Springer
New York
Berlin
Heidelberg
Barcelona
Hong Kong
London
Milan
Paris
Singapore
Tokyo

CHAOS

An Introduction to Dynamical Systems

KATHLEEN T. ALLIGOOD
George Mason University

TIM D. SAUER
George Mason University

JAMES A. YORKE
University of Maryland

Springer

Textbooks in Mathematical Sciences

Series Editors:

Thomas F. Banchoff
Brown University

Keith Devlin
St. Mary's College

Gaston Gonnet
ETH Zentrum, Zürich

Jerrold Marsden
California Institute of Technology

Stan Wagon
Macalester College

COVER: Rene Magritte, *Golconde 1953*. © 1996 C. Herscovici, Brussels/Artists Rights Society (ARS), New York. Used by permission of ARS.

Library of Congress Cataloging-in-Publication Data
Alligood, Kathleen T.
Chaos - an introduction to dynamical systems / Kathleen Alligood, Tim Sauer, James A. Yorke.
p. cm. — (Textbooks in mathematical sciences)
Includes bibliographical references and index.
1. Differentiable dynamical systems. 2. Chaotic behavior in systems. I. Sauer, Tim. II. Yorke, James A. III. Title. IV. Series.
QA614.8.A44 1996
003'.85—dc20 95-51304
CIP

Printed on acid-free paper.

Production managed by Frank Ganz; manufacturing supervised by Jeffrey Taub.
Photocomposed by Integre Technical Publishing Co., Inc., Albuquerque, NM.
Printed and bound by R.R. Donnelley & Sons, Harrisonburg, VA.
Printed in the United States of America.

9 8 7 6 5 4 3 (Corrected third printing, 2000)

ISBN 0-387-94677-2 SPIN 10778875

Springer-Verlag New York Berlin Heidelberg
A member of BertelsmannSpringer Science+Business Media GmbH

Introduction

Background

Sir Isaac Newton brought to the world the idea of modeling the motion of physical systems with equations. It was necessary to invent calculus along the way, since fundamental equations of motion involve velocities and accelerations, which are derivatives of position. His greatest single success was his discovery that the motion of the planets and moons of the solar system resulted from a single fundamental source: the gravitational attraction of the bodies. He demonstrated that the observed motion of the planets could be explained by assuming that there is a gravitational attraction between any two objects, a force that is proportional to the product of masses and inversely proportional to the square of the distance between them. The circular, elliptical, and parabolic orbits of astronomy were

no longer fundamental determinants of motion, but were approximations of laws specified with differential equations. His methods are now used in modeling motion and change in all areas of science.

Subsequent generations of scientists extended the method of using differential equations to describe how physical systems evolve. But the method had a limitation. While the differential equations were sufficient to determine the behavior—in the sense that solutions of the equations did exist—it was frequently difficult to figure out what that behavior would be. It was often impossible to write down solutions in relatively simple algebraic expressions using a finite number of terms. Series solutions involving infinite sums often would not converge beyond some finite time.

When solutions could be found, they described very regular motion. Generations of young scientists learned the sciences from textbooks filled with examples of differential equations with regular solutions. If the solutions remained in a bounded region of space, they settled down to either (A) a steady state, often due to energy loss by friction, or (B) an oscillation that was either periodic or quasiperiodic, akin to the clocklike motion of the moon and planets. (In the solar system, there were obviously many different periods. The moon traveled around the earth in a month, the earth around the sun in about a year, and Jupiter around the sun in about 11.867 years. Such systems with multiple incommensurable periods came to be called quasiperiodic.)

Scientists knew of systems which had more complicated behavior, such as a pot of boiling water, or the molecules of air colliding in a room. However, since these systems were composed of an immense number of interacting particles, the complexity of their motions was not held to be surprising.

Around 1975, after three centuries of study, scientists in large numbers around the world suddenly became aware that there is a third kind of motion, a type (C) motion, that we now call "chaos". The new motion is erratic, but not simply quasiperiodic with a large number of periods, and not necessarily due to a large number of interacting particles. It is a type of behavior that is possible in very simple systems.

A small number of mathematicians and physicists were familiar with the existence of a third type of motion prior to this time. James Clerk Maxwell, who studied the motion of gas molecules in about 1860, was probably aware that even a system composed of two colliding gas particles in a box would have neither motion type A nor B, and that the long term behavior of the motions would for all practical purposes be unpredictable. He was aware that very small changes in the initial motion of the particles would result in immense changes in the trajectories of the molecules, even if they were thought of as hard spheres.

Maxwell began his famous study of gas laws by investigating individual collisions. Consider two atoms of equal mass, modeled as hard spheres. Give the atoms equal but opposite velocities, and assume that their positions are selected at random in a large three-dimensional region of space. Maxwell showed that if they collide, all directions of travel will be equally likely after the collision. He recognized that small changes in initial positions can result in large changes in outcomes. In a discussion of free will, he suggested that it would be impossible to test whether a leopard has free will, because one could never compute from a study of its atoms what the leopard would do. But the chaos of its atoms is limited, for, as he observed, "No leopard can change its spots!"

Henri Poincaré in 1890 studied highly simplified solar systems of three bodies and concluded that the motions were sometimes incredibly complicated. (See Chapter 2). His techniques were applicable to a wide variety of physical systems. Important further contributions were made by Birkhoff, Cartwright and Littlewood, Levinson, Kolmogorov and his students, among others. By the 1960s, there were groups of mathematicians, particularly in Berkeley and in Moscow, striving to understand this third kind of motion that we now call chaos. But only with the advent of personal computers, with screens capable of displaying graphics, have scientists and engineers been able to see that important equations in their own specialties had such solutions, at least for some ranges of parameters that appear in the equations.

In the present day, scientists realize that chaotic behavior can be observed in experiments and in computer models of behavior from all fields of science. The key requirement is that the system involve a nonlinearity. It is now common for experiments whose previous anomalous behavior was attributed to experiment error or noise to be reevaluated for an explanation in these new terms. Taken together, these new terms form a set of unifying principles, often called dynamical systems theory, that cross many disciplinary boundaries.

The theory of dynamical systems describes phenomena that are common to physical and biological systems throughout science. It has benefited greatly from the collision of ideas from mathematics and these sciences. The goal of scientists and applied mathematicians is to find nature's unifying ideas or laws and to fashion a language to describe these ideas. It is critical to the advancement of science that exacting standards are applied to what is meant by knowledge. Beautiful theories can be appreciated for their own sake, but science is a severe taskmaster. Intriguing ideas are often rejected or ignored because they do not meet the standards of what is knowledge.

The standards of mathematicians and scientists are rather different. Mathematicians prove theorems. Scientists look at realistic models. Their approaches are

somewhat incompatible. The first papers showing chaotic behavior in computer studies of very simple models were distasteful to both groups. The mathematicians feared that nothing was proved so nothing was learned. Scientists said that models without physical quantities like charge, mass, energy, or acceleration could not be relevant to physical studies. But further reflection led to a change in viewpoints. Mathematicians found that these computer studies could lead to new ideas that slowly yielded new theorems. Scientists found that computer studies of much more complicated models yielded behaviors similar to those of the simplistic models, and that perhaps the simpler models captured the key phenomena.

Finally, laboratory experiments began to be carried out that showed unequivocal evidence of unusual nonlinear effects and chaotic behavior in very familiar settings. The new dynamical systems concepts showed up in macroscopic systems such as fluids, common electronic circuits and low-energy lasers that were previously thought to be fairly well understood using the classical paradigms. In this sense, the chaotic revolution is quite different than that of relativity, which shows its effects at high energies and velocities, and quantum theory, whose effects are submicroscopic. Many demonstrations of chaotic behavior in experiments are not far from the reader's experience.

In this book we study this field that is the uncomfortable interface between mathematics and science. We will look at many pictures produced by computers and we try to make mathematical sense of them. For example, a computer study of the driven pendulum in Chapter 2 reveals irregular, persistent, complex behavior for ten million oscillations. Does this behavior persist for one billion oscillations? The only way we can find out is to continue the computer study longer. However, even if it continues its complex behavior throughout our computer study, we cannot guarantee it would persist forever. Perhaps it stops abruptly after one trillion oscillations; we do not know for certain. We can prove that there exist initial positions and velocities of the pendulum that yield complex behavior forever, but these choices are conceivably quite atypical. There are even simpler models where we know that such chaotic behavior does persist forever. In this world, pictures with uncertain messages remain the medium of inspiration.

There is a philosophy of modeling in which we study idealized systems that have properties that can be closely approximated by physical systems. The experimentalist takes the view that only quantities that can be measured have meaning. Yet we can prove that there are beautiful structures that are so infinitely intricate that they can never be seen experimentally. For example, we will see immediately in Chapters 1 and 2 the way chaos develops as a physical parameter like friction is varied. We see infinitely many periodic attractors appearing with infinitely many periods. This topic is revisited in Chapter 12, where we show

how this rich bifurcation structure, called a cascade, exists with mathematical certainty in many systems. This is a mathematical reality that underlies what the experimentalist can see. We know that as the scientist finds ways to make the study of a physical system increasingly tractable, more of this mathematical structure will be revealed. It is there, but often hidden from view by the noise of the universe. All science is of course dependent on simplistic models. If we study a vibrating beam, we will generally not model the atoms of which it is made. If we model the atoms, we will probably not reflect in our model the fact that the universe has a finite age and that the beam did not exist for all time. And we do not include in our model (usually) the tidal effects of the stars and the planets on our vibrating beam. We ignore all these effects so that we can isolate the implications of a very limited list of concepts.

It is our goal to give an introduction to some of the most intriguing ideas in dynamics, the ideas we love most. Just as chemistry has its elements and physics has its elementary particles, dynamics has its fundamental elements: with names like attractors, basins, saddles, homoclinic points, cascades, and horseshoes. The ideas in this field are not transparent. As a reader, your ability to work with these ideas will come from your own effort. We will consider our job to be accomplished if we can help you learn what to look for in your own studies of dynamical systems of the world and universe.

About the Book

As we developed the drafts of this book, we taught six one semester classes at George Mason University and the University of Maryland. The level is aimed at undergraduates and beginning graduate students. Typically, we have used parts of Chapters 1–9 as the core of such a course, spending roughly equal amounts of time on iterated maps (Chapters 1–6) and differential equations (Chapters 7–9). Some of the maps we use as examples in the early chapters come from differential equations, so that their importance in the subject is stressed. The topics of stable manifolds, bifurcations, and cascades are introduced in the first two chapters and then developed more fully in the Chapters 10, 11, and 12, respectively. Chapter 13 on time series may be profitably read immediately after Chapter 4 on fractals, although the concepts of periodic orbit (of a differential equation) and chaotic attractor will not yet have been formally defined.

The impetus for advances in dynamical systems has come from many sources: mathematics, theoretical science, computer simulation, and experimen-

tal science. We have tried to put this book together in a way that would reflect its wide range of influences.

We present elaborate dissections of the proofs of three deep and important theorems: The Poincaré-Bendixson Theorem, the Stable Manifold Theorem, and the Cascade Theorem. Our hope is that including them in this form tempts you to work through the nitty-gritty details, toward mastery of the building blocks as well as an appreciation of the completed edifice.

Additionally, each chapter contains a special feature called a Challenge, in which other famous ideas from dynamics have been divided into a number of steps with helpful hints. The Challenges tackle subjects from period-three implies chaos, the cat map, and Sharkovskii's ordering through synchronization and renormalization. We apologize in advance for the hints we have given, when they are of no help or even mislead you; for one person's hint can be another's distraction.

The Computer Experiments are designed to present you with opportunities to explore dynamics through computer simulation, the venue through which many of these concepts were first discovered. In each, you are asked to design and carry out a calculation relevant to an aspect of the dynamics. Virtually all can be successfully approached with a minimal knowledge of some scientific programming language. Appendix B provides an introduction to the solution of differential equations by approximate means, which is necessary for some of the later Computer Experiments.

If you prefer not to work the Computer Experiments from scratch, your task can be greatly simplified by using existing software. Several packages are available. *Dynamics: Numerical Explorations* by H.E. Nusse and J.A. Yorke (Springer-Verlag 1994) is the result of programs developed at the University of Maryland. *Dynamics*, which includes software for Unix and PC environments, was used to make many of the pictures in this book. The web site for *Dynamics* is `www.ipst.umd.edu/dynamics`. We can also recommend *Differential and Difference Equations through Computer Experiments* by H. Kocak (Springer-Verlag, 1989) for personal computers. A sophisticated package designed for Unix platforms is *dstool*, developed by J. Guckenheimer and his group at Cornell University. In the absence of special purpose software, general purpose scientific computing environments such as Matlab, Maple, and Mathematica will do nicely.

The Lab Visits are short reports on carefully selected laboratory experiments that show how the mathematical concepts of dynamical systems manifest themselves in real phenomena. We try to impart some flavor of the setting of the experiment and the considerable expertise and care necessary to tease a new secret from nature. In virtually every case, the experimenters' findings far surpassed what we survey in the Lab Visit. We urge you to pursue more accurate and detailed discussions of these experiments by going straight to the original sources.

ACKNOWLEDGEMENTS

In the course of writing this book, we received valuable feedback from colleagues and students too numerous to mention. Suggestions that led to major improvements in the text were made by Clark Robinson, Eric Kostelich, Ittai Kan, Karen Brucks, Miguel San Juan, and Brian Hunt, and from students Leon Poon, Joe Miller, Rolando Castro, Guocheng Yuan, Reena Freedman, Peter Calabrese, Michael Roberts, Shawn Hatch, Joshua Tempkin, Tamara Gibson, Barry Peratt, and Ed Fine.

We offer special thanks to Tamer Abu-Elfadl, Peggy Beck, Marty Golubitsky, Eric Luft, Tom Power, Mike Roberts, Steve Schiff, Myong-Hee Sung, Bill Tongue, Serge Troubetzkoy, and especially Mark Wimbush for pointing out errors in the first printing.

Kathleen T. Alligood
Tim D. Sauer
Fairfax, VA

James A. Yorke
College Park, MD

Contents

INTRODUCTION v

1 ONE-DIMENSIONAL MAPS 1

1.1 One-Dimensional Maps 2
1.2 Cobweb Plot: Graphical Representation of an Orbit 5
1.3 Stability of Fixed Points 9
1.4 Periodic Points 13
1.5 The Family of Logistic Maps 17
1.6 The Logistic Map $G(x) = 4x(1 - x)$ 22
1.7 Sensitive Dependence on Initial Conditions 25
1.8 Itineraries 27
CHALLENGE 1: PERIOD THREE IMPLIES CHAOS 32
EXERCISES 36
LAB VISIT 1: BOOM, BUST, AND CHAOS IN THE BEETLE CENSUS 39

2 TWO-DIMENSIONAL MAPS 43
2.1 Mathematical Models 44
2.2 Sinks, Sources, and Saddles 58
2.3 Linear Maps 62
2.4 Coordinate Changes 67
2.5 Nonlinear Maps and the Jacobian Matrix 68
2.6 Stable and Unstable Manifolds 78
2.7 Matrix Times Circle Equals Ellipse 87
CHALLENGE 2: COUNTING THE PERIODIC ORBITS OF LINEAR MAPS ON A TORUS 92
EXERCISES 98
LAB VISIT 2: IS THE SOLAR SYSTEM STABLE? 99

3 CHAOS 105
3.1 Lyapunov Exponents 106
3.2 Chaotic Orbits 109
3.3 Conjugacy and the Logistic Map 114
3.4 Transition Graphs and Fixed Points 124
3.5 Basins of Attraction 129
CHALLENGE 3: SHARKOVSKII'S THEOREM 135
EXERCISES 140
LAB VISIT 3: PERIODICITY AND CHAOS IN A CHEMICAL REACTION 143

4 FRACTALS 149
4.1 Cantor Sets 150
4.2 Probabilistic Constructions of Fractals 156
4.3 Fractals from Deterministic Systems 161
4.4 Fractal Basin Boundaries 164
4.5 Fractal Dimension 172
4.6 Computing the Box-Counting Dimension 177
4.7 Correlation Dimension 180
CHALLENGE 4: FRACTAL BASIN BOUNDARIES AND THE UNCERTAINTY EXPONENT 183
EXERCISES 186
LAB VISIT 4: FRACTAL DIMENSION IN EXPERIMENTS 188

5 CHAOS IN TWO-DIMENSIONAL MAPS 193
5.1 Lyapunov Exponents 194
5.2 Numerical Calculation of Lyapunov Exponents 199
5.3 Lyapunov Dimension 203
5.4 A Two-Dimensional Fixed-Point Theorem 207
5.5 Markov Partitions 212
5.6 The Horseshoe Map 216
CHALLENGE 5: COMPUTER CALCULATIONS AND SHADOWING 222
EXERCISES 226
LAB VISIT 5: CHAOS IN SIMPLE MECHANICAL DEVICES 228

6 CHAOTIC ATTRACTORS 231
6.1 Forward Limit Sets 233
6.2 Chaotic Attractors 238
6.3 Chaotic Attractors of Expanding Interval Maps 245
6.4 Measure 249
6.5 Natural Measure 253
6.6 Invariant Measure for One-Dimensional Maps 256
CHALLENGE 6: INVARIANT MEASURE FOR THE LOGISTIC MAP 264
EXERCISES 266
LAB VISIT 6: FRACTAL SCUM 267

7 DIFFERENTIAL EQUATIONS 273
7.1 One-Dimensional Linear Differential Equations 275
7.2 One-Dimensional Nonlinear Differential Equations 278
7.3 Linear Differential Equations in More than One Dimension 284
7.4 Nonlinear Systems 294
7.5 Motion in a Potential Field 300
7.6 Lyapunov Functions 304
7.7 Lotka-Volterra Models 309
CHALLENGE 7: A LIMIT CYCLE IN THE VAN DER POL SYSTEM 316
EXERCISES 321
LAB VISIT 7: FLY VS. FLY 325

8 PERIODIC ORBITS AND LIMIT SETS 329
8.1 Limit Sets for Planar Differential Equations 331
8.2 Properties of ω-Limit Sets 337
8.3 Proof of the Poincaré-Bendixson Theorem 341
CHALLENGE 8: TWO INCOMMENSURATE FREQUENCIES FORM A TORUS 350
EXERCISES 353
LAB VISIT 8: STEADY STATES AND PERIODICITY IN A SQUID NEURON 355

9 CHAOS IN DIFFERENTIAL EQUATIONS 359
9.1 The Lorenz Attractor 359
9.2 Stability in the Large, Instability in the Small 366
9.3 The Rössler Attractor 370
9.4 Chua's Circuit 375
9.5 Forced Oscillators 376
9.6 Lyapunov Exponents in Flows 379

CHALLENGE 9: SYNCHRONIZATION OF CHAOTIC ORBITS 387
EXERCISES 393
LAB VISIT 9: LASERS IN SYNCHRONIZATION 394

10 STABLE MANIFOLDS AND CRISES 399
10.1 The Stable Manifold Theorem 401
10.2 Homoclinic and Heteroclinic Points 409
10.3 Crises 413
10.4 Proof of the Stable Manifold Theorem 422
10.5 Stable and Unstable Manifolds for Higher Dimensional Maps 430
CHALLENGE 10: THE LAKES OF WADA 432
EXERCISES 440
LAB VISIT 10: THE LEAKY FAUCET: MINOR IRRITATION OR CRISIS? 441

11 BIFURCATIONS 447
11.1 Saddle-Node and Period-Doubling Bifurcations 448
11.2 Bifurcation Diagrams 453
11.3 Continuability 460
11.4 Bifurcations of One-Dimensional Maps 464
11.5 Bifurcations in Plane Maps: Area-Contracting Case 468
11.6 Bifurcations in Plane Maps: Area-Preserving Case 471
11.7 Bifurcations in Differential Equations 478
11.8 Hopf Bifurcations 483
CHALLENGE 11: HAMILTONIAN SYSTEMS AND THE LYAPUNOV CENTER THEOREM 491
EXERCISES 494
LAB VISIT 11: IRON + SULFURIC ACID ⟶ HOPF BIFURCATION 496

12 CASCADES 499
12.1 Cascades and 4.669201609. . . 500
12.2 Schematic Bifurcation Diagrams 504
12.3 Generic Bifurcations 510
12.4 The Cascade Theorem 518
CHALLENGE 12: UNIVERSALITY IN BIFURCATION DIAGRAMS 525
EXERCISES 531
LAB VISIT 12: EXPERIMENTAL CASCADES 532

13 STATE RECONSTRUCTION FROM DATA 537
13.1 Delay Plots from Time Series 537
13.2 Delay Coordinates 541
13.3 Embedology 545
CHALLENGE 13: BOX-COUNTING DIMENSION AND INTERSECTION 553

A MATRIX ALGEBRA 557
A.1 Eigenvalues and Eigenvectors 557
A.2 Coordinate Changes 561
A.3 Matrix Times Circle Equals Ellipse 563

B COMPUTER SOLUTION OF ODES 567
B.1 ODE Solvers 568
B.2 Error in Numerical Integration 570
B.3 Adaptive Step-Size Methods 574

ANSWERS AND HINTS TO SELECTED EXERCISES 577

BIBLIOGRAPHY 587

INDEX 595

CHAPTER ONE

One-Dimensional Maps

THE FUNCTION $f(x) = 2x$ is a rule that assigns to each number x a number twice as large. This is a simple mathematical model. We might imagine that x denotes the population of bacteria in a laboratory culture and that $f(x)$ denotes the population one hour later. Then the rule expresses the fact that the population doubles every hour. If the culture has an initial population of 10,000 bacteria, then after one hour there will be $f(10{,}000) = 20{,}000$ bacteria, after two hours there will be $f(f(10{,}000)) = 40{,}000$ bacteria, and so on.

A **dynamical system** consists of a set of possible states, together with a rule that determines the present state in terms of past states. In the previous paragraph, we discussed a simple dynamical system whose states are population levels, that change with time under the rule $x_n = f(x_{n-1}) = 2x_{n-1}$. Here the variable n stands for time, and x_n designates the population at time n. We will

require that the rule be **deterministic**, which means that we can determine the present state (population, for example) *uniquely* from the past states.

No randomness is allowed in our definition of a deterministic dynamical system. A possible mathematical model for the price of gold as a function of time would be to predict today's price to be yesterday's price plus or minus one dollar, with the two possibilities equally likely. Instead of a dynamical system, this model would be called a **random**, or **stochastic**, **process**. A typical realization of such a model could be achieved by flipping a fair coin each day to determine the new price. This type of model is not deterministic, and is ruled out by our definition of dynamical system.

We will emphasize two types of dynamical systems. If the rule is applied at discrete times, it is called a **discrete-time** dynamical system. A discrete-time system takes the current state as input and updates the situation by producing a new state as output. By the state of the system, we mean whatever information is needed so that the rule may be applied. In the first example above, the state is the population size. The rule replaces the current population with the population one hour later. We will spend most of Chapter 1 examining discrete-time systems, also called maps.

The other important type of dynamical system is essentially the limit of discrete systems with smaller and smaller updating times. The governing rule in that case becomes a set of differential equations, and the term **continuous-time** dynamical system is sometimes used. Many of the phenomena we want to explain are easier to describe and understand in the context of maps; however, since the time of Newton the scientific view has been that nature has arranged itself to be most easily modeled by differential equations. After studying discrete systems thoroughly, we will turn to continuous systems in Chapter 7.

1.1 ONE-DIMENSIONAL MAPS

One of the goals of science is to predict how a system will evolve as time progresses. In our first example, the population evolves by a single rule. The output of the rule is used as the input value for the next hour, and the same rule of doubling is applied again. The evolution of this dynamical process is reflected by composition of the function f. Define $f^2(x) = f(f(x))$ and in general, define $f^k(x)$ to be the result of applying the function f to the initial state k times. Given an initial value of x, we want to know about $f^k(x)$ for large k. For the above example, it is clear that if the initial value of x is greater than zero, the population will grow without bound. This type of expansion, in which the population is multiplied by

a constant factor per unit of time, is called **exponential growth**. The factor in this example is 2.

WHY STUDY MODELS?

We study models because they suggest how real-world processes behave. In this chapter we study extremely simple models.

Every model of a physical process is at best an idealization. The goal of a model is to capture some feature of the physical process. The feature we want to capture now is the patterns of points on an orbit. In particular, we will find that the patterns are sometimes simple, and sometimes quite complicated, or "chaotic", even for simple maps.

The question to ask about a model is whether the behavior it exhibits is *because* of its simplifications or if it captures the behavior *despite* the simplifications. Modeling reality too closely may result in an intractable model about which little can be learned. Model building is an art. Here we try to get a handle on possible behaviors of maps by considering the simplest ones.

The fact that real habitats have finite resources lies in opposition to the concept of exponential population increase. From the time of Malthus (Malthus, 1798), the fact that there are limits to growth has been well appreciated. Population growth corresponding to multiplication by a constant factor cannot continue forever. At some point the resources of the environment will become compromised by the increased population, and the growth will slow to something less than exponential.

In other words, although the rule $f(x) = 2x$ may be correct for a certain range of populations, it may lose its applicability in other ranges. An improved model, to be used for a resource-limited population, might be given by $g(x) = 2x(1 - x)$, where x is measured in millions. In this model, the initial population of 10,000 corresponds to $x = .01$ million. When the population x is small, the factor $(1 - x)$ is close to one, and $g(x)$ closely resembles the doubling function $f(x)$. On the other hand, if the population x is far from zero, then $g(x)$ is no longer proportional to the population x but to the product of x and the "remaining space" $(1 - x)$. This is

a nonlinear effect, and the model given by $g(x)$ is an example of a **logistic** growth model.

Using a calculator, investigate the difference in outcomes imposed by the models $f(x)$ and $g(x)$. Start with a small value, say $x = 0.01$, and compute $f^k(x)$ and $g^k(x)$ for successive values of k. The results for the models are shown in Table 1.1. One can see that for $g(x)$, there is computational evidence that the population approaches an eventual limiting size, which we would call a steady-state population for the model $g(x)$. Later in this section, using some elementary calculus, we'll see how to verify this conjecture (Theorem 1.5).

There are obvious differences between the behavior of the population size under the two models, $f(x)$ and $g(x)$. Under the dynamical system $f(x)$, the starting population size $x = 0.01$ results in arbitrarily large populations as time progresses. Under the system $g(x)$, the same starting size $x = 0.01$ progresses in a strikingly similar way at first, approximately doubling each hour. Eventually, however, a limiting size is reached. In this case, the population saturates at $x = 0.50$ (one-half million), and then never changes again.

So one great improvement of the logistic model $g(x)$ is that populations can have a finite limit. But there is a second improvement contained in $g(x)$. If

n	$f^n(x)$	$g^n(x)$
0	0.0100000000	0.0100000000
1	0.0200000000	0.0198000000
2	0.0400000000	0.0388159200
3	0.0800000000	0.0746184887
4	0.1600000000	0.1381011397
5	0.3200000000	0.2380584298
6	0.6400000000	0.3627732276
7	1.2800000000	0.4623376259
8	2.5600000000	0.4971630912
9	5.1200000000	0.4999839039
10	10.2400000000	0.4999999995
11	20.4800000000	0.5000000000
12	40.9600000000	0.5000000000

Table 1.1 Comparison of exponential growth model $f(x) = 2x$ to logistic growth model $g(x) = 2x(1 - x)$.
The exponential model explodes, while the logistic model approaches a steady state.

we use starting populations other than $x = 0.01$, the same limiting population $x = 0.50$ will be achieved.

⇨ COMPUTER EXPERIMENT 1.1

Confirm the fact that populations evolving under the rule $g(x) = 2x(1 - x)$ prefer to reach the population 0.5. Use a calculator or computer program, and try starting populations x_0 between 0.0 and 1.0. Calculate $x_1 = g(x_0), x_2 = g(x_1)$, etc. and allow the population to reach a limiting size. You will find that the size $x = 0.50$ eventually "attracts" any of these starting populations.

Our numerical experiment suggests that this population model has a natural built-in carrying capacity. This property corresponds to one of the many ways that scientists believe populations should behave—that they reach a steady-state which is somehow compatible with the available environmental resources. The limiting population $x = 0.50$ for the logistic model is an example of a fixed point of a discrete-time dynamical system.

Definition 1.1 A function whose domain (input) space and range (output) space are the same will be called a **map**. Let x be a point and let f be a map. The **orbit** of x under f is the set of points $\{x, f(x), f^2(x), \ldots\}$. The starting point x for the orbit is called the **initial value** of the orbit. A point p is a **fixed point** of the map f if $f(p) = p$.

For example, the function $g(x) = 2x(1 - x)$ from the real line to itself is a map. The orbit of $x = 0.01$ under g is $\{0.01, 0.0198, 0.0388, \ldots\}$, and the fixed points of g are $x = 0$ and $x = 1/2$.

1.2 COBWEB PLOT: GRAPHICAL REPRESENTATION OF AN ORBIT

For a map of the real line, a rough plot of an orbit—called a cobweb plot—can be made using the following graphical technique. Sketch the graph of the function f together with the diagonal line $y = x$. In Figure 1.1, the example $f(x) = 2x$ and the diagonal are sketched. The first thing that is clear from such a picture is the location of fixed points of f. At any intersection of $y = f(x)$ with the line $y = x$,

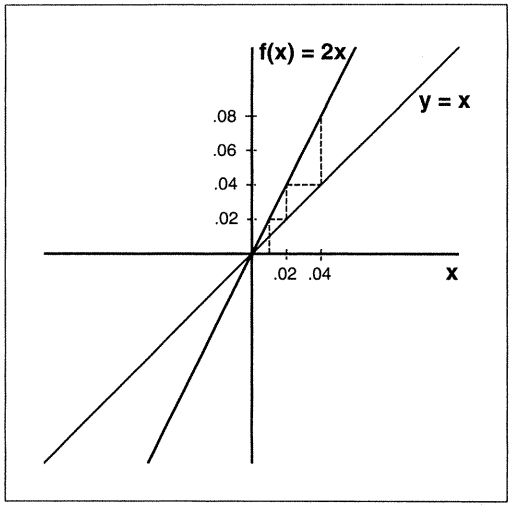

Figure 1.1 An orbit of $f(x) = 2x$.
The dotted line is a cobweb plot, a path that illustrates the production of a trajectory.

the input value x and the output $f(x)$ are identical, so such an x is a fixed point. Figure 1.1 shows that the only fixed point of $f(x) = 2x$ is $x = 0$.

Sketching the orbit of a given initial condition is done as follows. Starting with the input value $x = .01$, the output $f(.01)$ is found by plotting the value of the function above .01. In Figure 1.1, the output value is .02. Next, to find $f(.02)$, it is necessary to consider .02 as the new input value. In order to turn an output value into an input value, draw a horizontal line from the input–output pair $(.01, .02)$ to the diagonal line $y = x$. In Figure 1.1, there is a vertical dotted line segment starting at $x = .01$, representing the function evaluation, and then a horizontal dotted segment which effectively turns the output into an input so that the process can be repeated.

Then start over with the new value $x = .02$, and draw a new pair of vertical and horizontal dotted segments. We find $f(f(.01)) = f(.02) = .04$ on the graph of f, and move horizontally to move output to the input position. Continuing in this way, a graphical history of the orbit $\{.01, .02, .04, \ldots\}$ is constructed by the path of dotted line segments.

EXAMPLE 1.2

A more interesting example is the map $g(x) = 2x(1 - x)$. First we find fixed points by solving the equation $x = 2x(1 - x)$. There are two solutions, $x = 0$

COBWEB PLOT

A cobweb plot illustrates convergence to an attracting fixed point of $g(x) = 2x(1 - x)$. Let $x_0 = 0.1$ be the initial condition. Then the first iterate is $x_1 = g(x_0) = 0.18$. Note that the point (x_0, x_1) lies on the function graph, and (x_1, x_1) lies on the diagonal line. Connect these points with a horizontal dotted line to make a path. Then find $x_2 = g(x_1) = 0.2952$, and continue the path with a vertical dotted line to (x_1, x_2) and with a horizontal dotted line to (x_2, x_2). An entire orbit can be mapped out this way.

In this case it is clear from the geometry that the orbit we are following will converge to the intersection of the curve and the diagonal, $x = 1/2$. What happens if instead we start with $x_0 = 0.8$? These are examples of simple cobweb plots. They can be much more complicated, as we shall see later.

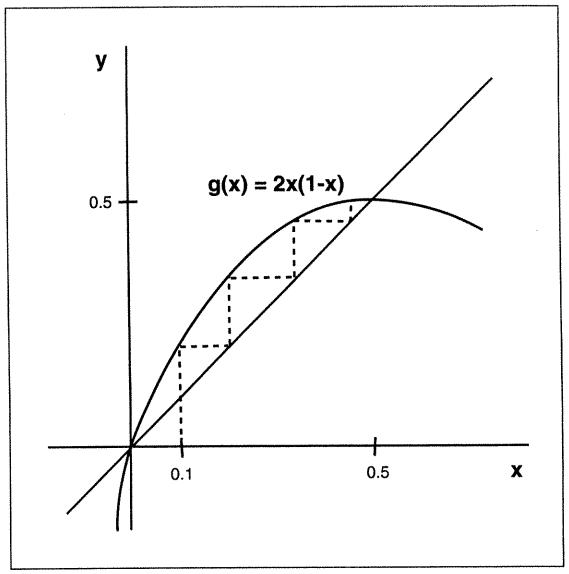

Figure 1.2 A cobweb plot for an orbit of $g(x) = 2x(1 - x)$.
The orbit with initial value .1 converges to the sink at .5.

and $x = 1/2$, which are the two fixed points of g. Contrast this with a linear map which, except for the case of the identity $f(x) = x$, has only one fixed point $x = 0$. What is the behavior of orbits of g? The graphical representation of the orbit with initial value $x = 0.1$ is drawn in Figure 1.2. It is clear from the figure that the orbit, instead of diverging to infinity as in Figure 1.1, is converging to the fixed point $x = 1/2$. Thus the orbit with initial condition $x = 0.1$ gets stuck, and cannot move beyond the fixed point $x = 0.5$. A simple rule of thumb for following the graphical representation of an orbit: If the graph is above the diagonal line $y = x$, the orbit will move to the right; if the graph is below the line, the orbit moves to the left.

EXAMPLE 1.3

Let f be the map of $\mathbb{R}$ given by $f(x) = (3x - x^3)/2$. Figure 1.3 shows a graphical representations of two orbits, with initial values $x = 1.6$ and 1.8, respectively. The former orbit appears to converge to the fixed point $x = 1$ as the map is iterated; the latter converges to the fixed point $x = -1$.

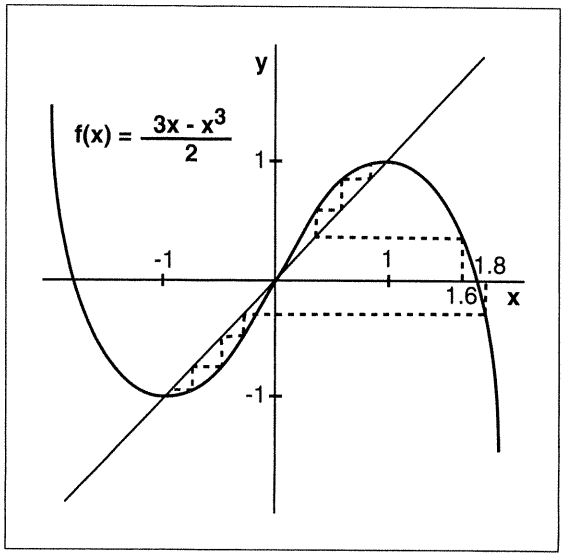

Figure 1.3 A cobweb plot for two orbits of $f(x) = (3x - x^3)/2$.
The orbit with initial value 1.6 converges to the sink at 1; the orbit with initial value 1.8 converges to the sink at -1.

Fixed points are found by solving the equation $f(x) = x$. The map has three fixed points, namely -1, 0, and 1. However, orbits beginning near, but not precisely on, each of the fixed points act differently. You may be able to convince yourself, using the graphical representation technique, that initial values near -1 stay near -1 upon iteration by the map, and that initial values near 1 stay near 1. On the other hand, initial values near 0 depart from the area near 0. For example, to four significant digits, $f(.1) = 0.1495, f^2(.1) = 0.2226, f^5(.1) = 0.6587$, and so on. The problem with points near 0 is that f magnifies them by a factor larger than one. For example, the point $x = .1$ is moved by f to approximately .1495, a magnification factor of 1.495. This magnification factor turns out to be approximately the derivative $f'(0) = 1.5$.

1.3 STABILITY OF FIXED POINTS

With the geometric intuition gained from Figures 1.1, 1.2, and 1.3, we can describe the idea of stability of fixed points. Assuming that the discrete-time system exists to model real phenomena, not all fixed points are alike. A stable fixed point has the property that points near it are moved even closer to the fixed point under the dynamical system. For an unstable fixed point, nearby points move away as time progresses. A good analogy is that a ball at the bottom of a valley is stable, while a ball balanced at the tip of a mountain is unstable.

The question of stability is significant because a real-world system is constantly subject to small perturbations. Therefore a steady state observed in a realistic system must correspond to a stable fixed point. If the fixed point is unstable, small errors or perturbations in the state would cause the orbit to move away from the fixed point, which would then not be observed.

Example 1.3 gave some insight into the question of stability. The derivative of the map at a fixed point p is a measure of how the distance between p and a nearby point is magnified or shrunk by f. That is, the points 0 and .1 begin exactly .1 units apart. After applying the rule f to both points, the distance separating the points is changed by a factor of approximately $f'(0)$. We want to call the fixed point 0 "unstable" when points very near 0 tend to move away from 0.

The concept of "near" is made precise by referring to all real numbers within a distance ϵ of p as the **epsilon neighborhood** $N_\epsilon(p)$. Denote the real line by $\mathbb{R}$. Then $N_\epsilon(p)$ is the interval of numbers $\{x \in \mathbb{R} : |x - p| < \epsilon\}$. We usually think of ϵ as a small, positive number.

Definition 1.4 Let f be a map on $\mathbb{R}$ and let p be a real number such that $f(p) = p$. If all points sufficiently close to p are attracted to p, then p is called a

sink or an **attracting fixed point**. More precisely, if there is an $\epsilon > 0$ such that for all x in the epsilon neighborhood $N_\epsilon(p)$, $\lim_{k\to\infty} f^k(x) = p$, then p is a sink. If all points sufficiently close to p are repelled from p, then p is called a **source** or a **repelling fixed point**. More precisely, if there is an epsilon neighborhood $N_\epsilon(p)$ such that each x in $N_\epsilon(p)$ except for p itself eventually maps outside of $N_\epsilon(p)$, then p is a source.

In this text, unless otherwise stated, we will deal with functions for which derivatives of all orders exist and are continuous functions. We will call this type of function a **smooth** function.

Theorem 1.5 *Let f be a (smooth) map on $\mathbb{R}$, and assume that p is a fixed point of f.*

1. *If $|f'(p)| < 1$, then p is a sink.*
2. *If $|f'(p)| > 1$, then p is a source.*

Proof: PART 1. Let a be any number between $|f'(p)|$ and 1; for example, a could be chosen to be $(1 + |f'(p)|)/2$. Since

$$\lim_{x\to p} \frac{|f(x) - f(p)|}{|x - p|} = |f'(p)|,$$

there is a neighborhood $N_\varepsilon(p)$ for some $\varepsilon > 0$ so that

$$\frac{|f(x) - f(p)|}{|x - p|} < a$$

for x in $N_\varepsilon(p)$, $x \neq p$.

In other words, $f(x)$ is closer to p than x is, by at least a factor of a (which is less than 1). This implies two things: First, if $x \in N_\epsilon(p)$, then $f(x) \in N_\varepsilon(p)$; that means that if x is within ε of p, then so is $f(x)$, and by repeating the argument, so are $f^2(x), f^3(x)$, and so forth. Second, it follows that

$$|f^k(x) - p| \leq a^k|x - p| \tag{1.1}$$

for all $k \geq 1$. Thus p is a sink.

✎ EXERCISE T1.1

Show that inequality (1.1) holds for $k = 2$. Then carry out the mathematical induction argument to show that it holds for all $k = 1, 2, 3, \ldots$

✎ EXERCISE T1.2

Use the ideas of the proof of Part 1 of Theorem 1.5 to prove Part 2.

Note that the proof of part 1 of Theorem 1.5 says something about the rate of convergence of $f^k(x)$ to p. The fact that $|f^k(x) - p| \leq a^k|x - p|$ for $a < 1$ is described by saying that $f^k(x)$ converges exponentially to p as $k \to \infty$.

Our definition of a fixed-point sink requires only that the sink attract some epsilon neighborhood $(p - \epsilon, p + \epsilon)$ of nearby points. As far as the definition is concerned, the radius ϵ of the neighborhood, although nonzero, could be extremely small. On the other hand, sinks often attract orbits from a large set of nearby initial conditions. We will refer to the set of initial conditions whose orbits converge to the sink as the basin of the sink.

With Theorem 1.5 and our new terminology, we can return to Example 1.2, an example of a logistic model. Setting $x = g(x) = 2x(1 - x)$ shows that the fixed points are 0 and $1/2$. Taking derivatives, we get $g'(0) = 2$ and $g'(1/2) = 0$. Theorem 1.5 shows that $x = 1/2$ is a sink, which confirms our suspicions from Table 1.1. On the other hand, $x = 0$ is a source. Points near 0 are repelled from 0 upon application of g. In fact, points near 0 are repelled at an exponential magnification factor of approximately 2 (check this number with a calculator). These points are attracted to the sink $x = 1/2$.

What is the basin of the sink $x = 1/2$ in Example 1.2? The point 0 does not belong, since it is a fixed point. Also, 1 does not belong, since $g(1) = 0$ and further iterations cannot budge it. However, all initial conditions from the interval $(0, 1)$ will produce orbits that converge to the sink. You should sketch a graph of $g(x)$ as in Figure 1.1 and use the idea of the cobweb plot to convince yourself of this fact.

There is a second way to show that the basin of $x = 1/2$ is $(0, 1)$, which is quicker and trickier but far less general. That is to use algebra (not geometry) to compare $|g(x) - 1/2|$ to $|x - 1/2|$. If the former is smaller than the latter, it means the orbit is getting closer to $1/2$. The algebra says:

$$\begin{aligned} |g(x) - 1/2| &= |2x(1 - x) - 1/2| \\ &= 2|x - 1/2||x - 1/2| \end{aligned} \tag{1.2}$$

Now we can see that if $x \in (0, 1)$, the multiplier $2|x - 1/2|$ is smaller than one. Any point x in $(0, 1)$ will have its distance from $x = 1/2$ decreased on each iteration by g. Notice that the algebra also tells us what happens for initial conditions outside of $(0, 1)$: they will never converge to the sink $x = 1/2$. Therefore the

basin of the sink is exactly the open interval $(0, 1)$. Informally, we could also say that the basin of infinity is $(-\infty, 0) \cup (1, \infty)$, since the orbit of each initial condition in this set tends toward $-\infty$.

Theorem 1.5 also clarifies Example 1.3, which is the map $f(x) = (3x - x^3)/2$. The fixed points are -1, 0, and 1, and the derivatives are $f'(-1) = f'(1) = 0$, and $f'(0) = 1.5$. By the theorem, the fixed points -1 and 1 are attracting fixed points, and 0 is a repelling fixed point.

Let's try to determine the basins of the two sinks. Example 1.3 is already significantly more complicated than Example 1.2, and we will have to be satisfied with an incomplete answer. We will consider the sink $x = 1$; the other sink has very similar properties by the symmetry of the situation.

First, cobweb plots (see Figure 1.3) convince us that the interval $I_1 = (0, \sqrt{3})$ of initial conditions belongs to the basin of $x = 1$. (Note that $f(\sqrt{3}) = f(-\sqrt{3}) = 0$.) So far it is similar to the previous example. Have we found the entire basin? Not quite. Initial conditions from the interval $I_2 = [-2, -\sqrt{3})$ map to $(0, 1]$, which we already know are basin points. (Note that $f(-2) = 1$.) Since points that map to basin points are basin points as well, we know that the set $[-2, -\sqrt{3}) \cup (0, \sqrt{3})$ is included in the basin of $x = 1$. Now you may be willing to admit that the basin can be quite a complicated creature, because the graph shows that there is a small interval I_3 of points to the right of $x = 2$ that map into the interval $I_2 = [-2, -\sqrt{3})$, and are therefore in the basin, then a small interval I_4 to the left of $x = -2$ that maps into I_3, and so forth ad infinitum. These intervals are all separate (they don't overlap), and the gaps between them consist of similar intervals belonging to the basin of the other sink $x = -1$. The intervals I_n get smaller with increasing n, and all of them lie between $-\sqrt{5}$ and $\sqrt{5}$. Since $f(\sqrt{5}) = -\sqrt{5}$ and $f(-\sqrt{5}) = \sqrt{5}$, neither of these numbers is in the basin of either sink.

✎ EXERCISE T1.3

Solve the inequality $|f(x) - 0| > |x - 0|$, where $f(x) = (3x - x^3)/2$. This identifies points whose distance from 0 increases on each iteration. Use the result to find a large set of initial conditions that do not converge to any sink of f.

There is one case that is not covered by Theorem 1.5. The stability of a fixed point p cannot be determined solely by the derivative when $|f'(p)| = 1$ (see Exercise 1.2).

So far we have seen the important role of fixed points in determining the behavior of orbits of maps. If the fixed point is a sink, it provides the final state for the orbits of many nearby initial conditions. For the linear map $f(x) = ax$ with $|a| < 1$, the sink $x = 0$ attracts all initial conditions. In Examples 1.2 and 1.3, the sinks attract large sets of initial conditions.

✎ EXERCISE T1.4

Let p be a fixed point of a map f. Given some $\epsilon > 0$, find a geometric condition under which all points x in $N_\epsilon(p)$ are in the basin of p. Use cobweb plot analysis to explain your reasoning.

1.4 PERIODIC POINTS

Changing a, the constant of proportionality in the logistic map $g_a(x) = ax(1 - x)$, can result in a picture quite different from Example 1.2. When $a = 3.3$, the fixed points are $x = 0$ and $x = 23/33 = .\overline{69} = .696969\ldots$, both of which are repellers. Now that there are no fixed points around that can attract orbits, where do they go? Use a calculator to convince yourself that for almost every choice of initial condition, the orbit settles into a pattern of alternating values $p_1 = .4794$ and $p_2 = .8236$ (to four decimal place accuracy). Some typical orbits are shown in Table 1.2. The orbit with initial condition 0.2 is graphed in Figure 1.4. This figure shows typical behavior of an orbit converging to a period-2 sink $\{p_1, p_2\}$. It is attracted to p_1 every two iterates, and to p_2 on alternate iterates.

There are actually two important parts of this fact. First, there is the apparent coincidence that $g(p_1) = p_2$ and $g(p_2) = p_1$. Another way to look at this is that $g^2(p_1) = p_1$; thus p_1 is a fixed point of $h = g^2$. (The same could be said for p_2.) Second, this periodic oscillation between p_1 and p_2 is stable, and attracts orbits. This fact means that periodic behavior will show up in a physical system modeled by g. The pair $\{p_1, p_2\}$ is an example of a periodic orbit.

Definition 1.6 Let f be a map on $\mathbb{R}$. We call p a **periodic point of period k** if $f^k(p) = p$, and if k is the smallest such positive integer. The orbit with initial point p (which consists of k points) is called a **periodic orbit of period k**. We will often use the abbreviated terms **period-k point** and **period-k orbit**.

Notice that we have defined the period of an orbit to be the *minimum* number of iterates required for the orbit to repeat a point. If p is a periodic point of period 2 for the map f, then p is a fixed point of the map $h = f^2$. However, the converse is not true. A fixed point of $h = f^2$ may also be a fixed point of a lower

n	$g^n(x)$	$g^n(x)$	$g^n(x)$
0	0.2000	0.5000	0.9500
1	0.5280	0.8250	0.1568
2	0.8224	0.4764	0.4362
3	0.4820	0.8232	0.8116
4	0.8239	0.4804	0.5047
5	0.4787	0.8237	0.8249
6	0.8235	0.4792	0.4766
7	0.4796	0.8236	0.8232
8	0.8236	0.4795	0.4803
9	0.4794	0.8236	0.8237
10	0.8236	0.4794	0.4792
11	0.4794	0.8236	0.8236
12	0.8236	0.4794	0.4795
13	0.4794	0.8236	0.8236
14	0.8236	0.4794	0.4794

Table 1.2 Three different orbits of the logistic model $g(x) = 3.3x(1 - x)$.
Each approaches a period-2 orbit.

iterate of f, specifically f, and so may not be a periodic point of period two. For example, if p is a fixed point of f, it will be a fixed point of f^2 but not, according to our definition, a period-two point of f.

EXAMPLE 1.7

Consider the map defined by $f(x) = -x$ on $\mathbb{R}$. This map has one fixed point, at $x = 0$. Every other real number is a period-two point, because f^2 is the identity map.

✎ EXERCISE T1.5

The map $f(x) = 2x^2 - 5x$ on $\mathbb{R}$ has fixed points at $x = 0$ and $x = 3$. Find a period-two orbit for f by solving $f^2(x) = x$ for x.

What about the stability of periodic orbits? As in the fixed point case, points near the periodic orbit can be trapped or repelled by the orbit. The key fact is that a periodic point for f is a fixed point for f^k. We can use Theorem 1.5 to investigate the stability of a periodic orbit. For a period-k orbit, we apply Theorem 1.5 to the map f^k instead of f.

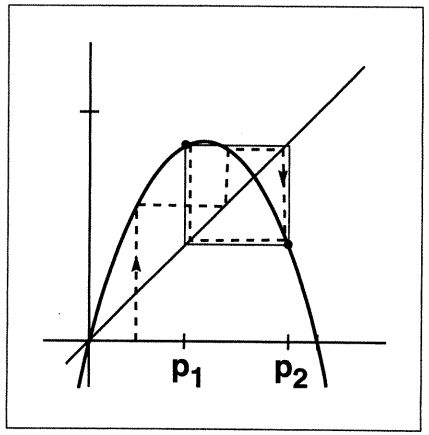

Figure 1.4 Orbit converging to a period-two sink.
The dashed lines form a cobweb plot showing an orbit which moves toward the sink orbit $\{p_1, p_2\}$.

Definition 1.8 Let f be a map and assume that p is a period-k point. The period-k orbit of p is a **periodic sink** if p is a sink for the map f^k. The orbit of p is a **periodic source** if p is a source for the map f^k.

It is helpful to review the chain rule of calculus, which shows how to expand the derivative of a composition of functions:

$$(f \circ g)'(x) = f'(g(x))g'(x) \tag{1.3}$$

Our current interest in the chain rule is for $f = g$, in which case we have $(f^2)'(x) = f'(f(x))f'(x)$. If x happens to be a period-two point for f, the chain rule is saying something quite simple: the derivative of f^2 at a point of a period-two orbit is simply the product of the derivatives of f at the two points in the orbit. In particular, the derivative of f^2 is the same, when evaluated at either point of the orbit. This agreement means that it makes sense to talk about the stability of a period-two orbit.

Now the period-two behavior of $g(x) = 3.3x(1 - x)$ we found in Table 1.2 can be completely explained. The periodic orbit $\{.4794, .8236\}$ will be a sink as long as the derivative $(g^2)'(p_1) = g'(p_1)g'(p_2) = (g^2)'(p_2)$ is smaller than 1 in absolute value. An easy calculation shows this number to be $g'(.4794)g'(.8236) = -0.2904$.

If instead we consider yet another version of the logistic map, $g(x) = 3.5x(1 - x)$, the situation is again changed. The fixed points are $x = 0$ and $x = 5/7$. Checking derivatives, $g'(0) = 3.5$ and $g'(5/7) = -1.5$, so they are sources. The orbit $\{3/7, 6/7\}$ is a period-two orbit for g. Check that $(g^2)'$ at each of the orbit points is $-5/4$, so that this period-two orbit repels nearby points. Now where do points end up?

COMPUTER EXPERIMENT 1.2

Write a computer program with the goal of redoing Table 1.2 for the logistic map $g_a(x) = ax(1 - x)$, using $a = 3.5$. What periodic behavior wins out in the long run? Try several different initial conditions to explore the basin of the attracting periodic behavior. Then try different values of $a < 3.57$ and report your results.

Now that we have some intuition from period-two orbits, we note that the situation is essentially the same for higher periods. Let $\{p_1, \ldots, p_k\}$ denote a period-k orbit of f. The chain rule says that

$$\begin{aligned}(f^k)'(p_1) &= (f(f^{k-1}))'(p_1)\\ &= f'(f^{k-1}(p_1))(f^{k-1})'(p_1)\\ &= f'(f^{k-1}(p_1))f'(f^{k-2}(p_1))\cdots f'(p_1)\\ &= f'(p_k)f'(p_{k-1})\cdots f'(p_1). \qquad (1.4)\end{aligned}$$

STABILITY TEST FOR PERIODIC ORBITS

The periodic orbit $\{p_1, \ldots, p_k\}$ is a sink if

$$|f'(p_k)\cdots f'(p_1)| < 1$$

and a source if

$$|f'(p_k)\cdots f'(p_1)| > 1.$$

This formula tells us that the derivative of the kth iterate f^k of f at a point of a period-k orbit is the product of the derivatives of f at the k points of the orbit. In particular, stability is a collective property of the periodic orbit, in that $(f^k)'(p_i) = (f^k)'(p_j)$ for all i and j.

1.5 THE FAMILY OF LOGISTIC MAPS

We are beginning to get an overall view of the family $g_a(x) = ax(1 - x)$ associated with the logistic model. When $0 \le a < 1$, the map has a sink at $x = 0$, and we will see later that every initial condition between 0 and 1 is attracted to this sink. (In other words, with small reproduction rates, small populations tend to die out.) The graph of the map is shown in Figure 1.5(a).

If $1 < a < 3$, the map, shown in Figure 1.5(b), has a sink at $x = (a - 1)/a$, since the magnitude of the derivative is less than 1. (Small populations grow to a steady state of $x = (a - 1)/a$.) For a greater than 3, as in Figure 1.5(c), the fixed point $x = (a - 1)/a$ is unstable since $|g_a'(x)| > 1$, and a period-two sink takes its place, which we saw in Table 1.2 for $a = 3.3$. When a grows above $1 + \sqrt{6} \approx 3.45$, the period-two sink also becomes unstable.

EXERCISE T1.6

Verify the statements in the previous paragraph by solving for the fixed points and period-two points of $g_a(x)$ and evaluating their stability.

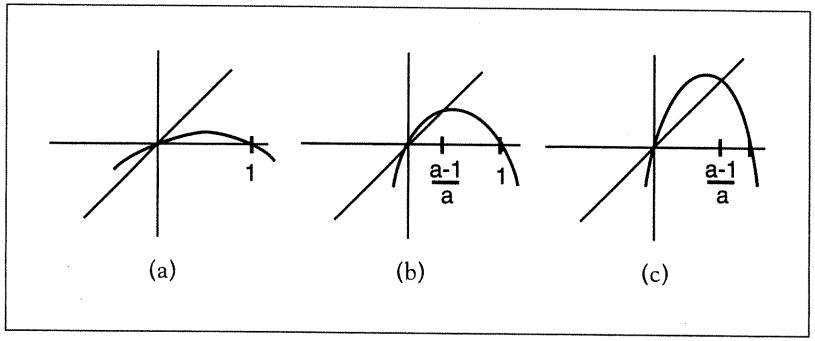

Figure 1.5 The logistic family.
(a) The origin attracts all initial conditions in $[0, 1]$. (b) The fixed point at $(a - 1)/a$ attracts all initial conditions in $(0, 1)$. (c) The fixed point at $(a - 1)/a$ is unstable.

⇨ COMPUTER EXPERIMENT 1.3

Use your logistic map program to investigate the long-run behavior of g_a for a near $a_* = 1 + \sqrt{6}$. Repeat Table 1.2 for values of a slightly smaller than a_*. What qualitative or quantitative conclusions can be made about the speed of convergence to the period-two orbit as a gets closer to a_*? What happens to iterations beginning at a period-two point for a slightly larger than a_*?

For slightly larger values of a, the story of the periodic points of $g_a(x)$ becomes significantly more complicated. Many new periodic orbits come into existence as a is increased from 3.45 to 4. Figure 1.6 shows the limiting behavior of orbits for values of a in the range $1 < a < 4$. This computer-generated picture was made by repeating the following procedure: (1) Choose a value of a, starting with $a = 1$, (2) Choose x at random in [0,1], (3) Calculate the orbit of x under $g_a(x)$, (4) Ignore the first 100 iterates and plot the orbit beginning with iterate 101. Then increment a and begin the procedure again. The points that are plotted will (within the resolution of the picture) approximate either fixed or periodic sinks or other attracting sets. This figure is called a **bifurcation diagram** and shows the birth, evolution, and death of attracting sets. The term "bifurcation" refers to significant changes in the set of fixed or periodic points or other sets of dynamic interest. We will study bifurcations in detail in Chapter 11.

We see, for example, that the vertical slice $a = 3.4$ of Figure 1.6 intersects the diagram in the two points of a period-two sink. For a slightly larger than 3.45, there appears to be a period-four sink. In fact, there is an entire sequence of periodic sinks, one for each period 2^n, $n = 1, 2, 3, \ldots$. Such a sequence is called a "period-doubling cascade". The phenomenon of cascades is the subject of Chapter 12. Figure 1.7 shows portions of the bifurcation diagram in detail. Magnification near a period-three sink, in Figure 1.7(b) hints at further period-doublings that are invisible in Figure 1.6.

For other values of the parameter a, the orbit appears to randomly fill out the entire interval [0, 1], or a subinterval. A typical cobweb plot formed for $a = 3.86$ is shown in Figure 1.8. These attracting sets, called "chaotic attractors", are harder to describe than periodic sinks. We will try to unlock some of their secrets in later chapters. As we shall see, it is a characteristic of chaotic attractors that they can abruptly appear or disappear, or change size discontinuously. This phenomenon, called a "crisis", is apparent at various a values. In particular, at $a = 4$, there is a

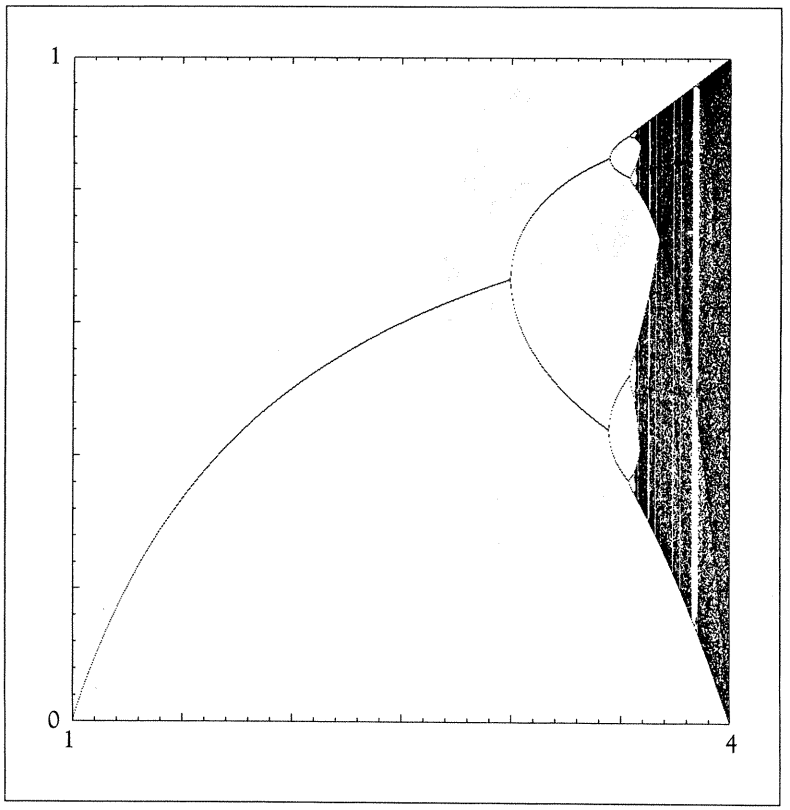

Figure 1.6 Bifurcation diagram of $g_a(x) = ax(1 - x)$.
The fixed point that exists for small values of a gives way to a period-two orbit at the "bifurcation point" $a = 3$, which in turn leads to more and more complicated orbits for larger values of a. Notice that the fixed point is only plotted while it is a sink. When the period-two orbit appears, the fixed point is no longer plotted because it does not attract orbits. See Lab Visit 12 for laboratory versions.

crisis at which the chaotic attractor disappears. For $a > 4$, there is no attracting set.

The successive blow-ups of the bifurcation diagrams reveal another interesting feature, that of "periodic windows". The period-three window, for example, is apparent in Figure 1.7(a) and is shown in magnified form in Figure 1.7(b). This refers to a set of parameter values for which there is a periodic sink, in this case a period-three sink. Since a period-three point of g_a is a fixed point of the third iterate g_a^3, the creation of the period-three sink can be seen by viewing the de-

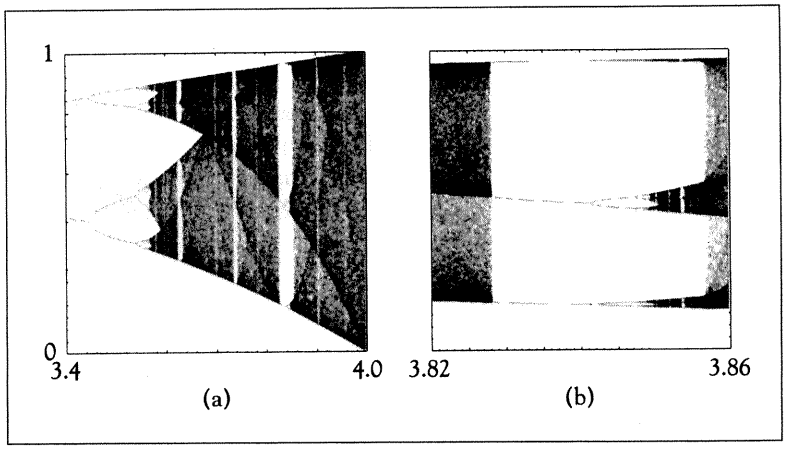

Figure 1.7 Magnifications of the logistic bifurcation diagram.
(a) Horizontal axis is $3.4 \leq a \leq 4.0$ (b) Horizontal axis is $3.82 \leq a \leq 3.86$.

velopment of the graph of g_a^3 as a moves from 3.82 to 3.86. This development is shown in Figure 1.9.

In Figure 1.9(a), the period-three orbit does not exist. This parameter value $a = 3.82$ corresponds to the left end of Figure 1.7(b). In Figure 1.9(b), the period-three orbit has been formed. Of course, since each point of a period-three orbit of g

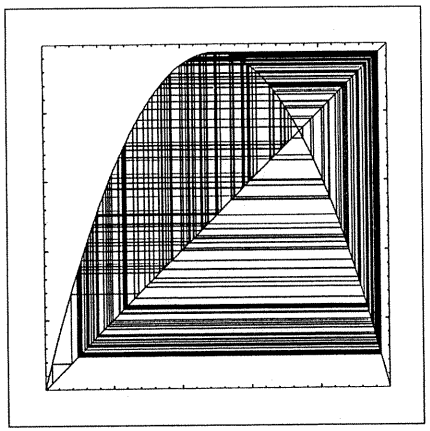

Figure 1.8 Cobweb plot for the logistic map.
A single orbit of the map $g(x) = 3.86x(1 - x)$ shows complicated behavior.

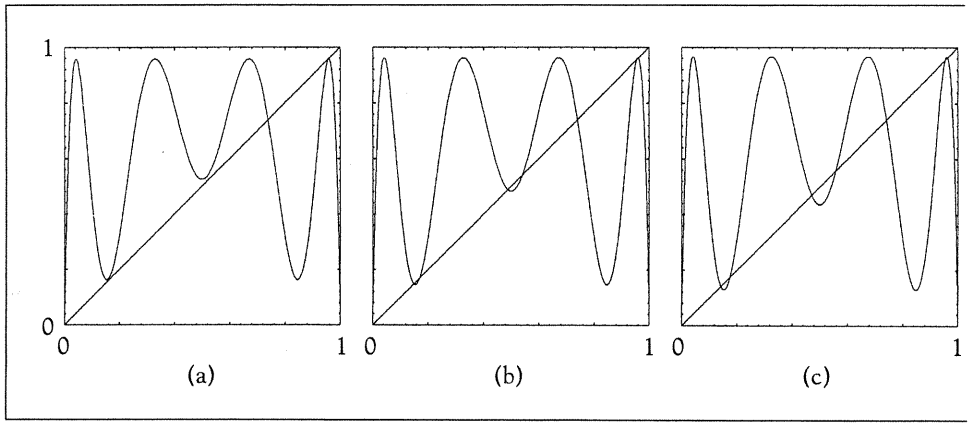

Figure 1.9 Graphs of the third iteration $g^3(x)$ of the logistic map $g_a(x) = ax(1 - x)$.
Three different parameter values are shown: (a) $a = 3.82$ (b) $a = 3.84$ (c) $a = 3.86$.

is a fixed point of g^3, the period-three orbit will appear as three intersections with the diagonal $y = x$. As you can see from the figure, the shape of the graph forces two period-three orbits to be created simultaneously. This is called a saddle-node bifurcation, or alternatively, a tangency bifurcation. The "node" is the sink, which is the set of three points at which the graph intersects the diagonal in negative slope. (Can you explain why the three negative slopes are exactly equal? Use the chain rule.) The fact that it is a sink corresponds to the fact that the negative slopes are between -1 and 0. The "saddle" is a period-three source consisting of the three upward sloping points. A vertical slice through the middle of Figure 1.7(b) shows that all initial conditions are attracted to the period-three sink. In Figure 1.9(c), the period-three sink has turned into a source. This parameter value $a = 3.86$ corresponds to the right side of Figure 1.7(b).

There are many more features of Figure 1.7 that we have to leave unexplained for now. The demise of the period-three sink as an attractor coincides with a so-called period-doubling bifurcation, which creates a period-six sink, which then meets a similar fate. There are periodic windows of arbitrarily high period. We will try to unlock some of the deeper mysteries of bifurcations in Chapter 11.

What happens to the bifurcation diagram if different x values are selected? (Recall that for each a, the orbit of one randomly chosen initial x is computed.) Surprisingly, nothing changes. The diagram looks the same no matter what initial

condition is picked at random between 0 and 1, since there is at most one attracting fixed or periodic orbit at each parameter value. As we shall see, however, there are many unstable, hence unseen, periodic orbits for larger a.

1.6 THE LOGISTIC MAP $G(x) = 4x(1 - x)$

In the previous sections we studied maps from the logistic family $g(x) = ax(1 - x)$. For $a = 2.0, 3.3$, and 3.5, we found the existence of sinks of period 1, 2, and 4, respectively. Next, we will focus on one more case, $a = 4.0$, which is so interesting that it gets its own section. The reason that it is so interesting is that it has no sinks, which leads one to ask where orbits end up.

The graph of $G(x) = g_4(x) = 4x(1 - x)$ is shown in Figure 1.10(a). Although the graph is a parabola of the type often studied in elementary precalculus courses, the map defined by G has very rich dynamical behavior. To begin with, the diagonal line $y = x$ intersects $y = G(x) = 4x(1 - x)$ in the points $x = 0$ and $x = 3/4$, so there are two fixed points, both unstable. Does G have any other periodic orbits?

One way to look for periodic orbits is to sketch the graph of $y = G^n(x)$. Any period-two point, for example, will be a fixed point of $G^2(x)$. Therefore we can find periodic points graphically.

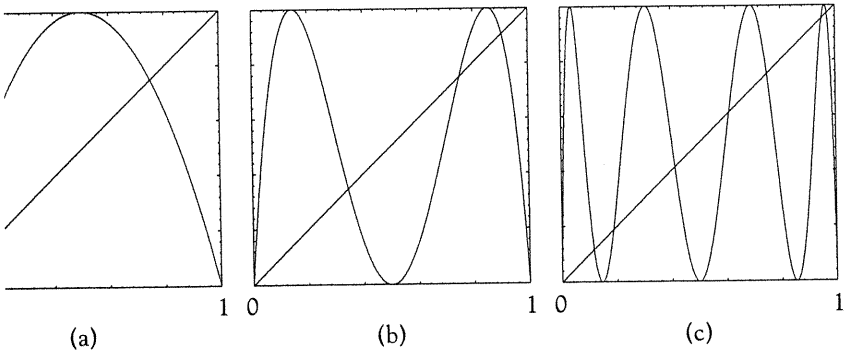

Figure 1.10 Graphs of compositions of the logistic map.
(a) the logistic map $G(x) = 4x(1 - x)$. (b) The map $G^2(x)$. (c) The map $G^3(x)$.

The graph of $y = G^2(x)$ is shown in Figure 1.10(b). It is not hard to verify by hand the general shape of the graph. First, note that the image of [0,1] under G is [0,1], so the graph stays entirely within the unit square. Second, note that $G(1/2) = 1$ and $G(1) = 0$ implies that $G^2(1/2) = 0$. Further, since $G(a_1) = 1/2$ for some a_1 between 0 and $1/2$, it follows that $G^2(a_1) = 1$. Similarly, there is another number a_2 such that $G^2(a_2) = 1$.

It is clear from Figure 1.10(b) that G^2 has four fixed points, and therefore G has four points that have period either one or two. Two of these points are already known to us—they are fixed points for G. The new pair of points, p_1 and p_2, make up a period-two orbit: that is, $G(p_1) = p_2$ and $G(p_2) = p_1$. This reasoning should have you convinced that the period-two orbit exists. The next exercise asks you to explicitly find p_1 and p_2.

✎ EXERCISE T1.7

Find the period-two orbit of $G(x) = 4x(1 - x)$.

Does G have any period-three points? There is a point b_1 between 0 and a_1 for which $G(b_1) = a_1$. This implies that $G^3(b_1) = 1$. The same holds for three other points in [0,1], so $y = G^3(x)$ has four relative maxima of height 1 in [0,1]. Since $G(1) = 0$, G^3 has roots at $x = 0, a_1, 1/2, a_2$, and 1, which separate the maxima. The graph of G^3 is shown in Figure 1.10(c).

The map G^3 has eight fixed points, two of which were known to be the fixed points 0 and $3/4$ of G. The period-two points of G are not fixed points of G^3. (Why not?) There remain six more points to account for, which must form two period-three orbits. You should be able to prove to yourself in a similar way that G^4 has $16 = 2^4$ fixed points, all in [0, 1]. With each successive iteration of G, the number of fixed points of the iterate is doubled. In general, we see that G^k has 2^k fixed points, all in [0, 1]. Of course, for $k > 1$, G has fewer than 2^k points of period-k. (Remember that the definition of period-k for the point p is that k is the smallest positive integer for which $f^k(p) = p$.) For example, $x = 0$ is a period one point and therefore not a period-k point for $k > 1$, although it is one of the 2^k fixed points of G^k.

✎ EXERCISE T1.8

Let $G(x) = 4x(1 - x)$. Prove that for each positive integer k, there is an orbit of period-k.

Period k	Number of fixed points of G^k	Number of fixed points of G^k due to lower period orbits	Orbits of period k
1	2	0	2
2	4	2	1
3	8	2	2
4	16	4	3
⋮	⋮	⋮	⋮

Table 1.3 The periodic table for the logistic map.
The nth iterate of the map $G(x) = 4x(1 - x)$ has 2^n fixed points, which are periodic orbits for G.

The number of orbits of the map for each period can be tabulated in the map's **periodic table** . For the logistic map it begins as shown in Table 1.3. The first column is the period k, and the second column is the number of fixed points of f^k, which is 2^k, as seen in Figure 1.10. The third column keeps track of fixed points of G^k which correspond to orbits of lower period than k. When these are subtracted away from the entry in the second column, the result is the number of period-k points, which is divided by k to get the number of period-k orbits.

✎ EXERCISE T1.9

Let $G(x) = 4x(1 - x)$.

(a) Decide whether the fixed points and period-two points of G are sinks.

(b) Continue the periodic table for G begun in Table 1.3. In particular, how many periodic orbits of (minimum) period k does G have, for each $k \leq 10$?

Is this what we mean by "chaos"? Not exactly. The existence of infinitely many periodic orbits does not in itself imply the kind of unpredictability usually associated with chaotic maps, although it does hint at the rich structure present. Chaos is identified with nonperiodicity and sensitive dependence on initial conditions, which we explore in the next section.

1.7 SENSITIVE DEPENDENCE ON INITIAL CONDITIONS

EXAMPLE 1.9

Consider the map $f(x) = 3x \pmod 1$ on the unit interval. The notation $y \pmod 1$ stands for the number $y + n$, where n is the unique integer that makes $0 \leq y + n < 1$. For example, $14.92 \pmod 1 = .92$ and $-14.92 \pmod 1 = .08$. For a positive number y, this is the fractional part of y. See Figure 1.11(a) for a graph of the map. Because of the breaks at $x = 1/3, 2/3$, this function is not continuous.

This map is not continuous, however the important property that we are interested in is not caused by the discontinuity . It may be more natural to view f as a map on the circle of circumference one. Glue together the ends of the unit interval to form a circle, as in Figure 1.11(b). If we consider $f(x)$ as a map from this circle to itself, it is a continuous map. In Figure 1.11(b), we show the image of the subinterval $[0, 1/2]$ on the circle. Whether we think of f as a discontinuous map on the unit interval or as a continuous map on the circle makes no difference for the questions we will try to answer below.

We call a point x **eventually periodic** with period p for the map f if for some positive integer N, $f^{n+p}(x) = f^n(x)$ for all $n \geq N$, and if p is the smallest

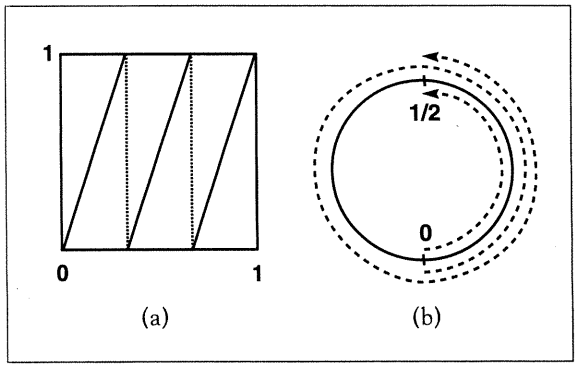

Figure 1.11 The 3x mod 1 map.
(a) The map $f(x) = 3x \pmod 1$ is discontinuous on the unit interval. (b) When the points 0 and 1 are identified, turning the unit interval into a circle, the map is continuous. The inner dashed semicircle is the subinterval $[0, 1/2]$, and the outer dashed curve is its image under the map. If x and y are two points that are close together on the circle, then $f(x)$ and $f(y)$ will be 3 times further apart than x and y.

such positive integer. This says exactly that the orbit of x eventually maps directly onto a periodic orbit. For example, $x = 1/3$ is an eventually periodic point, since it maps under f to the period one orbit 0.

EXERCISE T1.10

Show that a point x is eventually periodic for Example 1.9 if and only if x is a rational number.

EXERCISE T1.11

Construct the periodic table for f in Example 1.9 (follow the form given by Table 1.3).

The 3x mod 1 map demonstrates the main characteristic of chaos: sensitive dependence on initial conditions. This refers to the property that pairs of points, which begin as close together as desired, will eventually move apart. Table 1.4 shows the beginning of two separate orbits whose initial conditions differ by .0001. In fact, no matter how close they begin, the difference between two nearby orbits is—as measured on the circle—magnified by a factor of 3 on each iteration. This idea is important enough to be assigned a formal definition.

n	$f^n(x_0)$	$f^n(y_0)$
0	0.25	0.2501
1	0.75	0.7503
2	0.25	0.2509
3	0.75	0.7527
4	0.25	0.2581
5	0.75	0.7743
6	0.25	0.3229
7	0.75	0.9687
8	0.25	0.9061
9	0.75	0.7183
10	0.25	0.1549

Table 1.4 Comparison of the orbits of two nearly equal initial conditions under the 3x mod 1 map.
The orbits become completely uncorrelated in fewer than 10 iterates.

Definition 1.10 Let f be a map on $\mathbb{R}$. A point x_0 has **sensitive dependence on initial conditions** if there is a nonzero distance d such that some points arbitrarily near x_0 are eventually mapped at least d units from the corresponding image of x_0. More precisely, there exists $d > 0$ such that any neighborhood N of x_0 contains a point x such that $|f^k(x) - f^k(x_0)| \geq d$ for some nonnegative integer k. Sometimes we will call such a point x_0 a **sensitive point.**

Ordinarily, the closer x is to x_0, the larger k will need to be. The point x will be sensitive if it has neighbors as close as desired that eventually move away the prescribed distance d for some sufficiently large k.

EXERCISE T1.12

Consider the $3x$ mod 1 map of the unit interval [0, 1]. Define the distance between a pair of points x, y to be either $|x - y|$ or $1 - |x - y|$, whichever is smaller. (We are measuring with the "circle metric", in the sense of Figure 1.11, corresponding to the distance between two points on the circle.)

(a) Show that the distance between any pair of points that lie within $1/6$ of one another is tripled by the map. (b) Find a pair of points whose distance is not tripled by the map. (c) Show that to prove sensitive dependence for any point, d can be taken to be any positive number less than $1/2$ in Definition 1.10, and that k can be chosen to be the smallest integer greater than $\ln(d/|x - x_0|)/\ln 3$.

EXERCISE T1.13

Prove that for any map f, a source has sensitive dependence on initial conditions.

1.8 ITINERARIES

The fact that the logistic map $G(x) = 4x(1 - x)$ has periodic orbits of every period is one indication of its complicated dynamics. An even more important reflection of this complexity is sensitive dependence on initial conditions, which is the hallmark of chaos.

In this section we will show for the logistic map $G = g_4$ that for any initial point in the unit interval and any preset distance $\delta > 0$, no matter how small, there is a second point within δ units of the first so that their two orbits will map at least $d = 1/4$ units apart after a sufficient number of iterations. Since $1/4$

unit is 25% of the length of the unit interval, it is fair to say that the two initial conditions which began very close to one another are eventually moved by the map so they are no longer close, by any reasonable definition of "close".

In order to investigate sensitive dependence, we introduce the concept of the **itinerary** of an orbit. This is a bookkeeping device that allows much of the information of an orbit to be coded in terms of discrete symbols.

For the logistic map, assign the symbol L to the left subinterval $[0, 1/2]$, and R to the right subinterval $[1/2, 1]$. Given an initial value x_0, we construct its itinerary by listing the subintervals, L or R, that contain x_0 and all future iterates. For example, the initial condition $x_0 = 1/3$ begets the orbit $\{\frac{1}{3}, \frac{8}{9}, \frac{32}{81}, \ldots\}$, whose itinerary begins LRL.... For the initial condition $x_0 = \frac{1}{4}$, the orbit is $\{\frac{1}{4}, \frac{3}{4}, \frac{3}{4}, \ldots\}$,

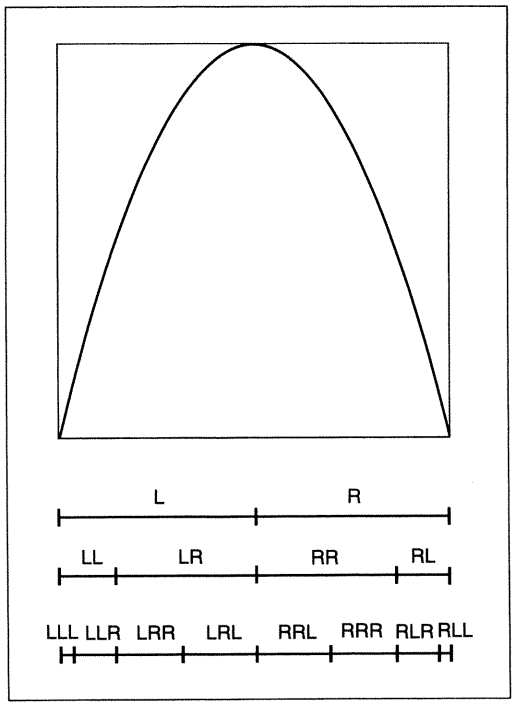

Figure 1.12 Schematic itineraries for $G(x) = 4x(1 - x)$.
The rules: (1) an interval ending in L splits into two subintervals ending in LL and LR if there are an even number of R's; the order is switched if there are an odd number of R's, (2) an interval ending in R splits into two subintervals ending in RL and RR if there are an even number of R's; the order is switched if there are an odd number of R's

which terminates in the fixed point $x = \frac{3}{4}$. The itinerary for this orbit is LRR..., which we abbreviate by $\mathsf{L}\overline{\mathsf{R}}$; the overbar indicates that the R repeats indefinitely.

Notice that there is a special orbit, or group of orbits, for which the itinerary is not uniquely defined. That is because the intervals L and R overlap at $x = 1/2$. In particular, consider the initial condition $x_0 = 1/2$. The corresponding orbit is $\{1/2, 1, 0, 0, \ldots\}$, which can be assigned itinerary $\mathsf{RR}\overline{\mathsf{L}}$ or $\mathsf{LR}\overline{\mathsf{L}}$. This particular orbit (and some others like it) are assigned two different names under this naming system. Except for the case of orbits which land precisely on $x = 1/2$ at some point of the orbit (and therefore end up mapping onto the fixed point 0), the itinerary is uniquely defined.

Once we are given this way of assigning an itinerary to each orbit, we can map out, on the unit interval, the locations of points that have certain itineraries. Of course, an itinerary is in general an infinite sequence, but we could ask: what is the set of points whose itinerary begins with, say, LR? These points share the property of beginning in the L subinterval and being mapped to the R subinterval by one iterate of the map. This set, which we could call the LR set, is shown in Figure 1.12, along with a few other similar sets.

We would like to identify the sets of all initial points whose itineraries begin with a specified sequence of symbols. For example, the set of initial conditions whose itinerary begins with LR forms a subinterval of the unit interval. The subintervals in Figure 1.12 give information about the future behavior of the initial conditions lying in them. Another example is the subinterval marked LRL, which consists of orbits that start out in the interval $\mathsf{L} = [0, 1/2]$, whose first iterate lands in $\mathsf{R} = [1/2, 1]$, and whose second iterate lands in $[0, 1/2]$. For example, $x = 1/3$ lies in LRL. Likewise, $x = 1/4$ lies in LRR because its first and second iterate are in R.

✎ EXERCISE T1.14

(a) Find a point that lies in the subinterval LLR. (You are asked for a specific number.) (b) For each subinterval corresponding to a sequence of length 3, find a point in the subinterval.

You may see some patterns in the arrangement of the subintervals of Figure 1.12. It turns out that the rule for dividing an interval, say LR, into its two subintervals is the following: Count the number of R's in the sequence (one in this case). If odd, the interval is divided into LRR, LRL in that order. If even, the L subinterval precedes the R subinterval. With this information, the reader can continue Figure 1.12 schematically to finer and finer levels.

✎ EXERCISE T1.15

Continue the schematic diagram in Figure 1.12 to subintervals corresponding to length 4 sequences.

✎ EXERCISE T1.16

Let x_0 be a point in the subinterval RLLRRRLRLR. (a) Is x_0 less than, equal to, or greater than 1/2? (b) Same question for $f^6(x_0)$.

A graphical way of specifying the possible itineraries for the logistic map is shown in Figure 1.13. We call this the **transition graph** for the subintervals L and R. An arrow from L to R, for example, means that the image $f(L)$ contains the interval R. For every path through this graph with directed edges (arrows), there exists an orbit with an itinerary satisfying the sequence of symbols determined by the path. It is clear from Figure 1.12 that the image of each of the intervals L and R contains both L and R, so the transition graph for the logistic map is fully connected (every symbol is connected to every other symbol by an arrow). Since the graph is fully connected, all possible sequences of L and R are possible.

The concept of itineraries makes it easy to explain what we mean by sensitive dependence on initial conditions. In specifying the first k symbols of the itinerary, we have 2^k choices. If k is large, then most of the 2^k subintervals are forced to be rather small, since the sum of their lengths is 1. It is a fact (that we will prove in Chapter 3) that each of the 2^k subintervals is shorter than $\pi/2^{k+1}$ in length.

Consider any one of these small subintervals for a large value of k, corresponding to some sequence of symbols $S_1 \cdots S_k$, where each S_i is either R or L. This subinterval in turn contains subintervals corresponding to the sequences $S_1 \cdots S_k$LL, $S_1 \cdots S_k$LR, $S_1 \cdots S_k$RR and $S_1 \cdots S_k$RL. If we choose one point from each, we have four initial conditions that lie within $\pi/2^{k+1}$ (since they all lie

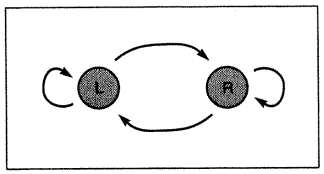

Figure 1.13 Transition graph for the logistic map $G(x) = 4x(1 - x)$.
The leftmost arrow tells us that f maps the interval L over itself, i.e., that $f(L)$ contains L. The top arrow says that $f(L)$ contains R, and so forth.

in $S_1 \cdots S_k$), but which map k iterates later to subintervals LL, LR, RR, and RL, respectively. (If this step isn't clear, it may help to recall Exercise T1.16.) In Figure 1.12, the width of the LR and RR subintervals are greater than 1/4, so that the LR and RL subintervals, for example, lie over 1/4 unit apart.

It is now possible to see why *every* point in [0, 1] has sensitive dependence on initial conditions under the logistic map G. To find a neighbor close to x_0 that eventually separates by a distance of at least $d = 1/4$, identify which subinterval of level $k + 2$ that x_0 belongs to, say $S_1 \cdots S_k$LR. Then it is always possible to identify a subinterval within $\pi/2^{k+1}$ which maps 1/4 unit away after k iterates, such as $S_1 \cdots S_k$RL. Therefore every point exhibits sensitive dependence with neighbors that are arbitrarily close.

We illustrate for $k = 1000$: There is a pair of initial conditions within $2^{-1001} \approx 10^{-300}$ that eventually are mapped at least 1/4 unit apart. This is an expansion of 1000 factors of 2 in 1000 iterates, for an average multiplicative expansion rate of approximately 2. In Chapter 3 we will introduce the term "Lyapunov number", which will quantify the average multiplicative separation rate of a map, which is in this case $2^{1000/1000} \approx 2$ per iterate. The fact that this number is greater than 1 will mean that repeated expansion is occurring.

The impact of sensitive dependence is that changes in initial measurements or errors in calculation along the orbit can result in quite different outcomes. The consequences of this behavior were not fully appreciated until the advent of computers and the computer simulation of dynamical models.

⇨ COMPUTER EXPERIMENT 1.4

Use a computer program to illustrate sensitive dependence for the logistic map $G(x) = 4x(1 - x)$. Start with two different initial conditions that are very close together, and iterate G on each. The two orbits should stay near one another for a while, and then separate for good. By collecting statistics on your experiments, try to quantify how many iterations are required for the points to move apart, say 1/2 unit, when the initial separation is .01, .001, etc. Does the location of the initial pair matter?

☞ CHALLENGE 1

Period Three Implies Chaos

IN CHAPTER 1 we have studied periodic orbits for continuous maps and the idea of sensitive dependence on initial conditions. In Challenge 1 you will prove the fact that the existence of a period-three orbit alone implies the existence of a large set of sensitive points. The set is infinite (in fact, uncountably infinite, a concept we will study in more detail later in the book). This surprising fact was discovered by T.Y. Li and J.A. Yorke (Li and Yorke, 1975).

A chaotic orbit is a bounded, non-periodic orbit that displays sensitive dependence. When we give a precise definition of chaos, we will find that the discussion is simplified if we require a stronger definition of sensitivity, namely that chaotic orbits separate exponentially fast from their neighbors as the map is iterated.

A much simpler fact about continuous maps is that the existence of a period-three orbit implies that the map has periodic orbits of all periods (all integers). See Exercise T3.10 of Chapter 3. This fact doesn't say anything directly about sensitive dependence, although it guarantees that the map has rather complicated dynamical behavior.

We show a particular map f in Figure 1.14 that has a period-three orbit, denoted $\{A, B, C\}$. That is, $f(A) = B, f(B) = C$, and $f(C) = A$. We will discover that there are infinitely many points between A and C that exhibit sensitive dependence on initial conditions. To simplify the argument, we will use an assumption that is explicitly drawn into Figure 1.14: the map $f(x)$ is **unimodal**, which means that it has only one critical point. (A **critical point** for a function $f(x)$ is a point for which $f'(x) = 0$ or where the derivative does not exist.) This assumption, that $f(x)$ has a single maximum, is not necessary to prove sensitive dependence—in fact sensitive dependence holds for any continuous map with a period-three orbit. The final step of Challenge 1 asks you to extend the reasoning to this general case.

The existence of the period-three orbit in Figure 1.14 and the continuous nature of f together guarantee that the image of the interval $[A, B]$ covers $[B, C]$; that is, that $f[A, B] \supseteq [B, C]$. Furthermore, $f[B, C] = [A, C]$. We will try to repeat our analysis of itineraries, which was successful for the logistic map, for this new map. Let the symbol L represent the interval $[A, B]$, and R represent $[B, C]$.

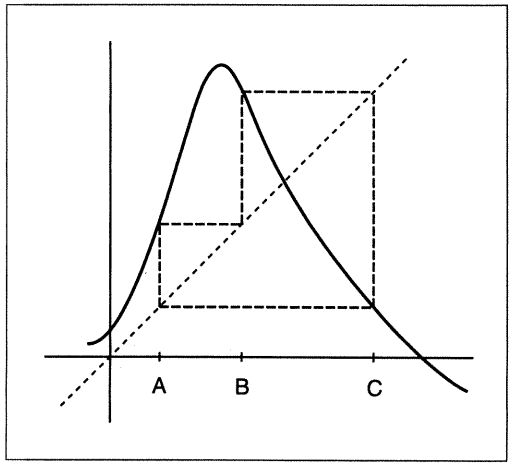

Figure 1.14 A map with a period-three orbit.
The dashed lines follow a cobweb orbit, from A to B to C to A.

Unlike the logistic map example, notice that in the itinerary of an orbit, L must be followed by R, although R can be followed by either L or R. Some of the itinerary subintervals are shown schematically in Figure 1.15.

A second difference from the logistic map example is that there may be gaps in the interval, as shown in Figure 1.15. For example, points just to the left of B are mapped to the right of C, and therefore out of the interval $[A, C]$; we do not include these points in our analysis. (A more sophisticated analysis might include these points, but we can demonstrate sensitive dependence without considering these orbits.) To simplify our analysis, we will not assign an itinerary to points that map outside $[A, C]$.

The corresponding transition graph is shown in Figure 1.16. The transition graph tells us that every finite sequence of the symbols L and R corresponds to a subinterval of initial conditions x, as long as there are never two consecutive L's in the sequence. (This follows from the fact that the left-hand interval $[A, B]$ does not map over itself.)

The proof that period three implies chaos is given below in outline form. In each part, you are expected to fill in a reason or argument.

Step 1 Let d denote the length of the subinterval RR. Denote by $J = S_1 \cdots S_k R$ any subinterval that ends in R. (Each S_i denotes either R or L.) Show

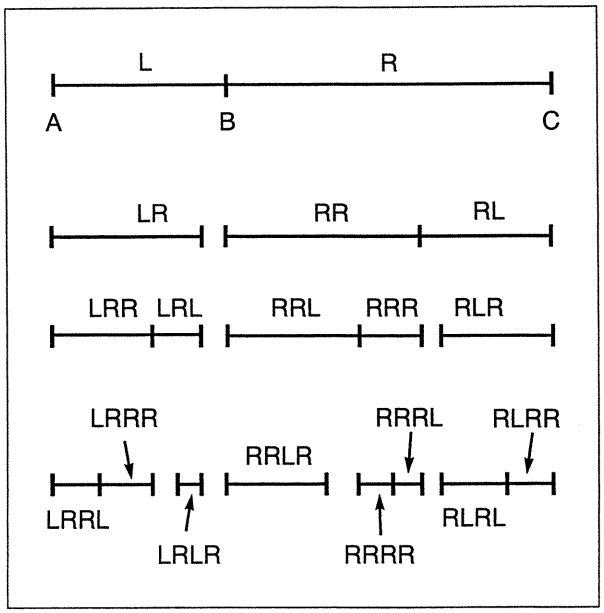

Figure 1.15 Schematic itineraries for period-three map.
The rules: (1) an interval ending in R splits into two subintervals ending in RR and RL; the order is switched if there is an even number of R's, (2) an interval ending in L contains a shorter subinterval ending in LR, and a gap on the left (for an odd number of R's) or the right (for an even number).

that J contains a pair of points, one in each of the subintervals $S_1 \cdots S_k RRL$ and $S_1 \cdots S_k RLR$, that eventually map at least d units apart.

Step 2 Show that inside J there are 3 subintervals of type $S_1 \cdots S_k RS_{k+2} S_{k+3} R$. Explain why at least 2 of them must have length that is less than half the length of J.

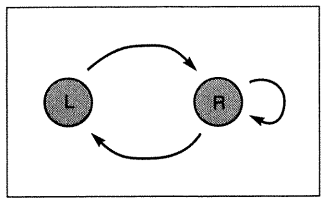

Figure 1.16 Transition graph for map with period-three orbit.
The three arrows imply that $f(L) \supseteq R$, $f(R) \supseteq R$, and $f(R) \supseteq L$.

Step 3 Combine the two previous steps. Show that a subinterval J of form $S_1 \cdots S_k R$ must contain a subinterval J_1 of form $S_1 \cdots S_k RS_{k+2} S_{k+3} RLR$ with the following property. Each point x of J_1 has a neighbor y within length$(J)/2$ whose pairwise distance upon further iteration eventually exceeds d.

Step 4 Let $h = C - A$ be the length of the original interval. Show that for each positive integer k there are 2^k disjoint subintervals (denoted by sequences of $5k + 1$ symbols) of length less than $2^{-k}h$, each of which contain a point that has sensitive dependence on initial conditions. Therefore, there are infinitely many sensitive points.

Step 5 Quantify the number of sensitive points you have located in the following way. Show that there is a one-to-one correspondence between the sensitive points found in Step 4 and binary numbers between 0 and 1 (infinite sequences of form $.a_1a_2a_3\cdots$, where each a_i is a 0 or 1). This means that the set of sensitive points is uncountable, a concept we will meet in Chapter 4.

Step 6 Our argument is based on Figure 1.14, where $f(A) = B, f(B) = C, f(C) = A$, and where $A < B < C$. How many other distinct "cases" need to be considered? Does a similar argument work for these cases? What changes are necessary?

Step 7 Explain how to modify the arguments above to work for the case where f is any continuous map with a period-three orbit. (Begin by identifying one-piece subintervals of $[A, B]$ and $[B, C]$ that are mapped onto $[A, B]$ and $[B, C]$.)

Postscript. The subintervals described in the previous argument, although many in number, may comprise a small proportion of all points in the interval [A, C]. For example, the logistic map $g(x) = 3.83x(1 - x)$ has a period-three sink. Since there is a period-three orbit (its stability does not matter), we now know that there are many points that exhibit sensitive dependence with respect to their neighbors. On the other hand, the orbits of most points in the unit interval converge to one or another point in this periodic attractor under iteration by g^3. These points do *not* exhibit sensitive dependence. For example, points that lie a small distance from one of the points p of the period-three sink will be attracted toward p, as we found in Theorem 1.5. The distances between points that start out near p decrease by a factor of approximately $|(g^3)'(p)|$ with each three iterates. These nearby points do not separate under iteration. There are, however, infinitely many points whose orbits do not converge to the period-three sink. It is these points that exhibit sensitive dependence.

EXERCISES

1.1. Let $l(x) = ax + b$, where a and b are constants. For which values of a and b does l have an attracting fixed point? A repelling fixed point?

1.2. (a) Let $f(x) = x - x^2$. Show that $x = 0$ is a fixed point of f, and describe the dynamical behavior of points near 0.

(b) Let $g(x) = \tan x$, $-\pi/2 < x < \pi/2$. Show that $x = 0$ is a fixed point of g, and describe the dynamical behavior of points near 0.

(c) Give an example of a function h for which $h'(0) = 1$ and $x = 0$ is an attracting fixed point.

(d) Give an example of a function h for which $h'(0) = 1$ and $x = 0$ is a repelling fixed point.

1.3. Let $f(x) = x^3 + x$. Find all fixed points of f and decide whether they are sinks or sources. You will have to work without Theorem 1.5, which does not apply.

1.4. Let $x_1 < \cdots < x_8$ be the eight fixed points of $G^3(x)$, where $G(x) = 4x(1 - x)$, as in 1.10(c). Clearly, $x_1 = 0$.

(a) For which i is $x_i = 3/4$?

(b) Group the remaining six points into two orbits of three points each. It may help to consult Figure 1.10(c). The most elegant solution (that we know of) uses the chain rule.

1.5. Is the period-two orbit of the map $f(x) = 2x^2 - 5x$ on $\mathbb{R}$ a sink, a source, or neither? See Exercise T1.5.

1.6. Define the map $f(x) = 2x \pmod 1$ on the unit interval $[0, 1]$. Let L denote the subinterval $[0, 1/2]$ and R the subinterval $[1/2, 1]$.

(a) Draw a chart of the itineraries of f as in Figure 1.12.

(b) Draw the transition graph for f.

(c) Establish sensitive dependence for orbits under this map. Show that each point has neighbors arbitrarily near that eventually map at least $1/2$ unit apart.

1.7. Define the tent map on the unit inverval $[0, 1]$ by

$$T(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1/2 \\ 2(1 - x) & \text{if } 1/2 \le x \le 1 \end{cases}.$$

(a) Divide the unit interval into two appropriate subintervals and repeat parts (a)–(c) of Exercise 1.6 for this map.

(b) Complete a periodic table for f, similar to the one in Table 1.3, for periods less than or equal to 10. In what ways, if any, does it differ from the periodic table for the logistic map G?

1.8. Let $f(x) = 4x(1 - x)$. Prove that there are points in $I = [0, 1]$ that are not fixed points, periodic points, or eventually periodic points of f.

1.9. Define $x_{n+1} = (x_n + 2)/(x_n + 1)$.

(a) Find $L = \lim_{n\to\infty} x_n$ for $x_0 \ge 0$.

(b) Describe the set of all negative x_0 for which the limit $\lim_{n\to\infty} x_n$ either exists and is not equal to the L in part (a) or does not exist (for example, $x_0 = -1$).

1.10. For the map $g(x) = 3.05x(1 - x)$, find the stability of all fixed points and period-two points.

1.11. Let f be a one-to-one smooth map of the real line to itself. One-to-one means that if $f(x_1) = f(x_2)$, then $x_1 = x_2$. A function f is called **increasing** if $x_1 < x_2$ implies $f(x_1) < f(x_2)$, and **decreasing** if $x_1 < x_2$ implies $f(x_1) > f(x_2)$.

(a) Show that f is increasing for all x or f is decreasing for all x.

(b) Show that every orbit $\{x_0, x_1, x_2, \ldots\}$ of f^2 satisfies either $x_0 \ge x_1 \ge x_2 \ge \ldots$ or $x_0 \le x_1 \le x_2 \le \ldots$.

(c) Show that every orbit of f^2 either diverges to $+\infty$ or $-\infty$ or converges to a fixed point of f^2.

(d) What does this imply about convergence of the orbits of f?

1.12. The map $g(x) = 2x(1 - x)$ has negative values for large x. Population biologists sometimes prefer maps that are positive for positive x.

(a) Find out for what value of a the map $h(x) = axe^{-x}$ has a **superstable fixed point** x_0, which means that $h(x_0) = x_0$ and $h'(x_0) = 0$.

(b) Investigate the orbit starting at $x_0 = 0.1$ for this value of a using a calculator. How does the behavior of this orbit differ if a is increased by 50%?

(c) What is the range of $a \ge 1$ for which $h(x)$ has a positive sink?

1.13. Let $f : [0, \infty) \to [0, \infty)$ be a smooth function, $f(0) = 0$, and let $p > 0$ be a fixed point such that $f'(p) \ge 0$. Assume further that $f'(x)$ is decreasing. Show that all positive x_0 converge to p under f.

1.14. Let $f(x) = x^2 + x$. Find all fixed points of f. Where do nonfixed points go under iteration by f?

1.15. Prove the following explicit formula for any orbit $\{x_0, x_1, x_2, \ldots\}$ of the logistic map $G(x) = 4x(1 - x)$:

$$x_n = \frac{1}{2} - \frac{1}{2}\cos(2^n \arccos(1 - 2x_0)).$$

Caution: Not all explicit formulas are useful.

1.16. Let f be a map defined on the real line, and assume p is a fixed point. Let $\epsilon > 0$ be a given number. Find a condition that guarantees that every initial point x in the interval $(p - \epsilon, p + \epsilon)$ satisfies $f^n(x) \to p$ as $n \to \infty$.

1.17. Let $f(x) = 4x(1 - x)$. Prove that LLL . . . L, the interval of initial values x in $[0, 1]$ such that $0 < f^i(x) < 1/2$ for $0 \le i < k$, has length $[1 - \cos(\pi/2^k)]/2$.

☞ LAB VISIT I

Boom, Bust, and Chaos in the Beetle Census

DAMAGE DUE TO flour beetles is a significant cost to the food processing industry. One of the major goals of entomologists is to gain insight into the population dynamics of beetles and other insects, as a way of learning about insect physiology. A commercial application of population studies is the development of strategies for population control.

A group of researchers recently designed a study of population fluctuation in the flour beetle *Tribolium*. The newly hatched larva spends two weeks feeding before entering a pupa stage of about the same length. The beetle exits the pupa stage as an adult. The researchers proposed a discrete map that models the three separate populations. Let the numbers of larvae, pupae, and adults at any given time t be denoted L_t, P_t, and A_t, respectively. The output of the map is three numbers: the three populations L_{t+1}, P_{t+1}, and A_{t+1} one time unit later. It is most convenient to take the time unit to be two weeks. A typical model for the three beetle populations is

$$\begin{aligned} L_{t+1} &= bA_t \\ P_{t+1} &= L_t(1 - \mu_l) \\ A_{t+1} &= P_t(1 - \mu_p) + A_t(1 - \mu_a), \end{aligned} \tag{1.5}$$

where b is the birth rate of the species (the number of new larvae per adult each time unit), and where μ_l, μ_p, and μ_a are the death rates of the larva, pupa, and adult, respectively.

We call a discrete map with three variables a three-dimensional map, since the state of the population at any given time is specified by three numbers L_t, P_t, and A_t. In Chapter 1, we studied one-dimensional maps, and in Chapter 2 we move on to higher dimensional maps, of which the beetle population model is an example.

Tribolium adds an interesting twist to the above model: cannibalism caused by overpopulation stress. Under conditions of overcrowding, adults will eat pupae

Costantino, R.F., Cushing, J.M., Dennis, B., Desharnais, R.A., Experimentally induced transitions in the dynamic behavior of insect populations. Nature **375**, 227–230 (1995).

and unhatched eggs (future larvae); larvae will also eat eggs. Incorporating these refinements into the model yields

$$\begin{aligned} L_{t+1} &= bA_t \exp(-c_{ea}A_t - c_{el}L_t) \\ P_{t+1} &= L_t(1-\mu_l) \\ A_{t+1} &= P_t(1-\mu_p)\exp(-c_{pa}A_t) + A_t(1-\mu_a). \end{aligned} \tag{1.6}$$

The parameters $c_{el} = 0.012$, $c_{ea} = 0.009$, $c_{pa} = 0.004$, $\mu_l = 0.267$, $\mu_p = 0$, and $b = 7.48$ were determined from population experiments. The mortality rate of the adult was determined from experiment to be $\mu_a = 0.0036$.

The effect of calling the exterminator can be modeled by artificially changing the adult mortality rate. Figure 1.17 shows a bifurcation diagram from Equations (1.6). The horizontal axis represents the mortality rate μ_a. The asymptotic value of L_t—found by running the model for a long time at a fixed μ_a and recording the resulting larval population—is graphed vertically.

Figure 1.17 suggests that for relatively low mortality rates, the larval population reaches a steady state (a fixed point). For $\mu_a > .1$ (representing a death rate of 10% of the adults over each 2 week period), the model shows oscillation between two widely-different states. This is a "boom-and-bust" cycle, well-known to population biologists. A low population (bust) leads to uncrowded living con-

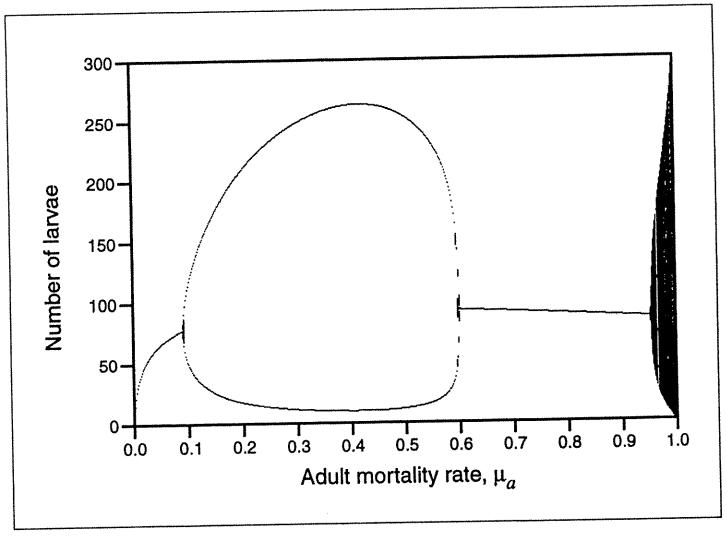

Figure 1.17 Bifurcation diagram for the model equations (1.6).
The bifurcation parameter is μ_a, the adult mortality rate.

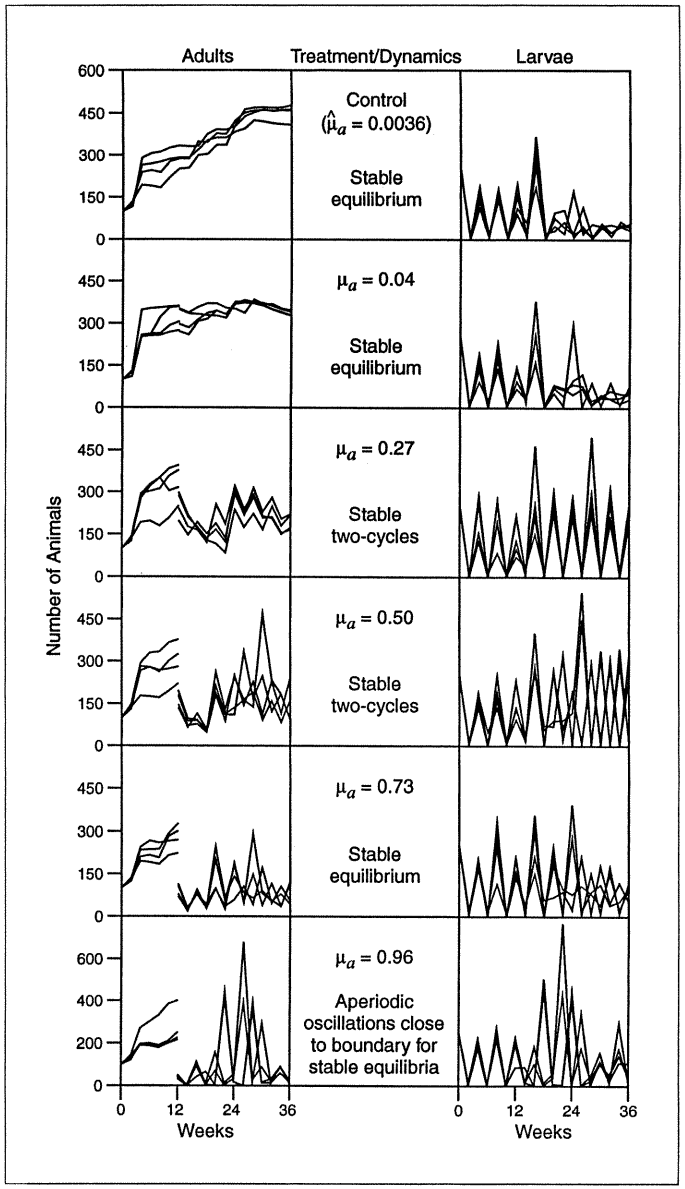

Figure 1.18 Population as a function of time.
Four replicates of the experiment for each of six different adult mortality rates are plotted together.

ditions and runaway growth (boom) at the next generation. At this point the limits to growth (cannibalism, in this system) take over, leading to a catastrophic decline and repeat of the cycle.

The period-doubling bifurcation near $\mu_a = 0.1$ is followed by a period-halving bifurcation at $\mu_a \approx 0.6$. For very high adult mortality rates (near 100%), we see the complicated, nonperiodic behavior.

The age-stratified population model discussed above is an interesting mathematical abstraction. What does it have to do with real beetles? The experimenters put several hundred beetles and 20 grams of food in each of several half-pint milk bottles. They recorded the populations for 18 consecutive two-week periods. Five different adult mortality rates, $\mu_a = 0.0036$ (the natural rate), 0.04, 0.27, 0.50, 0.73, and 0.96 were enforced in different bottles, by periodically removing the requisite number of adult beetles to artificially reach that rate. Each of the five experiments was replicated in four separate bottles.

Figure 1.18 shows the population counts taken from the experiment. Populations of adults from the four separate bottles are graphed together in the boxes on the left. The four curves in the box are the adult population counts for the four bottles as a function of time. The boxes on the right are similar but show the population counts for the larvae. During the first 12 weeks, the populations were undisturbed, so that the natural adult mortality rate applied; after that, the artificial mortality rates were imposed by removing or adding adult beetles as needed.

The population counts from the experiment agree remarkably well with the computer simulations from Figure 1.18. The top two sets of boxes represent $\mu_a = 0.0036$ and 0.04, which appear experimentally to be sinks, or stable equilibria, as predicted by Figure 1.18. The period-two sink predicted also can be seen in the populations for $\mu_a = 0.27$ and 0.50. For $\mu_a = 0.96$, the populations seem to be governed by aperiodic oscillations.

CHAPTER TWO

Two-Dimensional Maps

IN CHAPTER 1 we developed the fundamental properties of one-dimensional dynamics. The concepts of periodic orbits, stability, and sensitive dependence of orbits are most easily understood in that context.

In this chapter we will begin the transition from one-dimensional to many dimensional dynamics. The discussion centers around two-dimensional maps, since much of the new phenomena present in higher dimensions appears there in its simplest form. For example, we will expand our classification of one-dimensional fixed points as sinks and sources to include saddles, which are contracting in some directions and expanding in others.