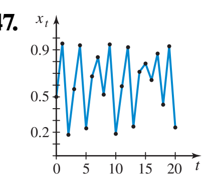
43.



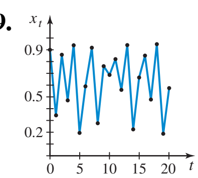
45.



47.

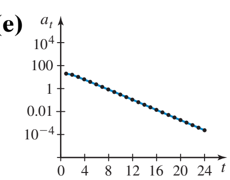


49.



51. **(a)** $C_{t+1} = 0.923C_t$ **(b)** $C_t = 33.8(0.923)^t$ **(c)** After $72.67 \approx 73$ hours **53.** **(a)** $a_{t+1} = 0.6a_t + 20 \cdot (0.2)^t$ **(b)** 20.00, 16.00, 10.40, 6.40, 3.87, 2.33 **(c)** 20, at $t = 1$ **(d)** $a_7 = 1.40$, 0.84, 0.50, 0.30, 0.18, 0.11, 0.07, 0.04, 0.02, 0.01, 0.01, 0.01, 0, 0, 0, 0, 0, 0

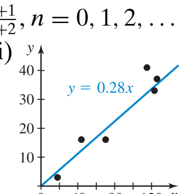
(e)



(c) $a_{t+1} = 20(0.7)^t$ **(d)** No. **55.** **(a)** 30% **(b)** $a_{t+1} = 0.7a_t$ **57.** **(a)** 15 mg/ml **(b)** $c_{t+1} = c_t - 15$, $c_2 = 20$ **(c)** 30% **(d)** $c_{t+1} = 0.7c_t$, $c_2 = 24.5$ **(e)** Zeroth-order **59.** First-order **61.** Zeroth-order **63.** **(a)** The amount of hormone in the blood one day from now is the amount in the blood now, plus 20μg added, minus 4% of what is there today. **(b)** $a_{t+1} = 0.96a_t + 20$ **(c)** 20, 39.2, 57.6, 75.3, 92.3, 108.6μg **(d)** 500μg

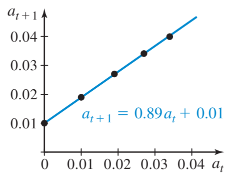
Chapter 2 Review Problems

1. 0 **3.** 40 **5.** ∞ **7.** 1 **9.** 0 **11.** $a_n = \frac{2n+1}{2n+2}$, $n = 0, 1, 2, \ldots$ **13.** $a_n = \frac{n+1}{(n+1)^2+1}$, $n = 0, 1, 2, \ldots$ **15.** **(a)** (i)



(ii) $N_{t+1} = 1.06N_t$ (iii) 138, 146, 155, 164, 174, 184, 195, 207, 220, 233 (iv) 9 years **(b)** (i) Add r for the r wolves introduced; the 22% death rate becomes a 30% death rate (ii) $r = 5$: 132, 135, 137, 139, 142, 144, 146, 148, 150, 152 **(c)** $N_{t+1} = 1.06N_t + r$; earlier **17.** **(a)** 10, 20, 37, 67, 112, 161 fish **(b)** 200 fish **(c)** $200(1-p)$ fish

19. **(a)** $a_{t+1} = (1-p)a_t + 0.01$ **(b)**



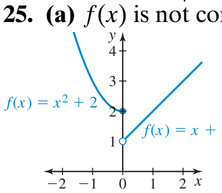
The line has slope 0.89; since $\frac{a_t - a_{t-1}}{a_{t+1}-a_t} \approx \frac{a_t}{a_{t+1}} \approx 1 - p$, this gives $p \approx 0.11$ **(c)** 0.046, 0.051, 0.055, 0.059, 0.063 **(d)** The equation with no added insulin is $a_{t+5} = 0.89a_{t+4}$ with $a_5 = 0.040$, giving $a_6 = 0.036, 0.032, 0.028, 0.025, a_{10} = 0.022$.

Section 3.1

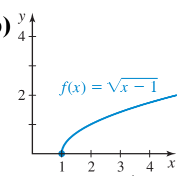
1. 0 **3.** -1 **5.** $\frac{3}{2}\sqrt{2} \approx 2.1213$ **7.** $\frac{4}{3}\sqrt{3} \approx 2.3094$ **9.** 1 **11.** 0 **13.** 7 **15.** 0 **17.** 1 **19.** ∞ **21.** $-\infty$ **23.** ∞ **25.** ∞ **27.** ∞ **29.** 0 **31.** $\frac{1}{2}$ **33.** $-\infty; -\infty$ **35.** 1 **37.** **(a)** ≈ 0.3679 **39.** -9 **41.** 54 **43.** $\frac{53}{3}$ **45.** $\frac{47}{2}$ **47.** $\frac{28}{3}$ **49.** 2 **51.** 4 **53.** $-\frac{1}{4}$ **55.** -5

Section 3.2

1. $f(1) = 2$ **3.** $f(2) = 13$ **5.** $f(2) = 3$ **7.** $a = 6$ **9.** $x = 3$ **11.** $x = -1$ **13.** $f(5/2) = 2$. See Example 4 for $k = 3$. **15.** $x \in \mathbf{R}$ **17.** $x \neq 1$ **19.** $x \in \mathbf{R}$ **21.** $x \neq -1$ **23.** $\{x \in \mathbf{R} : x \neq \frac{1}{4} + \frac{k}{2}, k \in \mathbf{Z}\}$ **25.** **(a)** $f(x)$ is not continuous at $x = 0$ **(b)** $c = 2$



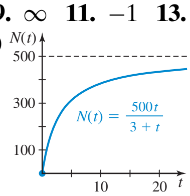
27. **(b)**



(c) No **29.** $\frac{1}{2}$ **31.** 1 **33.** 3 **35.** 1 **37.** 1 **39.** 1 **41.** 2 **43.** $\frac{1}{4}$ **45.** $\frac{1}{6}$ **47.** 0 **49.** **(b)** $N_c = A$ **51.** **(a)** $T_c \approx 42.6977$

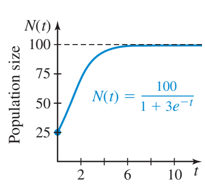
Section 3.3

1. 0 **3.** ∞ **5.** -2 **7.** ∞ **9.** ∞ **11.** -1 **13.** 0 **15.** 1 **17.** 0 **19.** $\frac{3}{2}$ **21.** $\frac{3}{2}$ **23.** 0 **27.** **(a)**



(b) 500 **(c)** 250

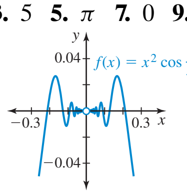
29. **(a)**



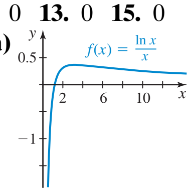
(b) 100

Section 3.4

1. 1 **3.** 5 **5.** π **7.** 0 **9.** 1 **11.** 0 **13.** 0 **15.** 0 **17.** **(a)**



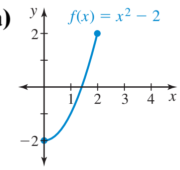
19. **(a)**



(b) $x \geq e$

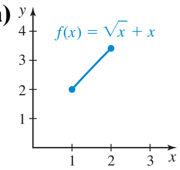
Section 3.5

1. **(a)**

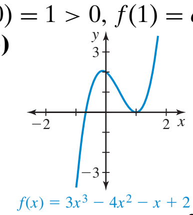


(b) $f(0) = -2, f(2) = 2$

3. **(a)**



(b) $f(1) = 2 < 3 < f(2) = \sqrt{2} + 2$ **5.** Let $f(x) = e^{-x} - x^2$ $f(0) = 1 > 0$, $f(1) = e^{-1} - 1 < 0$ **7.** $x \approx 0.70$ **9.** **(a)** $x \approx -0.67$ **(b)**

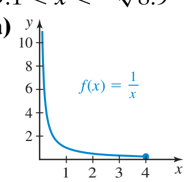


(c) No

11. If $f(x)$ has largest degree term ax^3, either $\lim_{x\to\infty} f(x) = \infty$ and $\lim_{x\to\infty} f(x) = -\infty$ (if $a > 0$) or $\lim_{x\to\infty} f(x) = -\infty$ and $\lim_{x\to\infty} f(x) = \infty$ (if $a < 0$). Thus $f(\mathbf{R}) = \mathbf{R}$. **13.** If $f(x) = x^2 - 5$, $f(0) = -5 < 0$, $f(-3) = f(3) = 4 > 0$. **15.** **(b)** ≈ 0.4429 **17.** ≈ -1.61803 **19.** ≈ 2.60585

Section 3.6

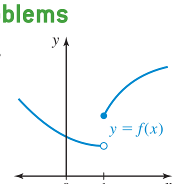
1. $-0.505 < x < -0.495$ **3.** $\sqrt{8.9} < x < \sqrt{9.1}$ or $-\sqrt{9.1} < x < -\sqrt{8.9}$ **5.** **(b)** $1.95 < x < 2.05$ **7.** **(a)**



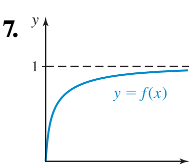
(b) $0 < x < \frac{1}{5}$ **9.** $\delta = \frac{\epsilon}{2}$ **11.** $\delta = \epsilon^{1/5}$ **13.** $\delta = \sqrt{\frac{3}{M}}$ **15.** $\delta = \frac{1}{\sqrt[4]{M}}$ **17.** $\delta = \frac{1}{M}$ **19.** $\delta = -\frac{1}{M}$ **21.** $\delta = \frac{\epsilon}{|m|}$

Chapter 3 Review Problems

1. $x \in \mathbf{R}$ **3.** $x \in \mathbf{R}$ **5.**



7.

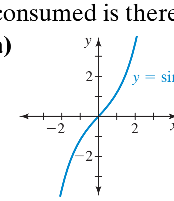


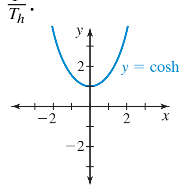
9. $f(-2) = -2$, $\lim_{x\to-2^+} f(x) = -2$, $\lim_{x\to-2^-} f(x) = -3$ **11.** $a = 1.24 \times 10^6, k = 5$ **13.** **(a)** To plot $g(t)$, note that

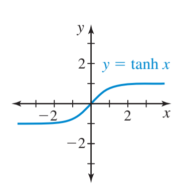
$$g(t) = \begin{cases} 1 & \text{for } \frac{1}{6} + 2k \leq x \leq \frac{5}{6} + 2k, k = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

(b) $s(t)$ is continuous, $g(t)$ is not continuous **15.** **(a)** No value of A **(b)** 3.6 **17.** **(c)** $T_h = 0$: $g(N) = aTN = f(N)$ **(d)** $\lim_{N\to\infty}(aTN) = \infty$; $\lim_{N\to\infty} \frac{aTN}{1+aT_hN} = \frac{T}{T_h}$. If prey density is infinite, and predator has no handling time, then a predator can catch infinitely many prey. Handling time means that a predator can consume at most one prey in time T_h. Maximum number of prey consumed is therefore $\frac{T}{T_h}$.

19. **(a)**







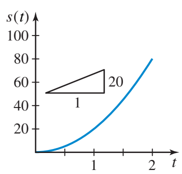
(b) $\lim_{x\to\infty}\sinh x=\infty$, $\lim_{x\to-\infty}\sinh x=-\infty$, $\lim_{x\to\infty}\cosh x=\infty$, $\lim_{x\to-\infty}\cosh x=\infty$, $\lim_{x\to\infty}\tanh x=1$, $\lim_{x\to-\infty}\tanh x=-1$

Section 4.1

1. 0 **3.** 2 **5.** 0 **7.** 0 **9.** 0 **11.** -2 **13.** -3 **15.** $c=2k+1$, $k\in\mathbf{Z}$ **17.** $-2h$ **19.** $\sqrt{4+h}-2$ **21.** **(a)** $f'(-1)=-4$ **(b)** $y=-4x-2$ **23.** **(a)** $f'(2)=-12$ **(b)** $y=\frac{1}{12}x-\frac{43}{6}$ **25.** $y'=\frac{1}{2\sqrt{x}}$, $\frac{\sqrt{2}}{4}$ **27.** $y=1$ **29.** $y=\frac{1}{4}x+1$ **31.** $y=-\frac{1}{7}x-\frac{29}{7}$ **33.** $y=-\frac{1}{4}x+\frac{5}{4}$ **35.** $f(x)=x^2$ **37.** $f(x)=\frac{1}{x^2+1}$, $a=2$

Section 4.2

1. The rate at which the fish grows **3.** The change in heart rate if the mammal's mass increases by one unit. **5.** The increase in the number of cars exiting if one extra car is added to the freeway. **7.** The rate of increase of the mammal's mass over time **9.** The rate at which the height of water collected changes over time **11.** **(a)**



(b) 20 km/hr **(c)** 20 km/hr; instantaneous velocity shown by triangle in graph. **13.** **(a)** $s\left(\frac{3}{4}\right)=30$, $s(1)=\frac{160}{3}$ **(b)** $\frac{280}{3}$ **(c)** $v\left(\frac{3}{4}\right)=80$, $\left|v\left(\frac{3}{4}\right)\right|=80$ **15.** **(a)** $t=\frac{2v}{g}$ **(d)** **(i)** 5m **(ii)** From $t=0$ to $t=1$ **(iii)** From $t=1$ to $t=2$ **(iv)** 0 **17.** **(a)** rN **(c)** Larger **(d)** The same **19.** $\frac{d[C]}{dt}=k[A][B]$; $[C]=x$, $[A]=a-x$, $[B]=b-x$. **21.** **(b)** $\frac{d[C]}{dt}=k[A]$, $\frac{d[A]}{dt}=-k[A]$ **(c)** B is created by decay of A. So $[B]=$ amount of A that has decayed; $[B]=a-[A]$. **23.** B **25.** No, $f(x)$ could have a jump discontinuity at $x=c$. **27.** $x=-5$ **29.** $x=3$ **31.** $x=3$ **33.** $x=-3$ **35.** $x=\pm\sqrt{1/2}$ **37.** $x=1$ **39.** $x=0$

Section 4.3

1. $12x^2-7$ **3.** $-10x^4+7$ **5.** $-4-10x$ **7.** $35s^6+6s^2-5$ **9.** $-\frac{4}{3}t^3+4$ **11.** $2x\sin\frac{\pi}{3}=x\sqrt{3}$ **13.** $-12x^3\tan\frac{\pi}{6}=-4\sqrt{3}x^3$ **15.** $3t^2e^{-2}+1$ **17.** $3s^2e^3$ **19.** $60x^2-24x^5+72x^7$ **21.** $3\pi x^2+\frac{1}{\pi}$ **23.** $3ax^2$ **25.** $2ax$ **27.** $2rs$ **29.** $3rs^2x^2-r$ **31.** $4(b-1)N^3-\frac{2N}{b}$ **33.** a^3-3at^2 **35.** $V_0\gamma$ **37.** $1-\frac{2N}{K}$ **39.** $2rN-3\frac{r}{K}N^2$ **41.** $8\frac{2\pi^5}{15}\frac{k^4}{c^2h^3}T^3$ **43.** $191x-y+377=0$ **45.** $3x-y-6=0$ **47.** $8x-\sqrt{2}y-18=0$ **49.** $x-2y+7=0$ **51.** $x-24y+73\sqrt{3}=0$ **53.** $x+3y+5=0$ **55.** $2ax-y-a=0$ **57.** $(a^2+2)y-4ax+4a=0$ **59.** $\frac{1}{3a}x+y+a+\frac{1}{3a}=0$ **61.** $2a(a+1)y+\frac{1}{2}(a+1)^2x-8a^2-(a+1)^2=0$ **63.** $(0,0)$ **65.** $\left(\frac{3}{2},\frac{9}{4}\right)$ **67.** $(0,0)$ and $\left(\frac{2}{9},-\frac{4}{243}\right)$ **69.** $(0,0)$, $\left(-\frac{1}{2},-\frac{17}{96}\right)$, and $\left(4,-\frac{160}{3}\right)$ **71.** $(0,4)$ **73.** $\left(\frac{1}{4},-\frac{3}{8}\right)$ **75.** $\left(\frac{1}{3}\sqrt{3},\frac{7}{9}\sqrt{3}+2\right)$, $\left(-\frac{1}{3}\sqrt{3},-\frac{7}{9}\sqrt{3}+2\right)$
77. Tangent line: $y=2x-1$ **79.** $y=2ax-a^2$ and $y=-2ax-a^2$ **81.** $P'(x)$ is a polynomial of degree 3. **85.** Rates of growth; 3.28, 2.95, 2.00 mm/week. Rate of growth decreases.

Section 4.4

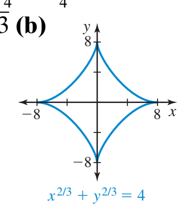
1. $f'(x)=3x^2+10x-3$ **3.** $f'(x)=-105x^6+30x^4+75x^2-10$ **5.** $f'(x)=x(2x+3x^2)+(\frac{1}{2}x^2-1)(2+6x)$ **7.** $f'(x)=\frac{4}{5}x^3$ **9.** $f'(x)=6(3x-1)$ **11.** $f'(x)=-12(1-2x)$ **13.** $g'(s)=2(4s-5)(2s^2-5s)$ **15.** $g'(t)=6(4t-20t^3)(2t^2-5t^4)$ **17.** $y=-16x+17$ **19.** $y=-56x-64$ **21.** $y=-\frac{1}{22}x+\frac{133}{22}$ **23.** $y=\frac{1}{5}x+2$ **25.** $f'(x)=9x^2-52x+61$ **27.** $f'(x)=(6x^2-12x+1)(1-x^2)-2x(x-3)(2x^2+1)$ **29.** $f'(x)=a(4x+1)$ **31.** $f'(x)=4ax$ **33.** $g'(t)=2a(at+1)$ **35.** 11 **37.** 2 **39.** -27 **41.** $y'=3f'(x)g(x)+3f(x)g'(x)$ **43.** $y'=f'(x)g(x)+f(x)g'(x)+4g(x)g'(x)$ **45.** $H'(p)=2-4p$; $p=\frac{1}{2}$ **47.** **(a)** $N>0$, $N-a<0$, $1-N/K>0$ so $f(N)<0$ **(b)** $N>0$, $N-a>0$, $1-N/K>0$ so $f(N)>0$ **(c)** $f'(0)=-ar<0$, $f'(K)=-rK(K-a)<0$ **(d)** $f'(a)=ra(1-a/K)>0$ **49.** $f'(x)=\frac{4}{(x+1)^2}$ **51.** $2\frac{3x^2+3x-2}{(2x+1)^2}$ **53.** $f'(x)=\frac{2x^3-3x^2+3}{(1-x)^2}$ **55.** $h'(t)=\frac{t^2+2t-4}{(t+1)^2}$ **57.** $f'(s)=\frac{2s^2-4s+4}{(1-s)^2}$ **59.** $f'(x)=\frac{1}{2\sqrt{x}}(x-1)+\sqrt{x}$ **61.** $f'(x)=\frac{\sqrt{3}}{2\sqrt{x}}(x^2-1)+2x\sqrt{3x}$ **63.** $f'(x)=3x^2-\frac{3}{x^4}$ **65.** $f'(x)=4x+\frac{x-6}{x^3}$ **67.** $g'(s)=\frac{2s^{-1/3}-s^{-2/3}-1}{3(s^{2/3}-1)^2}$ **69.** $f'(x)=(-2)\left(\sqrt{2x}+\frac{2}{\sqrt{x}}\right)+(1-2x)\left(\frac{1}{\sqrt{2x}}-\frac{1}{x^{3/2}}\right)$ **71.** $y=\frac{3}{5}$ **73.** $y=\frac{7}{8}-\frac{19}{16}(x-2)$ **75.** $f'(x)=\frac{3a}{(3+x)^2}$ **77.** $f'(x)=\frac{8ax}{(4+x^2)^2}$ **79.** **(a)** $f'(P)=\frac{nP^{n-1}k^n}{(k^n+P^n)^2}$ **81.** $h'(t)=\frac{\sqrt{a}}{2\sqrt{t}}(t-a)+\sqrt{at}+a$ **83.** -4 **85.** $\frac{20}{9}$ **87.** 1 **89.** $y'=\frac{f'(x)g(x)-2f(x)g'(x)}{g(x)^3}$ **91.** $y'=\frac{1}{2\sqrt{x}}f(x)g(x)+\sqrt{x}f'(x)g(x)+\sqrt{x}f(x)g'(x)$ **93.** $y=-\frac{cx}{x_1^2}+2\frac{c}{x_1}$; intercepts x-axis at $x=2x_1$

Section 4.5

1. $2(x-3)$ **3.** $-24x(1-3x^2)^3$ **5.** $\frac{x}{\sqrt{x^2+3}}$ **7.** $\frac{-3x^2}{2\sqrt{1-x^3}}$ **9.** $-\frac{12x^2}{(x^3-1)^5}$ **11.** $-\frac{2x+3}{(2x^2-1)^{3/2}}$ **13.** $\frac{1-3x}{\sqrt{2x-1}(x-1)^3}$ **15.** $\frac{1}{n}\left(s+s^{1/n}\right)^{-1+1/n}\left(1+\frac{1}{n}s^{-1+1/n}\right)$ **17.** $\frac{-9t^2}{(t-3)^4}$ **19.** $(r^2-r)^2(r+3r^3)^{-5}[3(2r-1)(r+3r^3)-4(1+9r^2)(r^2-r)]$ **21.** $-\frac{4}{5}x^3(3-x^4)^{-4/5}$ **23.** $\frac{2x-2}{7(x^2-2x+1)^{6/7}}$ **25.** $g'(s)=\frac{3}{2}(3s^7-7s)^{1/2}(21s^6-7)$ **27.** $\frac{2}{5}(3t+\frac{3}{t})^{-3/5}(3-\frac{3}{t^2})$ **29.** $3a(ax+1)^2$ **31.** $g'(N)=\frac{bk}{(k+N)^2}$ **33.** $g'(T)=-3a(T_0-T)^2$ **35.** **(a)** $\frac{2x}{x^2+3}$ **(b)** $\frac{1}{2(x-1)}$ **37.** $\frac{f'(x)g(x+1)-f(x)g'(x+1)}{g(x+1)^2}$ **39.** $\frac{2f'(2x)(g(2x)+2x)-f(2x)(2g'(2x)+2)}{(g(2x)+2x)^2}$ **41.** $y'=4(\sqrt{x^3-3x}+3x)^3\left(\frac{3x^2-3}{2\sqrt{x^3-3x}}+3\right)$ **43.** $y'=36x(3x^2-1)^2(1+(3x^2-1)^3)$ **45.** $y'=3\left(\frac{2x+1}{3(x^3-1)^3-1}\right)^2\frac{6(x^3-1)^3-2-27x^2(2x+1)(x^3-1)^2}{(3(x^3-1)^3-1)^2}$ **47.** **(a)** $\frac{dc}{dt}=-\frac{0.128k}{(1+2\sqrt{t})^{1.128}\sqrt{t}}$ **(b)** $\frac{dc}{dL}=-0.410\frac{k}{r}\left(\frac{L}{r}\right)^{-1.410}$

Section 4.6

1. $\frac{dy}{dx}=-\frac{x}{y}$ **3.** $\frac{dy}{dx}=-\left(\frac{y}{x}\right)^{1/4}$ **5.** $\frac{dy}{dx}=4\sqrt{xy}-\frac{y}{x}$ **7.** $\frac{dy}{dx}=\frac{x}{y}$ **9.** **(a)** $y=\frac{4}{3}x-\frac{25}{3}$ **(b)** $y=-\frac{3}{4}x$ **11.** **(a)** $y=\frac{3}{4}x-\frac{9}{4}$ **(b)** $y=-\frac{4}{3}x+\frac{136}{9}$ **13.** **(a)** $\frac{dy}{dx}=(27)^{1/6}=\sqrt{3}$ **(b)**



15. $-\frac{2}{3}\sqrt{3}$ **17.** $-\frac{3}{4}$ **19.** $\frac{dV}{dt}=3x^2\frac{dx}{dt}$ **21.** $\frac{dS}{dt}=8\pi r\frac{dr}{dt}$ **23.** $\frac{dh}{dt}=-\frac{1}{100\pi}$ m/min ≈ 0.0032 m/min **25.** $3\sqrt{61}\,\frac{\text{mi}}{\text{hr}}$ for both $t=20$ min and $t=40$ min **27.** $\frac{dE}{dt}=\frac{3}{4}cM^{-1/4}\frac{dM}{dt}$ **29.** $\frac{dS}{dt}=\frac{2}{a}\frac{dV}{dt}$

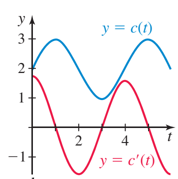
Section 4.7

1. $f'(x)=3x^2-6x$, $f''(x)=6x-6$ **3.** $g'(x)=-(x+1)^{-2}$, $g''(x)=2(x+1)^{-3}$ **5.** $g'(t)=\frac{9t^2+2}{2\sqrt{3t^3+2t}}$, $g''(t)=\frac{27t^4+36t^2-4}{4(3t^3+2t)^{3/2}}$ **7.** $f'(s)=\frac{3}{2}s^{1/2}$, $f''(s)=\frac{3}{4}s^{-1/2}$ **9.** $g'(t)=-\frac{5}{2}t^{-7/2}-\frac{1}{2}t^{-1/2}$, $g''(t)=\frac{35}{4}t^{-9/2}+\frac{1}{4}t^{-3/2}$ **11.** $f'(x)=6x^5$, $f''(x)=30x^4$, $f'''(x)=120x^3$, $f^{(4)}(x)=360x^2$, $f^{(5)}(x)=720x$, $f^{(6)}(x)=720$, $f^{(7)}=\cdots=f^{(10)}(x)=0$ **13.** $p(x)=3x^2+2x+3$

15. **(a)** velocity: $v_0 - gt$, acceleration: $-g$ **(b)** $t = \frac{v_0}{g}$; up right before; down right after. **17.** **(a)** $r = \sqrt[6]{2a/b}$ **(b)** $V''(r) = 156ar^{-14} - 42br^{-8}$ $V''(\sqrt[6]{\frac{2a}{b}}) = \frac{18b^{7/3}}{2^{1/3}a^{4/3}} > 0$

Section 4.8

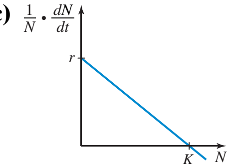
1. $f'(x) = 2\cos x + \sin x$ **3.** $f'(x) = 3\cos x - 5\sin x$
5. $f'(x) = -\sin(x+1)$ **7.** $f'(x) = 3\cos(3x)$
9. $f'(x) = 6\cos(3x+1)$ **11.** $f'(x) = 4\sec^2 4x$
13. $f'(x) = 4\sec(1+2x)\tan(1+2x)$ **15.** $f'(x) = 6x\cos(x^2)$
17. $f'(x) = 4x\sin(x^2-3)\cos(x^2-3)$
19. $f'(x) = 12x\sin x^2\cos x^2$ **21.** $f'(x) = -8x\sin x^2 + 4\sin x\cos x$
23. $f'(x) = -8\sin x\cos x - 8x^3\sin x^4$
25. $f'(x) = -4x\sec^2(1-x^2)$ **27.** $f'(x) = \frac{\cos\sqrt{x}}{2\sqrt{x}}$
29. $f'(x) = \frac{2x\cos(2x^2-1)}{\sqrt{\sin(2x^2-1)}}$ **31.** $g'(s) = \frac{-\sin s}{2\sqrt{\cos s}}$
33. $g'(t) = \frac{2\cos(2t)(\cos(6t)-1)+6\sin(6t)(\sin(2t)+1)}{(\cos(6t)-1)^2}$
35. $f'(x) = \frac{2x\sec(x^2-1)[\tan(x^2-1)+\cot(x^2+1)]}{\csc(x^2+1)}$
37. $f'(x) = 2\cos(2x-1)\cos(3x+1) - 3\sin(2x-1)\sin(3x+1)$
39. $f'(x) =$
$6x\sec^2(3x^2-1)\cot(3x^2+1) - 6x\csc^2(3x^2+1)\tan(3x^2-1)$
41. $f'(x) = \sec^2 x$ **43.** $f'(x) = 0$ **45.** $g'(x) = \frac{-6x\cos(3x^2-1)}{\sin^2(3x^2-1)}$
47. $g'(x) = 20x\cot(1-5x^2)\csc^2(1-5x^2)$
49. $h'(x) = -\frac{6\sec^2(2x)-3}{(\tan(2x)-x)^2}$ **51.** $h'(s) = 3\sin s\cos s(\sin s - \cos s)$
53. $f'(x) = \frac{2(1+x^2)\cos(2x)-2x\sin(2x)}{(1+x^2)^2}$ **55.** $f'(x) = -\frac{1}{x^2}\sec^2(\frac{1}{x})$
57. $f'(x) = \frac{2\sin x\cos x^2-2x\cos x\sin x^2}{\cos^3 x}$ **59.** $x = \frac{3}{2} + 3k, k \in \mathbf{Z}$
61. Write $\frac{d}{dx}\cos x = \lim_{h\to 0}\frac{\cos(x+h)-\cos x}{h}$ and use the identity for $\cos(x+h)$. **63.** $\frac{d}{dx}\sec x = \frac{d}{dx}\left(\frac{1}{\cos x}\right) = \frac{(0)(\cos x)-(1)(-\sin x)}{\cos^2 x} = \frac{\sin x}{\cos^2 x} = \frac{1}{\cos x}\cdot\frac{\sin x}{\cos x} = \sec x\cdot\tan x$. **65.** $f'(x) = \frac{x\cos\sqrt{x^2+1}}{\sqrt{x^2+1}}$
67. $f'(x) = (\cos\sqrt{3x^2+3x})\frac{6x+3}{2\sqrt{3x^2+3x}}$
69. $f'(x) = 4x\sin(x^2-1)\cos(x^2-1)$
71. $f'(x) = 2\cos 2x + 2\sin x\cos x$ **73.** **(a)** $\frac{dc}{dt} = \frac{\pi}{2}\cos(\frac{\pi}{2}t)$
(b)



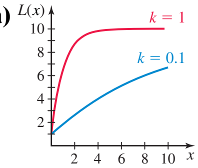
(c) **(i)** $\frac{dc}{dt} = 0$ **(ii)** increasing **(iii)** $c(t)$ has a horizontal tangent and either a maximum or a minimum.

Section 4.9

1. $f'(x) = 3e^{3x}$ **3.** $f'(x) = -12e^{1-3x}$
5. $f'(x) = (-4x+3)e^{-2x^2+3x-1}$ **7.** $f'(x) = 28x(x^2+1)e^{7(x^2+1)^2}$
9. $f'(x) = e^x(1+x)$ **11.** $f'(x) = xe^{-x}(2-x)$
13. $f'(x) = \frac{e^x(1+x^2-2x)-2x}{(1+x^2)^2}$ **15.** $f'(x) = \frac{2e^x-2e^{-x}-2}{(2+e^x)^2}$
17. $f'(x) = 3\cos(3x)e^{\sin(3x)}$ **19.** $f'(x) = 2xe^{\sin(x^2-1)}\cos(x^2-1)$
21. $f'(x) = e^x\cos(e^x)$ **23.** $f'(x) = (2e^{2x}+1)\cos(e^{2x}+x)$
25. $f'(x) = (1-\cos x)\exp(x-\sin x)$
27. $g'(s) = 2s\exp[\sec s^2]\sec s^2\tan s^2$
29. $f'(x) = (\sin x + x\cos x)e^{x\sin x}$
31. $f'(x) = (-3)(2x+\sec^2 x)e^{x^2+\tan x}$ **33.** $f'(x) = (\ln 2)2^x$
35. $f'(x) = (\ln 2)2^{x+1}$ **37.** $f'(x) = \frac{\ln 5}{\sqrt{2x-1}}5^{\sqrt{2x-1}}$
39. $f'(x) = 2x(\ln 2)2^{x^2+1}$ **41.** $h'(t) = (2t)(\ln 2)2^{t^2-1}$
43. $f'(x) = \frac{\ln 2}{2\sqrt{x}}2^{\sqrt{x}}$ **45.** $f'(x) = \frac{x\ln 2}{\sqrt{x^2-1}}2^{\sqrt{x^2-1}}$ **47.** $h'(t) = \frac{\ln 5}{2\sqrt{t}}5^{\sqrt{t}}$
49. $g'(x) = -2\sin x 2^{2\cos x}\ln 2$ **51.** $g'(r) = \frac{\ln 3}{5r^{4/5}}3^{r^{1/5}}$ **53.** 2 **55.** 0
57. $2 + 2\ln 2(x-1)$ **59.** **(a)** $N(0) = 1$ **61.** $\frac{dN}{dt} = N(t)\ln 2$ which implies that $\frac{dN}{dt}$ is proportional to $N(t)$ **63.** **(a)** $\frac{dN}{dt} = \frac{rK(K-1)e^{-rt}}{(1+(K-1)e^{-rt})^2}$
(c)



65. **(a)**



(b) L_∞ is the limiting size and L_0 is the initial size. **(c)** The fish with $k = 1$ reaches $L = 5$ more quickly. **(d)** As the fish ages, its rate of growth decreases. **(e)** The larger the value of k, the more quickly the fish grows and reaches its limiting size.
67. $\frac{dW}{dt} = -4W(t)$ **69.** $\frac{dW}{dt} = -\frac{\ln 2}{5}W(t)$ **71.** **(a)** $W(4) = 6e^{-12}$ **(b)** half-life: $\frac{\ln 2}{3}$ **73.** **(a)** $\frac{dW}{dt} = (\ln\frac{2}{5})W(t)$ **(b)** $W(3) = 5(\frac{2}{5})^3$ **(c)** half-life: $\frac{\ln 2}{\ln 5-\ln 2}$

Section 4.10

1. $\frac{d}{dx}f^{-1}(x) = x$ **3.** $\frac{d}{dx}f^{-1}(x) = \frac{\sqrt{2}}{4\sqrt{x-2}}$ **5.** $f^{-1}(x) = \left(\frac{3-x}{2}\right)^{1/3}$, $x \le 3$, $\frac{d}{dx}f^{-1}(x) = -\frac{1}{6}\left(\frac{2}{3-x}\right)^{2/3}$ **7.** $\frac{d}{dx}f^{-1}(0) = \frac{1}{4}$ **9.** $\frac{d}{dx}f^{-1}(2) = 4$
11. $\frac{d}{dx}f^{-1}(2) = \frac{1}{3}$ **13.** $\frac{d}{dx}f^{-1}(\pi)$ does not exist; vertical tangent
15. $\frac{d}{dx}f^{-1}(0) = \frac{2}{3}$ **17.** $\frac{d}{dx}f^{-1}(-\ln 2) = \frac{1}{\sqrt{3}}$ **19.** $\frac{d}{dx}f^{-1}(1) = 1$
21. $\frac{d}{dx}f^{-1}(1) = \frac{1}{2}$ **23.** $f'(x) = \frac{1}{x+1}$ **25.** $f'(x) = \frac{-2}{1-2x}$
27. $f'(x) = \frac{2}{x}$ **29.** $f'(x) = \frac{6x^2-1}{2x^3-x}$ **31.** $f'(x) = 2(\ln x)\frac{1}{x}$
33. $f'(x) = \frac{8\ln x}{x}$ **35.** $f'(x) = \frac{x}{x^2+1}$ **37.** $f'(x) = \frac{1}{x(x+1)}$
39. $f'(x) = \frac{-1}{1-x} - \frac{2}{1+2x}$ **41.** $f'(x) = \left(1-\frac{1}{x}\right)\exp[x-\ln x]$
43. $f'(x) = \cot x$ **45.** $f'(x) = \frac{2x\sec^2(x^2)}{\tan(x^2)}$ **47.** $f'(x) = 1 + \ln x$
49. $f'(x) = \frac{1-\ln x}{x^2}$ **51.** $h'(t) = \cos(\ln(3t))\frac{1}{t}$ **53.** $f'(x) = \frac{2x}{x^2-3}$ for $|x| \neq \sqrt{3}$ **55.** $f'(x) = \frac{-2x}{(\ln 10)(1-x^2)}$ **57.** $f'(x) = \frac{3x^2-3}{(\ln 10)(x^3-3x)}$
59. $f'(u) = \frac{4u^3}{(\ln 3)(3+u^4)}$ **63.** **(a)** $H'(t) = 3.2371t - 0.3458t\ln t$
(b) $H'(8) \approx 20.1445$, $H'(36) \approx 71.9273$. Rate of growth dH/dt is larger later in development **(c)** $\frac{1}{H(8)}H'(8) \approx 0.3547$, $\frac{1}{H(36)}H'(36) \approx 0.0522$. Relative rate of growth $\frac{1}{H}\frac{dH}{dt}$ is larger earlier in development **65.** $\frac{dy}{dx} = 2x^x(\ln x + 1)$
67. $\frac{dy}{dx} = (\ln x)^x[\ln(\ln x) + \frac{1}{\ln x}]$ **69.** $\frac{dy}{dx} = 2x^{-1+\ln x}\ln x$
71. $\frac{dy}{dx} = x^{-2+1/x}(1-\ln x)$ **73.** $\frac{dy}{dx} = \left[x^x(\ln x+1)\ln x + x^{x-1}\right]x^{x^x}$
75. $\frac{dy}{dx} = x^{\cos x}\left[\frac{\cos x}{x} - (\sin x)(\ln x)\right]$
77. $\frac{1}{y}\frac{dy}{dx} = 2 + \frac{27}{9x-2} - \frac{x}{2(x^2+1)} - \frac{9x^2}{4(3x^3-7)}$

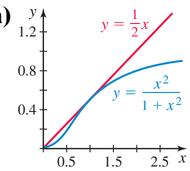
Section 4.11

1. $\sqrt{65} \approx 8.0625$, error $= 2.42\times 10^{-4}$
3. $\sqrt[3]{124} \approx 5 - \frac{1}{75} \approx 4.987$, error $\approx 3.57\times 10^{-5}$
5. $(0.99)^{25} \approx 0.75$, error ≈ 0.0278 **7.** $\sin\left(\frac{\pi}{2}+0.02\right) \approx 1$, error $\approx 2.00\times 10^{-4}$ **9.** $\ln(1.01) \approx 0.01$, error $\approx 4.97\times 10^{-5}$
11. $f(x) \approx 1 - x$ **13.** $f(x) \approx \frac{3}{2} - \frac{1}{2}x$ **15.** $f(x) \approx 1 - 2x$
17. $f(x) \approx x$ **19.** $f(x) \approx \frac{1}{\ln 10}(x-1)$ **21.** $f(x) \approx 1 + x$
23. $f(x) \approx 1 - x$ **25.** $f(x) \approx x$ **27.** $f(x) \approx 1 - nx$ **29.** $f(x) \approx 1$
31. 100.3 **33.** $B(1.1) \approx 5.005$ **35.** [1.8, 2.2] **37.** [10.8, 13.2]
39. [5.91, 8.87] **41.** $\pm 6\%$ **43.** $\pm 0.668\%$ **45.** $\pm 9\%$ **47.** $\pm 2.4\%$
49. $\Delta R = |k(2x-(a+b))\Delta x|$

Chapter 4 Review Problems

1. $f'(x) = -12x^3 - \frac{1}{x^{3/2}}$ **3.** $h'(t) = \frac{1}{3}\left(\frac{1+t}{1-t}\right)^{2/3}\frac{-2}{(1+t)^2} = -\frac{2}{3}\cdot\frac{1}{(1-t)^{2/3}(1+t)^{4/3}}$ **5.** $f'(x) = 2e^{2x}\sin\left(\frac{\pi}{2}x\right) + e^{2x}\frac{\pi}{2}\cos\left(\frac{\pi}{2}x\right)$
7. $f'(x) = -x + 2x\ln x$ **9.** $f'(x) = -xe^{-x^2/2}$, $f''(x) = e^{-x^2/2}(x^2-1)$ **11.** $h'(x) = \frac{1}{(x+1)^2}$, $h''(x) = -\frac{2}{(x+1)^3}$
13. $\frac{dy}{dx} = \frac{\cos x+y^2-2xy}{x^2-2xy}$ **15.** $\frac{dy}{dx} = 1 - x + y$ **17.** $\frac{dy}{dx} = -\frac{x}{y}$, $\frac{d^2y}{dx^2} = -\frac{16}{y^3}$
19. $\frac{dy}{dx} = \frac{1}{x\ln x}$, $\frac{d^2y}{dx^2} = -\frac{\ln x+1}{(x\ln x)^2}$ **21.** $5.70\frac{\text{ft}}{\text{sec}}$ **23.** **(a)** $\frac{dy}{dx} = f'(x)e^{f(x)}$

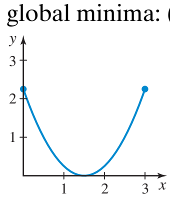
(b) $\frac{dy}{dx} = \frac{f'(x)}{f(x)}$ **(c)** $\frac{dy}{dx} = 2f(x)f'(x)$ **25. (a)**



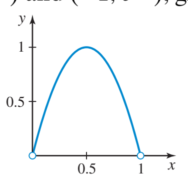
(b) $c = 1, y = \frac{1}{2}x$
27. $y - \frac{1}{2}e^{-(\pi/3)^2} = -e^{-(\pi/3)^2}\left(\frac{1}{2}\sqrt{3} + \frac{\pi}{3}\right)\left(x - \frac{\pi}{3}\right)$
29. $y = \sqrt{3} + \frac{2}{\sqrt{3}}(x-2), y = -\sqrt{3} - \frac{2}{\sqrt{3}}(x-2)$
31. $p(x) = 2x^2 + 4x + 8$ **33. (a)** $\approx 3.3\%$ increase in U
(b) $\approx 1.2\%$ increase in U (much smaller effect than for a spherical diatom)

Section 5.1

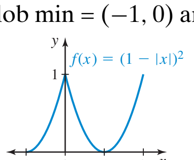
1. global minimum: $(0, 0)$; global maximum: $(1, 2)$ **3.** global minimum: $\left(2 - \frac{\pi}{2}, -1\right)$; global maximum: $(\pi, \sin(\pi - 2))$
5. global minimum: $(0, 0)$; global maximum: $(-1, 1)$ and $(1, 1)$
7. global minima: $(1, e^{-1})$ and $(-1, e^{-1})$; global maximum: $(0, 1)$
9.



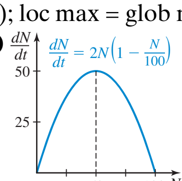
11.



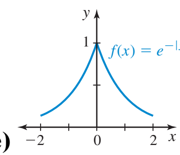
13. local maximum = global maximum = $(-1, 5)$, no local and global minima **15.** local minimum = global minimum = $(0, -2)$, local maximum = global maximum = $(-1, -1)$ and $(1, -1)$
17. local minimum = global minimum = $(1, 0)$, local maximum = global maximum = $(5, 5\ln 5)$ **19.** $f'(0) = 0$; $f(x)$ has a local minimum at $x = 0$ **21.** $f'(0) = 0$; $f(x) = -x^2$ has a local maximum at $x = 0$ **23.** $f'(0) = 0$, but $x = 0$ is not a local extremum: $f(x) > 0$ for $x < 0$ and $f(x) > 0$ for $x > 0$
25. $f'(-1) = 0$; $f(x)$ has a local minimum at $x = -1$
27. $f(0) = 0$ and $f(x) > 0$ for $x \neq 0$, If it exists $f'(0) = \lim_{h\to 0} \frac{f(h)-f(0)}{h}$. Now $f(0) = 0$ and if $h > 0$, $f(h) = h$. So $\lim_{h\to 0+} \frac{f(h)-f(0)}{h} = \lim_{h\to 0+} \frac{h}{h} = 1$. While if $h < 0$, $f(h) = -h$, so $\lim_{h\to 0-} \frac{f(h)-f(0)}{h} = \lim_{h\to 0-} \frac{(-h)}{h} = -1$. So the left limit and right limit do not agree as $h \to 0$. Since there are different limits, the derivative does not exist. **29.** $f(2) = f(-2) = 0$ and $f(x) < 0$ for $x \neq \pm 2$; $x = \pm 2$ are global maxima of the function. However for $x > 2$, $f(x) = -(x^2 - 4)$ so $\lim_{h\to 0+}\left[\frac{f(2+h)-f(2)}{h}\right] = \lim_{h\to 0+}\left[\frac{-((2+h)^2-4)-0}{h}\right] = \lim_{h\to 0+}\frac{(-4h-h^2)}{h} = -4$. Meanwhile if $-2 < x < 2$, $f(x) = x^2 - 4$ so $\lim_{h\to 0-}\left[\frac{f(2+h)-f(2)}{h}\right] = \lim_{h\to 0-}\left[\frac{(2+h)^2-4-0}{h}\right] = \lim_{h\to 0-}\left[\frac{4h+h^2}{h}\right] = 4$. Hence the limit $\lim_{h\to 0}\left[\frac{f(2+h)-f(2)}{h}\right]$ is not defined (we get different limits if $h \to 0+$ or if $h \to 0-$), so the derivative is not defined at $x = 2$. Similarly we can show that f is not differentiable at $x = -2$.
31. loc min = glob min = $(-1, 0)$ and $(1, 0)$; loc max = glob max = $(0, 1)$ and $(2, 1)$



33. (a)



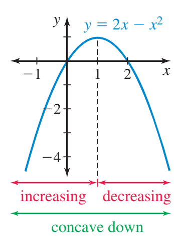
$\frac{dN}{dt}$ is maximal for $N = 50$ **(b)** $f'(N) = r - \frac{2r}{K}N$
35. (a) Slope: 1 **(b)** $c = 1/\sqrt{3}$; guaranteed by the Mean-Value Theorem. **37. (a)** Slope: 0 **(b)** $c = 0$; guaranteed by the Mean-Value Theorem. **39.** Slope of secant from $(-1, 1)$ to $(2, -2)$ is -1; guaranteed by the Mean-Value Theorem.
41. $f(0) = f(2) = 0$; guaranteed by Rolle's theorem.
43. (a) Slope: $a + b$
(b) Apply the Mean-Value Theorem. Midpoint $= a + \frac{b-a}{2} = \frac{a+b}{2}$.
45. Use the Mean-Value Theorem to find c such that $f'(c) = \frac{f(b)-f(a)}{b-a}$. **47. (a)** 1 m/s **(b)** $\frac{t}{5}$, $0 < t < 10$ **(c)** $t = 5$ s
49. $v(t)$ varies from 10 to 50; use the Mean-Value Theorem to argue that the slope of the secant must be between 10 and 50.
51. $0 \le N(5) \le 150$
53. (a) $f'(x) < 2$ implies, by the MVT: $f(x) - f(0) < 2(x - 0)$, i.e. $f(x) < 2x$. Similarly, $(x) > 1$ implies $f(x) > x$. **(b)** Since $x \le f(x) \le 2x$, setting $x = 1$ gives $1 \le f(1) \le 2$, so $f(1)$ cannot be 3 **(c)** $1 \le f(1) \le 2$ **55. (a)** $f(-2) = f(2) = e^{-2}$.
(b) $f'(x) = -e^{-x}$ for $x > 0$ and $f'(x) = e^x$ for $x < 0$; not differentiable at $x = 0$. **(c)** $-e^{-x}$ is never zero. **(d)** f is not differentiable everywhere. **(e)**



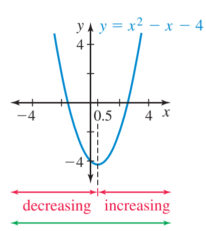
57. (a) Differentiate $C(t)$ using the product rule **(b)** Substitute $M'(t) = 1 - M(t)$ into the result from (a) and simplify
(c) $C(t)$ is a constant since its derivative is zero. Since $C(0) = M(0)e^0 - e^0 = -1$, $C(t) = -1$ everywhere **(d)** Substitute -1 for $C(t)$ in the definition of $C(t)$ and simplify.

Section 5.2

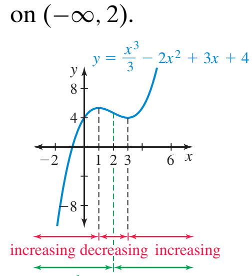
1. $y' = 2 - 2x$, $y'' = -2$; $y' > 0$ and y is increasing on $(-\infty, 1)$; $y' < 0$ and y is decreasing on $(1, \infty)$; $y'' < 0$ and y is concave down.



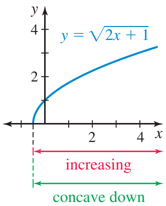
3. $y' = 2x - 1$, $y'' = 2$; $y' > 0$ and y is increasing on $(1/2, \infty)$; $y' < 0$ and y is decreasing on $(-\infty, 1/2)$; $y'' > 0$ and y is concave up.



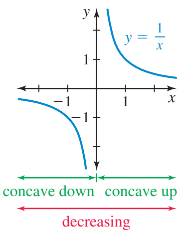
5. $y' = x^2 - 4x + 3$, $y'' = 2x - 4$; $y' > 0$ and y is increasing on $(-\infty, 1) \cup (3, \infty)$; $y' < 0$ and y is decreasing on $(1, 3)$; $y'' > 0$ and y is concave up on $(2, \infty)$; $y'' < 0$ and y is concave down on $(-\infty, 2)$.



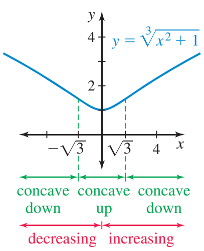
7. $y' = \frac{1}{\sqrt{2x+1}}, x > -1/2$; $y'' = -\frac{1}{(2x+1)^{3/2}}, x > -1/2$; $y' > 0$ and y is increasing; $y'' < 0$ and y is concave down.



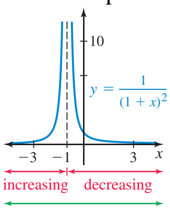
9. $y' = -\frac{1}{x^2}, x \neq 0$; $y'' = \frac{2}{x^3}, x \neq 0$; $y' < 0$ and y is decreasing for $x \neq 0$; $y'' < 0$ and y is concave down for $x < 0$; $y'' > 0$ and y is concave up for $x > 0$.



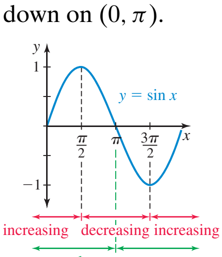
11. $y' = \frac{1}{3}(x^2+1)^{-2/3}2x$, $y'' = \frac{6-2x^2}{9(x^2+1)^{5/3}}$; $y' > 0$ and y is increasing on $(0, \infty)$; $y' < 0$ and y is decreasing on $(-\infty, 0)$; $y'' > 0$ and y is concave up on $(-\sqrt{3}, \sqrt{3})$; $y'' < 0$ and y is concave down on $(-\infty, -\sqrt{3}) \cup (\sqrt{3}, \infty)$.



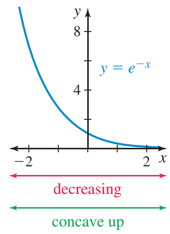
13. $y' = -\frac{2}{(1+x)^3}$, $y'' = \frac{6}{(1+x)^4}$; $y' > 0$ and y is increasing on $(-\infty, -1)$; $y' < 0$ and y is decreasing on $(-1, \infty)$; $y'' > 0$ and y is concave up for $x \neq -1$.



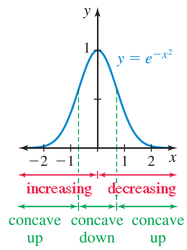
15. $y' = \cos x$, $y'' = -\sin x$; $y' > 0$ and y is increasing on $(0, \pi/2) \cup (3\pi/2, 2\pi)$; $y' < 0$ and y is decreasing on $(\pi/2, 3\pi/2)$; $y'' > 0$ and y is concave up on $(\pi, 2\pi)$; $y'' < 0$ and y is concave down on $(0, \pi)$.



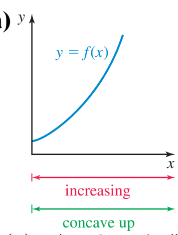
17. $y' = -e^{-x}$, $y'' = e^{-x}$; $y' < 0$ and y is decreasing for $x \in \mathbf{R}$; $y'' > 0$ and y is concave up for $x \in \mathbf{R}$.



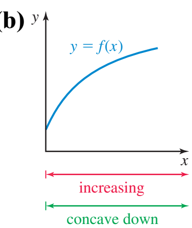
19. $y' = -2xe^{-x^2}$, $y'' = e^{-x^2}(4x^2-2)$; $y' > 0$ and y is increasing for $x < 0$; $y' < 0$ and y is decreasing for $x > 0$; $y'' > 0$ and y is concave up for $x < -1/\sqrt{2}$ and $x > 1/\sqrt{2}$; $y'' < 0$ and y is concave down for $-1/\sqrt{2} < x < 1/\sqrt{2}$.



21. **(a)**



(b)



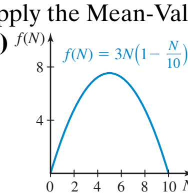
(c) In (a), $y' > 0$ and $y'' > 0$; in (b), $y' > 0$ and $y'' < 0$.

23. Assume there are two such solutions, and use Rolle's theorem to find a point between them where $f'(x) = 0$.

25. $f(-1) = -3 < 0$, $f(0) = 1 > 0$; $f'(x) > 0$ for $-1 < x < 0$.

27. Apply the Mean-Value Theorem to any two points in $[a, b]$.

29. **(a)**



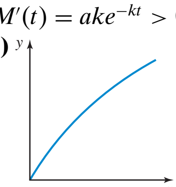
(b) $f'(N) = \frac{3}{5}(5 - N)$; increasing on $[0, 5)$, decreasing on $(5, 10]$. **(c)** $f'(N) = \frac{3}{K}(K - 2N)$

31. $r'(C) = \frac{(1.35)\cdot(0.022)}{(C+0.022)^2}$; yes.

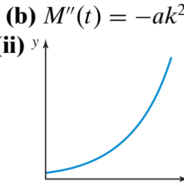
33. **(a)** $f'(P) = \frac{30}{(P+30)^2}$; $f''(P) = -\frac{60}{(P+30)^3}$. **(b)** $f'(P) > 0$ for all P, $f''(P) > 0$ for $2P^3 < 30^3$, so $f(P)$ increases at an increasing rate.

35. **(a)** $M'(t) = ake^{-kt} > 0$ **(b)** $M''(t) = -ak^2e^{-kt} < 0$

37. **(a)(i)**



(ii)



(b) Optimal strategy is polycarpy if $S'' < 0$ and monocarpy otherwise.

39. **(a)** $p'(R) = \frac{2k^2R}{(k^2+R^2)^2} > 0$ **(b)** Analyze numerator of $p''(R) = \frac{2-6R^2}{(R^2+1)^3}$ **(c)** p is concave up and increasing

41. $y' = -\frac{x}{y}$; decreasing

43. $y' = \frac{y^2}{x(x+y)}$; note that $x, y > 0$; increasing
45. **(a)** $F(r) = e^{-r}(6e^{-r} - 2)$; $6e^{-r} - 2 = 0$ when $r = \ln 3$ **(b)** Because attracting force is zero **(c)** $P'(f) = \frac{e^{-f}}{(1+e^{-f})^2}$; $P''(f) = \frac{e^{-2f} - e^{-f}}{(1+e^{-f})^3}$; $e^{-2f} < e^{-f}$ since $f > 0$ **(d)** P is concave down so increases at a decreasing rate
47. **(a)** kW^{a-1}; $G'(W) = (a-1)kW^{a-2}$ **(b)** $a < 1$ for $G'(W) < 0$.
49. **(a)** $\frac{dY}{dX} = abX^{a-1}$; $a > 0$ **(b)** $\frac{d}{dX}\frac{Y}{X} = (a-1)bX^{a-2}$; $0 < a < 1$; concave down **(c)** Skull size becomes smaller compared to body size

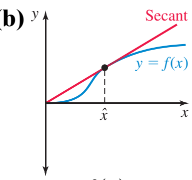
Section 5.3

1. local max: $(-2, 9)$ and $(3, 4)$; local min: $(1, 0)$; y is decreasing on $(-2, 1)$ and increasing on $(1, 3)$. **3.** local max: $(2, \sqrt{3})$; local min: $(1/2, 0)$; y is increasing on $(1/2, 2)$. **5.** local min: $(0, 0)$; local max: $(1, e^{-1})$; y is increasing on $[0, 1]$ **7.** no extrema; y is increasing for all $x \in \mathbf{R}$. **9.** local max: $(0, 1)$; local min: $(-1, e^{-1})$ and $(1, e^{-1})$; y is increasing on $(-1, 0)$ and decreasing on $(0, 1)$.
11. local max: $(0, 1)$; y is increasing on $(-\infty, 0)$, and decreasing on $(0, \infty)$. **13.** local max: $(-1, 13/6)$; local min: $(0, 2)$; y is increasing on $(-\infty, -1) \cup (0, \infty)$; y is decreasing on $(-1, 0)$.
15. local max: $(2/3, 4/27)$; local min: $(0, 0)$; increasing on $(0, 2/3)$; decreasing on $(-\infty, 0) \cup (2/3, \infty)$ **17.** local min: $(0, 1)$; no other extrema; increasing for $x > 0$; decreasing for $x < 0$
19. If $f'(b) > 0$ then f is increasing just to the left of b; if $f'(b) < 0$ then f is decreasing just to the left of b.
21. $g'(x) = f'(x)e^{f(x)}$ and $e^{f(x)} > 0$; $g'(x)$ so goes from increasing to decreasing at $x = c$, just as $f(x)$ does **23.** $f''(x) = 20(x-3)^2$; no inflection points because although $f''(3) = 0$, $f''(x)$ does not change sign at $x = 3$. **25.** $f''(x) = (x-2)e^{-x}$, $x = 2$
27. $f''(x) = \frac{2(1-3x^2)}{(x^2+1)^3}$; $x = \pm\frac{\sqrt{3}}{3}$ **29.** $f''(x) = \frac{1}{x}$; no inflection points **31.** **(b)** $N'(t) = \frac{600e^{-2t}}{(1+3e^{-2t})^2} > 0$ **(c)** Use $\lim_{t\to\infty} e^{-2t} = 0$
(d) $N'' = \frac{600(1-3e^{-2t})(-2e^{-2t})}{(1+3e^{-2t})^3}$ goes from positive to negative when $t = \frac{1}{2}\ln 3$, i.e., N has an inflection point there. When $t = \frac{1}{2}\ln 3$, $N = 50$. **33.** **(b)** $M'(t) = ke^{-kt}(3e^{-2kt} - 1)$; $t = \frac{\ln 3}{2k}$
(c) $M''(t) = k^2e^{-kt}(1 - 9e^{-2kt})$; $t = \frac{\ln 3}{k}$; down to up **(d)** No

Section 5.4

1. 20 in. **3.** 4 **5.** 4 cm^2 **7.** Need to maximize $A = \frac{1}{2}ab$ when $a^2 + b^2 = 10$. Maximizing A is equivalent to maximizing $A^2 = \frac{1}{4}a^2b^2 = \frac{1}{4}a^2(10 - a^2)$. A^2 is maximized when $a^2 = 5$ (i.e., when $b^2 = 10 - a^2 = 5 = a^2$). **9.** **(b)** $(6/5, 2/5)$ **(c)** local minimum at $x = 6/5$ as in (b). **11.** $\sqrt{2}$ **13.** $g'(x) = 2f(x)f'(x)$ has the same sign change as $f'(x)$.
15. $g'(x) = -f'(x)e^{-f(x)}$ has the opposite sign as $f'(x)$ everywhere.
17. **(a)** \$200 **(b)** \$250
19. **(a)** $\frac{710}{r} + 2\pi r^2$ **(b)** $S(r) \to \infty$ **(c)** $S(r) \to \infty$ **(d)** $r = \left(\frac{355}{2\pi}\right)^{1/3} \approx 3.837$ **21.** **(a)** $r = \sqrt{2}, \theta = 2$ **(b)** $r = \sqrt{10}, \theta = 2$
23. **(a)** $V(r) \to \infty$, $V(r) \to 0$ **(b)** $r = \left(\frac{2}{A}\right)^{1/3}$ **(c)** No, since then $V'(r) < 0$ for all r. **25.** **(a)** $w'(0) > 0$ and $w'(1) > 0$ and w' linear, so positive on $[0, 1]$ so w is increasing on $[0, 1]$ **(b)** $w'(0) < 0$ and $w'(1) < 0$, so negative on $[0, 1]$, so w is decreasing on $[0, 1]$ **(c)** If S confers fitness more than s, then eventually essentially all members will have an S gene, and similarly for the other direction.

27. **(a)** $w'(x) = \frac{R}{x}\left(f'(x) - \frac{f(x)}{x}\right)$ **(b)**



Secant from $(0, 0)$ to $(\hat{x}, f(\hat{x}))$ is tangent to curve $y = f(x)$ at $x = \hat{x}$. **(c)** Use the product rule **(d)** When $x = \hat{x}$ then $f'(x) = \frac{f(x)}{x}$ and $\frac{d}{dx}\frac{f(x)}{x} = 0$.
29. **(a)** $r = \left(\frac{0.142 f^2}{2\pi c}\right)^{1/5}$ **(b)** Recall that $f = f_1 + f_2$ and use (a)

Section 5.5

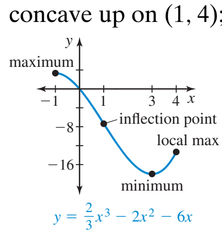
1. 10 **3.** -7 **5.** $-\frac{1}{6}$ **7.** 0 **9.** $-\infty$ **11.** ∞ **13.** 0 **15.** $\frac{\ln 5}{\ln 7}$ **17.** $-\frac{\ln 3}{\ln 2}$ **19.** 1 **21.** 0 **23.** 1 **25.** 0 **27.** 0 **29.** 0 **31.** 0 **33.** ∞ **35.** 0 **37.** $-\infty$ **39.** 0 **41.** 1 **43.** 1 **45.** e^3 **47.** e^{-2} **49.** e^{-1} **51.** 0 **53.** ∞ **55.** 0 **57.** 0 **59.** 1 **61.** $\frac{\ln a}{\ln b}$

63. $\lim_{x\to\infty}\left(1 + \frac{1}{x^p}\right)^x = \begin{cases} \infty & \text{if } 0 < p < 1 \\ e & \text{if } p = 1 \\ 1 & \text{if } p > 1 \end{cases}$

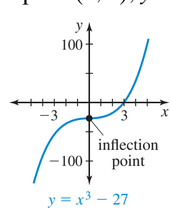
65. **(a)** Apply L'Hôpital's rule to $\frac{t^{k-1}}{e^{t^k}}$ **(b)** For $k \geq 1$
67. **(a)** Compute the one-sided limits **(b)** Analyze $H'(p) = \ln\frac{1-p}{p}$

Section 5.6

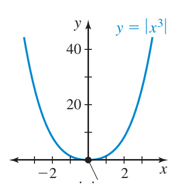
1. local min: $(3, -18)$; absolute min: $(3, -18)$; local max: $(4, -40/3)$, $(-1, 10/3)$; absolute max: $(-1, 10/3)$; inflection point at $x = 1$; y is increasing on $(3, 4)$; y is decreasing on $(-1, 3)$; y is concave up on $(1, 4)$; y is concave down on $(-1, 1)$;



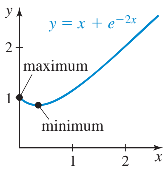
3. local min: $(-5, -152)$; absolute min: $(-5, -152)$; local max: $(5, 98)$; absolute max: $(5, 98)$; inflection point at $x = 0$; y is increasing on $(-5, 0) \cup (0, 5)$; y is never decreasing; y is concave up on $(0, 5)$; y is concave down on $(-5, 0)$;



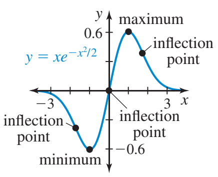
5. local min: $(0, 0)$; absolute min: $(0, 0)$; No local or absolute max; No inflection points; y is increasing on $(0, \infty)$; y is decreasing on $(-\infty, 0)$; y is concave up on $(-\infty, \infty)$;



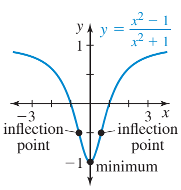
7. local min: $\left(\frac{1}{2}\ln 2, \frac{1+\ln 2}{2}\right)$; absolute min: $\left(\frac{1}{2}\ln 2, \frac{1+\ln 2}{2}\right)$; local max: $(0, 1)$; No absolute max; No inflection points; y is increasing on $\left(\frac{1}{2}\ln 2, \infty\right)$; y is decreasing on $\left(0, \frac{1}{2}\ln 2\right)$; y is concave up everywhere;



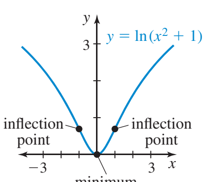
9. local min: $(-1, -e^{-1/2})$; absolute min: $(-1, -e^{-1/2})$; local max: $(1, e^{-1/2})$; absolute max: $(1, e^{-1/2})$; inflection points at $x = 0$, $x = \pm\sqrt{3}$; y is increasing on $(-1, 1)$; y is decreasing on $(-\infty, -1) \cup (1, \infty)$; y is concave up on $(-\sqrt{3}, 0) \cup (\sqrt{3}, \infty)$; y is concave down on $(-\infty, -\sqrt{3}) \cup (0, \sqrt{3})$;



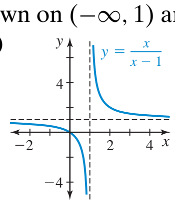
11. local min: $(0, -1)$; absolute min: $(0, -1)$; No local max; No absolute max; inflection points at $x = \pm\frac{1}{\sqrt{3}}$; y is increasing on $(0, \infty)$; y is decreasing on $(-\infty, 0)$; y is concave up on $(-1/\sqrt{3}, 1/\sqrt{3})$; y is concave down on $(-\infty, -1/\sqrt{3}) \cup (1/\sqrt{3}, \infty)$;



13. local min: $(0, 0)$; absolute min: $(0, 0)$; No local max; No absolute max; inflection points at $x = \pm 1$; y is increasing on $(0, \infty)$; y is decreasing on $(-\infty, 0)$; y is concave up on $(-1, 1)$; y is concave down on $(-\infty, -1) \cup (1, \infty)$;

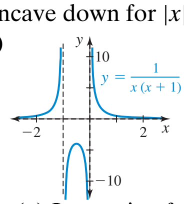


15. **(c)** decreasing for all $x \neq 1$; no extrema **(d)** $f(x)$ is concave down on $(-\infty, 1)$ and concave up on $(1, \infty)$; no inflection points **(e)**

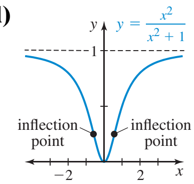


17. **(b)** Increasing for $x < -\frac{1}{2}$ and $x \neq -1$, decreasing for $x > -\frac{1}{2}$ and $x \neq 0$. Local maximum at $x = -\frac{1}{2}$ **(c)** Concave up for $|x| > 1$, concave down for $|x| < 1$ **(d)** Both limits are zero.

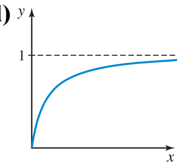
(e)



19. **(a)** Increasing for $x > 0$, decreasing for $x < 0$ **(b)** Concave down for $|x| > 1/\sqrt{3}$, concave up for $|x| < 1/\sqrt{3}$. Inflection points at $\pm 1/\sqrt{3}$. **(c)** Both limits are 1; horizontal asymptote is $y = 1$ **(d)**

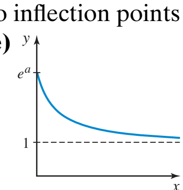


21. **(a)** $f'(x) = \frac{a}{(a+x)^2}$; $f(x)$ is increasing for $x > 0$ **(b)** $f''(x) = -\frac{2a}{(a+x)^3} < 0$; f is concave down for all x; there are no inflection points **(c)** $\lim_{x\to\infty} \frac{x}{a+x} = 1$; there is a horizontal asymptote at $y = 1$ **(d)**



(e) The general shape is the same, but approaches the limit more slowly as a increases.

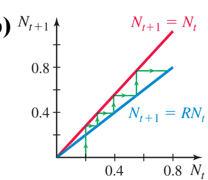
23. **(a)** Decreasing everywhere; $M'(t) = -ae^{-t}\exp(ae^{-t})$ **(b)** $t = 0$ is a local maximum. **(c)** $M''(t) = ae^{-t}\exp(ae^{-t})(1 + ae^{-t})$; concave up everywhere, no inflection points **(d)** 1, horizontal asymptote is $y = 1$ **(e)**



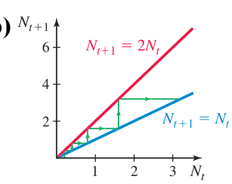
(f) The general shape is the same, but the curve approaches the asymptote more quickly as a increases.

Section 5.7

1. **(a)** $N = 0$; unstable **(b)**



3. **(a)** $N = 0$; unstable **(b)**

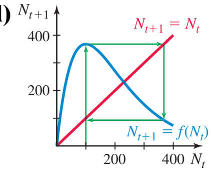


5. $-\frac{1}{3}$, stable; 2, unstable

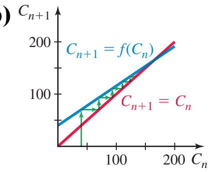
7. 1, stable; 4, unstable **9.** $-\frac{1}{2}$, $\frac{1}{2}$ both unstable **11.** 0, unstable; 1, stable **13.** **(a)** $x = 0$ is locally stable, $x = 1$ is unstable, $x = 4$ is locally stable **(b)** **(i)** 0 **(ii)** 4 **15.** **(a)** 10, 18.18, 30.77, 47.06, 64, 78.05, 87.67, 93.43, 96.6, 98.27, 99.13 **(b)** 150, 120, 109.09, 104.35, 102.13, 101.05, 100.52, 100.26, 100.13, 100.07, 100.03 **(c)** 0, unstable; 100, stable **(d)** Both sequences approach 100

17. **(a)** 0, unstable; 50, stable **(b)** 10, 14, 19.04, 24.93, 31.18, 37.05, 41.85, 45.26, 47.41, 48.64, 49.3; approach equilibrium **19.** **(a)** 0, unstable; 250, unstable **(b)** 10, 34, 107.44, 260.61, 232.97, 272.65, 210.89, 293.37, 166.15, 305.47, 136.04; does not approach any equilibrium. **21.** **(a)** $\frac{1}{a}(-1 + \sqrt{R_0})$ **(b)** $f' = \frac{2}{\sqrt{R_0}} - 1$ at the equilibrium point so $0 < f' < 1$ if $R_0 > 1$. **23.** **(a)** Otherwise the population will not grow when it is not resource limited. **(b)** 0, $\frac{1}{a}\ln R_0$ **(d)** $f' = 1 - \ln R_0$ at the nontrivial equilibrium point so $|f'| < 1$ if $0 < \ln R_0 < 2$. **25.** **(a)** 100, 367.879, 92.902, 366.908, 93.561, 367.083, 93.442, 367.053, 93.463, 367.058, 93.459, 367.057, 93.46, 367.057, 93.46, 367.057, 93.46, 367.057, 93.46, 367.057, 93.46

(b) 0 is the other equilibrium point **(c)** Population will oscillate between two values **(d)**



27. **(a)** $\frac{40}{1-0.7575^T}$, stable **(b)**

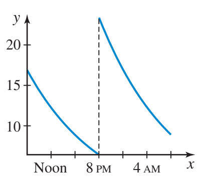


Section 5.8

1. $\sqrt{7} \approx 2.645751$ **3.** 2.153292 **5.** 0.652919
7. 1, 0.465571
9. 0.876726
11. **(a)** 78.8385 **(b)** 98.4606
13. **(a)** $N_0 \approx 316.9$, $N_1 \approx -306.9$, $r \approx 0.965$; $N(t) = 316.9 - 306.9e^{-0.965t}$ **(b)** 316.9, or (i.e., approx 317)
15. $a \approx 3.306$, $b \approx 2.19$, $c \approx 1.096$

Section 5.9

1. **(a)** 22162 **(b)** Must be that t_b is wrong **(c)** $b \approx 1.237$, $t_b \approx 0.560$
3. **(a)** $t_b \approx 1.26$ **(b)** ≈ 3.26 hours **(c)** Decrease **(d)** Stays the same.
5. **(a)** $t_m = \ln 2$ **(b)** Since $b = m$, they must both be 4. So the antibiotic must actually reduce the cell division time (i.e., increase the birth rate)
7. **(a)** 5657 **(b)** between 5091 and 6223
9. **(a)** $\frac{dM}{dt} = -0.08M(t)$, $M(0) = 33.8$ **(b)** 4.96 ng/ml **(c)** ≈ 8.66 hours **(d)** Since absorption of half the amount happens rapidly **(e)** 6.47 ng/ml **(f)** $M(t) = \begin{cases} 16.9e^{-0.08t} & 0 \le t < 12 \\ 23.37e^{-0.08(t-12)} & 12 \le t \le 24 \end{cases}$
(g)



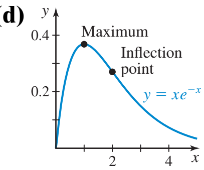
(h) 6.47 ng/ml
11. **(c)** 1.128 g/L **(d)** 8.06 hours **(e)** 3.76 hours
13. **(b)** $c_0 = 40$, $k_1 \approx 0.278$

Section 5.10

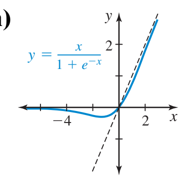
1. $F(x) = \frac{1}{3}x^3 - 2x^2 + C$ **3.** $F(x) = \frac{1}{3}x^3 + \frac{3}{2}x^2 - 4x + C$
5. $F(x) = \frac{1}{3}x^3 - x + C$ **7.** $F(x) = x^4 - x^2 + 3x + C$
9. $F(x) = x - \ln|x| - \frac{1}{x} + C$ **11.** $F(x) = x - \frac{1}{x} + C$
13. $F(x) = \ln|1+x| + C$ **15.** $F(x) = x^5 - \frac{5}{3x^3} + C$
17. $F(x) = \frac{1}{2}\ln|1+2x| + C$ **19.** $F(x) = -\frac{1}{3}e^{-3x} + C$
21. $F(x) = e^{2x} + C$ **23.** $F(x) = -\frac{1}{2}e^{-2x} + C$
25. $F(x) = -\frac{1}{2}\cos(2x) + C$
27. $F(x) = -3\cos(x/3) + 3\sin(x/3) + C$
29. $F(x) = -\frac{4}{\pi}\cos(\pi x/2) - \frac{6}{\pi}\sin(\pi x/2) + C$
31. $F(x) = \frac{1}{2}\tan(2x) + C$ **33.** $F(x) = -3\ln\left|\cos\frac{x}{3}\right| + C$
35. $F(x) = \frac{3}{2}x + \frac{1}{4}\sin(2x) + C$ **37.** $-\frac{1}{6}x^{-6} + \frac{1}{2}x^6 - \frac{1}{2}\cos(2x) + C$
39. $\frac{1}{3}\tan(3x-1) + \frac{1}{2}x^2 - 3\ln|x| + C$ **41.** $\frac{1}{a(a+1)}e^{(a+1)x} + C$
43. $\frac{1}{a}\ln|ax+3| + C$ **45.** $\frac{1}{a}e^{ax} + C$ **47.** $y = 2\ln x - \frac{1}{2}x^2 + C$
49. $y = \frac{1}{2}x^2 + \frac{1}{3}x^3 + C$ **51.** $y = \frac{1}{2}t^2 - \frac{1}{3}t^3 + C$
53. $y = -2e^{-t/2} + C$ **55.** $y = -\frac{1}{\pi}\cos(\pi s) + C$
57. $y = -\ln(x-1) + C$ **59.** $y = x^3 + 1, x \ge 0$ **61.** $y = \frac{2}{3}x^{3/2} + \frac{4}{3}$
63. $N(t) = \ln t + 10, t \ge 1$ **65.** $W(t) = e^t, t \ge 0$
67. $W(t) = e^{t+1} + \frac{2}{3} - e$
69. $T(t) = 3 + \frac{1}{\pi} - \frac{1}{\pi}\cos(\pi t)$ **71.** $y = \frac{e^x - e^{-x}}{2}$
73. $y = \sqrt{2x-1}$
75. $y = -\ln(1-x)$
77. $x = \frac{1}{2}y^2 + \ln|y| - \frac{1}{2}$ **79.** $L(t) = 25 - 10e^{-0.1t}$, $L(0) = 15$
81. $N(t) = N_0e^{rt}$
83. **(a)** ≈ 1.630 mm **(b)** ≈ 15.380 mm **(c)** ≈ 38.4 mm

Chapter 5 Review Problems

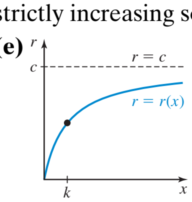
1. **(b)** Absolute maximum at $(1, e^{-1})$ **(c)** inflection point at $(2, 2e^{-2})$ **(d)**



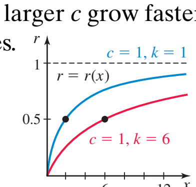
3. **(a)** $f'(x) = \frac{4}{(e^x+e^{-x})^2} > 0$, hence, $f(x)$ is strictly increasing. $\lim_{x\to-\infty} = \tanh x = \frac{-\infty}{\infty}$. Can't use l'Hôpital's rule, so rearrange: $\lim_{x\to-\infty}\tanh x = \lim_{x\to-\infty}\left[\frac{e^{2x}-1}{e^{2x}+1}\right] = \frac{0-1}{0+1} = -1$. Similarly $\lim_{x\to\infty}\tanh x = \lim_{x\to\infty}\left[\frac{1-e^{-2x}}{1+e^{-2x}}\right] = 1$.
(b) Domain is $x \in (-1, 1)$.
5. **(a)** Use L'Hôpital's rule **(e)** $c \approx -1.27846$.
(g) $x - f(x) = \frac{xe^{-x}}{1+e^{-x}} = \frac{x}{e^x+1}$ **(h)**

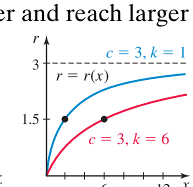


7. **(a)** Use l'Hôpital's rule **(b)** $r'(x) = \frac{ck}{(k+x)^2}$; $r'(x) > 0 \Rightarrow r(x)$ is strictly increasing so $r(x) < \lim_{x\to\infty}[r(x)]$ **(d)** $r''(x) = -\frac{2cx}{(k+x)^3}$
(e)

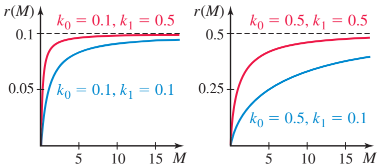


(f) Curves for larger k grow more slowly; curves for larger c grow faster and reach larger horizontal asymptotes.





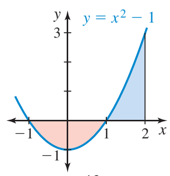
9. **(a)** $r'(q) = p(-b + \frac{c}{2}) + (b - c)$ **(i)** $r'(q) > 0$ if $p < \frac{b-c}{b-c/2}$, so $r(q)$ is maximized by $q = 1$ **(ii)** $r'(q) < 0$ if $p > \frac{b-c}{b-c/2}$ so $r(q)$ is maximized by $q = 0$. **(b)** $r'(p) = \left(b - \frac{c}{2}\right)(2 - 2p)$ $r'(p) > 0$ if $p \in [0, 1]$ so $r(p)$ is maximized by $p = 1$. **(c)** $r(p)$ is decreasing if $b < \frac{c}{2}$ and increasing if $b > \frac{c}{2}$, so cooperation works in the second case. If $b > \frac{c}{2}$ then net benefit of cooperation with another cooperator is positive. That is, two cooperators cooperating together have higher net benefit than two defectors defecting together. Whereas if $b < \frac{c}{2}$ then two defectors defecting together have higher net benefit than two cooperators who cooperate together.
11. **(a)** $r'(M) = \frac{k_0^2k_1}{(k_0+k_1M)^2} > 0$ **(c)** Show that $r'(M) \approx k_1$ **(d)** M small implies $r(M) \approx k_1M$, first order. M large, $r(M) \approx k_0$, zeroth order **(e)** $r(M) = k_0\left(\frac{M}{M+\frac{k_0}{k_1}}\right) < k_0$ **(f)** $r''(M) = -\frac{2k_0^2k_1}{(k_0+k_1M)^3} < 0$; concave down **(g)** Curves for larger k_1 increase more rapidly initially but have same horizontal asymptote. Curves for larger

k_0 increase at same rate initially but reach a larger horizontal asymptote. 

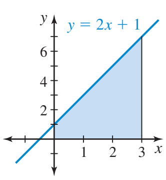
(h) Rate of elimination is negative of change in amount present **(i)** $t = 2 - M + \ln\frac{2}{M}$ **(j)** $t \approx 4.103$ **(k)** $M(2) \approx 0.853$

Section 6.1

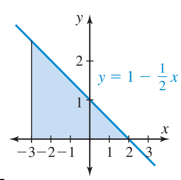
1. $\frac{7}{32}(\approx 0.2188)$ **3.** $\frac{25}{144}(\approx 0.1736)$ **5.** $-\frac{6}{25}(=-0.24)$ **7.** **(a)** $\frac{a}{n}$ **(d)** $\frac{a^3}{3}$ **9.** $\frac{32}{25}(=1.28)$ **11.** 1.5 **13.** $\frac{118}{25}=4.72$ **15.** **(a)** $\frac{b^2-a^2}{2}$ **(b)** $\frac{4}{3}(\approx 1.33)$, $\frac{17}{12}(\approx 1.42)$, true value $\frac{3}{2}(=1.5)$ **17.** $\int_1^2 \frac{x^2}{2}\,dx$ **19.** $\int_{-1}^1 \left(2x-\frac{1}{2}\right)\,dx$ **21.** $\int_0^1 2^x\,dx$ **23.** $\int_0^2 \frac{1}{e^{-x}+1}\,dx$
25. $\lim_{n\to\infty} \frac{2}{n}((x_0+1)^{1/3}+\cdots+(x_{n-1}+1)^{1/3})$
27. $\lim_{n\to\infty} \frac{e-1}{n}(\ln x_0+\ln x_1+\cdots+\ln x_{n-1})$
29. $\lim_{n\to\infty} \frac{5}{n}(x_0^3+x_1^3+\cdots+x_{n-1}^3)$ **31.**



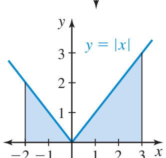
33.



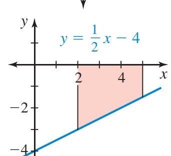
35.



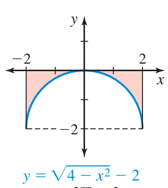
37. $\frac{13}{2}$



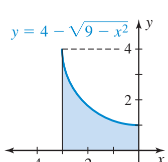
39. $-\frac{27}{4}$



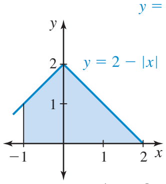
41. $2\pi - 8$



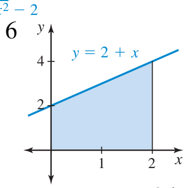
43. $12 - \frac{9\pi}{4}$



45. $\frac{7}{2}$



47. 6



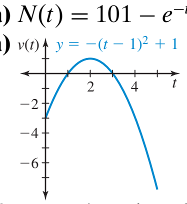
49. Proof **51.** 0 by property 1 **53.** 0 by symmetry **55.** 0 by symmetry
57. **(a)** 4 **(b)** $\frac{1}{2}$ **(c)** 0 **(d)** 4 **(e)** $\frac{175}{2}$
59. Use $x \ge x^2$ for $0 \le x \le 1$
61. Use $0 \le \sqrt{x} \le 3$ for $0 \le x \le 9$
63. Use $\frac{1}{2} \le \sin x \le 1$ in the given range
65. $\frac{\pi}{2}$
67. 5 **69.** **(a)** 8 mm **(b)** 10 mm

Section 6.2

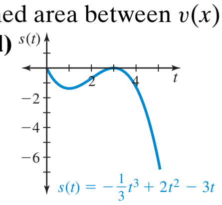
1. $\frac{dy}{dx}=2x^2$ **3.** $\frac{dy}{dx}=4x-3$ **5.** $\frac{dy}{dx}=\sqrt{1+2x}$ **7.** $\frac{dy}{dx}=\sqrt{\sin 2x}$
9. $\frac{dy}{dx}=xe^{4x}$ **11.** $\frac{dy}{dx}=\frac{1}{x+1}$ **13.** $\frac{dy}{dx}=\sin(x^2+1)$
15. $\frac{dy}{dx}=3(1+9x^2)$ **17.** $\frac{dy}{dx}=[2(1-4x)^2+1](-4)$
19. $\frac{dy}{dx}=2x\sqrt{x^2+1}$ **21.** $\frac{dy}{dx}=3(1+e^{3x})$
23. $\frac{dy}{dx}=(6x+1)[1+e^{3x^2+x}]$ **25.** $\frac{dy}{dx}=-(1+x)$
27. $\frac{dy}{dx}=-2[1+\cos 2x]$ **29.** $\frac{dy}{dx}=-\frac{1}{x^2}$ **31.** $\frac{dy}{dx}=-2x\sec(x^2)$
33. $\frac{dy}{dx}=2[1+(2x)^2]-(1+x^2)$
35. $\frac{dy}{dx}=3x^2\ln(x^3-3)-2x\ln(x^2-3)$
37. $\frac{dy}{dx}=(3x^2+1)((x+x^3)^2-1)+2x((2-x^2)^2-1)$
39. $x+x^3+C$ **41.** $\frac{1}{9}x^3-\frac{1}{4}x^2+C$ **43.** $\frac{1}{6}x^3+\frac{3}{2}x^2-\frac{1}{3}x+C$
45. $x^{3/2}\left(\frac{4}{5}x+\frac{2}{3}\right)+C$ **47.** $\frac{2}{7}x^{7/2}+C$ **49.** $\frac{2}{9}x^{9/2}+\frac{7}{9}x^{9/7}+C$
51. $\frac{2}{3}x^{3/2}+2\sqrt{x}+C$ **53.** $\frac{1}{3}x^3-x+C$ **55.** $\frac{1}{3}x^3+\frac{1}{2}x^2+C$
57. $\frac{1}{2}e^{2x}+C$ **59.** $-3e^{-x}+C$ **61.** $\frac{1}{2}e^{2x}+e^x+C$
63. $-\frac{1}{2}\cos(2x)+C$ **65.** $\frac{1}{3}\sin(3x)+C$ **67.** $-\frac{1}{2}\cos(2x-1)+C$
69. $\sec x+C$ **71.** $-\frac{1}{4}\cos(2x)+C$ **73.** $\sin x+\frac{x}{2}+\frac{1}{4}\sin(2x)+C$
75. $4\tan^{-1}x+C$ **77.** $\sin^{-1}x+C$ **79.** $\ln|x+2|+C$
81. $\frac{2}{3}x-\frac{1}{3}\ln|x|+C$ **83.** $\frac{1}{2}\ln|2x+1|+C$ **85.** $\frac{1}{2}\tan^{-1}\frac{x}{2}+C$
87. $2(x-\tan^{-1}x)+C$ **89.** $\frac{3^x}{\ln 3}+C$ **91.** $-\frac{4^{-x}}{\ln 4}+C$
93. $\frac{1}{3}x^3+\frac{2^x}{\ln 2}+C$ **95.** $\frac{2}{3}x^{3/2}+2e^{x/2}+C$ **97.** 24 **99.** $-\frac{1}{2}$ **101.** 3
103. $\frac{26}{3}$ **105.** $\frac{1}{2}$ **107.** $\frac{\sqrt{2}}{2}-1$ **109.** $\frac{\pi}{4}$ **111.** $\frac{\pi}{6}$ **113.** 0
115. $\frac{1}{3}(1-e^{-3})$ **117.** $\frac{1}{3}(e-e^{-1})$ **119.** 1 **121.** $\frac{1}{2}\ln 3$ **123.** $4x$
125. $f(x)=\cos x$; $C=0$

Section 6.3

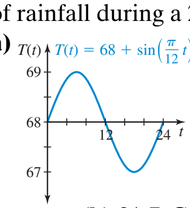
1. **(a)** $N(t)=101-e^{-t}$ **(b)** $1-e^{-5}$ **(c)** $\int_0^t \frac{dN(x)}{dx}\,dx$
3. **(a)**



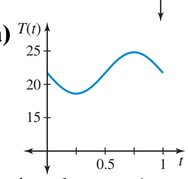
(b) Particle moves to the left for $0 \le t < 1$ and $3 < t \le 5$, and to the right for $1 < t < 3$.
(c) $s(t)=2t^2-\frac{1}{3}t^3-3t$; signed area between $v(x)$ and the horizontal axis from 0 to t **(d)**



The rightmost position: $s(0)=s(3)=0$; the leftmost position: $s(5)=-\frac{20}{3}$. **5.** Amount of growth or shrinkage from month 2 to month 7 **7.** Amount of biomass gained or lost from $t=1$ to $t=6$ **9.** **(a)** Total amount of rainfall during a 24-hour period **(b)** $1.5\ln 25$ **11.** $-\frac{2}{3}$ **13.** **(a)**



(b) 68 **15.** Use symmetry **17.** **(a)**

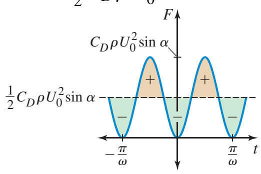


(b) 21.7°C **(d)** 20.05°C
19. The area is a triangle; $x=1$
21. **(a)** Use the Mean Value Theorem **(b)** $t=\frac{1}{4}$ or $t=\frac{3}{4}$
(d) 19.2 in. **23.** $\frac{9}{2}$. **25.** $2e$ **27.** $\frac{1}{3}$
29. $\frac{7}{3}-\ln 2$ **31.** $\frac{\pi}{4}-1+\frac{\sqrt{2}}{2}$ **33.** $\frac{3}{2}$ **35.** $\frac{2}{3}$ **37.** $\frac{1}{4}$ **39.** $\frac{1}{3}\pi r^2 h$
41. $\frac{256}{15}\pi$ **43.** $\frac{\pi}{3}$ **45.** $\frac{2\pi}{5}$ **47.** $\frac{2}{15}\pi$ **49.** $\frac{\pi}{2}(e^4+e^{-4}-2)$ **51.** $\frac{2}{3}\pi$
53. 8π **55.** $\frac{4\pi}{3}$ **57.** $\frac{2\pi}{35}$ **59.** $2\sqrt{5}$ **61.** $\frac{1}{27}(40^{3/2}-13^{3/2})$ **63.** $\frac{181}{9}$
65. $\int_{-1}^1 \sqrt{1+4x^2}\,dx$ **67.** $\int_0^1 \sqrt{1+e^{-2x}}\,dx$ **69.** $\frac{\pi}{2}$
71. $f'(a)=\frac{e^a-e^{-a}}{2}$

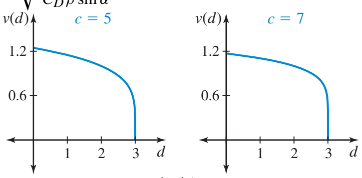
Chapter 6 Review Problems

1. $\frac{2}{5}x^{5/2}+C$ **3.** $\ln|x+1|+C$ **5.** $\tan^{-1}x+C$ **7.** $\frac{4}{3}$ **9.** $\frac{1}{2}$ **11.** 0
13. **(a)** U_0 is the maximum wing speed; period is $\frac{2\pi}{\omega}$ **(b)** Areas

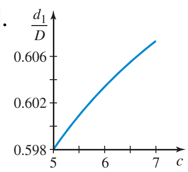
above and below the line $F = \frac{1}{2}C_D\rho U_0^2 \sin\alpha$ are the same, so line gives the average force.



(c) $U_0 \geq \sqrt{\frac{2W}{C_D \rho \sin\alpha}}$ **15.** 3.43
17. (a)



(b) $v'(d) = -\frac{1}{c}\left(\frac{D-d}{a}\right)^{-1+1/c} < 0$ so v is decreasing; $v(d)$ therefore takes its largest value when d takes it smallest allowed value ($d = 0$). Also $v(D) = \left(\frac{D-D}{a}\right)^{1/c} = 0$. **19.** Solve (6.28) for d_1 using $\overline{v}$ computed in Exercise 18. As shown in graph $d_1/D \approx 0.6$ for all $c \in [5, 7]$.



Section 7.1

1. $\frac{2}{3}(x^2+1)^{3/2} + C$ **3.** $\frac{6}{5}(1+x^2)^{5/4} + C$ **5.** $-\frac{5}{2}\cos(2x) + C$ **7.** $-\frac{7}{8}\cos(4x^2) + C$ **9.** $\frac{1}{2}e^{2x+3} + C$ **11.** $e^{x^2/2} + C$ **13.** $\frac{1}{2}\ln|x^2+4x| + C$ **15.** $3(x+2) - 6\ln|x+2| + C$ **17.** $\frac{2}{3}(x+2)^{3/2} + C$ **19.** $\frac{2}{3}(2x^2 - x + 2)^{3/2} + C$ **21.** $-\frac{1}{2}\ln|1+4x-2x^2| + C$ **23.** $\frac{3}{4}\ln|1+2x^2| + C$ **25.** $\frac{3}{2}e^{x^2} + C$ **27.** $\frac{1}{3}(\ln x)^3 + \ln x + C$ **29.** $\frac{1}{2}\sin^2 x + C$ **31.** $\frac{1}{3}(1+x^2)^{3/2} + C$ **33.** $\frac{1}{3}(\ln x)^3 + C$ **35.** $\frac{1}{5}(1+x^2)^{5/2} - \frac{1}{3}(1+x^2)^{3/2} + C$ **37.** $\ln|ax^2+bx+c| + C$ **39.** $\frac{1}{n+1}[g(x)]^{n+1} + C$ **41.** $\tan^{-1}(g(x)) + C$ **43.** $\frac{1}{3}(10^{3/2} - 1)$ **45.** $\frac{77}{1296}$ **47.** $\frac{1}{2}(e^{-1} - e^{-25})$ **49.** $\frac{3}{8}$ **51.** $\frac{1}{2}$ **53.** $2 - 2\ln 2$ **55.** $\frac{1}{2}$ **57.** $\frac{2}{9}(2\sqrt{2} - 1)$ **59.** $\ln|\sin x| + C$

Section 7.2

1. $x\sin x + \cos x + C$ **3.** $\frac{2}{3}x\sin 3x + \frac{2}{9}\cos 3x + C$ **5.** $-4x\cos\frac{x}{2} + 8\sin\frac{x}{2} + C$ **7.** $xe^x - e^x + C$ **9.** $x^2e^x - 2xe^x + 2e^x + C$ **11.** $\frac{1}{2}x^2\ln x - \frac{1}{4}x^2 + C$ **13.** $\frac{1}{2}x^2\ln(3x) - \frac{1}{4}x^2 + C$ **15.** $x\tan x + \ln|\cos x| + C$ **17.** $\frac{1}{2}\left(\sqrt{3} - \frac{\pi}{3}\right)$ **19.** $2\ln 2 - 1$ **21.** $2\ln 4 - \frac{3}{2}$ **23.** $1 - 2e^{-1}$ **25.** $\frac{1}{2}(e^{\pi/2} + 1)$ **27.** $\frac{1}{13}e^{-3x}(-3\cos 2x + 2\sin 2x) + C$ **29.** $\frac{1}{2}x[\sin(\ln x) - \cos(\ln x)] + C$ **31.** $\frac{1}{2}\sin x\cos x + \frac{1}{2}x + C$ **33. (b)** $x\arcsin x + \sqrt{1-x^2} + C$ **35. (b)** $\frac{1}{2}(\ln x)^2 + C$ **37. (b)** $-\left(\frac{1}{3}x^2 + \frac{2}{9}x + \frac{2}{27}\right)e^{-3x} + C$ **39.** $2e^{\sqrt{x}}(\sqrt{x} - 1) + C$ **41.** $-(x^2+2)e^{-x^2/2} + C$ **43.** $\frac{1}{2}(-x^2\cos x^2 + \sin x^2) + C$ **45.** $\sqrt{\pi}\sin\frac{\sqrt{\pi}}{2} + 2\cos\frac{\sqrt{\pi}}{2} - 2$ **47.** $-\frac{1}{2} + 3\ln 3$ **49.** $-\frac{1}{2}xe^{-2x} - \frac{1}{4}e^{-2x} + C$ **51.** $\frac{3}{7}(x+1)^{7/3} - \frac{3}{4}(x+1)^{4/3} + C$ **53.** $-\cos(x^2) + C$ **55.** $\frac{1}{4}\tan^{-1}\left(\frac{x}{4}\right) + C$ **57.** $\ln|\tan x + 1| + C$ **59.** $x(\sin x - \cos x) + \cos x + \sin x + C$ **61.** $\int \ln x\,dx = x\ln x - x + C$ **63.** $2e^2$ **65.** $\frac{\pi}{2}$ **67.** $\frac{1}{2}(e^{\pi/2} + 1)$ **69.** $-\ln|\ln N| + C$ **71.** $-\ln|1 - L| + C$

Section 7.3

1. $1 + \frac{1}{x+1}$ **3.** $2x + 1 - \frac{3}{x+2}$ **5.** $3x - 2 + \frac{2x}{x^2+1}$ **7.** $1 - \frac{x}{x^2+2x+1}$ **9.** $x + 3 + \frac{2x-2}{x^2+1}$ **11.** $x - 1 + \frac{x}{x^2+x}$ **13.** $x^3 + x^2 + x + 1 + \frac{2}{x-1}$ **15.** $\frac{1}{2}\left(1 - \frac{1}{2x+3}\right)$ **17.** $-\frac{3}{x} + \frac{5}{x+1}$ **19.** $\frac{1}{2x} + \frac{1}{2(x+2)}$ **21.** $\frac{4}{2x-5} + \frac{2}{3x+1}$ **23.** $\frac{2}{x-1} + \frac{3}{x+1}$ **25.** $-\frac{3}{3x-1} + \frac{1}{x+3}$ **27.** $-\frac{1}{2x} + \frac{3}{2(x-2)}$ **29.** $\frac{1}{3x} - \frac{1}{2(x-1)} + \frac{1}{6(x-3)}$ **31.** $\frac{1}{2}\ln|x-2| - \frac{1}{2}\ln|x| + C$ **33.** $\frac{1}{4}\ln|x-3| - \frac{1}{4}\ln|x+1| + C$ **35.** $x - \ln|x| - 3\ln|x+2| + C$ **37.** $2\ln|x-1| + \frac{1}{x-1} + C$ **39.** $\ln|x| + 2\ln|x-1| - \frac{3}{x-1} + C$ **41.** $-\frac{1}{2}\ln|x| + \frac{3}{2}\ln|x+2| + \frac{1}{x} + C$ **43.** $\ln|x-1| + \frac{1}{x-1} + C$ **45.** $\ln|x| + \ln|x+1| + \frac{2}{x+1} + C$ **47.** (d) $B = -1$ (e) $\ln|x-1| - \ln|x+1| + \frac{1}{x+1} + C$; (f) $\frac{1}{x} - \frac{2}{x^2} + \frac{1}{2(1+x)}$ **49.** $\frac{x}{x^2+4} - \frac{1}{x^2+1}$ **51.** $\frac{1}{x-1} + \frac{2}{x^2+1}$ **53.** $\frac{1}{x} + \frac{1}{x^2} + \frac{1}{x^2+1}$ **55.** $\arctan(x-1) + C$ **57.** $\frac{1}{3}\arctan\frac{x-2}{3} + C$ **59.** $\frac{1}{3}\arctan\frac{x}{3} + C$ **61.** $\frac{1}{2}\ln|x^2+4x+5| - 2\arctan(x+2) + C$ **63.** $\ln|x| - \frac{1}{2}\ln|x^2+1| + \arctan x + C$ **65.** $\ln|x| + \frac{1}{2}\ln|x^2+2x+5| - \arctan\frac{x+1}{2} + C$ **67.** $-\frac{1}{x} - \arctan x + C$ **69.** $\frac{1}{5}\ln|x-3| - \frac{1}{5}\ln|x+2| + C$ **71.** $\frac{1}{6}\ln|x-3| - \frac{1}{6}\ln|x+3| + C$ **73.** $\frac{1}{3}\ln|x-2| - \frac{1}{3}\ln|x+1| + C$ **75.** $x - 5\ln|x+2| + 2\ln|x+1| + C$ **77.** $x + 2\ln|x-2| - 2\ln|x+2| + C$ **79.** $\ln\frac{1372}{1125} \approx 0.198486$ **81.** $\frac{1}{2}\ln 2$ **83.** $-\ln 2$ **85.** $\frac{\pi}{4} - \frac{1}{2}\ln 2$

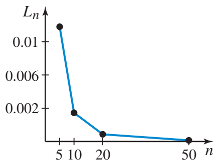
Section 7.4

1. infinite interval; 3 **3.** infinite interval; π **5.** infinite interval; $\frac{\pi}{4}$ **7.** infinite interval; 2 **9.** infinite interval; 0 **11.** integrand undefined at $x = 9$; 6 **13.** integrand undefined at $x = 2$; π **15.** integrand undefined at $x = 0$; -2 **17.** infinite interval; integral divergent **19.** infinite interval; $\frac{1}{2}$ **21.** integrand undefined at $x = 0$; integral divergent **23.** integrand undefined at $x = 1$; 0 **25.** integrand undefined at $x = 1$; integral divergent **27.** infinite interval; integral divergent **29.** integrand undefined at $x = \pm 1$; 0 **31.** infinite interval, and integrand undefined at $x = \pm 1$; integral divergent **33.** $c = 3$ **35.** Proof **37. (b)** $0 \leq \int_1^\infty e^{-x^2}\,dx \leq \lim_{z\to\infty}\int_1^z e^{-x}\,dx = e^{-1} < \infty$; convergent **39. (b)** $\int_1^\infty \frac{1}{\sqrt{1+x^2}}\,dx \geq \lim_{z\to\infty}\frac{1}{2}\int_1^z \frac{1}{x}\,dx = \infty$; divergent **41.** For $x \geq 1$: $0 \leq e^{-x^2/2} \leq e^{-x/2}$; convergent **43.** For $x \geq 1$: $\frac{1}{\sqrt{x+1}} \geq \frac{1}{2\sqrt{x}}$; divergent **45. (a)** Use l'Hôpital's rule **(b)** Show that $\lim_{x\to\infty}\frac{\ln x}{\sqrt{x}} = 0$ and use a graphing calculator to show that, for $x > 74.2$, $2\ln x \leq \sqrt{x}$ **(c)** For $x > 74.2$, use the fact that $e^{-\sqrt{x}} \leq e^{-2\ln x} = \frac{1}{x^2}$ and conclude that the integral is convergent **47. (a)** $\frac{1}{b}$; **(b)** $\frac{1}{b^2}$; **(c)** Break integral into $\int_0^1 t^{1/2}e^{-bt}dt + \int_1^\infty t^{1/2}e^{-bt}dt$. First integral is proper (integrand is continuous, interval is finite.) For $t \geq 1$, $t^{1/2}e^{-bt} < te^{-bt}$; $\int_1^\infty te^{-bt}\,dt$ converges from (b), so the second integral is also convergent.

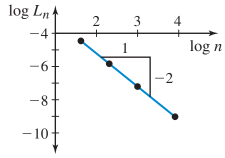
Section 7.5

1. 2.328 **3.** 1.43826 **5.** $M_4 \approx 0.6912$; error ≈ 0.0019 **7.** $M_4 = 62$; error $= 2$ **9.** $T_4 \approx 1.8111$ **11.** $T_3 \approx 1.9677$ **13.** $T_5 = 20.32$; error ≈ 0.32 **15.** $T_4 \approx 1.8195$; error ≈ 0.0661 **17. (a)** $M_{10} \approx 0.3325$; **(b)** $M_{20} \approx 0.3331$ **19. (a)** $M_{10} \approx 0.3098$; **(b)** $M_{20} \approx 0.3102$ **21. (a)** $M_{20} \approx 0.5205$; **(b)** $M_{40} \approx 0.5213$ **23. (a)** $T_{10} \approx 64.0875$; **(b)** $T_{20} \approx 63.8344$ **25. (a)** $T_{10} \approx 3.1157$; **(b)** $T_{20} \approx 3.1315$ **27. (a)** $T_{20} \approx 0.2207$; **(b)** $T_{40} \approx 0.2188$ **29.** $n = 82$ **31.** $n = 58$ **33.** $n = 92$ **35.** $n = 48$ **37. (a)** $M_5 \approx 0.245$; $T_5 \approx 0.26$; $|\int_0^1 x^3\,dx - M_5| = 0.005$; $|\int_0^1 x^3\,dx - T_5| = 0.01$ **(c)** $0.6433 \leq \int_0^1 \sqrt{x}\,dx \leq 0.6730$ **39. (a)** $T_5 \approx 0.8762$; **(b)** $T_{10} \approx 0.8675$; **(c)** $T_{20} \approx 0.8654$; **(d)** $T_{50} \approx 0.8648$; **(e)** $\int_0^2 e^{-x}\,dx = 1 - e^{-2} \approx 0.8647$;

(f) $L_5 \approx 0.01150$; $L_{10} \approx 0.00288$; $L_{20} \approx 0.00072$; $L_{50} \approx 0.000115$.



(g)



The graph is a straight line with slope -2; consistent with $\log L_n = b - 2\log n$ or $L_n \propto n^{-2}$.

Section 7.6

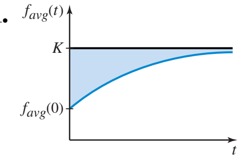
1. $L(x) = e + ex$ **3.** $L(x) = 1 - x$ **5.** $L(x) = x$
7. $P_5(x) = 1 + 3x + 3x^2 + x^3 = (x+1)^3$ **9.** $P_6(x) = \frac{120}{5!}x^5 = x^5$
11. $P_3(x) = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3$; $P_3(0.1) \approx 1.0488125$; $f(0.1) = \sqrt{1.1} \approx 1.0488088$; $|f(0.1) - P_3(0.1)| \approx 4 \times 10^{-6}$
13. $P_5(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}$; $P_5(1) \approx 0.8417$; $f(1) \approx 0.8415$; $|P_5(1) - f(1)| \approx 1.96 \times 10^{-4}$ **15.** $P_2(x) = x$; $P_2(0.1) = 0.1$; $f(0.1) \approx 0.10033$; $|P_2(0.1) - f(0.1)| \approx 3.35 \times 10^{-4}$
17. **(a)** $P_3(x) = x - \frac{x^3}{3!}$ **(b)** $\lim_{x\to 0} \frac{P_3(x)}{x} = 1$ and $P_3(x)$ approximates $f(x) = \frac{\sin x}{x}$ at $x = 0$
19. $P_3(x) = 1 + \frac{1}{2}(x-1) - \frac{1}{8}(x-1)^2 + \frac{1}{16}(x-1)^3$; $P_3(2) \approx 1.4375$; $f(2) \approx 1.4142$; $|P_3(2) - f(2)| \approx 0.023$
21. $P_3(x) = -\left(x - \frac{\pi}{2}\right) + \frac{1}{6}\left(x - \frac{\pi}{2}\right)^3$; $P_3\left(\frac{\pi}{3}\right) \approx 0.499674$; $f\left(\frac{\pi}{6}\right) = 0.5$; $\left|P_3\left(\frac{\pi}{6}\right) - f\left(\frac{\pi}{2}\right)\right| \approx 3.26 \times 10^{-4}$
23. $P_3(x) = e^2 + e^2(x-2) + \frac{e^2}{2}(x-2)^2 + \frac{e^2}{6}(x-2)^3$; $P_3(2.1) \approx 8.1661$; $f(2.1) \approx 8.1662$; $|P_3(2.1) - f(2.1)| \approx 3.14 \times 10^{-5}$ **25.** With $f(N) = rN\left(1 - \frac{N}{K}\right)$, $f'(N) = r - \frac{2r}{K}N$, so $L(N) = f(0) + f'(0)(N - 0) = rN$
27. $n = 10$ **29.** $n = 2$ **31.** $f'(x) = \alpha(1+x)^{\alpha-1}$, $f''(x) = \alpha(\alpha-1)(1+x)^{\alpha-2}$, and so forth. Evaluating at $x = 0$ gives the series on the right; equality results from the assumed fact about convergence.

Section 7.7

1. $\frac{x}{2} - \frac{3}{4}\ln|2x+3| + C$
3. $\frac{1}{2}\left(x\sqrt{x^2+16} + 16\ln|x + \sqrt{x^2+16}|\right) + C$ **5.** $\frac{1}{8}e^2 + \frac{3}{8}$
7. $\frac{2}{9}e^3 + \frac{1}{9}$ **9.** $\frac{1}{4}e^{\pi/2}(1+\sqrt{3}) + \frac{1}{4}(1-\sqrt{3})$
11. $-2e^{-x/2}(x^2 + 4x + 7) + C$ **13.** $\frac{1}{20}[10x - 6 + \sin(10x - 6)] + C$
15. $\frac{1}{2}\left((x+1)\sqrt{x^2+2x+2} + \ln|x + 1 + \sqrt{x^2+2x+2}|\right) + C$
17. $\frac{4e^{2x+1}}{16+\pi^2}\left[2\sin\left(\frac{\pi}{2}x\right) - \frac{\pi}{2}\cos\left(\frac{\pi}{2}x\right)\right] + C$ **19.** $\frac{3}{4}(\ln 2)^2$
21. $\frac{x}{2}(\sin(\ln(3x)) - \cos(\ln(3x))) + C$

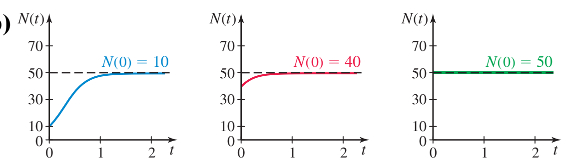
Chapter 7 Review Problems

1. $\frac{2}{9}(1+x^3)^3 + C$ **3.** $-(x+1)e^{-x} + C$
5. $\frac{6}{7}(1+\sqrt{x})^{7/3} - \frac{3}{2}(1+\sqrt{x})^{4/3} + C$ **7.** $x\sin x + \cos x + C$
9. $\frac{1}{2}x^2\ln x - \frac{1}{4}x^2 + C$ **11.** $e^{\sin x} + C$ **13.** $\frac{1}{6}\ln\left|\frac{x+3}{x-3}\right| + C$
15. $-\ln|\cos x| + C$ **17.** $\frac{e^x}{5}(\sin 2x - 2\cos 2x) + C$
19. $\frac{1}{2}\sin(x^2+1) - \frac{1}{2}x^2\cos(x^2+1) + C$ **21.** $-\frac{1}{4}\sin 2x + \frac{x}{2} + C$
23. $\ln\left|\frac{x-1}{x}\right| + C$ **25.** $x - 2\ln|x+2| + C$ **27.** $\frac{1}{2}\ln|x^2+4| + C$
29. $\frac{1}{2}x^2 + 3x + 4\ln|x-1| + C$ **31.** $\frac{1}{3}(2\sqrt{2} - 1)$ **33.** $1 - e^{-1/2}$
35. $\frac{\pi}{8}$ **37.** 4 **39.** $-\frac{1}{4}$ **41.** $e - e^{\sqrt{2}/2}$ **43.** $\frac{\pi}{6}$ **45.** divergent
47. divergent **49.** 2 **51.** **(a)** $M_4 \approx 3.2894$ **(b)** $T_4 \approx 3.0928$
53. **(a)** $M_5 \approx 0.7481$ **(b)** $T_5 \approx 0.7444$ **55.** $P_3(x) = 2x - \frac{4}{3}x^3$
57. $P_3(x) = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3$ **59.** **(a)** Use the substitution $u = \frac{x^2}{2}$ and show that the integrals from $-\infty$ to 0 and from 0 to ∞ are negatives of one another; **(b)** Use the definition of $|x|$; $d = \frac{2}{\sqrt{2\pi}}$; **(c)** $T_5 \approx 0.6762$ **61.**



Section 8.1

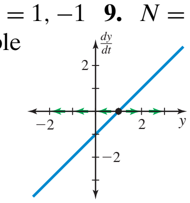
1. $y = \frac{1}{2}t^2 - \cos t + 1$ **3.** $y = \ln|x|$ **5.** $x(t) = 2 - \ln|1 - t|$
7. $s(t) = \frac{2}{3}(t+1)^{3/2} + \frac{1}{3}$ for $t \geq -1$ **9.** $V(t) = t + \sin t + 5$
11. $t = \frac{M_0}{k_0}$ **13.** $y = 2e^{2t}$ **15.** $x(t) = 3e^{2-2t}$
17. $h(s) = \frac{1}{2}(9e^{2s} - 1)$ **19.** $N(t) = 20e^{0.3t}$; $N(5) = 20e^{1.5} \approx 89.634$
21. **(a)** $N(t) = Ce^{rt}$ **(b)** $\ln N(t) = \ln C + (r\ln e)t = rt + \ln C$. To determine r, graph $N(t)$ on a semilog graph; the slope is then r. **(c)** 1. Obtain data at various points in time. 2. Plot on semilog paper. 3. Determine the slope of the resulting straight line. 4. The slope is r. **23.** **(a)** $M(t) = M_0e^{-k_1t}$ **(b)** No **(c)** $t = \frac{1}{2}\ln 10 \approx 1.1513$ **25.** **(a)** $L_\infty = 123$; $k = \frac{1}{27}\ln\frac{244}{123} \approx 0.0254$ **(b)** $L(10) = 123 - 122e^{-0.254} \approx 28.37$ in. **(c)** $t = \frac{1}{0.0254}\ln\frac{122}{12.3} \approx 90.33$ months **27.** **(a)** $C(t) = C_0e^{-k_1t}$ **(b)** $k_1 = \frac{1}{3}\ln\frac{90}{34} \approx 0.3245$ **(c)** $C(t) = \frac{r}{k_1}(1 - e^{-k_1t})$ **(d)** $\frac{r}{k_1}$ **(e)** $r \approx 42.185\, ng/ml \cdot hr$ **29.** $y = \frac{2}{3e^{-x}-2}$ **31.** $y = \frac{2}{1+e^{2x}}$
33. $y = \frac{15e^{2t-2}}{5e^{2t-2}-2}$ **35.** $y = \frac{Ce^x}{1-Ce^x}$ **37.** $y = -1 \pm (C - 2x)^{-1/2}$
39. **(b)** **(i)** $y = \frac{2-2e^{4x}}{1+e^{4x}}$ **(ii)** $y = 2$ **(iii)** $y = \frac{2e^{4x}+6}{3-e^{4x}}$
41. $N(t) = \frac{200}{1+3e^{-0.34t}}$, $\lim_{t\to\infty} N(t) = 200$
43. **(a)** $\frac{dN}{dt} = 5N\left(1 - \frac{N}{50}\right)$
(b)



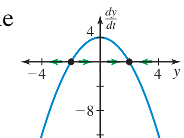
45. **(a)** $p(t) = \frac{p_0}{(1-p_0)e^{-st/2}+p_0}$, $t \geq 0$ **(b)** $\lim_{t\to\infty} e^{-st/2} = 0$ so $\lim_{t\to\infty} p(t) = 1$. A is more fit, so it eventually dominates. **(c)** $t = \frac{2}{0.01}\ln 9 \approx 439.4$ **47.** $y = \sqrt{x^2 + 2x + 4}$
49. $y = -1 + 3\exp[1 - e^{-x}]$ **51.** $y = -1 + 6|x - 1|$
53. $y = \sqrt{t^2 + 1}$ **55.** $\left|\frac{y}{y+1}\right| = \frac{1}{2}|t - 1|$
57. $\frac{1}{3}y^3 + \frac{1}{2}y^2 = \frac{1}{2}t^2 + t + \frac{5}{6}$ **59.** $(y+1)^{3/2} = (t+1)^{3/2} + 2\sqrt{2} - 1$
61. $N(24) = 5e^{48} \approx 3.5 \times 10^{21}$ **63.** S grows at relative rate of 2.85 times relative growth rate of L; $\frac{1}{S}\frac{dS}{dt} = \frac{2.85}{L}\frac{dL}{dt}$

Section 8.2

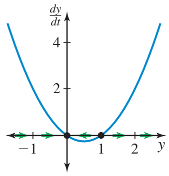
1. $y = 0, 1, -1$ **3.** $x = 1, 2$ **5.** $y = 2$ **7.** $x = 1, -1$ **9.** $N = 0$
11. $N = \pi k$, k an integer **13.** $y = 1$, unstable



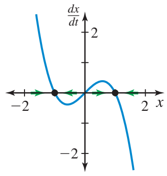
15. $y = 2$, stable; $y = -2$, unstable



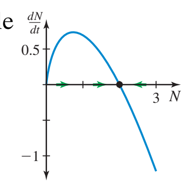
17. $y = 0$, stable; $y = 1$, unstable



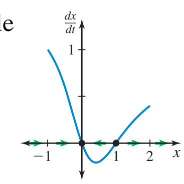
19. $x = 1, -1$, stable; $x = 0$, unstable



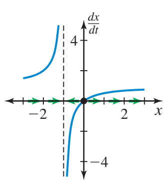
21. $N = 2$, stable



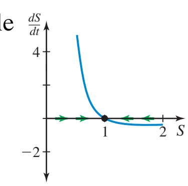
23. $x = 0$, stable; $x = 1$, unstable



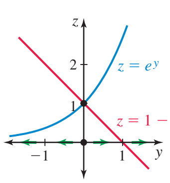
25. $x = 0$, unstable



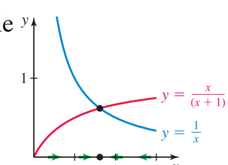
27. $S = 1$, stable



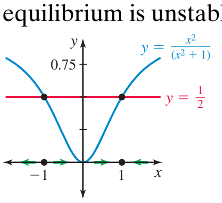
29. $y = 0$, unstable



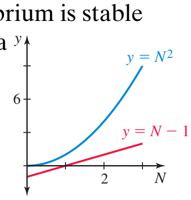
31. one equilibrium, stable



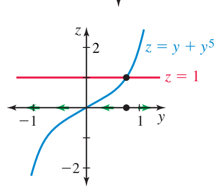
33. two equilibria; negative equilibrium is unstable, positive equilibrium is stable



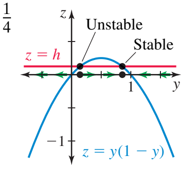
35. no equilibria



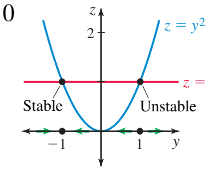
37. one equilibrium, unstable



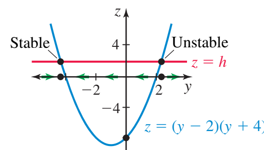
39. there are equilibria for $h < \frac{1}{4}$



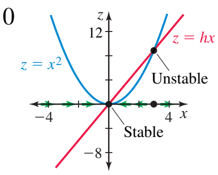
41. there are two equilibria for $h > 0$



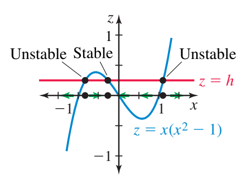
43. there are two equilibria for $h < 9$



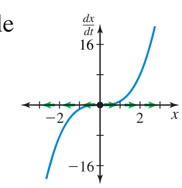
45. there are two equilibria for $h \neq 0$



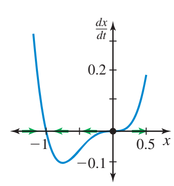
47. there are three equilibria for $-\frac{1}{\sqrt{3}} < h < \frac{1}{\sqrt{3}}$ if $|h| > \frac{1}{\sqrt{3}}$ there is one equilibrium.



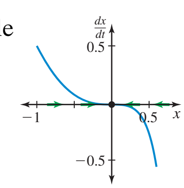
49. unstable



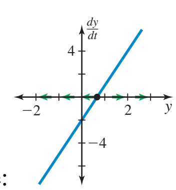
51. semi-stable **53.** unstable

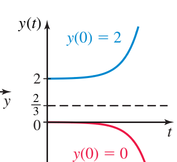


55. stable

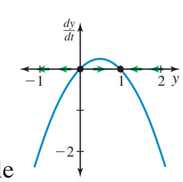


57. $y = \frac{2}{3}$ unstable:

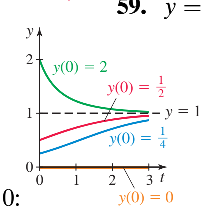




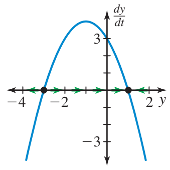
59. $y = 0$, unstable; $y = 1$, stable

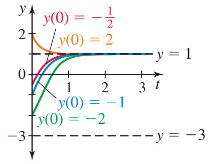


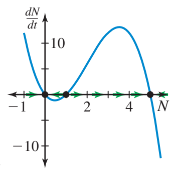
at 0:

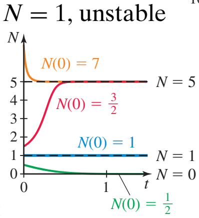


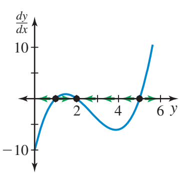
61. $y = -3$,

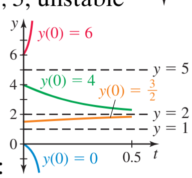
unstable; $y = 1$, stable


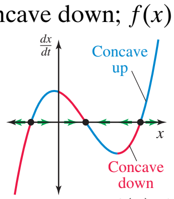
solutions converge to 1 unless $y(0) < -3$, when they diverge to $-\infty$:


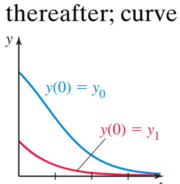
63. $N = 0, 5$, stable; $N = 1$, unstable


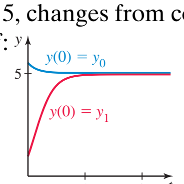
solutions converge to 5 or to 0:


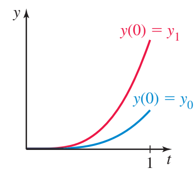
65. $y = 2$, stable; $y = 1, 5$, unstable


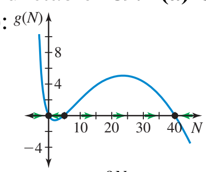
solutions converge to $2, -\infty$, or ∞:


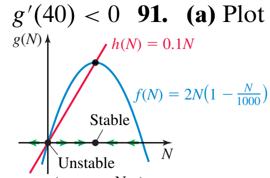
67. Mimic the proof in the text
69. **(a)** Use the chain rule **(b)** $f(x) > 0$, $f'(x) > 0$ implies $\frac{d^2x}{dt^2} > 0$, i.e. $x(t)$ is concave up; $f(x) > 0$, $f'(x) < 0$ implies $\frac{d^2x}{dt^2} < 0$, i.e. $x(t)$ is concave down **(c)** $f(x) < 0$, $f'(x) > 0$ implies concave down; $f(x) < 0$, $f'(x) < 0$ implies concave up
(d)


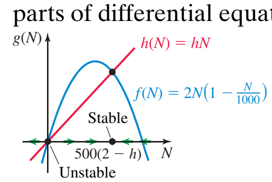
(e) concavity of $x(t)$ changes at local extremum of $f(x)$, i.e., extrema are inflection points of $x(t)$.
71. 0 is stable, 1 is unstable, both solution curves converge to 0; curve for y_0 is concave down for $y > \frac{1}{2}$ and concave up thereafter; curve for y_1 is concave up everywhere:


73. 5 is stable, −1 is unstable. Curve for y_0 converges to $y = 5$, concave up everywhere. Curve for y_1 converges to $y = 5$, changes from concave up to concave down at extremum of f:


75. −1 is unstable. Both solutions diverge to $+\infty$:


77. $y = \frac{2}{3}$, stable **79.** $y = 0$, stable; $y = 2, 3$, unstable **81.** $N = 2$, stable **83.** $y = 0$, stable; $y = 1$, unstable
85. $x = 0$, unstable **87.** **(a)** $x = 0$ unstable; $x = h$ stable
(b) $x = 0$ stable; $x = h$ unstable **89.** **(a)** Using graphing calculator to plot $g(N)$:


You can also sketch $f(N) = N\left(1 - \frac{N}{50}\right)$ and $h(N) = \frac{9N}{5+N}$ and look for points of intersection. **(b)** $N = 0$, $N = 5$, $N = 40$ **(c)** $N = 0$, $N = 40$ stable; $N = 5$ unstable **(d)** $g'(N) = 1 - \frac{N}{25} - \frac{45}{(N+5)^2}$; $g'(0) < 0$, $g'(5) > 0$, $g'(40) < 0$ **91.** **(a)** Plot $f(N) = 2N\left(1 - \frac{N}{1000}\right)$ and $h(N) = 0.1N$.


Curves intersection when: $2N\left(1 - \frac{N}{1000}\right) = 0.1N$ i.e. $N = 0$ (unstable) or $N = 950$ (stable).
(b) In general curves $f(N) = rN\left(1 - \frac{N}{K}\right)$ and $h(N) = hN$ intersect at $\hat{N} = 0$ and $\hat{N} = K(r - h)$. First equilibrium is trivial. Second equilibrium is nontrivial if $\hat{N} > 0$, i.e. if $h < r$.
(c) (i) $g'(N) = (r - h) - \frac{2rN}{K} = (2 - h) - \frac{4N}{1000}$. $g'(0) = 2 - h > 0$ so $\hat{N} = 0$ is unstable. $g'(500(2 - h)) = (2 - h)(1 - 2) = h - 2 < 0$ so $\hat{N} = 500(2 - h)$ is stable. (ii) For graphical approach plot two parts of differential equation separately.


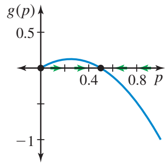
93. $g(N) = -\frac{r}{K}(N^3 - (a + K)N^2 + aKN)$; $g'(0) = -ra < 0$; $g'(a) = -\frac{r}{K}(a^2 - aK) > 0$; $g'(K) = -\frac{r}{K}(K^2 - aK) < 0$; $N = 0, K$ stable; $N = a$ unstable

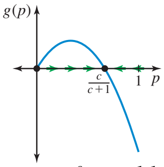
Section 8.3

1. **(a)** $C = C_I - (C_I - C_0)e^{-qt/V}$; $\lim_{t\to\infty} e^{-qt/V} = 0$ so $\lim_{t\to\infty} C = C_I$ **(b)** If $\frac{q}{V}$ is larger, $e^{-qt/V}$ goes to zero more rapidly.
3. **(a)** $\frac{dC}{dt} = \frac{1}{2000}(3 - C)$ **(b)** $C = 3$ is only equilibrium, stable.
(c) $C(t) = 3 - 3e^{-t/2000}$; $\lim_{t\to\infty} C(t) = 3$ **5.** **(a)** Rate at which molecules flow out is $V\frac{dC}{dt}$; use given equation and simplify
(b) $C = C_\infty$ is only equilibrium, stable **(c)** amount r is added in each unit of time.

$$V\frac{dC}{dt} = \text{rate at which molecule is added} - \text{rate at which molecule flows out}$$
$$= r - k(C - C_\infty)$$

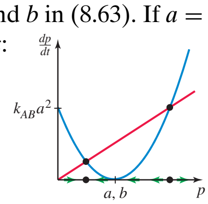
(d) $C = C_\infty + \frac{r}{k}$ is only equilibrium, stable since $g'(C) < 0$
7. With $C_I = C_0$ and $C(0) = C_0 + C_I$, we have $C(t) = C_0 + C_I e^{-qt/V}$. Want to find T_R such that $C(T_R) = C_0 + pC_I$. Solve the resulting equation. $p = \frac{1}{e}$ gives $T_R = \frac{V}{q}$; in general, $T_R = \frac{V}{q}\ln\frac{1}{p}$ **9.** **(a)** $\frac{dN}{dt} = q - fN$ **(b)** $N = \frac{q}{f}$ is only equilibrium,

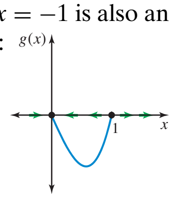
stable **(c)** $f = \frac{2}{3}$ **11.** **(a)** 

(b) $p = 0$, unstable; $p = \frac{1}{2}$, stable **(c)** $g'(0) = 1 > 0$, unstable; $g'(1/2) = -1 < 0$, stable **13.** **(a)** 

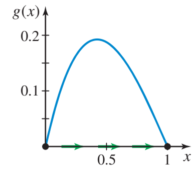
$p = 0$, unstable; $p = \frac{c}{c+1}$, stable **(b)** $\frac{c}{c+1}$ is always a stable equilibrium; Levins model will have nontrivial equilibrium only for some values of the colonization rate, c. **15.** **(a)** $p = 0$, $p = \frac{c-1}{c+2}$ **(b)** $c > 1$ **(c)** $g'\left(\frac{c-1}{c+2}\right) = 1 - c < 0$ **17.** **(a)** $p = 0$, $p = \frac{c-2}{c-1}$ **(b)** $c > 2$ **(c)** $g'\left(\frac{c-2}{c-1}\right) = 2 - c < 0$ **19.** **(a)** $\frac{dp}{dt}$ is rate of patch colonization, $cp(1 - p - D)$ is the number of new patches colonized, mp is number of patches that go extinct **(b)** $g(p) = -2p^2 + 1.4p$, $p = 0$, $p = 0.7$. $p = 0.7$ is stable. **21.** $k_{AB}(a - x)(b - x)$ is the rate at which C is produced. The rate is proportional to the amount of A and B remaining. $-k_C x$ is that the rate of C decomposes. It is proportional to the amount of C present. **23.**

$$g(x) = 2(2 - x)(3 - x) - x$$
$$= 12 - 11x + 2x^2$$

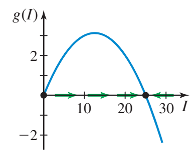
Roots are $x = 3/2$, $x = 4$. Need $0 \le x \le 2$ so only $\hat{x} = 3/2$ is admissible as an equilibrium $g'(x) = -11 + 4x$, so $g'(3/2) = -5 < 0$, so $x = 3/2$ is stable. **25.** If $a > b$, just reverse the roles of a and b in (8.63). If $a = b$, then $\frac{dx}{dt} = k_{AB}(a - x)^2 - k_C x$: 

27. **(a)** Amount of A is now constant, so C is produced at a rate $k_{AB}a(b - x)$. Amount of C that decomposes does not change. **(b)** $x = \frac{k_{AB}}{k_{AB}a + k_C} ab$ **(c)** Derivative of (8.65) is $-k_{AB}a - k_c < 0$, so equilibrium is stable **29.** Equilibrium is when $k_{AB}(a - x)(b - x) = k_C x$. Doubling both k_{AB} and k_C does not change the solutions. **31.** If $b = \frac{c}{2}$ then $\frac{dx}{dt} = -\frac{c}{2}knx(1 - x)$, so $x = 0$ and $x = 1$ are equilibria; 0 is stable, 1 is unstable **33.** $x = 0$, $x = 1$, ($x = -1$ is also an root but we need $x \in [0, 1]$); only $x = 0$ is stable: 

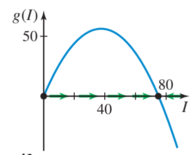
35. **(a)** $x = 0$, $x = 1$. $g'(x) = \left(b - \frac{c}{2}\right)kn(1 - 2x)$, so 0 is unstable, 1 is stable with $b > \frac{c}{2}$ **(b)** $b < \frac{c}{2}$ means $g'(0) < 0$, $g'(1) > 0$ so 0 is stable, 1 is unstable **(c)** $b > \frac{c}{2}$ means benefit derived by each of two cooperators exceeds cost of cooperation, so cooperation is rewarded, so cooperators will dominate. $b < \frac{c}{2}$ means that cheater are rewarded, so they will dominate. **(d)** $b = \frac{c}{2}$ gives $\frac{dx}{dt} = 0$, so $x(t) = x(0)$. No net benefit for cooperation or cheating, so mix will not change **37.** $x = 0$, $x = 1$, $x = \frac{b}{c}$. $g'(0), g'(1) > 0$; $g'\left(\frac{b}{c}\right) < 0$, so only $\frac{b}{c}$ is stable **39.** $x = 0$ unstable, $x = 1$ stable



41. $I = 0$ unstable, $I = 25$ stable



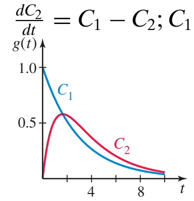
43. $I = 0$ unstable, $I = 100$ stable



45. **(a)** Quarantining reduces b to zero: $\frac{dI}{dt} = -cI$ **(b)** $I(t) = I_0 e^{-ct}$ **(c)** $t = \frac{1}{c}\ln 2$ **(d)** $c = \frac{1}{7}\ln 2 \approx 0.099$ **47.** **(a)** b is number of individuals you come in contact with, not affected by education. c is recovery rate, also not affected. k is rate at which contacted people acquire disease, which should be lowered by handwashing, since infected individuals will have less virus on their hands. **(b)** Need $c > \frac{bk}{2}$, so $k < 0.06$

Section 8.4

1. $y = \frac{C + \ln t}{t}$ **3.** $y = -t - 2 + Ce^t$ **5.** $y = \frac{2x^3 + 3x^2 - 12x + C}{6(x+2)}$ **7.** $y = \frac{x^2 + x\ln x + Cx}{x+1}$ **9.** $x = 1 + Ce^{t - t^2/2}$ **11.** $y = \frac{Ce^x}{x}$ **13.** Integrating factors; $y = -\frac{1}{2}t^3 + Ct$ **15.** Separation of variables; $y = -\frac{2}{t^2 + 2t + C}$ **17.** Separation of variables; $y = \sin t + C$ **19.** Integrating factors; $y = -t^2 - 2 + Ce^{t^2/2}$ **21.** Separation of variables; $y = -1 + Ce^{x + x^2/2}$ **23.** Separation of variables or integrating factors; $y = \frac{Ce^x}{x+1}$ **25.** **(a)** C_1 is as in text. Substitute V_1 for V_2, and equation for C_1, in differential equation for C_2 and solve (integrating factor e^{qt/V_1}). **(b)** $\lim_{t\to\infty} e^{-qt/V_1} = 0$, so $\lim_{t\to\infty} C_1(t) = C_\infty$; $\lim_{t\to\infty} te^{-qt/V_1} = 0$ by L'Hôpital, so $\lim_{t\to\infty} C_2(t) = C_\infty$ **27.** $\frac{dC_1}{dt} = -\frac{1}{3}C_1$, $\frac{dC_2}{dt} = C_1 - C_2$; $C_1(t) = e^{-t/3}$, $C_2(t) = \frac{3}{2}e^{-t/3} - \frac{3}{2}e^{-t}$

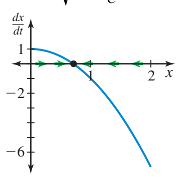


29. **(a)** The volumes are constant, but concentrations are different, so different amounts of solute flows from tank 1 to tank 2, than flows from 2 to 1, so concentrations change **(b)** Flows into and out of tank 2 are the same as in the text, though they go to different places, so differential equation for C_2 is the same. For C_1, inflow comes from tank 2, so just replace C_∞ with C_2 in differential equation for C_1. **(c)** Add $\frac{dC_1}{dt} = \frac{q}{V_1}(C_2 - C_1)$ and $\frac{dC_2}{dt} = \frac{q}{V_1}(C_1 - C_2)$, giving $\frac{dC}{dt} = 0$ **(d)** $C_1(t) = \frac{C}{2} + Ke^{-2qt/V_1}$, $C_2(t) = \frac{C}{2} - Ke^{-2qt/V_1}$ where K is a constant **(e)** $\lim_{t\to\infty} e^{-2qt/V_1} = 0$, so $\lim_{t\to\infty} C_1 = \lim_{t\to\infty} C_2 = \frac{C}{2}$ **(f)** Compute $\frac{dC}{dt} = \frac{d}{dt}\left(\frac{V_1C_1 + V_2C_2}{V_1 + V_2}\right)$ to get zero **(g)** $C_1(t) = C + Ke^{-qt(V_1+V_2)/(V_1V_2)}$, $C_2(t) = C - \frac{V_1}{V_2}Ke^{-qt(V_1+V_2)/(V_1V_2)}$ **(h)** $\lim_{t\to\infty} e^{-qt(V_1+V_2)/(V_1V_2)} = 0$, so $\lim_{t\to\infty} C_1 = \lim_{t\to\infty} C_2 = C$

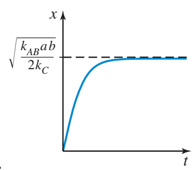
31. $\frac{fg_0}{k-f}\left(\left(\frac{k}{f}\right)^{f/(f-k)} - \left(\frac{k}{f}\right)^{k/(f-k)}\right)$

Chapter 8 Review Problems

1. **(a)** $x = 2 + Ce^{-t}$ **(b)** $\frac{1}{2}y^2 + y + \ln|y-1| = x + C$ **(c)** $y = 1 + Ce^{x^2/2}$ **(d)** $y = \frac{x^3}{2} + Cx$ **3.** **(a)** $\frac{dT}{dt}$ is rate of change of temperature, and difference between object and ambient temperatures is $T_a - T$. k is the constant of proportionality **(b)** $T(t) = T_a + (T_0 - T_a)e^{-kt}$ **(c)** Between 22.2 and 9.8 hours prior to now, so between just after midnight and just after noon today **(d)** $k \approx 0.0896$ **(e)** About 9.9 hours before now, or about 12:06 p.m. **5.** **(a)** $x = \sqrt{\frac{k_{AB}ab}{2k_C}}$, stable because $g'(x) = -4k_C x < 0$



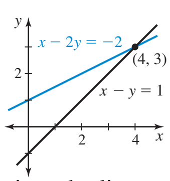
for any $x > 0$ **(b)**



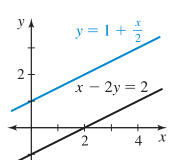
(c) Asymptotically approaches equilibrium from below (since $x(0) = 0$), concave down for all $t > 0$. **7.** **(a)** $x = 0$ stable, $x = 1$ unstable (reject $x = -\frac{1}{2}$) **(b)** If $0 \le x(0) < 1$ then $x(t)$ converges to $x = 0$; i.e. cheaters take over; If $x(0) = 1$, then $x(t)$ remains at $x = 1$ **(c)** Each receives cooperator-cooperator benefit of 3 per interaction; i.e. $3n$ per unit time (since each organism interacts with n others) **(d)** Each receives cheater-cheater benefit of 1 per interaction, or n per unit time.

Section 9.1

1. $\{(4, 3)\}$



3. No solution; the lines are parallel.



5. **(a)** $c = 10$ **(b)** $c \ne 10$ **(c)** No; lines must be parallel or identical **7.** Eliminate x_1 from the second equation and solve the system. **9.** $x = 2/5, y = -11/5$ **11.** $x = 9/17, y = -5/17$ **13.** No solution. **15.** Infinitely many solutions: $\{(x, y) : x = t, y = \frac{3}{2} - \frac{t}{2}; t \in \mathbf{R}\}$ **17.** **(a)** $a > 0$; higher values of BMI increase heart disease risk. $b < 0$; more physical activity decreases heart disease risk. **(b)** $a = 10, b = -3.5$ **(c)** For example, genetic predisposition, external stresses, dietary habits. **19.** $x = 1, y = 2, z = 3$ **21.** $x = 3, y = -1, z = 2$ **23.** $\{(x, y) : x = \frac{7}{4}(1-t), y = \frac{1-t}{2}, z = t; t \in \mathbf{R}\}$ **25.** $x = -1$, $y = -2, z = 3$ **27.** $\{(x, y, z) : x = 2 - t, y = 1 + t, z = t, t \in \mathbf{R}\}$ **29.** underdetermined; $\{(x, y, z) : x = 7 + t, y = t + 2, z = t, t \in \mathbf{R}\}$ **31.** overdetermined; no solution **33.** underdetermined; $\{(x, y, z) : x = \frac{10}{3} + \frac{13}{9}t, y = \frac{2}{3} + \frac{5}{9}t, z = t, t \in \mathbf{R}\}$ **35.** 750 gr of SL 24–4–8; 1000 gr of SL 21–7–12; $\frac{11{,}000}{17}$ gr of SL 17–0–0.

Section 9.2

1. $\begin{bmatrix} 1 & -3 \\ -4 & -5 \end{bmatrix}$ **3.** $D = \begin{bmatrix} 1 & 0 \\ 4 & 3 \end{bmatrix}$ **5.** Use the rules of matrix addition to calculate the left-hand side and the right-hand side. Show both sides are equal to $\begin{bmatrix} 0 & 1 \\ 1 & 4 \end{bmatrix}$

7. $\begin{bmatrix} 4 & -3 & 10 \\ -2 & 5 & -6 \\ 1 & 5 & 7 \end{bmatrix}$ **9.** $D = \begin{bmatrix} -3 & 1 & -9 \\ 0 & -4 & 1 \\ -3 & -1 & -5 \end{bmatrix}$

11. Use the rules of matrix addition to calculate the left-hand side and the right-hand side. Show both sides are equal to $\begin{bmatrix} 3 & -1 & 9 \\ 0 & 4 & -1 \\ 3 & 1 & 5 \end{bmatrix}$ **13.** If $A + B = C$ then $A + B + (-B) = C + (-B) = C - B$. But $A + B + (-B) = A + (B + (-B)) = A + 0 = A$, so $A = C - B$.

15. $A' = \begin{bmatrix} -1 & 3 \\ 0 & 1 \\ 0 & -4 \end{bmatrix}$

17. $(A + B)'_{ij} = (A + B)_{ji} = A_{ji} + B_{ji} = A'_{ij} + B'_{ij}$.

19. $(kA)'_{ij} = (kA)_{ji} = k(A_{ji}) = k(A'_{ij})$. **21.** **(a)** $\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$ **(b)** $\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$

23. $AC = \begin{bmatrix} -1 & -2 \\ 1 & 0 \end{bmatrix}, CA = \begin{bmatrix} 1 & 4 \\ -1 & -2 \end{bmatrix}$ so $AC \ne CA$

25. $(A + B)C = AC + BC = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ **27.** 3×2 **29.** **(a)** 1×4 **(b)** 4×4 **(c)** Not defined.

31. **(a)** $\begin{bmatrix} 7 & 3 & -3 & -3 \\ -4 & -2 & 2 & 0 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & -1 \\ 0 & -2 \\ 0 & 2 \\ -3 & -9 \end{bmatrix}$

33. $A^2 = \begin{bmatrix} 4 & -1 \\ 0 & 9 \end{bmatrix}, A^3 = \begin{bmatrix} 8 & 7 \\ 0 & -27 \end{bmatrix}, A^4 = \begin{bmatrix} 16 & -13 \\ 0 & 81 \end{bmatrix}$

35. Even powers are I_2; odd powers are B. **37.** Calculate AI_2 and I_2A. Compare to A.

39. $\begin{bmatrix} 2 & 3 & -1 \\ 0 & 3 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$

41. $\begin{bmatrix} 2 & -1 \\ -1 & 2 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 4 \end{bmatrix}$

43. $AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, BA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

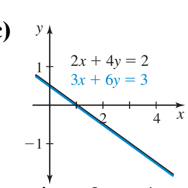
45. $A^{-1} = \begin{bmatrix} -3/5 & 1/5 \\ 2/5 & 1/5 \end{bmatrix}$ **47.** $A^{-1} = \begin{bmatrix} -3/5 & 1/5 \\ 2/5 & 1/5 \end{bmatrix}$ and show that $(A^{-1})^{-1} = A$.

49. C does not have an inverse because $\det C = 0$.

51. **(a)** $x_1 = 2, x_2 = 9$ **(b)** $A^{-1} = \begin{bmatrix} -1 & 0 \\ -2 & -1 \end{bmatrix}, x_1 = 2, x_2 = 9$

53. $\det A = 5$, A is invertible **55.** $\det A = 0$, A is not invertible **57.** **(a)** $\det A = 0$, A is not invertible **(b)** $2x + 4y = b_1$
$3x + 6y = b_2$

(c)



The matrix equation is equivalent to the equations $2x + 4y = 2$, $3x + 6y = 3$. These equations are represented by identical lines (see graph) so there are infinitely many solutions. **(d)** The system has no solutions when $\frac{b_1}{2} \ne \frac{b_2}{3}$, e.g., $b_1 = 1, b_2 = 0$.

59. $A^{-1} = \begin{bmatrix} -1 & -1 \\ 3 & 2 \end{bmatrix}$ **61.** $A^{-1} = \begin{bmatrix} 0 & 1/5 \\ 1/4 & 1/20 \end{bmatrix}$

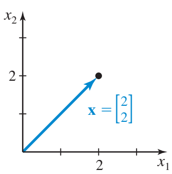
63. $\det A = 2$, A is invertible; $A^{-1} = \begin{bmatrix} 1 & 1/2 \\ 0 & 1/2 \end{bmatrix}$, $X = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

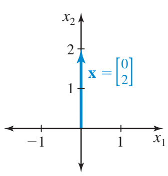
65. C^{-1} does not exist. $\mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}$ solves the equation for all $\{(x, y) : x = -3t, y = t, t \in \mathbf{R}\}$ **67.** $\begin{bmatrix} 1/4 & 1/4 & 0 \\ -1/8 & 3/8 & 1/2 \\ -3/8 & 1/8 & -1/2 \end{bmatrix}$

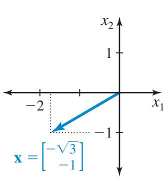
69. $\begin{bmatrix} -2/3 & -1/6 & -1/3 \\ 0 & -1/2 & 0 \\ -1/3 & 1/6 & 1/3 \end{bmatrix}$

Section 9.3

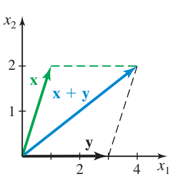
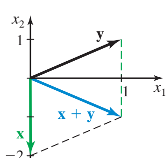
1. Both left and right hand sides are equal to:
(a) $\begin{bmatrix} 2(x_1 + y_1) + 2(x_2 + y_2) \\ -(x_1 + y_1) + 4(x_2 + y_2) \end{bmatrix}$ **(b)** $\begin{bmatrix} 2\lambda x_1 + 2\lambda x_2 \\ -\lambda x_1 + 4\lambda x_2 \end{bmatrix}$

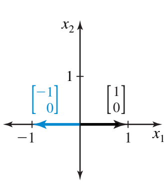
3.  length: $2\sqrt{2}$, angle: $\frac{\pi}{4}$

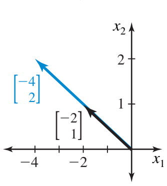
5.  length: 2, angle: $\frac{\pi}{2}$

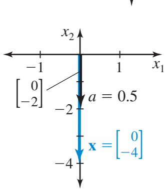
7.  length: 2, angle: $\frac{7\pi}{6}$ **9.** $\begin{bmatrix} 1 \\ \sqrt{3} \end{bmatrix}$ **11.** $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$

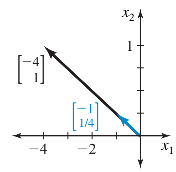
13. $x_1 = 3\cos 345^\circ \approx 2.898$, $x_2 = 3\sin 345^\circ \approx -0.776$
15. $x_1 = 3\cos 115^\circ \approx -1.268$, $x_2 = 3\sin 115^\circ \approx 2.719$

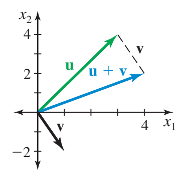
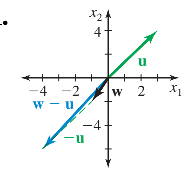
17.  $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ **19.**  $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

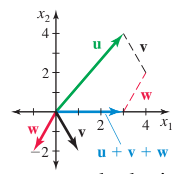
21.  $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ vectors cancel

23.  $\begin{bmatrix} -4 \\ 2 \end{bmatrix}$ increase length by factor 2

25.  $\begin{bmatrix} 0 \\ -2 \end{bmatrix}$ shrink by factor $\frac{1}{2}$

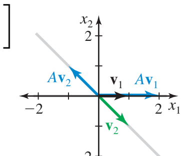
27.  $\begin{bmatrix} -1 \\ 1/4 \end{bmatrix}$ shrink by factor $\frac{1}{4}$

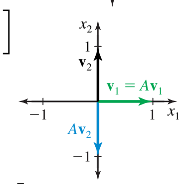
29.  $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ **31.**  $\begin{bmatrix} -4 \\ -6 \end{bmatrix}$

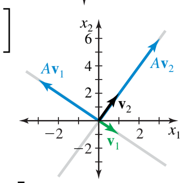
33.  $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ **35.** leaves $\mathbf{x}$ unchanged

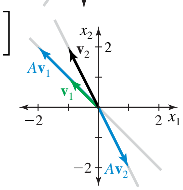
37. counterclockwise rotation by $\theta = \pi$ **39.** counterclockwise rotation by $\theta = \frac{\pi}{6}$ **41.** $\begin{bmatrix} -1/2 - \sqrt{3} \\ 1 - \sqrt{3}/2 \end{bmatrix}$ **43.** $\begin{bmatrix} 7\sqrt{2}/2 \\ -3\sqrt{2}/2 \end{bmatrix}$

45. $\begin{bmatrix} 1 + \sqrt{3}/2 \\ -\sqrt{3} + 1/2 \end{bmatrix}$ **47.** $\begin{bmatrix} -3 \\ -5 \end{bmatrix}$

49. $\lambda_1 = 2$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\lambda_2 = -1$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ 

51. $\lambda_1 = 1$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\lambda_2 = -1$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 

53. $\lambda_1 = -3$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\lambda_2 = 3$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ 

55. $\lambda_1 = 2$, $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\lambda_2 = -1$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ 

57. $\lambda_1 = 4, \lambda_2 = 3$ **59.** $\lambda_1 = 1, \lambda_2 = 2$ **61.** $\lambda_1 = a, \lambda_2 = b$
63. **(a)** Eigenvalues, λ, are roots of $\det(A - \lambda I) = (a - \lambda)(c - \lambda) = 0$. So eigenvalues are $\lambda_1 = a$, $\lambda_2 = c$. **(b)** Check that $A\mathbf{u}_1 = a\mathbf{u}_1$ and $A\mathbf{u}_2 = c\mathbf{u}_2$. **65.** The real parts of both eigenvalues are negative because $\operatorname{tr}(A) = -2 < 0$ and $\det(A) = 13 > 0$. **67.** The real parts of the eigenvalues are not both negative because $\det(A) = -10$ is not > 0. **69.** **(a)** $\mathbf{u}_1$ is an eigenvector for $\lambda_1 = -1$ because $A\mathbf{u}_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} = -\mathbf{u}_1$; $\mathbf{u}_2$ is an eigenvector for $\lambda_2 = 2$ because $A\mathbf{u}_2 = \begin{bmatrix} 2 \\ 6 \end{bmatrix} = 2\mathbf{u}_2$; since the

eigenvalues are unequal, the eigenvectors are linearly independent. **(b)** $\mathbf{x} = 2\mathbf{u}_1 - \mathbf{u}_2$ **(c)** $A^{20}\mathbf{x} = \begin{bmatrix} -1048574 \\ -3145728 \end{bmatrix}$

71. $\begin{bmatrix} -2 \\ 4 \end{bmatrix}$ **73.** $\begin{bmatrix} 2^{22} - 7 \cdot 3^{20} \\ -2^{22} + 2 \cdot 3^{20} \end{bmatrix}$

Section 9.4

1. $L = \begin{bmatrix} 0 & 2.4 & 1.3 \\ 0.2 & 0 & 0 \\ 0 & 0.7 & 0 \end{bmatrix}, N(2) = \begin{bmatrix} 1688 \\ 436 \\ 280 \end{bmatrix}$

3. $L = \begin{bmatrix} 0 & 0 & 4.6 & 3.7 \\ 0.7 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \end{bmatrix}, N(3) = \begin{bmatrix} 2507 \\ 869 \\ 467 \\ 52 \end{bmatrix}$

5. Four age classes; 40% of one-year olds survive until the end of the next breeding season; 3 is the average number of female offspring of a two-year-old. **7.** Four age classes; 40% of two-year olds survive until the end of the next breeding season; 2.5 is the average number of female offspring of a one-year-old. **9.** $q_0(t)$ and $q_1(t)$ seem to converge to 1.5; it appears that 60% of females will be age 0 in the stable age distribution. **11.** $q_0(t)$ and $q_1(t)$ oscillate between 0.4 and 3. **13. (a)** $\lambda_1 \approx 2.483$, $\lambda_2 \approx -0.483$ **(b)** The larger eigenvalue corresponds to the growth rate, 2.483. **(c)** About 89.2% are in age class 0, and 10.8% are in age class 1 in the stable age distribution. **15. (a)** $\lambda_1 = 0$, $\lambda_2 = 4$ **(b)** The larger eigenvalue, 4, corresponds to the growth rate. **(c)** 2/3 are in age class 0, and 1/3 are in age class 1 in the stable age distribution. **17. (a)** $\lambda_1 = \sqrt{0.45}$, $\lambda_2 = -\sqrt{0.45}$ **(b)** The larger eigenvalue, ≈ 0.671, corresponds to the growth rate. **(c)** 88.17% are in age class 0, and 11.83% are in age class 1 in the stable age distribution.

Section 9.5

1. (a) $\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$ **(b)** $\begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$ **(c)** $\begin{bmatrix} 6 \\ -3 \\ 0 \end{bmatrix}$ **3.** $\begin{bmatrix} -1 \\ -2 \end{bmatrix}$ **5.** $\begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix}$

7. $2\sqrt{2}$ **9.** $\sqrt{26}$ **11.** $\begin{bmatrix} 1/\sqrt{11} \\ 3/\sqrt{11} \\ -1/\sqrt{11} \end{bmatrix}$ **13.** $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ **15.** -5 **17.** 3

19. $\sqrt{5}$ **21.** $\sqrt{30}$ **23.** $\cos\theta = 4/5, \theta \approx 0.644$

25. $\cos\theta = 2/\sqrt{110}, \theta \approx 1.379$ **27.** $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ **29.** $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

31. (a,b) $\overline{PQ} = 3$, $\overline{QR} = 4$, $\overline{PR} = 5$, angle $QPR = \tan^{-1}(4/3) \approx 0.927$, angle $PRQ = \pi/2 - \tan^{-1}(4/3) \approx 0.644$, angle $RQP = \pi/2$. **33. (a)** $\overline{PQ} = \sqrt{2}$, $\overline{QR} = \sqrt{10}$, $\overline{PR} = \sqrt{10}$ **(b)** angle $QPR = \cos^{-1}(1/2\sqrt{5}) \approx 1.345 \approx 77.08°$, angle $PRQ = \cos^{-1}(9/10) \approx 0.451 \approx 25.842°$, angle $RQP = \cos^{-1}(1/2\sqrt{5}) \approx 1.345 \approx 77.08°$. **35.** $x + 2y = 4$ **37.** $4x + y = 2$ **39.** $x + y + z = 0$ **41.** $x = 0$ **43.** $x = 1 + 2t$ and $y = -1 + t$ for $t \in \mathbf{R}$ **45.** $x = -1 + t$ and $y = -2 - 2t$ for $t \in \mathbf{R}$ **47.** $x = -1 + 4t$, $y = 2 - 2t$; $x + 2y - 3 = 0$ **49.** $x = 1 + 3t$, $y = -3$; $y + 3 = 0$

51. $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix} + t\begin{bmatrix} 1 \\ -3/2 \end{bmatrix}, t \in \mathbf{R}$

53. $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} + t\begin{bmatrix} 1 \\ -2 \end{bmatrix}, t \in \mathbf{R}$

55. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + t\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, t \in \mathbf{R}$

57. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix} + t\begin{bmatrix} -1 \\ -2 \\ 4 \end{bmatrix}, t \in \mathbf{R}$

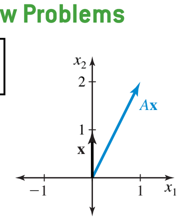
59. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \\ -1 \end{bmatrix} + t\begin{bmatrix} -3 \\ -4 \\ 4 \end{bmatrix}, t \in \mathbf{R}$

61. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} + t\begin{bmatrix} -7 \\ 5 \\ 0 \end{bmatrix}, t \in \mathbf{R}$ **63.** (5/4, −7/4, 13/4)

65. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, t \in \mathbf{R}$

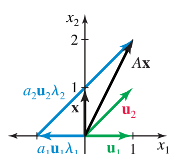
Chapter 9 Review Problems

1. (a) $A\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$



(b) $\lambda_1 = 1, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \lambda_2 = 2$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(c) $\mathbf{u}_1$ remains unchanged; $\mathbf{u}_2$ is stretched by a factor of 2. **(d)** $a_1 = -1, a_2 = 1$; $a_1\lambda_1\mathbf{u}_1 + a_2\lambda_2\mathbf{u}_2 = (-1)(1)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + (1)(2)\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ which is the same as the answer from (a).

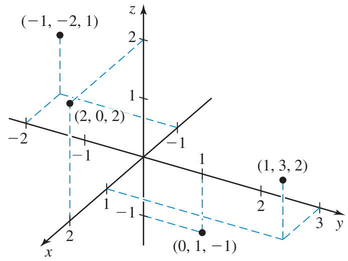


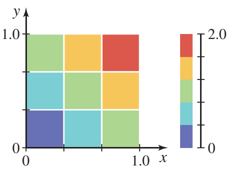
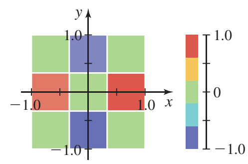
3. Growth rate: $\lambda_1 = 1$; a eigenvector is: $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

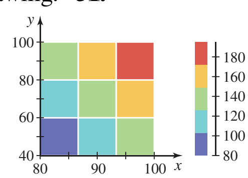
So 2/3 of organisms are in the first age category and 1/3 in the second age category. **5.** $\begin{bmatrix} -2 & 5 \\ -2 & 9 \end{bmatrix}$ **7.** 1. Gaussian elimination; 2. Write in matrix form $AX = B$ and find the inverse of A. Then compute $X = A^{-1}B$. **9.** $a = -3$ **11.** Largest eigenvalue is $\lambda = \frac{1}{8}(3 + \sqrt{32a + 9})$. Population will grow when $\lambda > 1$, i.e., when $\sqrt{32a + 9} > 5$, i.e., when $a > \frac{1}{2}$. **13.** Eigenvalues solve $\det(A - \lambda I) = 0 \Rightarrow (a - \lambda)(b - \lambda) = 0$. Eigenvalues are $\therefore a$ and b. **15. (a)** The growth rate of N_1 is the birth rate (3) less the death rate (0.3) and similarly for N_2. **(b)** $L = \begin{bmatrix} 2.7 & 0.1 \\ 0.2 & 2.6 \end{bmatrix}$ **(c)** $\lambda_1 = 2.8$ is the growth rate **(d)** $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is the stable age distribution, so the two populations equalize. **17. (a)** Two, since there are two unknowns. **(b)** $a \approx 0.203$, $b \approx 0.479$

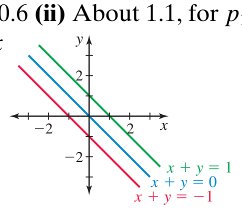
Section 10.1

1. (a) $BMI = \frac{m}{h^2}$ **(b)** $\frac{82}{1.75^2} = 26.78$ **(c)** largest ≈ 71.11, smallest ≈ 12.47 **3. (a) (i)** 0.756 **(ii)** 10.966 **(iii)** 10.179; Boston in January **(b)** ≈ 68.45; doesn't make sense since it is higher than the actual temperature **5.** $CO = HR \times SV$; liters/minute

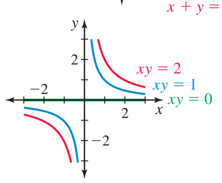
7. 

9. 7 **11.** $\sqrt{10}$ **13.** $e^{-1/10}$
15. $-e^{-2}$ **17.** $[0, 2]$ **19.** $(-\infty, 0]$ **21.** $[-5, 0]$ **23.** $[0, 1)$
25. Holding either x or y constant gives a parabola; consistent with drawing **27.** Holding x constant gives a line with negative slope; holding y constant gives a parabola; consistent with drawing **29.** Holding x constant gives negative exponential approaching zero; holding y constant gives a line; consistent with drawing. **31.**  **33.** 

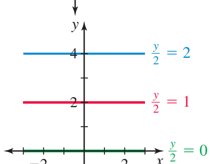
35.  **37.** **(a)** Each organism belongs to exactly one of the three species **(b)** $p_1 + p_2 + p_3 = 1$ results in $p_3 = 1 - p_1 - p_2$ **(c)** $0 \le p_3 \le 1$ gives $0 \le 1 - p_1 - p_2 \le 1$ or $0 \le p_1 + p_2 \le 1$ **(d)(i)** About 0.6 **(ii)** About 1.1, for $p_1 = p_2 = \frac{1}{3}$
39. Level curves are $y = c - x$



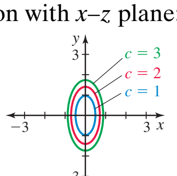
41. Level curves are $y = \frac{c}{x}$



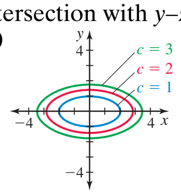
43. Level curves are $y = 2c$



45. **(a)** level curve: $x^2 + y^2 = c$, circle centered at origin with radius $\sqrt{c}$; intersection with x–z plane: $z = x^2$, intersection with y–z plane: $z = y^2$ **(b)**



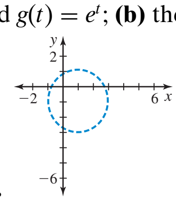
intersection with x–z plane: $z = 4x^2$,
intersection with y–z plane: $z = y^2$
(c)

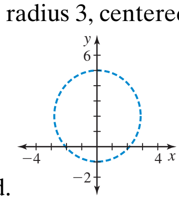


intersection with x–z plane: $z = \frac{1}{4}x^2$,
intersection with y–z plane: $z = y^2$
(d) When $a = 1$, the level curves are circles centered at the origin; surface is a collection of circles with increasing radii as z increases. When $a \ne 1$, level curves are ellipses with center at origin, oriented along y-axis for $a > 1$, x-axis for $a < 1$. So surfaces are elliptic paraboloids ("bowls") similar to Figure 10.29, but passing through (0,0,0). When $a > 1$, the bowls are elongated along the y-axis; when $a < 1$, the bowls are elongated along the x-axis. **47.** Day 180: 22 m; day 200: 18 m; day 220: 14 m **49.** **(a)** About 20°C **(b)** About 95°C **(c)** More than 90°C, but less than 100°C **(d)** Approximately circular region of cells is killed; height: 8, width: 8

Section 10.2

1. 1 **3.** -2 **5.** 18 **7.** $-\frac{1}{4}$ **9.** $\frac{5}{3}$ **11.** $-\frac{3}{2}$ **13.** $\frac{2}{3}$ **15.** Along positive x-axis: 2; along positive y-axis: -1 **17.** Along x-axis: 0; along y-axis: 0; along $y = x$: 2 **19.** Along $y = mx$, $m \ne 0$: 2; along $y = x^2$: 1; the limit does not exist.
21. 1. $f(x, y)$ is defined at (0, 0). 2. $\lim_{(x,y)\to(0,0)}(2x^2 + y^2 + 1)$ exists. 3. $f(0, 0) = 1 = \lim_{(x,y)\to(0,0)}(2x^2 + y^2 + 1)$ **23.** In Problem 17, we showed that $\lim_{(x,y)\to(0,0)} \frac{4xy}{x^2+y^2}$ does not exist. Hence, $f(x, y)$ is discontinuous at (0, 0) **25.** In Problem 19, we showed that $\lim_{(x,y)\to(0,0)} \frac{2xy}{x^3+yx}$ does not exist. Hence, $f(x, y)$ is discontinuous at (0, 0) **27.** **(a)** $h(x, y) = g[f(x, y)]$ with $f(x, y) = x^2 + y^2$ and $g(t) = \sin t$; **(b)** the function is continuous for all $(x, y) \in \mathbf{R}^2$ **29.** **(a)** $h(x, y) = g[f(x, y)]$ with $f(x, y) = xy$ and $g(t) = e^t$; **(b)** the function is continuous for all $(x, y) \in \mathbf{R}^2$

31.  $\{(x, y) : (x-1)^2 + (y+1)^2 < 4\}$ **33.** The boundary is a circle with radius 3, centered at (0, 2). The boundary is not included.

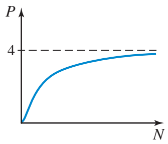


35. Choose $\delta^2 = \epsilon/2$.

Section 10.3

1. $\frac{\partial f}{\partial x} = 2xy + y^2$, $\frac{\partial f}{\partial y} = x^2 + 2xy$ **3.** $\frac{\partial f}{\partial x} = \frac{3}{2}y\sqrt{xy} - \frac{2y}{3(xy)^{1/3}}$, $\frac{\partial f}{\partial y} = \frac{3}{2}x\sqrt{xy} - \frac{2x}{3(xy)^{1/3}}$ **5.** $\frac{\partial f}{\partial x} = \cos(x + y)$, $\frac{\partial f}{\partial y} = \cos(x + y)$
7. $\frac{\partial f}{\partial x} = -4x\cos(x^2 - 2y)\sin(x^2 - 2y)$, $\frac{\partial f}{\partial y} = 4\cos(x^2 - 2y)\sin(x^2 - 2y)$ **9.** $\frac{\partial f}{\partial x} = e^{\sqrt{x+y}}\frac{1}{2\sqrt{x+y}}$, $\frac{\partial f}{\partial y} = e^{\sqrt{x+y}}\frac{1}{2\sqrt{x+y}}$ **11.** $\frac{\partial f}{\partial x} = e^x\sin(xy) + ye^x\cos(xy)$, $\frac{\partial f}{\partial y} = xe^x\cos(xy)$ **13.** $\frac{\partial f}{\partial x} = \frac{2}{2x+y}$, $\frac{\partial f}{\partial y} = \frac{1}{2x+y}$ **15.** $\frac{\partial f}{\partial x} = \frac{2x}{x^2+y}$, $\frac{\partial f}{\partial y} = \frac{1}{x^2+y} - \frac{1}{y}$ **17.** 6 **19.** $-e$ **21.** 1 **23.** $\frac{2}{9}$ **25.** $f_x(1, 1) = -2$, $f_y(1, 1) = -2$ **27.** $f_x(-2, 1) = 4$, $f_y(-2, 1) = -2$
29. $\frac{\partial P}{\partial a} = \frac{N}{(1+aT_hN)^2} > 0$: the number of prey items eaten increases with increasing attack rate.
31. $\frac{\partial f}{\partial x} = 2xz - y$, $\frac{\partial f}{\partial y} = z^2 - x$, $\frac{\partial f}{\partial z} = 2yz + x^2$
33. $\frac{\partial f}{\partial x} = 3x^2y^2z + \frac{2x}{yz}$, $\frac{\partial f}{\partial y} = 2x^3yz - \frac{x^2}{y^2z}$, $\frac{\partial f}{\partial z} = x^3y^2 - \frac{x^2}{yz^2}$
35. $\frac{\partial f}{\partial x} = e^{x+y+z}$, $\frac{\partial f}{\partial y} = e^{x+y+z}$, $\frac{\partial f}{\partial z} = e^{x+y+z}$
37. $\frac{\partial f}{\partial x} = \frac{1}{x+y+z}$, $\frac{\partial f}{\partial y} = \frac{1}{x+y+z}$, $\frac{\partial f}{\partial z} = \frac{1}{x+y+z}$
39. $2y^2$ **41.** e^y **43.** $2\sec^2(u + w)\tan(u + w)$ **45.** $-\sin y$
47. $-\frac{1}{(x+y)^2}$ **49.** **(a)** $\frac{\partial P}{\partial N} > 0$: the number of prey encounters per

predator increases as the prey density increases. **(b)** $\frac{\partial P}{\partial T_h} < 0$: the function decreases as the handling time T_h increases.



(c) **51.** **(a)** Increasing function of temperature **(b)** Increasing V decreases V when $T < 83.6$ **(c)** For $T < 83.6$

Section 10.4

1. $8 = 6x + 4y - z$ **3.** $z = 2x - y + 2$ **5.** $z - y = 0$ **7.** $z = 4e^2x - 3e^2$ **9.** $z = x + y - 1$ **11.** $f(x, y)$ is defined in an open disk centered at $(1, 1)$ and is continuous at $(1, 1)$ with continuous partial derivatives $(1, 2y)$ **13.** $f(x, y)$ is defined in an open disk centered at $(0, 0)$ and is continuous at $(0, 0)$ with continuous partial derivatives $(-\sin(x + y), -\sin(x + y))$ **15.** $f(x, y)$ is defined in an open disk centered at $(0, 1)$ and is continuous at $(0, 1)$ with continuous partial derivatives $(e^{-y}, -xe^{-y})$ **17.** $L(x, y) = x - 3y$ **19.** $L(x, y) = \frac{1}{2}x + 2y + \frac{1}{2}$ **21.** $L(x, y) = x + y$ **23.** $L(x, y) = x + \frac{1}{2}y - \frac{3}{2} + \ln 2$ **25.** $L(x, y) = 1 + x + y$, $L(0.1, 0.05) = 1.15$, $f(0.1, 0.05) \approx 1.1618$ **27.** $L(x, y) = \frac{1}{2}(x + 1)$, $L(1.1, 0.1) = 1.05$, $f(1.1, 0.1) \approx 1.0536$ **29.** $Df(x, y) = \begin{bmatrix} 1 & 1 \\ 2x & -2y \end{bmatrix}$ **31.** $Df(x, y) = \begin{bmatrix} e^{x-y} & -e^{x-y} \\ e^{x+y} & e^{x+y} \end{bmatrix}$ **33.** $Df(x, y) = \begin{bmatrix} -\sin(x - y) & \sin(x - y) \\ -\sin(x + y) & -\sin(x + y) \end{bmatrix}$ **35.** $Df(x, y) = \begin{bmatrix} 4xy + 1 & 2x^2 - 3 \\ e^x \sin y & e^x \cos y \end{bmatrix}$ **37.** $L(x, y) = \begin{bmatrix} 4x + 2y - 4 \\ -x - y + 3 \end{bmatrix}$ **39.** $L(x, y) = \begin{bmatrix} e^3(2x - y - 2) \\ x - y - 1 \end{bmatrix}$ **41.** $L(x, y) = \begin{bmatrix} \frac{1}{4}(2x - y + 2) \\ y - 2x + 2 \end{bmatrix}$ **43.** $L(1.1, 1.9) = \begin{bmatrix} -0.9 \\ 9.8 \end{bmatrix}$, $f(1.1, 1.9) = \begin{bmatrix} -0.88 \\ 9.83 \end{bmatrix}$ **45.** $L(1.9, -3.1) = \begin{bmatrix} 25 \\ -22.4 \end{bmatrix}$, $f(1.9, -3.1) = \begin{bmatrix} 25 \\ -22.382 \end{bmatrix}$ **47.** **(a)** Evaluate f and g at $(0.3, 50)$ **(b)** $\begin{bmatrix} 0.003v - 0.15 \\ -500u + 150 \end{bmatrix}$

Section 10.5

1. 22 **3.** $\dfrac{\frac{\pi}{3} + \frac{\sqrt{3}}{4}}{\sqrt{\frac{\pi^2}{9} + \frac{3}{4}}}$ **5.** 0 **7.** $\frac{dz}{dt} = \frac{\partial f}{\partial x}u'(t) + \frac{\partial f}{\partial y}v'(t)$ **9.** $-\frac{x}{y}$ **11.** $-\frac{y(1-2x^2)}{x(1-2y^2)}$ **13.** $\frac{dy}{dx} = -\frac{1}{\sqrt{1-x^2}}$ for $-1 \le x \le 1$ **15.** The growth rate decreases over time. **17.** **(a)** $\frac{d\rho}{dt} = \rho\left(-\frac{v\lambda}{T} + \frac{\lambda z}{T^2}\frac{dT}{dt}\right)$ **(b)** Set $\frac{d\rho}{dt} = 0$; then $\frac{dT}{dt} = \frac{Tv}{z}$. This has solution $T = C(1 + vt)$; $T(0) = 1$ gives $T(t) = 1 + vt$ **19.** $\frac{dm}{dk} = -\frac{m}{k} < 0$

Section 10.6

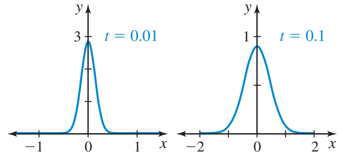
1. $\text{grad}\, f = \begin{bmatrix} 3x^2y^2 \\ 2x^3y \end{bmatrix}$ **3.** $\text{grad}\, f = \frac{1}{2\sqrt{x^3 - 3xy}}\begin{bmatrix} 3x^2 - 3y \\ -3x \end{bmatrix}$ **5.** $\text{grad}\, f = \frac{\exp[\sqrt{x^2+y^2}]}{\sqrt{x^2+y^2}}\begin{bmatrix} x \\ y \end{bmatrix}$ **7.** $\text{grad}\, f = \frac{x^2-y^2}{x^2+y^2}\begin{bmatrix} \frac{1}{x} \\ -\frac{1}{y} \end{bmatrix}$ **9.** $\frac{1}{\sqrt{2}}$ **11.** $\frac{1}{\sqrt{2}}$ **13.** $2\sqrt{5}$ **15.** $\frac{13}{\sqrt{2}}$ **17.** $-\frac{1}{4\sqrt{29}}$ **19.** $\text{grad}\, f(-1, 1) = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$ **21.** $\text{grad}\, f(5, 3) = \begin{bmatrix} \frac{5}{4} \\ -\frac{3}{4} \end{bmatrix}$ **23.** $\begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$ **25.** $\frac{1}{\sqrt{733}}\begin{bmatrix} 2 \\ -27 \end{bmatrix}$ **27.** Gradient is $\frac{1}{11\sqrt{11}}\begin{bmatrix} -12 \\ -4 \end{bmatrix}$, direction is $\begin{bmatrix} -3 \\ -1 \end{bmatrix}$ **29.** $-\text{grad}\, f(2, 3) = \begin{bmatrix} -4 \\ 6 \end{bmatrix}$

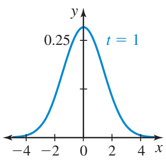
Section 10.7

1. $f(x, y)$ has a local minimum at $(0, -1)$ **3.** $f(x, y)$ has saddle points at $(2, 4)$ and $(-2, 4)$ **5.** $f(x, y)$ has a saddle point at $(0, 3)$ **7.** $f(x, y)$ has a local maximum at $(0, 0)$. **9.** $f(x, y)$ has saddle points at $(0, \pi/2 + k\pi)$ for $k \in \mathbf{Z}$ **11.** **(c)** Figure 10.86: $f(x, y)$ stays constant for fixed x; no local extremum at $(0, 0)$. Figure 10.87: saddle point at $(0, 0)$, Figure 10.88: local minimum at $(0, 0)$. **13.** Global maximum at $(1, 1)$; $f(1, 1) = 3$. Global minimum at $(-1, -1)$; $f(-1, -1) = -3$. **15.** Global maximum at $(\pm 1, 0)$; $f(\pm 1, 0) = 1$. Global minimum at $(0, \pm 1)$; $f(0, \pm 1) = -1$. **17.** Global maximum at $(0, -1)$ and $(0, 0)$; $f(0, -1) = f(0, 0) = 0$. Global minimum at $\left(1, -\frac{1}{2}\right)$; $f\left(1, -\frac{1}{2}\right) = -\frac{5}{4}$. **19.** Global maximum at $\left(\frac{2}{3}; \frac{2}{3}\right)$, and is equal to $\frac{8}{3}$. **21.** Global maximum at $(-2, 0)$; $f(-2, 0) = -5$. Global minimum at $(3, 0)$; $f(3, 0) = 20$. **23.** Global maximum at $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$; $f\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = 1 + \sqrt{2}$. Global minimum at $\left(-\frac{1}{2}, \frac{1}{2}\right)$; $f\left(-\frac{1}{2}, \frac{1}{2}\right) = -\frac{1}{2}$. **25.** Yes; if e.g., if there is a saddle point between the two maxima. **27.** Global maximum at $(N, P) = (1, 1)$. **29.** Maximum volume is $(2\sqrt{2})^3 = 16\sqrt{2}$ m^3. **31.** The minimum surface area is 96 m^2. **33.** The minimum distance is $1/\sqrt{3}$. **35.** **(a)** Use $p_3 = 1 - p_1 - p_2$ and $0 \le p_3 \le 1$ **(a)** Hessian at $\left(\frac{1}{3}, \frac{1}{3}\right)$ is $\begin{bmatrix} -6 & -3 \\ -3 & -6 \end{bmatrix}$ **37.** Global maxima at $\left(\pm\frac{\sqrt{63}}{8}, \frac{1}{8}\right)$; $f\left(\pm\frac{\sqrt{63}}{8}, \frac{1}{8}\right) = \frac{65}{16}$. Global minimum at $(0, -1)$; $f(0, -1) = -1$. **39.** Global maximum at $\left(\frac{1}{\sqrt{2}}, \frac{1}{2\sqrt{2}}\right)$ and $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right)$; $f\left(\frac{1}{\sqrt{2}}, \frac{1}{2\sqrt{2}}\right) = f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right) = \frac{1}{4}$. Global minima at $\left(\frac{1}{\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right)$ and $\left(-\frac{1}{\sqrt{2}}, \frac{1}{2\sqrt{2}}\right)$; $f\left(\frac{1}{\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right) = f\left(-\frac{1}{\sqrt{2}}, \frac{1}{2\sqrt{2}}\right) = -\frac{1}{4}$. **41.** Global minimum at $\left(\frac{12}{13}, -\frac{8}{13}\right)$; $f\left(\frac{12}{13}, -\frac{8}{13}\right) = \frac{16}{13}$. No global maximum. **43.** Global maximum at $\left(\pm\frac{1}{\sqrt{2}}, \frac{1}{6}\right)$; $f\left(\pm\frac{1}{\sqrt{2}}, \frac{1}{6}\right) = \frac{1}{12}$. No global minimum. **45.** Global maxima at $\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right)$; $f\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right) = \frac{1}{4}$. Global minima at $(\pm 1, 0)$, $(0\ \pm 1)$; $f(x, y) = 0$ at global minima. **47.** **(a)** $x = y = 2$ is a local minimum with value 4 **(b)** No **49.** As you travel along the constraint curve starting at Q, you intersect level curves with successively smaller values, starting with c_1 and ending with c_3, then values on the level curves start increasing again. **51.** **(a)** $3x_1 + 3x_2 = 10$ **(b)** $x_1 \approx 2.810$, $x_2 \approx 0.523$ **53.** **(a)** Each organism belongs to exactly one of the species **(b)** Equations are $-1 - \ln p_j = \lambda$ for all j, so all p_j are equal **(c)** **(i)** All terms except for $-1 \ln 1$ are 0, and $-1 \ln 1 = 0$ as well **(ii)** Use previous parts substituting $n - 1$ for n **55.** $m \approx 3.084$, $c \approx 1.194$ **57.** **(a)** Take logs; $\ln B = \ln c + a \ln M$ **(b)** $a \approx 0.700$, $\ln c \approx -3.684$, $c \approx 0.025$ **(c)** 0.7 is reasonably close to $\frac{3}{4}$ **59.** **(a)** $m \approx 1.922$, $c \approx -79.562$ **(b)** $T \approx 41.395$°C **61.** **(a)** Taking logs, $\ln N = \ln N_0 + rt$; $r \approx 0.122$, $\ln N_0 \approx 2.679$, $N_0 \approx 14.571$ **(b)** $t \approx 5.682$ **63.** **(a)** Invert both sides of equation and simplify **(b)** $\frac{1}{r}$ depends linearly on $\frac{1}{C}$; slope is $\frac{a}{k}$, vertical intercept $\frac{1}{k}$ **(c)** $k \approx 1.996$, $a \approx 0.609$

Section 10.8

1. Compute the derivatives **3. (a) (i)** All factors are positive **(ii)** $\frac{\partial c}{\partial x} = -\frac{x}{2Dt\sqrt{4\pi Dt}}\exp\left[-\frac{x^2}{4Dt}\right]$ **(iii)** Derivative in (ii) is only 0 at $x = 0$; c is an increasing function of x to the left and decreasing to the right, so maximum **(iv)** $\frac{\partial^2 c}{\partial x^2}$ is $\frac{1}{2Dt\sqrt{4\pi Dt}}\exp\left[-\frac{x^2}{4Dt}\right]\left(\frac{x^2}{2Dt} - 1\right)$ so $\frac{\partial^2 c}{\partial x^2} = 0$ when $x^2 = 2Dt$ **(b)**

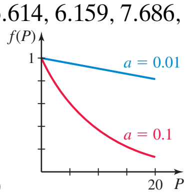




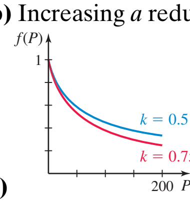
5. (a) $C = 0$ so $\frac{d^2C}{dx^2} = C(0) = C(1) = 0$ **(b)** $\frac{d^2C}{dx^2} = 0, C(0) = 1 - 0 = 1, C(1) = 1 - 1 = 0$ **(c)** $D \neq 0$, so divide through by D giving $0 = \frac{d^2C}{dx^2}$; doesn't change solutions, and this equation is independent of D **7.** Differentiate both expressions

Section 10.9

1. $N_t = 5, 10.5, 22.05, 46.305, 97.241, 204.205, 428.831, 900.544, 1891.14, 3971.4, 8339.94$; $P_t = 0$ for $t = 0, 1, 2, \dots, 10$
3. $N_t = b^t N_0$ **5.** $N_t =$ 5, 7.063, 10.411, 15.503, 23.148, 34.576, 51.559, 76.405, 110.403, 139.973, 75.312, 0.517, 0.009, 0.013, 0.019, 0.028; $P_t =$ 3, 0.874, 0.367, 0.228, 0.212, 0.294, 0.608, 1.869, 8.408, 51.264, 269.296, 224.9, 1.535, 0.001, 0, 0
7. $N_t = 5, 10.5, 22.05, 46.305, 97.241, 204.205, 428.831, 900.544, 1891.14, 3971.4, 8339.94$; $P_t = 0$ for $t = 0, 1, 2, \dots 10$
9. $N_t = b^t N_0$
11. $N_t =$ 100, 96.524, 52.658, 20.829, 10.84, 9.36, 11.1, 14.927, 20.961, 29.757, 41.845, 56.566, 68.48, 64.158, 41.631, 22.448, 14.298, 12.596, 14.137, 17.899, 23.768, 31.809, 41.488, 50.189, 51.895, 42.409
$P_t =$ 30, 106.951, 184.258, 116.315, 40.806, 13.799, 5.881, 3.447, 2.858, 3.37, 5.58, 12.404, 32.737, 77.123, 109.213, 79.996, 38.748, 17.703, 9.512, 6.614, 6.159, 7.686, 12.453, 24.084, 46.778, 70.868 **13. (a)**



(b) Increasing a reduces the chances of escaping parasitism. **15. (a)**

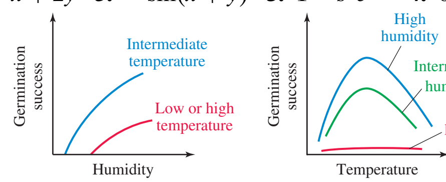


(b) Increasing k reduces the chances of escaping parasitism.
17. Stable **19.** Unstable **21.** Unstable **23.** Stable **25.** Unstable **27.** Eigenvalues: $\sqrt{3}/2, -\sqrt{3}/2$ **29.** Eigenvalues: $\frac{1}{6}(3 \pm i\sqrt{15})$ **31.** For $0 < a < 1$, $(0, 0)$ is locally stable. **33.** $(0, 0)$ is locally stable; $(1, 1)$ is unstable. **35.** $(0, 0)$ is locally stable if $-\frac{1}{2} < a < \frac{1}{2}$. **37. (a)** If $r > 1/2$, then $(r - 1/2, r - 1/2)$ is an equilibrium. **(b)** For $r > \frac{3}{2}$, the equilibrium $(r - 1/2, r - 1/2)$ is locally stable. **39.** $(0, 0)$ is unstable; $((40 \ln 4)/3, 10 \ln 4)$ is unstable. **41.** $(0, 0)$ is unstable; $(1000, 750)$ is locally stable.

Chapter 10 Review Problems

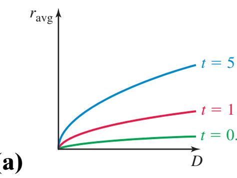
1. $x + 2y$ **3.** $-\sin(x + y)$ **5.** $1 - s^2 e^{-rs}$ **7.** 0

9.



11. (a) $\frac{\partial A_i}{\partial F} > 0, \frac{\partial A_i}{\partial D} < 0$ **(b)** $\frac{\partial A_e}{\partial F} > 0$: area covered by introduced species increased with the amount of fertilizer added. $\frac{\partial A_e}{\partial D} > 0$: area covered by introduced species increased with intensity of disturbance [note that this is the opposite of (a)].

13. $D\mathbf{f}(x, y) = \begin{bmatrix} 2x & -1 \\ 3x^2 & -2y \end{bmatrix}$ **15. (a)**

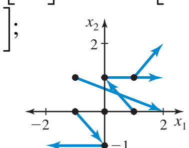


(b) Use (a): $r_{\text{avg}}^2 = \pi Dt$ and solve for D.
(c) $r_{\text{avg}} =$ arithmetic average $= \frac{1}{N}\sum_{i=1}^{N} d_i$, and use formula for D in (b). **17.** $L(x, y) = x$ **19.** $L(x, y) = 4x + 2y - 2$

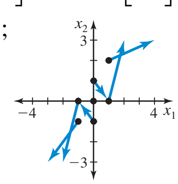
Section 11.1

1. $\frac{d\mathbf{x}}{dt} = \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}\mathbf{x}(t)$ **3.** $\frac{d\mathbf{x}}{dt} = \begin{bmatrix} -2 & 0 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}\mathbf{x}(t)$

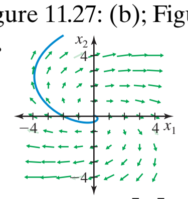
5. $(1, 0)$: $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$; $(0, 1)$: $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$; $(-1, 0)$: $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$; $(0, -1)$: $\begin{bmatrix} -2 \\ 0 \end{bmatrix}$; $(1, 1)$: $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$; $(0, 0)$: $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$; $(-1, 1)$: $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$;



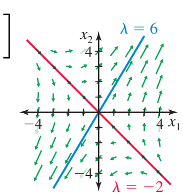
7. $(1, 0)$: $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$; $(0, 1)$: $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$; $(-1, 0)$: $\begin{bmatrix} -1 \\ -3 \end{bmatrix}$; $(0, -1)$: $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$; $(-1, -1)$: $\begin{bmatrix} -2 \\ 2 \end{bmatrix}$; $(0, 0)$: $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$; $(1, 2)$: $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$;



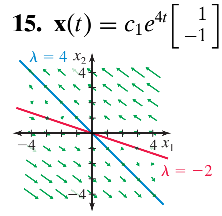
9. Figure 11.25: (d); Figure 11.26: (c); Figure 11.27: (b); Figure 11.28: (a)
11.



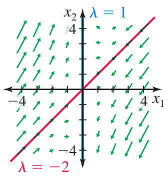
13. $\mathbf{x}(t) = c_1 e^{6t}\begin{bmatrix} 3 \\ 5 \end{bmatrix} + c_2 e^{-2t}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$



15. $\mathbf{x}(t) = c_1 e^{4t}\begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2 e^{-2t}\begin{bmatrix} 3 \\ -1 \end{bmatrix}$



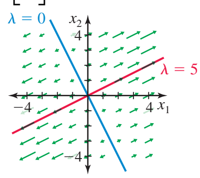
17. $\mathbf{x}(t) = c_1 e^{-2t}\begin{bmatrix}1\\1\end{bmatrix} + c_2 e^{t}\begin{bmatrix}0\\1\end{bmatrix}$



19. $\mathbf{x}(t) = e^{-3t}\begin{bmatrix}-5\\4\end{bmatrix} + e^{2t}\begin{bmatrix}0\\1\end{bmatrix}$ **21.** $\mathbf{x}(t) = e^{t}\begin{bmatrix}1\\1\end{bmatrix}$

23. $\mathbf{x}(t) = 2e^{2t}\begin{bmatrix}7\\2\end{bmatrix} - e^{-3t}\begin{bmatrix}1\\1\end{bmatrix}$ **25.** $\mathbf{x}(t) = -\frac{13}{8}e^{-3t}\begin{bmatrix}-1\\1\end{bmatrix} - \frac{3}{8}e^{5t}\begin{bmatrix}7\\1\end{bmatrix}$

27. (c) Use differentiation to find dx_1/dt and dx_2/dt.
29. unstable node **31.** saddle **33.** saddle **35.** saddle **37.** stable node **39.** stable node **41.** saddle **43.** unstable spiral **45.** stable spiral **47.** stable spiral **49.** unstable spiral **51.** center **53.** center **55.** stable spiral **57.** saddle **59.** stable spiral **61.** saddle **63.** stable node **65.** center **67. (a)** $\lambda_1 = 0$, $\mathbf{v}_1 = \begin{bmatrix}1\\-2\end{bmatrix}$; $\lambda_2 = 5$, $\mathbf{v}_2 = \begin{bmatrix}2\\1\end{bmatrix}$ **(b)** $x_1(t) = c_1 + 2c_2e^{5t}$, $x_2(t) = -2c_1 + c_2e^{5t}$ **(c)**



$\frac{dx_2}{dx_1} = \frac{1}{2}$; for each solution curve $x_2 = \frac{1}{2}x_1 + C$, for some constant C. Solution curves are therefore parallel to the eigenvector $\mathbf{v}_2$. Solution curves are all parallel lines, and each line travels away from a point on the line of eigenvectors along $\mathbf{v}_1$.

Section 11.2

1. $\frac{dx_1}{dt} = -0.7x_1 + 1.2x_2$ (0, 0) is a stable node.
$\frac{dx_2}{dt} = \quad 0.4x_1 - 1.2x_2$

3. $\frac{dx_1}{dt} = -2.5x_1 + 0.7x_2$ (0, 0) is a stable node.
$\frac{dx_2}{dt} = \quad 2.5x_1 - 0.8x_2$

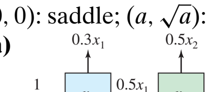
5. $\frac{dx_1}{dt} = -1.8x_1 + 0.6x_2$ (0, 0) is a stable node.
$\frac{dx_2}{dt} = \quad 1.7x_1 - 0.9x_2$

7. $\frac{dx_1}{dt} = -0.6x_1 + 0.5x_2$ (0, 0) is a stable node.
$\frac{dx_2}{dt} = \quad 0.1x_1 - 0.6x_2$

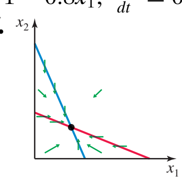
9. $a = 0.1, b = 0.3, c = 0.3, d = 0.2$ **11.** $a = 0, b = 0.1, c = 0.2, d = 0$ **13.** $a = 0.2, b = 1.1, c = 2.1, d = 1.2$ **15.** $a = 0.3, b = 0, c = 0.9, d = 0.2$ **17.** $a = 0, b = 0, c = 0.2, d = 0.3$ **19.** $\frac{dx_2}{dt} = 0.3x_1(t)$; $x_1(t) = 4e^{-0.3t}$, $x_2(t) = 4 - 4e^{-0.3t}$ **21. (a)** $a = 0.2, b = 0.1, c = 0, d = 0$ **(b)** The constant is the total area. **(e)** $x_2(t) = 20 - x_1(t) = \frac{40}{3} + \frac{14}{3}e^{-0.3t}$, $\lim_{t\to\infty}\frac{x_2(t)}{x_1(t)+x_2(t)} = \frac{2}{3}$ **23. (a)** Romeo's feelings for Juliet will always tend towards indifference—if $R > 0$, then $\frac{dR}{dt} < 0$, while if $R < 0$, then $\frac{dR}{dt} > 0$. So Romeo is afraid of either loving or hating Juliet too much. **(b)** $-c, -d$ **(c)** $\lim_{t\to\infty} R(t) = \lim_{t\to\infty} J(t) = 0$; both are indifferent. **25. (a)** $\Delta = -ck < 0$; saddle point. **(b)** $\lambda_1 = -c$: $\begin{bmatrix}1\\0\end{bmatrix}$; $R = 0$ carries J towards the origin; $\lambda_2 = k$: $\begin{bmatrix}a\\c+k\end{bmatrix}$; $J = \frac{a}{c+k}R$ carries values away from the origin. **(c)** Unless the initial condition has $R = 0$, the system will migrate towards either mutual love or mutual hatred, depending on Romeo's initial feelings towards Juliet. If initially Romeo is indifferent towards Juliet, the system will migrate to mutual indifference. **27. (a)** k **(b)** $J(t) = \frac{b}{k}(1 - e^{kt}) - 1$, $R(t) = e^{kt}$; $\lim_{t\to\infty} R(t) = \infty$, $\lim_{t\to\infty} J(t) = -\infty$. **(c)** Either Romeo loves Juliet while she hates him, or the reverse, depending on $R(0)$. **29.** Each person responds to the other depending on the other's opinion of them. That is, if Juliet feels positively towards Romeo, then Romeo will feel more positively towards Juliet, and vice versa. However, each person is also hesitant to respond; that is the meaning of the $-2J$ and $-0.1R$ terms—for example, if Juliet loves Romeo, the $-2J$ term will exert a negative influence on her further love for him. The coefficients mean that Juliet will respond more slowly to Romeo's love than he will to her love for him, but that Juliet is more hesitant to commit than he is. **31.** Romeo's feelings towards Juliet enhance both his feelings towards her and her feelings towards him. In Juliet's case, there is some hesitancy in committing towards either love or hatred, reflected in the $-2J$ term. In Romeo's case, his feelings towards Juliet are reinforced by her feelings for him. **33.** $x(t) = 4\sin(3t)$ **35.** $\frac{dx}{dt} = v$, $\frac{dv}{dt} = -\frac{1}{2}x$ **37.** $\frac{dx}{dt} = v$, $\frac{dv}{dt} = \frac{1}{2}x + 2v$

Section 11.3

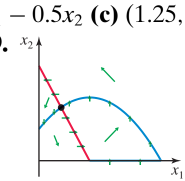
1. unstable spiral **3.** unstable spiral **5.** saddle **7.** (0, 0): unstable node; (1/2, 3/10): stable node; (1/2, 0): saddle; (0, 4/5): saddle **9.** (0, 0): unstable node; (0, 1): saddle; (1, 0): saddle; (1/2, 1): stable node **11.** (0, 0): unstable spiral; (2, −2): saddle **13.** (0, 0): saddle; $(a, \sqrt{a})$: unstable node; $(a, -\sqrt{a})$: stable node **15. (a)**



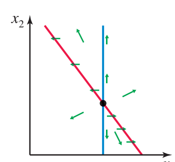
(b) $\frac{dx_1}{dt} = 1 - 0.8x_1$, $\frac{dx_2}{dt} = 0.5x_1 - 0.5x_2$ **(c)** (1.25, 1.25) is a stable node. **17.**



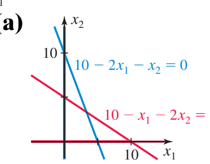
19.



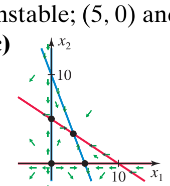
21.



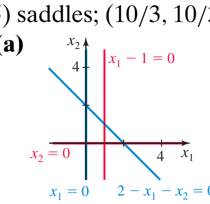
23. (a)



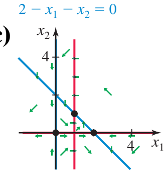
(b) (0, 0) unstable; (5, 0) and (0, 5) saddles; (10/3, 10/3) stable node
(c)



25. (a)

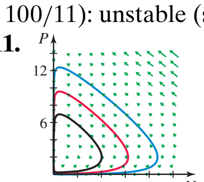


(b) (0, 0) saddle; (2, 0) saddle; (1, 1) stable spiral **(c)**

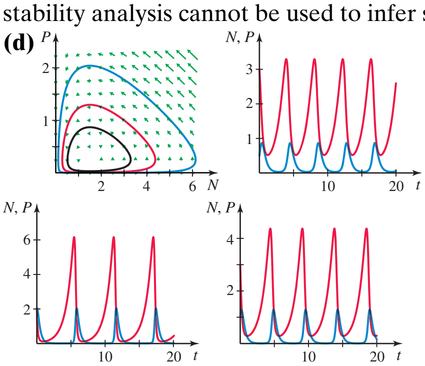


Section 11.4

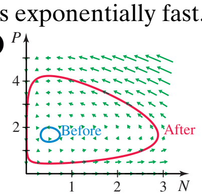
1. coexist **3.** founder control **5.** (0, 0): unstable (source); (18, 0): unstable (saddle); (0, 20): stable (sink) **7.** (0, 0): unstable (source); (35, 0): stable (sink); (0, 40): stable (sink); (85/11, 100/11): unstable (saddle) **9.** $(\alpha_{12}, \alpha_{21}) = (1/4, 7/18)$ **11.**

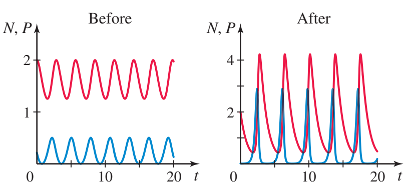


13. (a) trivial equilibrium: $(0,0)$; nontrivial equilibrium: $(3/2, 1/4)$ **(b)** $\lambda_1 = 1$, $\lambda_2 = -3$: $(0,0)$ is an unstable saddle **(c)** $\lambda_1 = i\sqrt{3}$, $\lambda_2 = -i\sqrt{3}$: purely imaginary eigenvalues, linear stability analysis cannot be used to infer stability of equilibrium. **(d)**

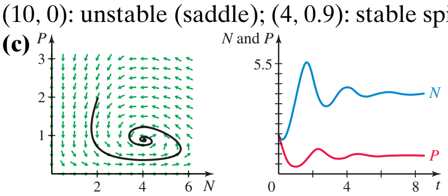


15. (a) $\frac{dN}{dt} = 5N$, $N(t) = N(0)e^{5t}$; in the absence of the predator, the insect species grows exponentially fast. **(b)** If $P(t) > 0$, then $N(t)$ stays bounded. **(c)**

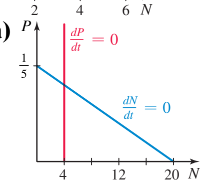




By spraying the field, the solution moves to a different cycle; this results in a much larger insect outbreak later in the year compared to before the spraying. **17. (a)** When $P = 0$, then $\frac{dN}{dt} = 3N\left(1 - \frac{N}{10}\right)$; equilibria: $\hat{N} = 0$ (unstable) and $\hat{N} = 10$ (locally stable). If $N(0) > 0$, then $\lim_{t\to\infty} N(t) = 10$ **(b)** $(0,0)$: unstable (saddle); $(10,0)$: unstable (saddle); $(4, 0.9)$: stable spiral **(c)**



19. (a)

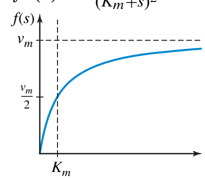


(b) $\tau = -1/5$, $\Delta = 32/5$; stable spiral **21.** $\hat{N} = \frac{d}{b}$ does not depend on r; hence, it remains unchanged if r changes. $\hat{P} = \frac{r}{a}(1 - \frac{d}{bK})$ is an increasing function of r; hence, the predator equilibrium increases when r increases. **23.** $\hat{N} = \frac{d}{b}$ is a decreasing function of b; hence, the prey abundance decreases as b increases. $\hat{P} = \frac{r}{a}(1 - \frac{d}{bK})$ is an increasing function of b; hence, the predator abundance increases as b increases.

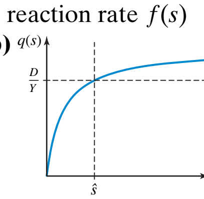
Section 11.5

1. predation; locally stable **3.** mutualism; locally stable **5.** mutualism; unstable **7.** competition; unstable **9.** $a_{ii} = \frac{\partial f_i}{\partial N_i} < 0$ means that species i self-regulates; if the number of individuals is perturbed from its equilibrium value, the population size will grow (if population is decreased) or shrink (if increased) to restore equilibrium.

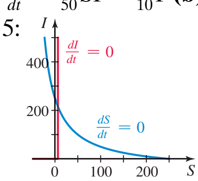
11. (a) $(N^*, P^*) = (d/b, r/a)$ **(b)** $\begin{bmatrix} 0 & -ad/b \\ rb/a & 0 \end{bmatrix}$ **(c)** Diagonal zeros: neither species either encourages or inhibits its own growth in the absence of other factors. Lower left: positive, so growth rate of second species (predators) increases if prey abundance increases. Upper right: negative, so growth rate of the first species (prey) decreases if predator abundance increases. **13. (a)** Trivial equilibrium only **(b)** Three equilibria **15.** Look at discriminant of quadratic $(V - 1/2)(V - 1) = -1/c$. **17.** $\frac{da}{dt} = -kab$, $\frac{db}{dt} = -kab$, $\frac{dc}{dt} = kab$ **19.** $\frac{ds}{dt} = -k_1se$, $\frac{de}{dt} = k_2c - k_1se$, $\frac{dc}{dt} = k_1se - k_2c$, $\frac{dp}{dt} = k_2c$ **21.** Add the two equations; result is zero. Note that $\frac{dx}{dt} + \frac{dy}{dt} = \frac{d(x+y)}{dt}$. **23.** Add the three equations **25. (a)** Use l'Hôpital's rule **(b)** $f(K_m) = v_m/2$ **(c) (i)** Since $v_m, K_m > 0$ **(ii, iii)** Compute $f'(s) = \frac{K_m v_m}{(K_m+s)^2} > 0$, $f''(s) = -\frac{2K_m v_m}{(K_m+S)^3} < 0$.



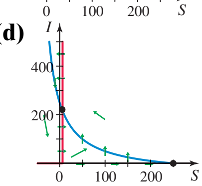
(d) From the graph, increasing the substrate s increases the reaction rate $f(s)$ **27. (a)** $(\hat{s}, Y(s_0 - \hat{s}))$ where $\hat{s} = \frac{DK_m}{Yv_m - D}$ **(b)**



$q(\hat{s}) = D/Y$, so we find $\hat{s}$ from the graph. Holding Y constant and increasing D increases $\hat{s}$, again from the graph. **29.** $(4,0)$ unstable; $(2,2)$ stable **31.** $(50,0)$ stable node **33. (a)** $\frac{dI}{dt} = \frac{1}{800}SI - \frac{1}{8}I$ **(b)** $(200,0)$ saddle; $(100, 200/27)$ stable spiral **35. (a)** $\frac{dS}{dt} = 250 - S - I - \frac{1}{50}SI$, $\frac{dI}{dt} = \frac{1}{50}SI - \frac{1}{10}I$ **(b)** For $\frac{dS}{dt}$: $I = \frac{50(250-S)}{50+S}$; for $\frac{dI}{dt}$: $I = 0$ or $S = 5$:



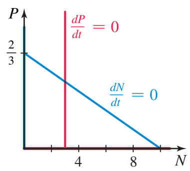
(c) $(250, 0)$ saddle, $(5, 2450/11)$ stable spiral **(d)**



37. (a) $\frac{dI}{dt} = \frac{kb}{N}SI - cI - mI$; $\frac{dS}{dt}$, $\frac{dR}{dt}$ unaffected **(b)** Number of people not constant so $S + I + R$ not conserved. **39. (a)** 1/3 **(b)** $\frac{dS}{dt} = -\frac{1}{200}SI + \frac{1}{10}(100 - S - I) + \frac{1}{3}I$, $\frac{dI}{dt} = \frac{1}{200}SI - \frac{2}{3}I$ **(c)** $c + m > kb$, so only equilibrium is $(100, 0)$ **(d)** stable node **41.** $\frac{dS}{dt} = aR - \frac{kb}{N}SI + mI \cdot \frac{S}{N}$, $\frac{dI}{dt} = \frac{kb}{N}SI - cI - mI + mI \cdot \frac{I}{N}$, $\frac{dR}{dt} = cI - aR + mI \cdot \frac{R}{N}$ **43. (a)** ≈ 95.3 **(b)** ≈ 39.96 **45. (a)** $S \geq 0, I \geq 0, R \geq 0, S + I + R = N$ **(b)** $(100, 0, 0)$, $(0, 100/11, 1000/11) \approx (0, 9.091, 90.909)$ **(c)** $\frac{dI}{dt} = \frac{1}{100}(-2I^2 + 190I + R - 2IR)$; $\frac{dR}{dt} = \frac{1}{10}I - \frac{1}{100}R$ **(d)** $(100, 0, 0)$ saddle, $(0, 100/11, 1000/11)$ stable spiral

Chapter 11 Review Problems

1. saddle **3.** center **5. (a)** $\frac{dx_1}{dt} = 1 - 0.6x_1 + 0.5x_2$, $\frac{dx_2}{dt} = 0.5x_1 - 0.7x_2$ **(b)** $\approx (4.118, 2.941)$ **(c)** stable node **7.** Compute the derivative, solve for Z, look at behavior depending on $r_1 > r_2$ or $r_1 < r_2$ **9. (a)** $\frac{dN}{dt} = N\left(2\left(1 - \frac{N}{10}\right) - 3P\right)$, $\frac{dP}{dt} = PN - 3P$



(b) stable spiral

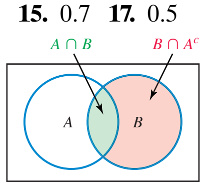
11. (a) Equilibrium at $p_1 = (c_1 - m_1)/c_1$, so meaningful and nontrivial if $c_1 > m_1$ **(b)** As Lotka-Volterra equations $K_1 = 1 - \frac{m_1}{c_1} > 1 - \frac{m_2}{c_2} = K_2$, and $\alpha_{12} = \alpha_{21} = 1$, so $K_1 > \alpha_{12}K_2$ and $K_2 < \alpha_{21}K_1$. **13. (a)** Set (11.95) to zero and solve for $q(s)$ assuming $x \neq 0$; substitute into (11.94) and solve for x. $\hat{x} > 0$ when $\hat{s} < s_0$. **(b)** Increasing uptake rate increases equilibrium abundance; increasing rate at which new medium enters decreases equilibrium abundance. **15. (a)** Setting the equations to zero and solving gives $(x, ax/D)$ is an equilibrium (for all values of x) **(b)** $\lambda_1 = 0$; eigenvector $\begin{bmatrix} D \\ a \end{bmatrix}$; $\lambda_2 = -(a + D)$; eigenvector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ **(c)** Compute $\frac{dx}{dt}$, $\frac{dy}{dt}$ using the eigenvalues and eigenvectors from above **(d)** Add the equations in (11.116). $x + y = A$ is $y = A - x$, which has slope -1; this is the direction vector of $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ **(e)** Isoclines are $-ax + Dy = 0 = ax - Dy$; slope is $\frac{a}{D}$; same direction as $\begin{bmatrix} D \\ a \end{bmatrix}$ **(f)** Solve $-(a + D)x + Dc = 0$ to get $\hat{x} = c\frac{D}{D+a}$; $f'(x) = -(a + D) < 0$, so stable

Section 12.1

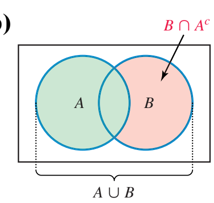
1. 40 **3.** 120 **5.** 84 **7.** $4^{9749} \approx 3.04 \times 10^{5869}$ **9.** 120 **11.** 5040 **13.** 358,800 **15.** 2730 **17.** 6! **19.** 120 **21.** 120 **23.** 1365 **25.** 126 **27.** $\frac{1000!}{980!} \approx 8.26 \times 10^{59}$ **29. (a)** exactly two red balls: 10; exactly two blue balls: 6; one of each: 20 **(b)** total: 36 **31.** 168, 168 **33.** $\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}$; to form a subset, for each element, we need to decide whether it should be in the subset. There are $2^3 = 8$ choices. **35.** 12 **37.** 30 **39.** 31 **41.** $\binom{60}{20}\binom{40}{20}\binom{20}{20}$ **43.** $x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$ **45.** $\binom{26}{4}\binom{26}{5}$ **47.** $4\binom{13}{2}\binom{4}{2}\binom{4}{2}\binom{11}{1}$ **49.** $\binom{13}{1}\binom{4}{4}\binom{12}{1}\binom{4}{1}$ **51.** 3!

Section 12.2

1. $\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$
3. $\Omega = \{(i, j) : 1 \leq i < j \leq 5\}$
5. $A \cup B = \{1, 2, 3, 5\}$, $A \cap B = \{1, 3\}$ **7.** $\{4, 6\}$ **9.** 0.6 **11.** 0.25 **13.** 0.3 **15.** 0.7 **17.** 0.5
19. (a)



(b)



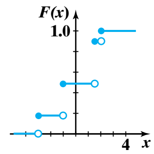
21. $\frac{3}{4}$ **23.** 3/8 **25.** $\frac{11}{36}$ **27.** $\frac{5}{12}$ **29.** $\frac{1}{4}$ **31.** $\frac{1}{2}$ **33.** $\frac{1}{8}$ **35.** 11/16 **37.** 1/2 **39.** 3/5 **41.** $\frac{1}{17}$ **43.** $\frac{12}{55}$ **45.** $1 - \frac{\binom{48}{4}}{\binom{52}{4}}$ **47.** $\frac{\binom{26}{13}}{\binom{52}{13}}$ **49.** $\frac{\binom{13}{2}\binom{4}{2}\binom{4}{2}\binom{11}{1}\binom{4}{1}}{\binom{52}{5}}$ **51. (a)** $\frac{\binom{N-100}{7}\binom{100}{3}}{\binom{N}{10}}$ **(b)** 333

Section 12.3

1. $\frac{13}{51}$ **3.** 13/50 **5.** $\frac{3}{5}$ **7.** $\frac{1}{2}$ **9.** $\frac{1}{6}$ **11.** $\frac{4}{7}$ **13.** 1/4 **15.** 0.8425 **17.** 0.7804 **19.** $\frac{5}{9}$
21. $P(\text{first card is an ace}) = P(\text{second card is an ace}) = \frac{1}{13}$ **23.** 0.3 **25.** $\frac{3}{4}$ **27.** A and B are independent. **29.** A and B are not independent. **31. (a)** $\frac{1}{8}$ **(b)** $\frac{7}{8}$ **(c)** $\frac{1}{2}$ **(d)** $\frac{7}{8}$ **33.** $\left(\frac{1}{4}\right)^{10}$ **35.** $1 - (0.9)^{10}$ **37.** 0.1624 **39.** $\frac{1}{3}$ **41.** $\frac{1}{3}$ **43. (a)** 1 **(b)** $\frac{1}{2}$ **(c)** $\frac{1}{3}$

Section 12.4

1. $P(X = 0) = \frac{1}{4}$, $P(X = 1) = \frac{1}{2}$, $P(X = 2) = \frac{1}{4}$
3. $P(X = 0) = \frac{6}{36}$, $P(X = 1) = \frac{10}{36}$, $P(X = 2) = \frac{8}{36}$, $P(X = 3) = \frac{6}{36}$, $P(X = 4) = \frac{4}{36}$, $P(X = 5) = \frac{2}{36}$
5. $P(X = 0) = \frac{\binom{3}{0}\binom{2}{2}}{\binom{5}{2}}$, $P(X = 1) = \frac{\binom{3}{1}\binom{2}{1}}{\binom{5}{2}}$, $P(X = 2) = \frac{\binom{3}{2}\binom{2}{0}}{\binom{5}{2}}$
7. $P(X = 0) = \frac{\binom{13}{0}\binom{39}{3}}{\binom{52}{3}}$, $P(X = 1) = \frac{\binom{13}{1}\binom{39}{2}}{\binom{52}{3}}$, $P(X = 2) = \frac{\binom{13}{2}\binom{39}{1}}{\binom{52}{3}}$, $P(X = 3) = \frac{\binom{13}{3}\binom{39}{0}}{\binom{52}{3}}$ **9.** $F(x) = \begin{cases} 0, & x < -3 \\ 0.2, & -3 \leq x < -1 \\ 0.5, & -1 \leq x < 1.5 \\ 0.9, & 1.5 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$



11. $P(X = -2) = 0.2$, $P(X = 0) = 0.1$, $P(X = 1) = 0.4$, $P(X = 2) = 0.3$ **13. (a)** $N = 55$ **(b)** $\frac{28}{55}$
15. (a)

k	12	13	14	15	16	17	18	19	20	21
p_k	1/25	4/25	5/25	3/25	1/25	5/25	2/25	1/25	2/25	1/25

(b) 15.84
17. (a)

k	2	3	4	5	6	7	8	9	10
p_k	1/25	1/25	2/25	3/25	4/25	2/25	6/25	3/25	3/25

(b) 6.84
19. (a) -0.4 **(b)** 1.0 **(c)** 1.4 **21.** $E(X) = -0.1$, $\text{var}(X) = 3.39$, s.d. $= \sqrt{3.39}$ **23. (a)** $E(X) = 5.5$ **(b)** $\text{var}(X) = 8.25$ **27. (a)** 0.1 **(b)** 0.5 **(c)** 0.4 **(d)** 0.2 **29. (a)** $E(X) = 0.75$, $E(Y) = 0.3$ **(b)** $E(X + Y) = 1.05$ **(c)** $\text{var}(X) = 1.7875$, $\text{var}(Y) = 3.01$ **(d)** $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) = 4.7975$ **31. (a)** Since $[X - E(X)]^2 \geq 0$, it follows that $E[X - E(X)]^2 \geq 0$; therefore, $\text{var}(X) \geq 0$. **(b)** Since $\text{var}(X) = E(X^2) - [E(X)]^2 \geq 0$, it follows that $E(X^2) \geq [E(X)]^2$. **33. (a)** $\binom{10}{5}(0.5)^{10}$ **(b)** $(0.5)^{10}\left[\binom{10}{8} + \binom{10}{9} + \binom{10}{10}\right]$ **(c)** $1 - (0.5)^{10}$
35. $P(X = k) = \binom{6}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{6-k}$, $k = 0, 1, 2, \ldots, 6$ **37.** 10/64
39. $\left(\frac{3}{5}\right)^3 + 3\left(\frac{2}{5}\right)\left(\frac{3}{5}\right)^2$ **41.** $(0.8)^{20}$ **43. (a)** 1 **(b)** $10 \cdot (0.9)^{10}$
45. 12.5 **47. (a)** $\frac{\binom{24}{6}\binom{12}{4}}{\binom{36}{10}}$ **(b)** $\binom{10}{6}\left(\frac{2}{3}\right)^6\left(\frac{1}{3}\right)^4$
49. $\frac{30!}{10!14!6!}(0.2)^{10}(0.35)^{14}(0.45)^6$
51. $\frac{40!}{20!10!8!2!}\left(\frac{9}{16}\right)^{20}\left(\frac{3}{16}\right)^{10}\left(\frac{3}{16}\right)^8\left(\frac{1}{16}\right)^2$ **53.** $\frac{6!}{2!2!2!}\left(\frac{6}{24}\right)^2\left(\frac{8}{24}\right)^2\left(\frac{10}{24}\right)^2$
55. $\frac{23!}{5!12!6!}\left(\frac{1}{4}\right)^5\left(\frac{1}{2}\right)^{12}\left(\frac{1}{4}\right)^6$ **57.** 5/16 **59.** $(3/4)^4$ **61.** 1/2, 1/4, 1/8
63. 1/8 **65.** 1/8 **67.** 15/16 **69.** $\left(\frac{14}{15}\right)^{19}$ **71.** $E(T) = 6$, $\text{var}(T) = 30$ **73. (a)** $\left(\frac{9}{10}\right)^6\frac{1}{10}$ **(b)** $\frac{9}{10}\frac{8}{9}\frac{7}{8}\frac{6}{7}\frac{5}{6}\frac{1}{5}$
75. (a) $(1 - p)^{k-1}p$ **(b)** $\binom{k-1}{1}p^2(1 - p)^{k-2}$
77.

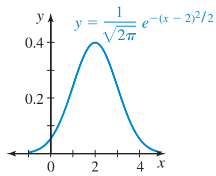
k	0	1	2	3
$P(X = k)$	e^{-2}	$2e^{-2}$	$2e^{-2}$	$\frac{4}{3}e^{-2}$

79. (a) $1 - 2e^{-1}$ **(b)** $e^{-1}\left(1 + \frac{1}{2} + \frac{1}{6}\right)$
81. $1 - e^{-1.5}\left[1 + 1.5 + \frac{(1.5)^2}{2} + \frac{(1.5)^3}{6}\right]$ **83.** $1 - 3e^{-2}$ **85.** e^{-7}
87. $1 - e^{-0.5}$ **89.** $1 - 4e^{-3}$ **91. (a)** $18e^{-6}$ **(b)** 1/4, 1/2, 1/4

93. $(2/3)^2$ assuming X and Y are independent
95. $P(X = 0) \approx e^{-0.5}$ **97.** **(a)** 0.5819 **(b)** 0.5820 **99.** $e^{-1.5}$

Section 12.5

1. $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)dx = 1$. Distribution function: $F(x) = 1 - e^{-3x}$ for $x \geq 0$, $F(x) = 0$ for $x \leq 0$ **3.** $c = \frac{1}{\pi}$
5. $E(X) = \frac{1}{2}$, $\text{var}(X) = \frac{1}{4}$ **7.** $E(X) = \frac{3}{2}$, $\text{var}(X) = \frac{3}{4}$
9. **(a)** If $a < 2$, $E(X) = \int_1^{\infty}(a-1)x^{-(a-1)}dx = \lim_{c\to\infty}\left[-x^{-(a-2)} \cdot \left(\frac{a-1}{a-2}\right)\right]_1^c = \infty$. If $a = 2$, $E(X) = \int_1^{\infty} x^{-1}dx = \infty$. **(b)** $E(X) = \frac{a-1}{a-2}$ **11.** **(d)**



13. 95%: (7.4, 18.2), 99%: (4.7, 20.9) **15.** 50% **17.** 99.5% **19.** 2.5%
21. **(a)** 0.6915 **(b)** 0.383 **(c)** 0.1587 **(d)** 0.0668 **23.** **(a)** $x = 3.56$ **(b)** $x = 1.5$ **(c)** $x = 0.5$ **(d)** $x = 1.34$ **25.** **(a)** 0.0228 **(b)** $x = 628$
27. 0.76% **29.** 0.8185 **33.** $E(|X|) = 2/\sqrt{2\pi}$ **35.** **(a)** 0.0735 **(b)** 0.6068 **(c)** 88 **(d)** 56 **37.** 0.1 **39.** $a = 1, b = 7$ **41.** $E(Y) = \frac{3}{4}$
43. $HHTHHTHHHT$ **45.** **(a)** $(1-x)^n$ **(b)** Use L'Hôpital's rule on $\lim_{n\to\infty} \ln(1 - \frac{x}{n})^n$. **47.** $E(X) = \frac{1}{\lambda}$ **49.** $e^{-4/3}$
51. **(a)** $e^{-20/27}$ **(b)** $e^{-20/27}$ **53.** **(a)** e^{-4} **(b)** $8e^{-4}$ **(c)** $32e^{-8}$ **(d)** $\frac{1}{4}$
55. 0.25 hour **57.** **(a)** $1 - e^{-2/3}$ **(b)** $P(N(5) = 1) = \frac{5}{3}e^{-5/3}$
59. **(a)** $1 - e^{-3/5}$ **(b)** $e^{-1/5}$ **61.** **(a)** 5 years **(b)** $\ln 2/0.2$ years
63. **(a)** $\exp\left[-(1.5 + 10e^{0.05} - 10)\right] \approx 0.134$. **(b)** $\exp\left[-(2.1 + 10e^{0.07} - 10)\right] - \exp\left[-(3 + 10e^{0.1} - 10)\right] \approx 0.042$.
65. Solution of $1.2x + (0.6)e^{0.5x} - 0.6 - \ln 2 = 0$ is approximately 0.451. **67.** **(a)** $\exp\left[-(2 \times 10^{-5})\frac{(50)^{2.5}}{2.5}\right] \approx 0.868$. **(b)** 0.1477
69. $x_m \approx 30.4$

Section 12.6

1. Exact probability: $e^{-3/2}$; Markov's inequality: $P(X \geq 3) \leq \frac{2}{3}$
5. Exact: $P(|X| \geq 1) = 1/2$; Chebyshev's inequality: $P(|X| \geq 1) \leq \frac{4}{3}$ **7.** $\frac{9}{25}$ **9.** $\frac{1}{n}\sum_{i=1}^n X_i$ converges to 0.9 as $n \to \infty$
11. Since $E(|X_i|) = \infty$, we cannot apply the law of large numbers as stated in Section 12.6. **13.** The sample size should be at least 380. **15.** 0.1711 **17.** **(a)** 0.0023 **(b)** 0.83
19. **(a)** 7.571×10^{-10} **(b)** 1.341×10^{-9} **(c)** 5.637×10^{-9}
21. **(a)** -11.2 **(b)** 0.579 **23.** 68 **25.** 385 **27.** **(a)** 0.3660 **(b)** 0.3679 **(c)** 0.243 **29.** **(a)** 0.1849 **(b)** 0.1755 **(c)** 0.1863
31. Likely not. **33.** **(a)** 0.6065, 0.3033, 0.0902 **(b)** 0.8397
35. 0.1429 **37.** 0.9515

Section 12.7

1. median: 15; sample mean: 16.2; sample variance: 180.2
3. median: 35.5; sample mean: 36; sample variance: 43.1
5. $\overline{X} = 11.93$; $S^2 = 3.389$ **7.** $\overline{X} = 5.69$; $S^2 = 3.465$
9.

Sample	*Sample Mean*
(1, 1)	1.0
(1, 6)	3.5
(1, 8)	4.5
(6, 1)	3.5
(6, 6)	6.0
(6, 8)	7.0
(8, 1)	4.5
(8, 6)	7.0
(8, 8)	8.0

15. **(a)** approximately normal with mean 1/3 and variance 1/450 **(b)** approximately normal **17.** **(c)** true values: $\mu = 0.5$, $\sigma^2 = \frac{1}{12}$
19. $\overline{X} = 16.2$; S.E. = 4.245 **21.** $[-0.3993, 0.5213]$
25. $\hat{p} = 0.72$; [0.651, 0.789] **27.** $y = 1.92x - 0.92$; $r^2 = 0.952$
31. $y = 0.207x + 0.488$; $r^2 = 0.841$

Chapter 12 Review Problems

1. **(a)** 0.431 **3.** 168,168,000 **5.** **(a)** $E(X) = 4.2$ **(b)** 0.0043 **(c)** 0.0215 **(d)** $1 - (0.9957)^5$ **7.** **(a)** 0.16 **(b)** 0.2 **9.** **(a)** $\mu = 170$, $\sigma = 5.102$ **(b)** 0.0795 **11.** 20% **13.** **(b)** $E(V) = \frac{n-1}{n}\sigma^2$

References

K. S. Albert and C. M. Gernaat (1984). Pharmacokinetics of ibuprofen. *The American Journal of Medicine* **77**: 40–46.

W. C. Allee (1931). *Animal Aggregations: A Study in General Sociology*. University of Chicago Press.

D. L. Altshuler and R. Dudley. (2003). Kinematics of hovering hummingbird flight along simulated and natural elevational gradients. *Journal of Experimental Biology* **206**: 3139–3147.

D. L. Altshuler, R. Dudley, S. M. Heredia, and J. A. McGuire (2010). Allometry of hummingbird lifting performance. *Journal of Experimental Biology* **213**: 725–734.

H. G. Andrewartha and L. C. Birch (1954). *The Distribution and Abundance of Animals*. University of Chicago Press.

T. E. Angelini, M. Roper, R. Kolter, D. A. Weitz, and M. P. Brenner (2009). *Bacillus subtilis* spreads by surfing on waves of surfactant. *Proceedings of the National Academy of Sciences* **106**: 18109–18113.

H. R. Babad, D. J. Nokes, N. J. Gay, E. Miller, P. Morgan-Capner, and R. M. Anderson (1995). Predicting the impact of measles vaccination in England and Wales: Model validation and analysis of policy options. *Epidemiology & Infection* **114**: 319–344.

N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler (2004). Bacterial persistence as a phenotypic switch. *Science* **305**: 1622–1625.

R. J. H. Beaverton and S. J. Holt. (1957). On the dynamics of exploited fish populations, Fishery Investigations Series II, Vol. XIX, Ministry of Agriculture. *Fisheries and Food* **1**: 957.

M. Begon and G. A. Parker (1986). Should egg size and clutch size decrease with age? *Oikos* **47**: 293–302.

T. S. Bellows, Jr. (1981). The descriptive properties of some models for density dependence. *Journal of Animal Ecology* **50**: 139–156.

M. J. Benton (1997). Models for the diversification of life. *TREE* **12**: 490–495.

M. J. Benton and D. Harper (1997). *Basic Paleontology*. Addison-Wesley and Longman.

W. H. Beyer, ed. (1991). *CRC Standard Probability and Statistics Tables and Formulae*. CRC Press.

W. J. Binder and B. V. Valkenburgh (2000). Bite strength and feeding in spotted hyenas (*Crocuta crocuta*). *Journal of Zoology, London* **252**: 273–283.

P. J. Bohlen, P. M. Groffman, C. T. Driscoll, T. J. Fahey, and T. G. Siccama (2001). Plant-soil-microbial interactions in a northern hardwood forest. *Ecology* **82**: 965–978.

R. Borchert (1994). Soil and stem water storage determine phenology and distribution of tropical dry forest trees. *Ecology* **75**: 1437–1449.

T. Boulinier, J. D. Nichols, and J. E. Hines (2001). Forest fragmentation and bird community dynamics: Inference at regional scales. *Ecology* **82**: 1159–1169.

G. E. Briggs and J. B. S. Haldane (1925). A note on the kinematics of enzyme actions. *Biochemical Journal* **19**: 338–339.

J. L. Brooks and S. I. Dodson (1965). Predation, body-size and composition of plankton. *Science* **150**: 28–35.

M. J. W. Burke and J. P. Grime (1996). An experimental study of plant community invasibility. *Ecology* **77**: 776–790.

M. L. Cain, S. W. Pacala, J. A. Silander, Jr., and M.-J. Fortin (1995). Neighborhood models of clonal growth in the white clover *Trifolium repens*. *American Naturalist* **145**: 888–917.

G. S. Campbell (1986). *An Introduction to Environmental Biophysics*. Springer.

T. Carlson (1913). Über geschwindigkeit und grösse der hefevermehrung in würze. *Biochemische Zeitschrift* **57**: 313–334.

P. Christiansen and J. S. Adolfssen (2005). Bite forces, canine strength and skull allometry in carnivores (Mammalia, Carnivora). *Journal of Zoology* **266**: 133–151.

H. E. Cline, K. Hynynen, C. J. Hardy, R. D. Watkins, J. F. Schenck, and F. A. Jolesz (1994). MR temperature mapping of focused ultrasound surgery. *Magnetic Resonance in Medicine* **31**: 628–636.

M. J. Crawley (1997). Plant–herbivore dynamics. In *Plant Ecology*, ed. M. J. Crawley, pp. 401–474. Blackwell Science.

J. W. Dalling, K. Winter, and J. D. Nason (2001). The unusual life history of *Alseis blackiana*: A shade-persistent pioneer tree? *Ecology* **82**: 933–945.

D. P. Dalzell (1944). On 22/7. *Journal of the London Mathematical Society* **19**: 133–134.

D. P. Dalzell (1971). On 22/7 and 355/113. *Eureka: The Archimedians' Journal* **34**: 1013.

C. Darwin (1859). *On the Origin of Species by Means of Natural Selection*. John Murray.

C. Darwin (1959). *The Origin of Species*. Reprinted by Penguin Classics (1985). Penguin Books.

R. Dawkins (2006). *The Selfish Gene*. Oxford University Press.

D. L. DeAngelis (1992). *Dynamics of Nutrient Cycling and Food Webs*. Chapman & Hall.

W. R. DeMott, R. D. Gulati, and K. Siewertsen (1998). Effects of phosphorus-deficient diets on the carbon and phosphorus balance of *Daphnia magna*. *Limnology and Oceanography* **43**: 1147–1161.

A. M. de Roos (1996). A gentle introduction to physiologically structured population models. In *Structured-Population Models in Marine, Terrestrial, and Freshwater Systems*, eds. S. Tuljapurkar and H. Caswell. Chapman & Hall.

J. DeWoody, C. A. Rowe, V. D. Hipkins, and K. E. Mock (2008). "Pando" lives: Molecular genetic evidence of a giant aspen clone in central Utah. *Western North American Naturalist* **68**: 493.

M. R. D'Orsogna, Y.-L. Chuang, A. L. Bertozzi, and L. S. Chayes (2006). Self-propelled particles with soft-core interactions: Patterns, stability, and collapse. *Physical Review Letters* **96**: 104302.

R. H. Dott, Jr., and R. L. Batten (1976). *Evolution of the Earth*, 2nd edition. McGraw-Hill.

J. E. Dowd and D. S. Riggs (1965). A comparison of estimates of Michaelis-Menten kinetic constants from various linear transformations. *Journal of Biological Chemistry* **240**: 863–869.

E. Dressaire, L. Yamada, B. Song, and M. Roper (2016). Mushrooms use convectively created airflows to disperse their spores. *Proceedings of the National Academy of Sciences* **113**: 2833–2838.

R. E. Druzinsky (1993). The time allometry of mammalian chewing movements: Chewing frequency scales with body mass in mammals. *Journal of Theoretical Biology* **160**: 427–440.

C. M. Duarte and S. Agustí (1998). The CO_2 balance of unproductive aquatic ecosystems. *Science* **281**: 234–236.

R. Durrett (2008). *Probability models for DNA sequence evolution*. Springer Science & Business Media.

P. J. Dyck, J. L. Davies, D. M.Wilson, F. J. Service, L. J. Melton III, and P. C. O'Brien (1999). Risk factors for severity of diabetic polyneuropathy. *Diabetes Care* **22**: 1479–1486.

G. P. Elliott, D. V. Merton, and P. W. Jansen (2001). Intensive management of a critically endangered species: The kakapo. *Biological Conservation* **99**: 121–133.

D. S. Falconer (1989). *Introduction to Quantitative Genetics*. Longman.

D. S. Falconer and T. F. C. Mackay (1996). *Introduction to Quantitative Genetics*, 4th edition. Longman.

B. A. Ferguson, T. A. Dresbach, C. G. Partes, G. M. Filip, and C. L. Schmitt (2003). Coarse-scale population structure of pathogenic *Armillaria* species in a mixed-conifer forest in the Blue Mountains of northeast Oregon. *Canadian Journal of Forest Research* **33**: 612–623.

D. Fisman, E. Khoo, and A. Tuite (2014). Early epidemic dynamics of the West African 2014 Ebola outbreak: Estimates derived with a simple two-parameter model. *PLOS Currents* **6**.

R. Fitzhugh (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophysics Journal* **1**: 445–466.

D. J. Futuyama (1995). *Science on Trial*, 2nd edition. Sinauer.

P. Gaastra (1959). Mededelinger van de Landbouwhogeschool Te Wageningen, The Netherlands. In *Plant Physiology* (1978), eds. F. B. Salisbury and C. W. Ross. Wadsworth.

G. F. Gause (1934). *The Struggle for Existence*. Reprinted by Hafner (1964). Williams & Wilkins.

L. A. Gavrilov and N. S. Gavrilova (2001). The reliability theory of aging and longevity. *Journal of Theoretical Biology* **213**: 527–545.

Z. M. Gliwicz (1990). Food threshold and body size in caldocerans. *Nature* **343**: 683–740.

M. A. F. Gomes, G. L.Vasconcelos, I. J. Tsang, and I. R. Tsang (1999). Scaling relations for diversity of languages. *Physica A: Statistical Mechanics and its Applications* **271**: 489–495.

V. D. Gooch, L. Freeman, and P. L. Lakin–Thomas (2004). Time-lapse analysis of the circadian rhythms of conidiation and growth rate in *Neurospora. Journal of Biological Rhythms* **19**: 493–503.

P. R. Grant (1982). Variation in the size and shape of Darwin's finch eggs. *Auk* **99**: 15–23.

P. R. Grant, I. Abbott, D. Schluter, R. L. Curry, and L. K. Abbott (1985). Variation in the size and shape of Darwin's finches. *Biological Journal of the Linnean Society* **25**: 1–39.

L. L. Greenhill, J. M. Swanson, K. Steinhoff, J. Fried, K. Posner, M. Lerner, ... and Tulloch, S. (2003). A pharmacokinetic/pharmacodynamic study comparing a single morning dose of Adderall to twice–daily dosing in children with ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry* **42**: 1234–1241.

K. J. Griffiths (1969). The importance of coincidence in the functional and numerical responses of two parasites of the European pine sawfly, *Neodiprion sertifer*. *Canadian Entomologist* **101**: 673–713.

J. B. S. Haldane (1957). The cost of natural selection. *Journal of Genetics* **55**: 511–524.

D. J. Hall (1964). An experimental approach to the dynamics of a natural population of *Daphnia galeata mendotae*. *Ecology* **45**: 94–111.

D. L. Hartl and A. G. Clark (1989). *Principles of Population Genetics*, 2nd edition. Sinauer.

A. Hartstone-Rose, J. M. Perry, and C. J. Morrow (2012). Bite force estimation and the fiber architecture of felid masticatory muscles. *The Anat. Rec.* **295**: 1336–1351.

R. W. Herschy (1995). *Streamflow Measurement*, 2nd edition. Taylor & Francis.

D. D. Ho, A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, M. Markowitz, et al. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**: 123–126.

T. Hojo and K. Pearson (1931). Distribution of the median, quartiles and interquartile distance in samples from a normal population. *Biometrika* **23**: 315–363.

A. A. Holazo, M. B. Winkler, and I. H. Patel (1988). Effects of age, gender and oral contraceptives on intramuscular midazolam pharmacokinetics. *The Journal of Clinical Pharmacology*, **28**: 1040–1045.

L. Holdridge, W. Grenke, W. Hatheway, T. Liang, and J. Tosi, Jr. (1971). *Forest Environments in Tropical Life Zones: A Pilot Study.* Pergamon Press.

C. S. Holling (1959). Some characteristics of simple types of predation and parasitism. *Canadian Entomologist* **91**: 385–398.

H. S. Horn (1971). The adaptive geometry of trees. *Monographs in Population Biology 3*. Princeton University Press.

W. Huang, R. T. Yen, M. McLaurine, and G. Bledsoe (1996). Morphometry of the human pulmonary vasculature. *Journal of Applied Physiology* **81**: 2123–2133.

S. B. Hulley and D. Grady (2004). The WHI estrogen-alone trial—do things look any better? *JAMA* **14**: 1769–1771.

C. E. Iselin, J. E. Robertson, and D. F. Paulson (1999). Radical perineal prostatectomy: Oncological outcome during a 20-year period. *Journal of Urology* **161**: 163–168.

Y. Iwasa, T. J. de Jong, and P. G. L. Klinkhamer (1995). Why pollinators visit only a fraction of the open flowers on a plant and its consequence for fitness curves of plants. *Journal of Evolutionary Biology* **8**: 439–453.

T. Y. James, J. Stenlid, A. Olsen, and H. Johannesson (2008). Evolutionary significance of imbalanced nuclear ratios within heterokaryons of the basidiomycete fungus *Heterobasidion parviporum. Evolution* **62**: 2279–2296.

T. H. Jukes and C. R. Cantor (1969). Evolution of protein molecules. In *Mammalian Protein Metabolism III*, ed. H. N. Munro, pp. 21–132. Academic Press.

P. M. Kareiva (1983). Local movement in herbivorous insects: Applying a passive diffusion model to mark-recapture field experiments. *Oecologia* **57**: 322–327.

M. J. Keeling and B. T. Grenfell (1997). Disease extinction and community size: Modeling the persistence of measles. *Science* **275**: 65–67.

D. J. Kempf, J. F. Denissen, E. McDonald, S. Vasavanonda, C. A. Flentge, B. E. Green, et al. (1995). ABT–538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. *Proceedings of the National Academy of Science USA* **92**: 2484–2498.

W. O. Kermack and A. G. McKendrick (1927). Contributions to the mathematical theory of epidemics. *Proceedings of the Royal Society A* **115**: 700–721.

W. O. Kermack and A. G. McKendrick (1932). Contributions to the mathematical theory of epidemics. II. The problem of endemicity. *Proceedings of the Royal Society A* **138**: 55–83.

W. O. Kermack and A. G. McKendrick (1933). Contributions to the mathematical theory of epidemics. III. Further studies of the problem of endemicity. *Proceedings of the Royal Society A* **141**: 94–122.

B. C. Kirkup and M. A. Riley (2004). Antibiotic-mediated antagonism leads to a bacterial game of rock-paper-scissor in vivo. *Nature* **428**: 412.

T. Kolokotrones, V. Savage, E. J. Deeds, and W. Fontana (2010). Curvature in metabolic scaling. *Nature* **464**: 753–756.

K. B. Krauskopf and D. K. Bird (1995). *Introduction to Geochemistry*, p. 268. McGraw-Hill.

A. K. Laird (1964). Dynamics of tumour growth. *British Journal of Cancer* **18**: 490.

H. Lamb (1911). XV. On the uniform motion of a sphere through a viscous fluid. *Philosphical Magazine* **21**: 112—121.

H. E. Landsberg (1969). *Weather and Health, an Introduction to Biometerology*. Doubleday.

Y. al-Lanqawi, T. A. Moreland, J. McEwen, F. Halliday, C. J. Durnin, and I. H. Stevenson (1992). Ethanol kinetics: Extent of error in back extrapolation procedures. *British Journal of Clinical Pharmacology* **34**: 316–321.

T. Lauritzen, S. Pramming, T. Deckert, and C. Binder (1983). Pharmacokinetics of continuous subcutaneous insulin infusion. *Diabetologia* **24**: 326–329.

A. C. Leopold and P. E. Kriedemann (1975). *Plant Growth and Development*. McGraw-Hill.

P. Leslie (1945). On the use of matrices in certain population mathematics. *Biometrika* **33**: 183–212.

R. Levins (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America* **15**: 237–240.

R. Levins (1970). Community equilibria and stability, and an extension of the competitive exclusion principle. *American Naturalist* **104**: 413–423.

D. C. Lloyd (1987). Selection of offspring size at independence and other size-versus-number strategies. *American Naturalist* **129**: 800–817.

N. K. Logothetis, J. Pauls, M. Augath, T. Trinath, and A. Oeltermann (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**: 150–157.

A. J. Lotka (1932). The growth of mixed populations: Two species competing for a common food supply. *Journal of the Washington Academy of Sciences* **22**: 461–469.

R. H. MacArthur and E. O. Wilson (1963). An equilibrium theory of insular zoogeography. *Evolution* **17**: 373–387.

T. F. C. Mackey (1984) Jumping genes meet abdominal bristles: Hybrid dysgenesis-induced quantitative variation in *Drosophila melanogaster*. *Genetics Research* **44**: 231–237.

C. H. Martin and P. C. Wainwright (2013). Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science* **339**: 208–211.

R. M. May (1975). *Stability and Complexity in Model Ecosystems*. Princeton University Press.

R. M. May (1976). Simple mathematical models with very complicated dynamics. *Nature* **261**: 459–461.

R. M. May (1978). Host–parasitoid systems in patchy environments: A phenomenological model. *Journal of Animal Ecology*, **47**: 833–843.

A. A. van der May and R. G. Bout (2004). Scaling of jaw muscle size and maximal bite force in finches. *Journal of Experimental Biology* **207**: 2745–2753.

K. A. McCullogh, J. S. Sperry, and F. R. Adler (2003). Water transport in plants obeys Murray's law. *Nature* **421**: 939.

J. J. McGough, J. Biederman, L. L. Greenhill, J. T. McCracken, T. J. Spencer, K. Posner, …and James M. Swanson (2003). Pharmacokinetics of SLI381 (Adderall XR), an extended–release formulation of Adderall. *Journal of the American Academy of Child & Adolescent Psychiatry* **42**: 684–691.

L. Michaelis and M. I. Menten (1913). Die Kinetik der Invertinwirkung. *Biochemische Zeitschrift* **49**: 333–369.

R. Milo, J. H. Hou, Michael Springer, M. P. Brenner, and M. W. Kirschner (2007). The relationship between evolutionary and physiological variation in hemoglobin. *Proceedings of the National Academy of Sciences USA* **104**: 16998–17003.

G. W. Molnar, H. J. Hurley, Jr., and R. Ford (1969). Application of Newton's law to body cooling. *Pflügers Archiv* **311**: 16–24.

J. L. Monod (1942). *Recherches sur la Croissance des Cultures Bacteriennes*. Hermann.

J. L. Monod (1949). The growth of bacterial cultures. *Annual Reviews in Microbiology* **3**: 371–394.

J. L. Monod (1950). La technique de culture continué: Théorie et applications. *Annales de l'Institut Pasteur* **79**: 390–410.

B. Moss (1980). *Ecology of Freshwaters*. Blackwell Scientific.

J. S. Nagumo, S. Arimoto, and S. Yoshizawa (1962). An active pulse transmission line simulating nerve axon. *Proceedings of the Institute of Radio Engineers* **50**: 2061–2071.

K. H. Nealson, T. Platt, and J. W. Hastings (1970). Cellular control of the synthesis and activity of the bacterial luminescent system *Journal of Bacteriology* **104**: 313–322.

S. Nee and R. M. May (1992). Dynamics of metapopulation: Habitat destruction and competitive coexistence. *Journal of Animal Ecology* **61**: 37–40.

A. J. Nicholson (1933). The balance of animal populations. *Journal of Animal Ecology* **2**: 131–178.

A. J. Nicholson and V. A. Bailey (1935). The balance of animal populations. *Proceedings of the Zoological Society of London* **3**: 551–598.

K. J. Niklas (1994). *Plant Allometry: The Scaling of Form and Process*. University of Chicago Press.

C. G. Nilsson, P. L. Lahteenmaki, T. Luukkainen, and D. N. Robertson (1986). Sustained intrauterine release of levonorgestrel over five years. *Fertility and Sterility* **45**: 805–807.

R. T. Oglesby (1977). Relationships of fish yield to lake phytoplankton, standing crop, production, and morphoedaphic factors. *Journal of the Fisheries Reserve Board of Canada* **34**: 2271–2279.

M. A. Nowak (2006). *Evolutionary Dynamics*. Harvard University Press.

R. Osczevski and M. Bluestein (2005). The new wind chill equivalent temperature chart. *Bulletin of the American Meteorological Society* **86**: 1453–1458.

J. G. Owen (1988). On productivity as a predictor of rodent and carnivore density. *Ecology* **69**: 1161–1165.

S. W. Pacala and M. Rees (1998). Models suggesting field experiments to test two hypotheses explaining successional diversity. *American Naturalist* **152**: 729–737.

R. Pearl and L. J. Reed (1920). On the rate of growth of the population of the United States since 1790 and its mathematical representation. *Proceedings of the National Academy of Sciences* **6**: 275–288.

G. W. Pierce (1949). *The Songs of Insects*. Harvard University Press.

A. Pisek, W. Larcher, W. Moser, and I. Pack (1969). Kardinale Temperaturbereiche der Photosynthese und Grenztemperaturendes Lebens der Blätter verschiedener Spermatophyten. III. Temperaturabhängigkeit und optimaler Temperaturbereichder Netto-Photosynthese. *Flora, Jena* **158**: 608–630.

S. D. Pletcher (1998). Mutation and the evolution of age-specific mortality rates: Experimental results and statistical developments. Ph.D. thesis, University of Minnesota.

L. F. Prescott (1980). Kinetics and metabolism of paracetamol and phenacetin. *British Journal of Clinical Pharmacology* **10**.

F. W. Preston (1962). The canonical distribution of commonness and rarity. *Ecology* **43**: 185–215, 410–431.

M. D. Rawlins, D. B. Henderson, and A. R. Hijab (1977). Pharmacokinetics of paracetamol (acetaminophen) after intravenous and oral administration. *European Journal of Clinical Pharmacology* **11**: 283–286.

P. M. Reis, S. Jung, J. M. Aristoff, and R. Stocker (2010). How cats lap: Water uptake by *Felis catus. Science* **330**: 1231–1234.

M. J. Reiss (1989). *The Allometry of Growth and Reproduction*. Cambridge University Press.

M. M. Robbins, M. Gray, E. Kagoda, and A. M. Robbins (2009). Population dynamics of the Bwindi mountain gorillas. *Biological Conservation* **142**: 2886–2895.

D. A. Roff (1992). *The Evolution of Life Histories*. Chapman & Hall.

M. L. Rosenzweig (1971). Paradox of enrichment: Destabilization of exploitation ecosystems in ecological time. *Science* **171**: 385–387.

M. L. Rosenzweig and Z. Abramsky (1993). How are diversity and productivity related? In *Species Diversity in Ecological Communities*, eds. R. E. Ricklefs and D. Schluter, pp. 52–65. University of Chicago Press.

J. Roughgarden (1996). *Theory of Population Genetics and Evolutionary Ecology*, 2nd edition. Prentice Hall.

M. D. Rowling, D. B. Henderson, and A. R. Hijab (1977). Pharmacokinetics of paracetamol (acetaminophen) after intravenous oral administration. *European Journal of Clinical Pharmacology* **11**: 283–286.

P. Royston and E. M. Wright (1998). How to construct "normal ranges" for fetal variables. *Ultrasound in Obstetrics & Gynecology* **11**: 30–38.

J. B. Rundle, D. L. Turcotte, R. Shcherbakov, W. Klein, and C. Sammis (2003). Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. *Reviews of Geophysics* **41**.

V. M. Savage, J. F. Gillooly, W. H. Woodruff, G. B. West, A. P. Allen, B. J. Enquist, and J. H. Brown (2004). The predominance of quarter-power scaling in biology. *Functional Ecology* **18:** 257–282.

F. E. Smith (1963). Population dynamics in *Daphnia magna* and a new model for population growth. *Ecology* **44**: 651–663.

J. S. Sperry, D. D. Smith, V. M. Savage, B. J. Enquist, K. A. McCulloh, P. B. Reich, . . . and E. I. Allmen (2012). A species-level model for metabolic scaling in trees I. Exploring boundaries to scaling space within and across species. *Functional Ecology* **26**: 1054–1065.

S. M. Stanley (1979). *Macroevolution: Pattern and Process*, p. 332. W. H. Freeman.

R. G. Steadman (1979). The assessment of sultriness. Part I: A temperature–humidity index based on human physiology and clothing science. *Journal of Applied Meteorology* **18**: 861–873.

R.W. Sterner (1997). Modeling interactions between food quality and quantity in homeostatic consumers. *Freshwater Biology* **38**: 473–482.

R.W. Sterner and J. J. Elser (2002). *Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere*. Princeton University Press.

A. Strandburg-Peshkin, C. R. Twomey, N. W. F. Bode, A. B. Kao, Y. Katz, C. C. Ioannou, et al. (2013). Visual sensory networks and effective information transfer in animal groups. *Current Biology* **23**: R709–R711.

S. H. Strogatz (2014). *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Westview Press.

H. Tanaka, K. D. Monahan, and D. R. Seals (2001). Age–predicted maximal heart rate revisited. *Journal of the American College of Cardiology* **37**: 153–156.

C. R. Taylor, G. M. O. Maloiy, E. R.Weibel, U. A. Langman, J. M. Z. Kaman, H. J. Seeherman, and B. N. C. Heglund (1980). Design of the mammalian respiratory system. III. Scaling maximum aerobic capacity to body mass: Wild and domestic mammals. *Respiration Physiology* **44**: 25–37.

W. R. Thompson (1924). La théorie mathématique de l'action des parasites entomophages et le facteur du hasard. *Annales de la Faculté des Sciences de Marseille* **2**: 69–89.

D. Tilman (1982). *Resource Competition and Community Structure*. Princeton University Press.

D. Tilman (1994). Competition and biodiversity in spatially structured habitats. *Ecology* **75**: 2–16.

M. M. Tilzer, W. Geller, U. Sommer, and H. H. Stable (1982). Kohlenstoffkreislauf und Nahrungsketten in der Freiwasserzone des Bodensees. *Konstanzer Blätter für Hochschulfragen* **73**: 51.

U.S. Bureau of the Census (1994). *Statistical Abstract of the United States*, 14th edition. Government Printing Office.

J. W. Valentine (1985). Biotic diversity and clade diversity. In *Phanerozoic Diversity Patterns: Profiles in Macroevolution*, ed. J. W. Valentine, pp. 419–424. Princeton University Press.

B. V. Valkenburgh (1990). Skeletal and dental predictions of body mass in carnivores. In *Body Size in Mammalian Paleobiology Estimation and Biological Implications,* ed. J. P. Damuth and B. MacFadden. Cambridge University Press.

B. O. Verburg, E. A. P. Steegers, M. De Ridder, R. J. M. Snijders, E. Smith, Albert Hofman, . . . and J. C. M. Witteman

(2008). New charts for ultrasound dating of pregnancy and assessment of fetal growth: Longitudinal data from a population-based cohort study. *Ultrasound in Obstetrics & Gynecology* **31**: 388–396.

E. Vieth (1989). Fitting piecewise linear regression functions to biological responses. *Journal of Applied Physiology* **67**: 390–396.

P. M. Vitousek and H. Farrington (1997). Nutrient limitation and soil development: Experimental test of a biogeochemical theory. *Biogeochemistry* **37**: 63–75.

S. Vogel (2004). Living in a physical world. *Journal of Biosciences* **29**: 391–397.

V. Volterra (1926). Fluctuations in the abundance of species considered mathematically. *Nature* **118**: 558–560.

T. D. Walker (1985). Diversification functions and the rate of taxonomic evolution. In *Phanerozoic Diversity Patterns: Profiles in Macroevolution*, ed. J. W. Valentine, pp. 311–334. Princeton University Press.

T. W. Walker and J. K. Syers (1976). The fate of phosphorus during pedogenesis. *Geoderma* **15**: 1–19.

T. D. Walker and J. W. Valentine (1984). Equilibrium models of evolutionary species diversity and the number of empty niches. *American Naturalist* **124**: 887–899.

G.-J. Wang, N. D. Volkow, J. Logan, N. R. Pappas, C. T. Wong, ... and J. S. Fowler (2001). Brain dopamine and obesity. *The Lancet* **357**: 354–357.

P. D. Ward (1992). *On Methuselah's Trail—Living Fossils and the Great Extinctions*. Freeman.

M. P. Wiedeman (1963). Dimensions of blood vessels from distributing artery to collecting vein. *Circulation Research* **12**: 375–378.

Women's Health Initiative Steering Committee (2004). Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: The Women's Health Initiative Randomized Controlled Trial. *JAMA* **291** (14), 1701–1712.

S. N. C. Woolley, D. P. Tittensor, P. K. Dunstan, G. Guillera-Arroita, J. J. Lahoz-Monfort, B. A. Wintle, ... and T. D. O'Hara. Deep-sea diversity patterns are shaped by energy availability. *Nature* **533**: 393–410.

S. Wright (1968). *Evolution of Populations, Vol. 1. Genetics and Biometrik Foundation.* University of Chicago Press.

Q. Yan, C. J. Van Der Gast, and Y. Yu (2012). Bacterial community assembly and turnover within the intestines of developing zebrafish. *PlOS One* **7**:e30603.

P. J. Yang, J. Pham, J. Choo, and D. L. Hu (2014). Duration of urination does not change with body size. *Proceedings of the National Academy of Science USA* **111**: 11932–11937.

E. K. Yeargers, R.W. Shonkwiler, and J. V. Herod (1996). *An Introduction to the Mathematics of Biology: With Computer Algebra Models*. Birkhäuser.

M. Zamir (1977). Shear forces and blood vessel radii in the cardiovascular system. *The Journal of General Physiology* **69**: 449–461.

Y. L. Zhang, M. L. Han, J. Vidyalakshmi, C. Y. Shee, and W. T. Ang (2009). Automatic control of mechanical forces acting on cell biomembranes using a vision-guided microrobotic system in computer microscopy. *Journal of Microscopy* **236**: 70–78.

Index

This page intentionally left blank

DERIVATIVES AND INTEGRALS

Basic Differentiation Rules

1. $\frac{d}{dx}[cu] = cu'$

2. $\frac{d}{dx}[u \pm v] = u' \pm v'$

3. $\frac{d}{dx}[uv] = uv' + vu'$

4. $\frac{d}{dx}\left[\frac{u}{v}\right] = \frac{vu' - uv'}{v^2}$

5. $\frac{d}{dx}[c] = 0$

6. $\frac{d}{dx}[x^n] = nx^{n-1}$

7. $\frac{d}{dx}[\ln x] = \frac{1}{x}$

8. $\frac{d}{dx}[e^x] = e^x$

9. $\frac{d}{dx}\log_a x = \frac{1}{x \ln a}$

10. $\frac{d}{dx}a^x = (\ln a)a^x$

11. $\frac{d}{dx}[\sin x] = \cos x$

12. $\frac{d}{dx}[\cos x] = -\sin x$

13. $\frac{d}{dx}[\tan x] = \sec^2 x$

14. $\frac{d}{dx}[\cot x] = -\csc^2 x$

15. $\frac{d}{dx}[\sec x] = \sec x \tan x$

16. $\frac{d}{dx}[\csc x] = -\csc x \cot x$

17. $\frac{d}{dx}[\arcsin x] = \frac{1}{\sqrt{1-x^2}}$

18. $\frac{d}{dx}[\arccos x] = -\frac{1}{\sqrt{1-x^2}}$

19. $\frac{d}{dx}[\arctan x] = \frac{1}{1+x^2}$

20. $\frac{d}{dx}[\operatorname{arccot} x] = -\frac{1}{1+x^2}$

Basic Integration Formulas

1. $\int kf(x)\,dx = k\int f(x)\,dx$

2. $\int [f(x) \pm g(x)]\,dx = \int f(x)\,dx \pm \int g(x)\,dx$

3. $\int dx = x + C$

4. $\int x^n\,dx = \frac{x^{n+1}}{n+1} + C, n \neq -1$

5. $\int \frac{dx}{x} = \ln|x| + C$

6. $\int e^x\,dx = e^x + C$

7. $\int \sin x\,dx = -\cos x + C$

8. $\int \cos x\,dx = \sin x + C$

9. $\int \tan x\,dx = -\ln|\cos x| + C$

10. $\int \cot x\,dx = \ln|\sin x| + C$

11. $\int \sec x\,dx = \ln|\sec x + \tan x| + C$

12. $\int \csc x\,dx = -\ln|\csc x + \cot x| + C$

13. $\int \sec^2 x\,dx = \tan x + C$

14. $\int \csc^2 x\,dx = -\cot x + C$

15. $\int \sec x \tan x\,dx = \sec x + C$

16. $\int \csc x \cot x\,dx = -\csc x + C$

17. $\int \frac{dx}{\sqrt{a^2 - x^2}} = \arcsin\frac{x}{a} + C$

18. $\int \frac{dx}{a^2 + x^2} = \frac{1}{a}\arctan\frac{x}{a} + C$

19. $\int \frac{dx}{x\sqrt{x^2 - a^2}} = \frac{1}{a}\operatorname{arcsec}\frac{|x|}{a} + C$

Algebra

Quadratic Formula
The solutions of the quadratic equation $ax^2 + bx + c = 0$ are given by

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Factorial notation
For each positive integer n,

$$n! = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1$$

By definition, $0! = 1$.

Radicals

$$\sqrt[n]{x^m} = \left(\sqrt[n]{x}\right)^m = x^{m/n}$$

Exponents

$$(ab)^r = a^r b^r \qquad a^r a^s = a^{r+s} \qquad x^{-n} = \frac{1}{x^n}$$

$$(a^r)^s = a^{rs} \qquad \frac{a^r}{a^s} = a^{r-s}$$

Binomial Formula

$$(x+y)^2 = x^2 + 2xy + y^2$$

$$(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$$

$$(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$$

In general,

$$(x+y)^n = x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{k}x^{n-k}y^k + \cdots + \binom{n}{n-1}xy^{n-1} + y^n$$

where the binomial coefficient $\binom{n}{m}$ is the integer $\frac{n!}{m!(n-m)!}$.

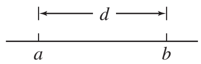
Special Factors

$x^2 - a^2 = (x-a)(x+a)$
$x^3 + a^3 = (x+a)(x^2 - ax + a^2)$
$x^3 - a^3 = (x-a)(x^2 + ax + a^2)$
$x^4 - a^4 = (x^2 - a^2)(x^2 + a^2)$

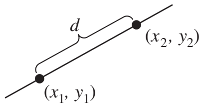
Geometry

Distance Formulas

Distance on the real number line: $d = |a - b|$

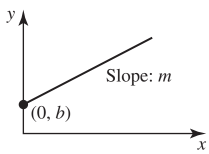


Distance in the coordinate plane:
$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$

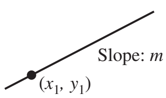


Equations of Lines and Circles

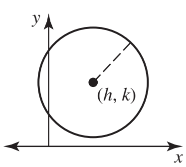
Slope–intercept equation: $y = mx + b$



Point–slope equation: $y - y_1 = m(x - x_1)$

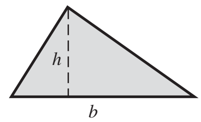


Circle with center (h, k) and radius r:
$(x - h)^2 + (y - k)^2 = r^2$

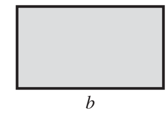


Areas and Volumes

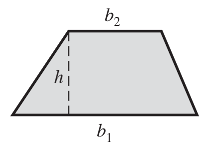
Triangle area: $A = \frac{1}{2}bh$



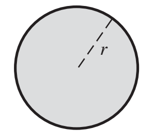
Rectangle area: $A = bh$



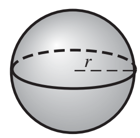
Trapezoid area: $A = \dfrac{b_1 + b_2}{2}h$



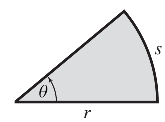
Circle area: $A = \pi r^2$
Circumference: $C = 2\pi r$



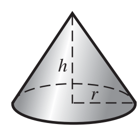
Sphere volume: $V = \frac{4}{3}\pi r^3$
Surface area: $A = 4\pi r^2$



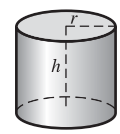
Sector of circle (θ in radians):
$A = \dfrac{\theta r^2}{2}$; $s = r\theta$



Cone volume: $V = \frac{1}{3}\pi r^2 h$
Curved surface area: $A = \pi r\sqrt{r^2 + h^2}$



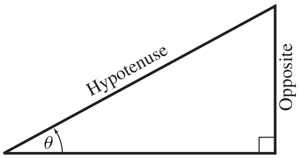
Cylinder volume: $V = \pi r^2 h$
Curved surface area: $A = 2\pi rh$



Trigonometry

Definition of the Six Trigonometric Functions

Right triangle definitions, where $0 < \theta < \pi/2$

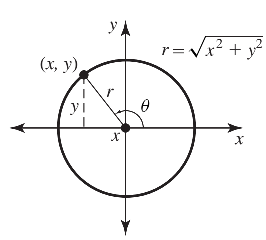


$$\sin\theta = \frac{\text{opp.}}{\text{hyp.}} \qquad \csc\theta = \frac{\text{hyp.}}{\text{opp.}}$$

$$\cos\theta = \frac{\text{adj.}}{\text{hyp.}} \qquad \sec\theta = \frac{\text{hyp.}}{\text{adj.}}$$

$$\tan\theta = \frac{\text{opp.}}{\text{adj.}} \qquad \cot\theta = \frac{\text{adj.}}{\text{opp.}}$$

Circular function definitions, where θ *is any angle*



$$\sin\theta = \frac{y}{r} \qquad \csc\theta = \frac{r}{y}$$

$$\cos\theta = \frac{x}{r} \qquad \sec\theta = \frac{r}{x}$$

$$\tan\theta = \frac{y}{x} \qquad \cot\theta = \frac{x}{y}$$