Π. Ιωαν
Δεκέμβριος 99.

NONLINEAR DYNAMICS AND CHAOS

With Applications to Physics, Biology, Chemistry, and Engineering

STEVEN H. STROGATZ

ADVANCED BOOK PROGRAM

PERSEUS BOOKS
Reading, Massachusetts

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book and Perseus Books was aware of a trademark claim, the designations have been printed in initial capital letters.

Library of Congress Cataloging-in-Publication Data
Strogatz, Steven H. (Steven Henry)
Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering / Steven H. Strogatz.
p. cm.
Includes bibliographical references and index.
ISBN 0-201-54344-3
1. Chaotic behavior in systems. 2. Dynamics. 3. Nonlinear theories. I. Title.
Q172.5.C45S767 1994
501'.1'85—dc20 93-6166
CIP

Cover design by Lynne Reed
Text design by Joyce C. Weston
Set in 10-point Times by Compset, Inc.
Cover art is a computer-generated picture of a scroll ring, from Strogatz (1985) with permission. Scroll rings are self-sustaining sources of waves in diverse excitable media, including heart muscle, neural tissue, and excitable chemical reactions (Winfree and Strogatz 1984, Winfree 1987b).

10 11 12 13 14 15—MA—03 02 01 00 99

CONTENTS

Part III. Chaos

PREFACE

This textbook is aimed at newcomers to nonlinear dynamics and chaos, especially students taking a first course in the subject. It is based on a one-semester course I've taught for the past several years at MIT and Cornell. My goal is to explain the mathematics as clearly as possible, and to show how it can be used to understand some of the wonders of the nonlinear world.

The mathematical treatment is friendly and informal, but still careful. Analytical methods, concrete examples, and geometric intuition are stressed. The theory is developed systematically, starting with first-order differential equations and their bifurcations, followed by phase plane analysis, limit cycles and their bifurcations, and culminating with the Lorenz equations, chaos, iterated maps, period doubling, renormalization, fractals, and strange attractors.

A unique feature of the book is its emphasis on applications. These include mechanical vibrations, lasers, biological rhythms, superconducting circuits, insect outbreaks, chemical oscillators, genetic control systems, chaotic waterwheels, and even a technique for using chaos to send secret messages. In each case, the scientific background is explained at an elementary level and closely integrated with the mathematical theory.

Prerequisites

The essential prerequisite is single-variable calculus, including curve-sketching, Taylor series, and separable differential equations. In a few places, multivariable calculus (partial derivatives, Jacobian matrix, divergence theorem) and linear algebra (eigenvalues and eigenvectors) are used. Fourier analysis is not assumed, and is developed where needed. Introductory physics is used throughout. Other scientific prerequisites would depend on the applications considered, but in all cases, a first course should be adequate preparation.

Possible Courses

The book could be used for several types of courses:

- A broad introduction to nonlinear dynamics, for students with no prior exposure to the subject. (This is the kind of course I have taught.) Here one goes straight through the whole book, covering the core material at the beginning of each chapter, selecting a few applications to discuss in depth and giving light treatment to the more advanced theoretical topics or skipping them altogether. A reasonable schedule is seven weeks on Chapters 1–8, and five or six weeks on Chapters 9–12. Make sure there's enough time left in the semester to get to chaos, maps, and fractals.
- A traditional course on nonlinear ordinary differential equations, but with more emphasis on applications and less on perturbation theory than usual. Such a course would focus on Chapters 1–8.
- A modern course on bifurcations, chaos, fractals, and their applications, for students who have already been exposed to phase plane analysis. Topics would be selected mainly from Chapters 3, 4, and 8–12.

For any of these courses, the students should be assigned homework from the exercises at the end of each chapter. They could also do computer projects; build chaotic circuits and mechanical systems; or look up some of the references to get a taste of current research. This can be an exciting course to teach, as well as to take. I hope you enjoy it.

Conventions

Equations are numbered consecutively within each section. For instance, when we're working in Section 5.4, the third equation is called (3) or Equation (3), but elsewhere it is called (5.4.3) or Equation (5.4.3). Figures, examples, and exercises are always called by their full names, e.g., Exercise 1.2.3. Examples and proofs end with a loud thump, denoted by the symbol ■.

Acknowledgments

Thanks to the National Science Foundation for financial support. For help with the book, thanks to Diana Dabby, Partha Saha, and Shinya Watanabe (students); Jihad Touma and Rodney Worthing (teaching assistants); Andy Christian, Jim Crutchfield, Kevin Cuomo, Frank DeSimone, Roger Eckhardt, Dana Hobson, and Thanos Siapas (for providing figures); Bob Devaney, Irv Epstein, Danny Kaplan, Willem Malkus, Charlie Marcus, Paul Matthews, Arthur Mattuck, Rennie Mirollo, Peter Renz, Dan Rockmore, Gil Strang, Howard Stone, John Tyson, Kurt Wiesenfeld, Art Winfree, and Mary Lou Zeeman (friends and colleagues who gave advice); and to my editor Jack Repcheck, Lynne Reed, Production Supervisor, and all the other helpful people at Perseus Books. Finally, thanks to my family and Elisabeth for their love and encouragement.

Steven H. Strogatz
Cambridge, Massachusetts

1

OVERVIEW

1.0 Chaos, Fractals, and Dynamics

There is a tremendous fascination today with chaos and fractals. James Gleick's book *Chaos* (Gleick 1987) was a bestseller for months—an amazing accomplishment for a book about mathematics and science. Picture books like *The Beauty of Fractals* by Peitgen and Richter (1986) can be found on coffee tables in living rooms everywhere. It seems that even nonmathematical people are captivated by the infinite patterns found in fractals (Figure 1.0.1). Perhaps most important of all, chaos and fractals represent hands-on mathematics that is alive and changing. You can turn on a home computer and create stunning mathematical images that no one has ever seen before.

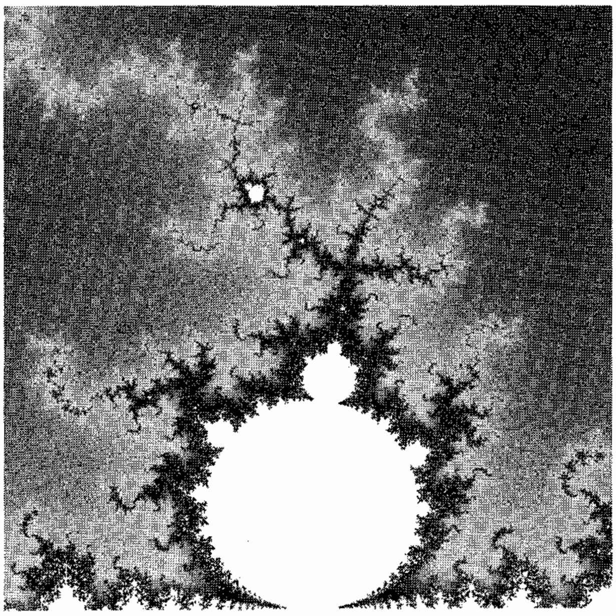

Figure 1.0.1

The aesthetic appeal of chaos and fractals may explain why so many people have become intrigued by these ideas. But maybe you feel the urge to go deeper—to learn the mathematics behind the pictures, and to see how the ideas can be applied to problems in science and engineering. If so, this is a textbook for you.

The style of the book is informal (as you can see), with an emphasis on concrete examples and geometric thinking, rather than proofs and abstract arguments. It is also an extremely "applied" book—virtually every idea is illustrated by some application to science or engineering. In many cases, the applications are drawn from the recent research literature. Of course, one problem with such an applied approach is that not everyone is an expert in physics *and* biology *and* fluid mechanics . . . so the science as well as the mathematics will need to be explained from scratch. But that should be fun, and it can be instructive to see the connections among different fields.

Before we start, we should agree about something: chaos and fractals are part of an even grander subject known as ***dynamics***. This is the subject that deals with change, with systems that evolve in time. Whether the system in question settles down to equilibrium, keeps repeating in cycles, or does something more complicated, it is dynamics that we use to analyze the behavior. You have probably been exposed to dynamical ideas in various places—in courses in differential equations, classical mechanics, chemical kinetics, population biology, and so on. Viewed from the perspective of dynamics, all of these subjects can be placed in a common framework, as we discuss at the end of this chapter.

Our study of dynamics begins in earnest in Chapter 2. But before digging in, we present two overviews of the subject, one historical and one logical. Our treatment is intuitive; careful definitions will come later. This chapter concludes with a "dynamical view of the world," a framework that will guide our studies for the rest of the book.

1.1 Capsule History of Dynamics

Although dynamics is an interdisciplinary subject today, it was originally a branch of physics. The subject began in the mid-1600s, when Newton invented differential equations, discovered his laws of motion and universal gravitation, and combined them to explain Kepler's laws of planetary motion. Specifically, Newton solved the two-body problem—the problem of calculating the motion of the earth around the sun, given the inverse-square law of gravitational attraction between them. Subsequent generations of mathematicians and physicists tried to extend Newton's analytical methods to the three-body problem (e.g., sun, earth, and moon) but curiously this problem turned out to be much more difficult to solve. After decades of effort, it was eventually realized that the three-body problem was essentially *impossible* to solve, in the sense of obtaining explicit formulas for the motions of the three bodies. At this point the situation seemed hopeless.

The breakthrough came with the work of Poincaré in the late 1800s. He introduced a new point of view that emphasized qualitative rather than quantitative questions. For example, instead of asking for the exact positions of the planets at all times, he asked "Is the solar system stable forever, or will some planets eventually fly off to infinity?" Poincaré developed a powerful *geometric* approach to analyzing such questions. That approach has flowered into the modern subject of dynamics, with applications reaching far beyond celestial mechanics. Poincaré was also the first person to glimpse the possibility of ***chaos,*** in which a deterministic system exhibits aperiodic behavior that depends sensitively on the initial conditions, thereby rendering long-term prediction impossible.

But chaos remained in the background in the first half of this century; instead dynamics was largely concerned with nonlinear oscillators and their applications in physics and engineering. Nonlinear oscillators played a vital role in the development of such technologies as radio, radar, phase-locked loops, and lasers. On the theoretical side, nonlinear oscillators also stimulated the invention of new mathematical techniques—pioneers in this area include van der Pol, Andronov, Littlewood, Cartwright, Levinson, and Smale. Meanwhile, in a separate development, Poincaré's geometric methods were being extended to yield a much deeper understanding of classical mechanics, thanks to the work of Birkhoff and later Kolmogorov, Arnol'd, and Moser.

The invention of the high-speed computer in the 1950s was a watershed in the history of dynamics. The computer allowed one to experiment with equations in a way that was impossible before, and thereby to develop some intuition about nonlinear systems. Such experiments led to Lorenz's discovery in 1963 of chaotic motion on a strange attractor. He studied a simplified model of convection rolls in the atmosphere to gain insight into the notorious unpredictability of the weather. Lorenz found that the solutions to his equations never settled down to equilibrium or to a periodic state—instead they continued to oscillate in an irregular, aperiodic fashion. Moreover, if he started his simulations from two slightly different initial conditions, the resulting behaviors would soon become totally different. The implication was that the system was *inherently* unpredictable—tiny errors in measuring the current state of the atmosphere (or any other chaotic system) would be amplified rapidly, eventually leading to embarrassing forecasts. But Lorenz also showed that there was structure in the chaos—when plotted in three dimensions, the solutions to his equations fell onto a butterfly-shaped set of points (Figure 1.1.1). He argued that this set had to be "an infinite complex of surfaces"—today we would regard it as an example of a fractal.

Lorenz's work had little impact until the 1970s, the boom years for chaos. Here are some of the main developments of that glorious decade. In 1971 Ruelle and Takens proposed a new theory for the onset of turbulence in fluids, based on abstract considerations about strange attractors. A few years later, May found examples of chaos in iterated mappings arising in population biology, and wrote an influential review article that stressed the pedagogical importance of studying simple nonlinear systems, to counterbalance the often misleading linear intuition fostered by traditional education. Next came the most surprising discovery of all, due to the physicist Feigenbaum. He discovered that there are certain universal laws governing the transition from regular to chaotic behavior; roughly speaking, completely different systems can go chaotic in the same way. His work established a link between chaos and

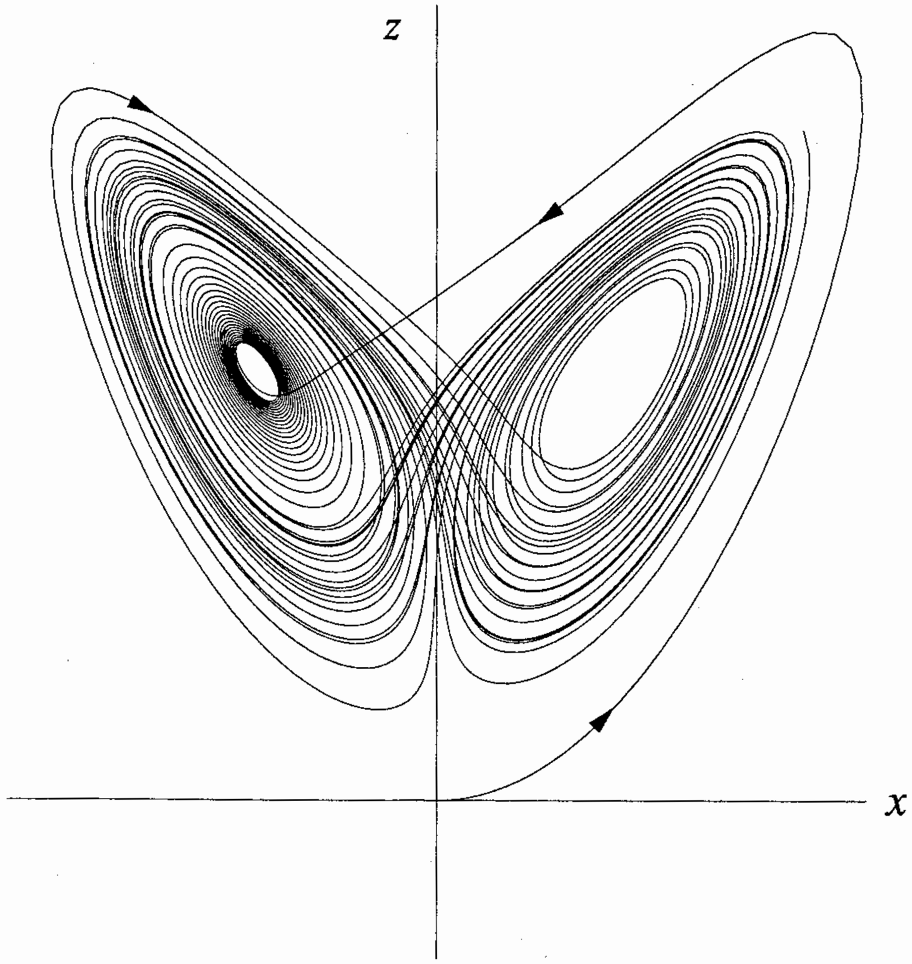

Figure 1.1.1

phase transitions, and enticed a generation of physicists to the study of dynamics. Finally, experimentalists such as Gollub, Libchaber, Swinney, Linsay, Moon, and Westervelt tested the new ideas about chaos in experiments on fluids, chemical reactions, electronic circuits, mechanical oscillators, and semiconductors.

Although chaos stole the spotlight, there were two other major developments in dynamics in the 1970s. Mandelbrot codified and popularized fractals, produced magnificent computer graphics of them, and showed how they could be applied in a variety of subjects. And in the emerging area of mathematical biology, Winfree applied the geometric methods of dynamics to biological oscillations, especially circadian (roughly 24-hour) rhythms and heart rhythms.

By the 1980s many people were working on dynamics, with contributions too numerous to list. Table 1.1.1 summarizes this history.

1.2 The Importance of Being Nonlinear

Now we turn from history to the logical structure of dynamics. First we need to introduce some terminology and make some distinctions.

Dynamics - A Capsule History		
1666	Newton	Invention of calculus, explanation of planetary motion
1700s		Flowering of calculus and classical mechanics
1800s		Analytical studies of planetary motion
1890s	Poincaré	Geometric approach, nightmares of chaos
1920–1950		Nonlinear oscillators in physics and engineering, invention of radio, radar, laser
1920–1960	Birkhoff Kolmogorov Arnol'd Moser	Complex behavior in Hamiltonian mechanics
1963	Lorenz	Strange attractor in simple model of convection
1970s	Ruelle &Takens	Turbulence and chaos
	May	Chaos in logistic map
	Feigenbaum	Universality and renormalization, connection between chaos and phase transitions
		Experimental studies of chaos
	Winfree	Nonlinear oscillators in biology
	Mandelbrot	Fractals
1980s		Widespread interest in chaos, fractals, oscillators, and their applications

Table 1.1.1

There are two main types of dynamical systems: ***differential equations*** and ***iterated maps*** (also known as difference equations). Differential equations describe the evolution of systems in continuous time, whereas iterated maps arise in problems where time is discrete. Differential equations are used much more widely in science and engineering, and we shall therefore concentrate on them. Later in the book we will see that iterated maps can also be very useful, both for providing simple examples of chaos, and also as tools for analyzing periodic or chaotic solutions of differential equations.

Now confining our attention to differential equations, the main distinction is between ordinary and partial differential equations. For instance, the equation for a damped harmonic oscillator

$$m\frac{d^2x}{dt^2} + b\frac{dx}{dt} + kx = 0 \qquad (1)$$

is an ordinary differential equation, because it involves only ordinary derivatives dx/dt and d^2x/dt^2. That is, there is only one independent variable, the time t. In contrast, the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

is a partial differential equation—it has both time t and space x as independent variables. Our concern in this book is with purely temporal behavior, and so we deal with ordinary differential equations almost exclusively.

A very general framework for ordinary differential equations is provided by the system

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, \dots, x_n) \\ &\vdots \\ \dot{x}_n &= f_n(x_1, \dots, x_n). \end{aligned} \tag{2}$$

Here the overdots denote differentiation with respect to t. Thus $\dot{x}_i \equiv dx_i/dt$. The variables $x_1, \dots, x_n$ might represent concentrations of chemicals in a reactor, populations of different species in an ecosystem, or the positions and velocities of the planets in the solar system. The functions $f_1, \dots, f_n$ are determined by the problem at hand.

For example, the damped oscillator (1) can be rewritten in the form of (2), thanks to the following trick: we introduce new variables $x_1 = x$ and $x_2 = \dot{x}$. Then $\dot{x}_1 = x_2$, from the definitions, and

$$\begin{aligned} \dot{x}_2 = \ddot{x} &= -\tfrac{b}{m}\dot{x} - \tfrac{k}{m}x \\ &= -\tfrac{b}{m}x_2 - \tfrac{k}{m}x_1 \end{aligned}$$

from the definitions and the governing equation (1). Hence the equivalent system (2) is

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\tfrac{b}{m}x_2 - \tfrac{k}{m}x_1. \end{aligned}$$

This system is said to be ***linear***, because all the x_i on the right-hand side appear to the first power only. Otherwise the system would be ***nonlinear***. Typical nonlinear terms are products, powers, and functions of the x_i, such as $x_1 x_2$, $(x_1)^3$, or $\cos x_2$.

For example, the swinging of a pendulum is governed by the equation

$$\ddot{x} + \tfrac{g}{L}\sin x = 0,$$

where x is the angle of the pendulum from vertical, g is the acceleration due to gravity, and L is the length of the pendulum. The equivalent system is nonlinear:

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -\tfrac{g}{L}\sin x_1.$$

Nonlinearity makes the pendulum equation very difficult to solve analytically. The usual way around this is to fudge, by invoking the small angle approximation $\sin x \approx x$ for $x << 1$. This converts the problem to a linear one, which can then be solved easily. But by restricting to small x, we're throwing out some of the physics, like motions where the pendulum whirls over the top. Is it really necessary to make such drastic approximations?

It turns out that the pendulum equation *can* be solved analytically, in terms of elliptic functions. But there ought to be an easier way. After all, the motion of the pendulum is simple: at low energy, it swings back and forth, and at high energy it whirls over the top. There should be some way of extracting this information from the system directly. This is the sort of problem we'll learn how to solve, using geometric methods.

Here's the rough idea. Suppose we happen to know a solution to the pendulum system, for a particular initial condition. This solution would be a pair of functions $x_1(t)$ and $x_2(t)$, representing the position and velocity of the pendulum. If we construct an abstract space with coordinates (x_1, x_2), then the solution $(x_1(t),\ x_2(t))$ corresponds to a point moving along a curve in this space (Figure 1.2.1).

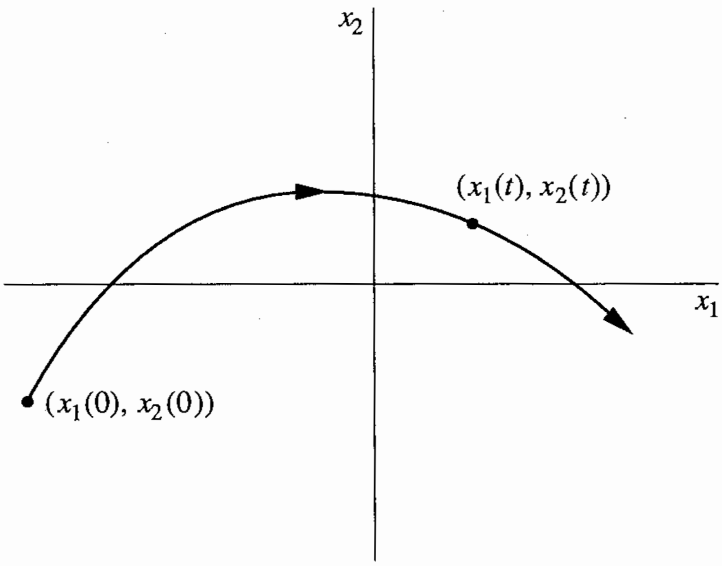

Figure 1.2.1

This curve is called a ***trajectory***, and the space is called the ***phase space*** for the system. The phase space is completely filled with trajectories, since each point can serve as an initial condition.

Our goal is to run this construction *in reverse:* given the system, we want to draw the trajectories, and thereby extract information about the solutions. In many cases, geometric reasoning will allow us to draw the trajectories *without actually solving the system*!

Some terminology: the phase space for the general system (2) is the space with coordinates $x_1, \ldots, x_n$. Because this space is n-dimensional, we will refer to (2) as an ***n*-dimensional system** or an ***n*th-order** system. Thus n represents the dimension of the phase space.

Nonautonomous Systems

You might worry that (2) is not general enough because it doesn't include any explicit *time dependence*. How do we deal with time-dependent or ***nonautonomous*** equations like the forced harmonic oscillator $m\ddot{x} + b\dot{x} + kx = F\cos t$? In this case too there's an easy trick that allows us to rewrite the system in the form (2). We let $x_1 = x$ and $x_2 = \dot{x}$ as before but now we introduce $x_3 = t$. Then $\dot{x}_3 = 1$ and so the equivalent system is

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \tfrac{1}{m}\left(-kx_1 - bx_2 + F\cos x_3\right) \\ \dot{x}_3 &= 1 \end{aligned} \tag{3}$$

which is an example of a *three*-dimensional system. Similarly, an nth-order time-dependent equation is a special case of an $(n+1)$-dimensional system. By this trick, we can always remove any time dependence by adding an extra dimension to the system.

The virtue of this change of variables is that it allows us to visualize a phase space with trajectories *frozen* in it. Otherwise, if we allowed explicit time dependence, the vectors and the trajectories would always be wiggling—this would ruin the geometric picture we're trying to build. A more physical motivation is that the ***state*** of the forced harmonic oscillator is truly three-dimensional: we need to know three numbers, x, $\dot{x}$, and t, to predict the future, given the present. So a three-dimensional phase space is natural.

The cost, however, is that some of our terminology is nontraditional. For example, the forced harmonic oscillator would traditionally be regarded as a second-order linear equation, whereas we will regard it as a third-order nonlinear system, since (3) is nonlinear, thanks to the cosine term. As we'll see later in the book, forced oscillators have many of the properties associated with nonlinear systems, and so there are genuine conceptual advantages to our choice of language.

Why Are Nonlinear Problems So Hard?

As we've mentioned earlier, most nonlinear systems are impossible to solve analytically. Why are nonlinear systems so much harder to analyze than linear ones? The essential difference is that *linear systems can be broken down into parts*. Then each part can be solved separately and finally recombined to get the answer. This idea allows a fantastic simplification of complex problems, and underlies such methods as normal modes, Laplace transforms, superposition arguments, and Fourier analysis. In this sense, a linear system is precisely equal to the sum of its parts.

But many things in nature don't act this way. Whenever parts of a system interfere, or cooperate, or compete, there are nonlinear interactions going on. Most of everyday life is nonlinear, and the principle of superposition fails spectacularly. If you listen to your two favorite songs at the same time, you won't get double the pleasure! Within the realm of physics, nonlinearity is vital to the operation of a laser, the formation of turbulence in a fluid, and the superconductivity of Josephson junctions.

1.3 A Dynamical View of the World

Now that we have established the ideas of nonlinearity and phase space, we can present a framework for dynamics and its applications. Our goal is to show the logical structure of the entire subject. The framework presented in Figure 1.3.1 will guide our studies thoughout this book.

The framework has two axes. One axis tells us the number of variables needed to characterize the state of the system. Equivalently, this number is the *dimension of the phase space*. The other axis tells us whether the system is linear or *nonlinear*.

For example, consider the exponential growth of a population of organisms. This system is described by the first-order differential equation

$$\dot{x} = rx$$

where x is the population at time t and r is the growth rate. We place this system in the column labeled "$n = 1$" because *one* piece of information—the current value of the population x—is sufficient to predict the population at any later time. The system is also classified as linear because the differential equation $\dot{x} = rx$ is linear in x.

As a second example, consider the swinging of a pendulum, governed by

$$\ddot{x} + \frac{g}{L}\sin x = 0.$$

In contrast to the previous example, the state of this system is given by *two* variables: its current angle x and angular velocity $\dot{x}$. (Think of it this way: we need the initial values of both x and $\dot{x}$ to determine the solution uniquely. For example, if we knew only x, we wouldn't know which way the pendulum was swinging.) Because two variables are needed to specify the state, the pendulum belongs in the $n = 2$ column of Figure 1.3.1. Moreover, the system is nonlinear, as discussed in the previous section. Hence the pendulum is in the lower, nonlinear half of the $n = 2$ column.

Number of variables →

Nonlinearity ↓	$n = 1$	$n = 2$	$n \geq 3$	$n >> 1$	Continuum
Linear	*Growth, decay, or equilibrium* Exponential growth RC circuit Radioactive decay	*Oscillations* Linear oscillator Mass and spring RLC circuit 2-body problem (Kepler, Newton)	Civil engineering, structures Electrical engineering	*Collective phenomena* Coupled harmonic oscillators Solid-state physics Molecular dynamics Equilibrium statistical mechanics	*Waves and patterns* Elasticity Wave equations Electromagnetism (Maxwell) Quantum mechanics (Schrödinger, Heisenberg, Dirac) Heat and diffusion Acoustics Viscous fluids
				The frontier	
Nonlinear	Fixed points Bifurcations Overdamped systems, relaxational dynamics Logistic equation for single species	Pendulum Anharmonic oscillators Limit cycles Biological oscillators (neurons, heart cells) Predator-prey cycles Nonlinear electronics (van der Pol, Josephson)	*Chaos* Strange attractors (Lorenz) 3-body problem (Poincaré) Chemical kinetics Iterated maps (Feigenbaum) Fractals (Mandelbrot) Forced nonlinear oscillators (Levinson, Smale) Practical uses of chaos Quantum chaos ?	Coupled nonlinear oscillators Lasers, nonlinear optics Nonequilibrium statistical mechanics Nonlinear solid-state physics (semiconductors) Josephson arrays Heart cell synchronization Neural networks Immune system Ecosystems Economics	*Spatio-temporal complexity* Nonlinear waves (shocks, solitons) Plasmas Earthquakes General relativity (Einstein) Quantum field theory Reaction-diffusion, biological and chemical waves Fibrillation Epilepsy Turbulent fluids (Navier-Stokes) Life

Figure 1.3.1

One can continue to classify systems in this way, and the result will be something like the framework shown here. Admittedly, some aspects of the picture are debatable. You might think that some topics should be added, or placed differently, or even that more axes are needed—the point is to think about classifying systems on the basis of their dynamics.

There are some striking patterns in Figure 1.3.1. All the simplest systems occur in the upper left-hand corner. These are the small linear systems that we learn about in the first few years of college. Roughly speaking, these linear systems exhibit growth, decay, or equilibrium when $n = 1$, or oscillations when $n = 2$. The italicized phrases in Figure 1.3.1 indicate that these broad classes of phenomena first arise in this part of the diagram. For example, an RC circuit has $n = 1$ and cannot oscillate, whereas an RLC circuit has $n = 2$ and can oscillate.

The next most familiar part of the picture is the upper right-hand corner. This is the domain of classical applied mathematics and mathematical physics where the linear partial differential equations live. Here we find Maxwell's equations of electricity and magnetism, the heat equation, Schrödinger's wave equation in quantum mechanics, and so on. These partial differential equations involve an infinite "continuum" of variables because each point in space contributes additional degrees of freedom. Even though these systems are large, they are tractable, thanks to such linear techniques as Fourier analysis and transform methods.

In contrast, the lower half of Figure 1.3.1—the nonlinear half—is often ignored or deferred to later courses. But no more! In this book we start in the lower left corner and systematically head to the right. As we increase the phase space dimension from $n = 1$ to $n = 3$, we encounter new phenomena at every step, from fixed points and bifurcations when $n = 1$, to nonlinear oscillations when $n = 2$, and finally chaos and fractals when $n = 3$. In all cases, a geometric approach proves to be very powerful, and gives us most of the information we want, even though we usually can't solve the equations in the traditional sense of finding a formula for the answer. Our journey will also take us to some of the most exciting parts of modern science, such as mathematical biology and condensed-matter physics.

You'll notice that the framework also contains a region forbiddingly marked "The frontier." It's like in those old maps of the world, where the mapmakers wrote, "Here be dragons" on the unexplored parts of the globe. These topics are not completely unexplored, of course, but it is fair to say that they lie at the limits of current understanding. The problems are very hard, because they are both large and nonlinear. The resulting behavior is typically complicated in *both space and time,* as in the motion of a turbulent fluid or the patterns of electrical activity in a fibrillating heart. Toward the end of the book we will touch on some of these problems—they will certainly pose challenges for years to come.

PART 1

ONE-DIMENSIONAL FLOWS

2

FLOWS ON THE LINE

2.0 Introduction

In Chapter 1, we introduced the general system

$$\begin{aligned}\dot{x}_1 &= f_1(x_1, \ldots, x_n)\\ &\vdots\\ \dot{x}_n &= f_n(x_1, \ldots, x_n)\end{aligned}$$

and mentioned that its solutions could be visualized as trajectories flowing through an n-dimensional phase space with coordinates $(x_1, \ldots, x_n)$. At the moment, this idea probably strikes you as a mind-bending abstraction. So let's start slowly, beginning here on earth with the simple case $n = 1$. Then we get a single equation of the form

$$\dot{x} = f(x).$$

Here $x(t)$ is a real-valued function of time t, and $f(x)$ is a smooth real-valued function of x. We'll call such equations ***one-dimensional*** or ***first-order systems***.

Before there's any chance of confusion, let's dispense with two fussy points of terminology:

1. The word *system* is being used here in the sense of a dynamical system, not in the classical sense of a collection of two or more equations. Thus a single equation can be a "system."
2. We do not allow f to depend explicitly on time. Time-dependent or "nonautonomous" equations of the form $\dot{x} = f(x,t)$ are more complicated, because one needs *two* pieces of information, x and t, to predict the future state of the system. Thus $\dot{x} = f(x,t)$ should really be regarded as a *two-dimensional* or *second-order* system, and will therefore be discussed later in the book.

2.1 A Geometric Way of Thinking

Pictures are often more helpful than formulas for analyzing nonlinear systems. Here we illustrate this point by a simple example. Along the way we will introduce one of the most basic techniques of dynamics: *interpreting a differential equation as a vector field.*

Consider the following nonlinear differential equation:

$$\dot{x} = \sin x. \qquad (1)$$

To emphasize our point about formulas versus pictures, we have chosen one of the few nonlinear equations that can be solved in closed form. We separate the variables and then integrate:

$$dt = \frac{dx}{\sin x},$$

which implies

$$t = \int \csc x \, dx$$

$$= -\ln|\csc x + \cot x| + C.$$

To evaluate the constant C, suppose that $x = x_0$ at $t = 0$. Then $C = \ln|\csc x_0 + \cot x_0|$. Hence the solution is

$$t = \ln\left|\frac{\csc x_0 + \cot x_0}{\csc x + \cot x}\right|. \qquad (2)$$

This result is exact, but a headache to interpret. For example, can you answer the following questions?

1. Suppose $x_0 = \pi/4$; describe the qualitative features of the solution $x(t)$ for all $t > 0$. In particular, what happens as $t \to \infty$?
2. For an *arbitrary* initial condition x_0, what is the behavior of $x(t)$ as $t \to \infty$?

Think about these questions for a while, to see that formula (2) is not transparent.

In contrast, a graphical analysis of (1) is clear and simple, as shown in Figure 2.1.1. We think of t as time, x as the position of an imaginary particle moving along the real line, and $\dot{x}$ as the velocity of that particle. Then the differential equation $\dot{x} = \sin x$ represents a ***vector field*** on the line: it dictates the velocity vector $\dot{x}$ at each x. To sketch the vector field, it is convenient to plot $\dot{x}$ versus x, and then draw arrows on the x-axis to indicate the corresponding velocity vector at each x. The arrows point to the right when $\dot{x} > 0$ and to the left when $\dot{x} < 0$.

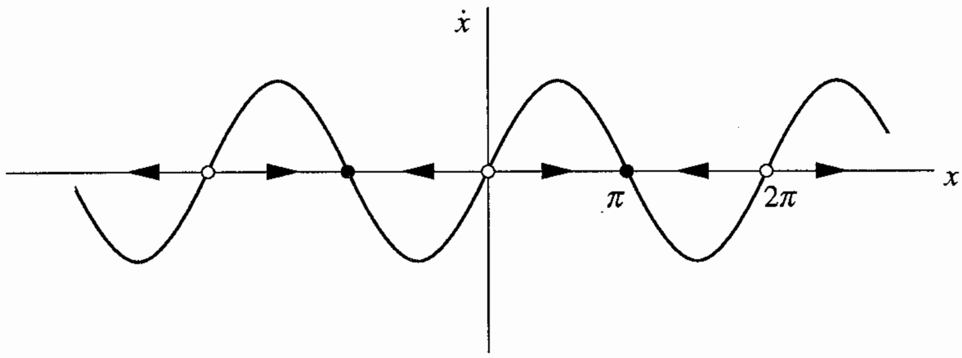

Figure 2.1.1

Here's a more physical way to think about the vector field: imagine that fluid is flowing steadily along the x-axis with a velocity that varies from place to place, according to the rule $\dot{x} = \sin x$. As shown in Figure 2.1.1, the ***flow*** is to the right when $\dot{x} > 0$ and to the left when $\dot{x} < 0$. At points where $\dot{x} = 0$, there is no flow; such points are therefore called ***fixed points***. You can see that there are two kinds of fixed points in Figure 2.1.1: solid black dots represent ***stable*** fixed points (often called *attractors* or *sinks,* because the flow is toward them) and open circles represent ***unstable*** fixed points (also known as *repellers* or *sources*).

Armed with this picture, we can now easily understand the solutions to the differential equation $\dot{x} = \sin x$. We just start our imaginary particle at x_0 and watch how it is carried along by the flow.

This approach allows us to answer the questions above as follows:

1. Figure 2.1.1 shows that a particle starting at $x_0 = \pi/4$ moves to the right faster and faster until it crosses $x = \pi/2$ (where $\sin x$ reaches its maximum). Then the particle starts slowing down and eventually approaches the stable fixed point $x = \pi$ from the left. Thus, the qualitative form of the solution is as shown in Figure 2.1.2.

   Note that the curve is concave up at first, and then concave down; this corresponds to the initial acceleration for $x < \pi/2$, followed by the deceleration toward $x = \pi$.

2. The same reasoning applies to any initial condition x_0. Figure 2.1.1 shows that if $\dot{x} > 0$ initially, the particle heads to the right and asymptotically approaches the nearest stable fixed point. Similarly, if $\dot{x} < 0$ initially, the particle approaches the nearest stable fixed point to its left. If $\dot{x} = 0$, then x remains constant. The qualitative form of the solution for any initial condition is sketched in Figure 2.1.3.

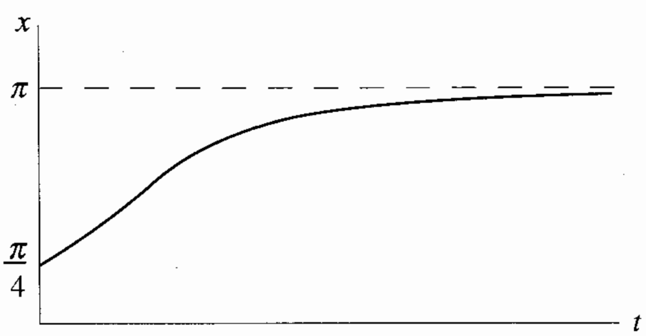

Figure 2.1.2

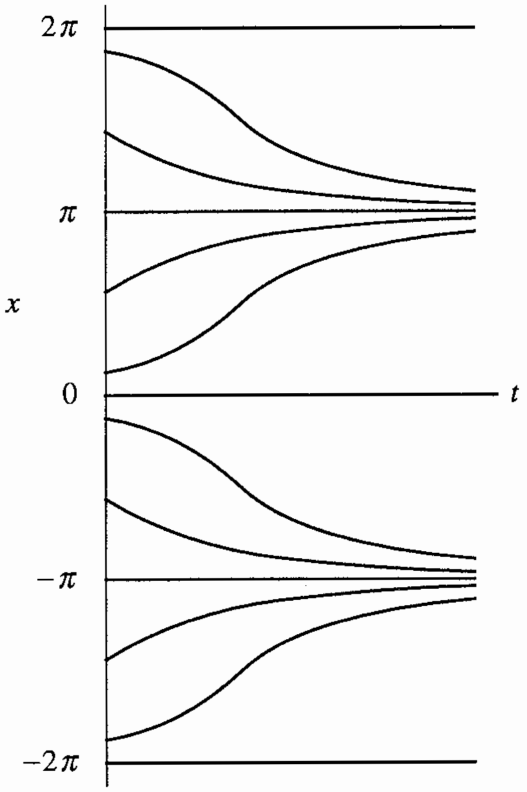

Figure 2.1.3

In all honesty, we should admit that a picture can't tell us certain *quantitative* things: for instance, we don't know the time at which the speed $|\dot{x}|$ is greatest. But in many cases *qualitative* information is what we care about, and then pictures are fine.

2.2 Fixed Points and Stability

The ideas developed in the last section can be extended to any one-dimensional system $\dot{x} = f(x)$. We just need to draw the graph of $f(x)$ and then use it to sketch the vector field on the real line (the x-axis in Figure 2.2.1).

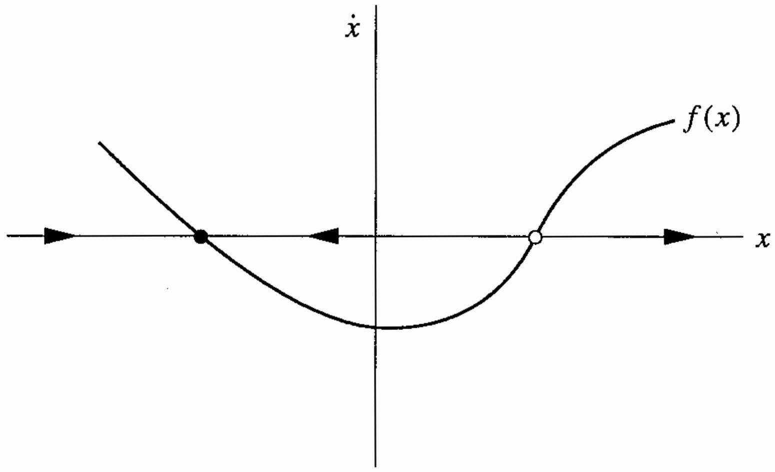

Figure 2.2.1

As before, we imagine that a fluid is flowing along the real line with a local velocity $f(x)$. This imaginary fluid is called the phase fluid, and the real line is the phase space. The flow is to the right where $f(x) > 0$ and to the left where $f(x) < 0$. To find the solution to $\dot{x} = f(x)$ starting from an arbitrary initial condition x_0, we place an imaginary particle (known as a ***phase point***) at x_0 and watch how it is carried along by the flow. As time goes on, the phase point moves along the x-axis according to some function $x(t)$. This function is called the ***trajectory*** based at x_0, and it represents the solution of the differential equation starting from the initial condition x_0. A picture like Figure 2.2.1, which shows all the qualitatively different trajectories of the system, is called a ***phase portrait.***

The appearance of the phase portrait is controlled by the fixed points x^*, defined by $f(x^*) = 0$; they correspond to stagnation points of the flow. In Figure 2.2.1, the solid black dot is a stable fixed point (the local flow is toward it) and the open dot is an unstable fixed point (the flow is away from it).

In terms of the original differential equation, fixed points represent ***equilibrium*** solutions (sometimes called steady, constant, or rest solutions, since if $x = x^*$ initially, then $x(t) = x^*$ for all time). An equilibrium is defined to be stable if all sufficiently small disturbances away from it damp out in time. Thus stable equilibria are represented geometrically by stable fixed points. Conversely, unstable equilibria, in which disturbances grow in time, are represented by unstable fixed points.

EXAMPLE 2.2.1:

Find all fixed points for $\dot{x} = x^2 - 1$, and classify their stability.

Solution: Here $f(x) = x^2 - 1$. To find the fixed points, we set $f(x^*) = 0$ and solve for x^*. Thus $x^* = \pm 1$. To determine stability, we plot $x^2 - 1$ and then sketch the vector field (Figure 2.2.2). The flow is to the right where $x^2 - 1 > 0$ and to the left where $x^2 - 1 < 0$. Thus $x^* = -1$ is stable, and $x^* = 1$ is unstable. ■

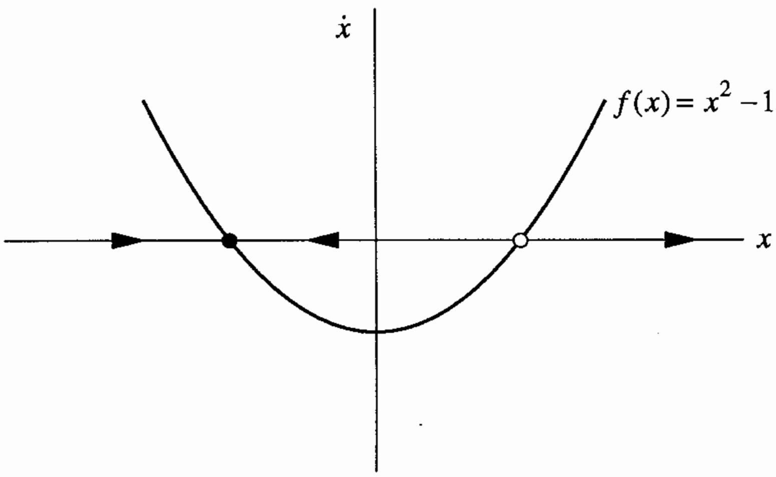

Figure 2.2.2

Note that the definition of stable equilibrium is based on *small* disturbances; certain large disturbances may fail to decay. In Example 2.2.1, all small disturbances to $x^* = -1$ will decay, but a large disturbance that sends x to the right of $x = 1$ will *not* decay—in fact, the phase point will be repelled out to $+\infty$. To emphasize this aspect of stability, we sometimes say that $x^* = -1$ is ***locally stable***, but not globally stable.

EXAMPLE 2.2.2:

Consider the electrical circuit shown in Figure 2.2.3. A resistor R and a capacitor C are in series with a battery of constant dc voltage V_0. Suppose that the switch is closed at $t = 0$, and that there is no charge on the capacitor initially. Let $Q(t)$ denote the charge on the capacitor at time $t \geq 0$. Sketch the graph of $Q(t)$.

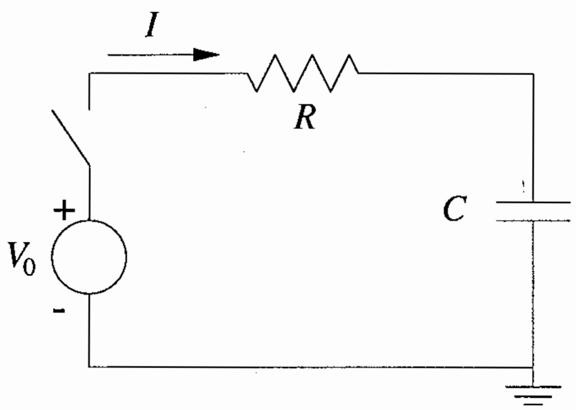

Figure 2.2.3

Solution: This type of circuit problem is probably familiar to you. It is governed by linear equations and can be solved analytically, but we prefer to illustrate the geometric approach.

First we write the circuit equations. As we go around the circuit, the total voltage drop must equal zero; hence $-V_0 + RI + Q/C = 0$, where I is the current flowing through the resistor. This current causes charge to accumulate on the capacitor at a rate $\dot{Q} = I$. Hence

$$-V_0 + R\dot{Q} + Q/C = 0 \quad \text{or}$$

$$\dot{Q} = f(Q) = \frac{V_0}{R} - \frac{Q}{RC}.$$

The graph of $f(Q)$ is a straight line with a negative slope (Figure 2.2.4). The corresponding vector field has a fixed point where $f(Q) = 0$, which occurs at $Q^* = CV_0$. The flow is to the right where $f(Q) > 0$ and to the left where $f(Q) < 0$. Thus the flow is always toward Q^*—it is a *stable* fixed point. In fact, it is ***globally stable***, in the sense that it is approached from *all* initial conditions.

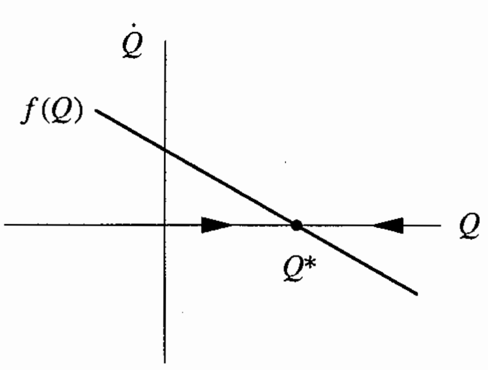

Figure 2.2.4

To sketch $Q(t)$, we start a phase point at the origin of Figure 2.2.4 and imagine how it would move. The flow carries the phase point monotonically toward Q^*. Its speed

$\dot{Q}$ decreases linearly as it approaches the fixed point; therefore $Q(t)$ is increasing and concave down, as shown in Figure 2.2.5. ■

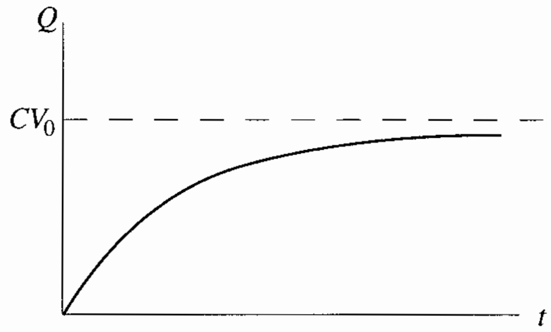

Figure 2.2.5

EXAMPLE 2.2.3:

Sketch the phase portrait corresponding to $\dot{x} = x - \cos x$, and determine the stability of all the fixed points.

Solution: One approach would be to plot the function $f(x) = x - \cos x$ and then sketch the associated vector field. This method is valid, but it requires you to figure out what the graph of $x - \cos x$ looks like.

There's an easier solution, which exploits the fact that we know how to graph $y = x$ and $y = \cos x$ *separately*. We plot both graphs on the same axes and then observe that they intersect in exactly one point (Figure 2.2.6).

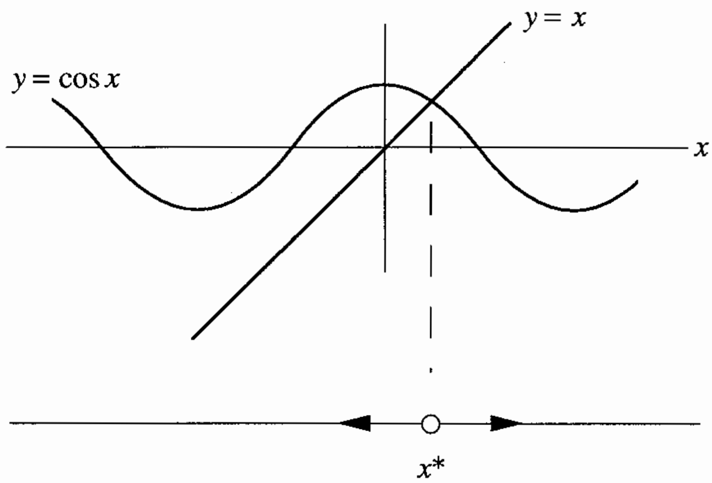

Figure 2.2.6

This intersection corresponds to a fixed point, since $x^* = \cos x^*$ and therefore $f(x^*) = 0$. Moreover, when the line lies above the cosine curve, we have $x > \cos x$ and so $\dot{x} > 0$: the flow is to the right. Similarly, the flow is to the left where the line is below the cosine curve. Hence x^* is the only fixed point, and it is unstable. Note that we can classify the stability of x^*, even though we don't have a formula for x^* itself! ■

2.3 Population Growth

The simplest model for the growth of a population of organisms is $\dot{N} = rN$, where $N(t)$ is the population at time t, and $r > 0$ is the growth rate. This model predicts exponential growth: $N(t) = N_0 e^{rt}$, where N_0 is the population at $t = 0$.

Of course such exponential growth cannot go on forever. To model the effects of overcrowding and limited resources, population biologists and demographers often assume that the per capita growth rate $\dot{N}/N$ decreases when N becomes sufficiently large, as shown in Figure 2.3.1. For small N, the growth rate equals r, just as before. However, for populations larger than a certain ***carrying capacity*** K, the growth rate actually becomes negative; the death rate is higher than the birth rate.

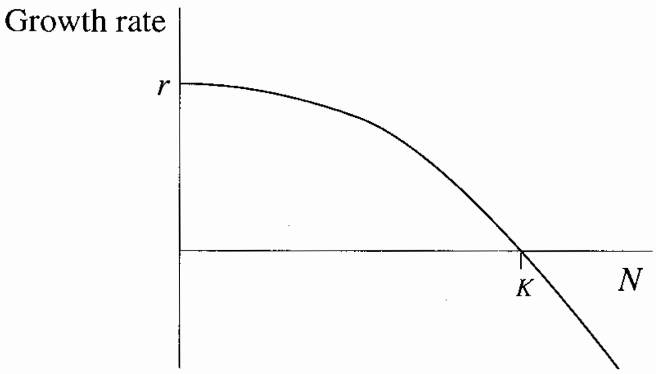

Figure 2.3.1

A mathematically convenient way to incorporate these ideas is to assume that the per capita growth rate $\dot{N}/N$ decreases *linearly* with N (Figure 2.3.2).

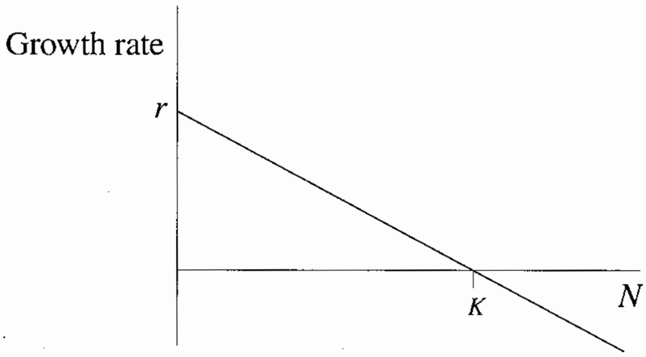

Figure 2.3.2

This leads to the ***logistic equation***

$$\dot{N} = rN\left(1 - \frac{N}{K}\right)$$

first suggested to describe the growth of human populations by Verhulst in 1838. This equation can be solved analytically (Exercise 2.3.1) but once again we prefer a graphical approach. We plot $\dot{N}$ versus N to see what the vector field looks like. Note that we plot only $N \geq 0$, since it makes no sense to think about a negative population (Figure 2.3.3). Fixed points occur at $N^* = 0$ and $N^* = K$, as found by setting $\dot{N} = 0$ and solving for N. By looking at the flow in Figure 2.3.3, we see that $N^* = 0$ is an unstable fixed point and $N^* = K$ is a stable fixed point. In biological terms, $N = 0$ is an unstable equilibrium: a small population will grow exponentially fast and run away from $N = 0$. On the other hand, if N is disturbed slightly from K, the disturbance will decay monotonically and $N(t) \to K$ as $t \to \infty$.

In fact, Figure 2.3.3 shows that if we start a phase point at *any* $N_0 > 0$, it will always flow toward $N = K$. Hence *the population always approaches the carrying capacity.*

The only exception is if $N_0 = 0$; then there's nobody around to start reproducing, and so $N = 0$ for all time. (The model does not allow for spontaneous generation!)