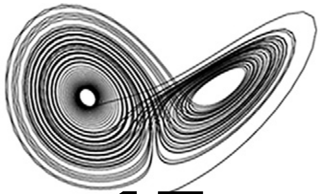


15
Discrete Dynamical Systems

Our goal in this chapter is to begin the study of discrete dynamical systems. As we have seen at several stages in this book, it is sometimes possible to reduce the study of the flow of a differential equation to that of an iterated function, namely, a Poincaré map. This reduction has several advantages. First and foremost, the Poincaré map lives on a lower dimensional space, which therefore makes visualization easier. Secondly, we do not have to integrate to find "solutions" of discrete systems. Rather, given the function, we simply iterate the function over and over to determine the behavior of the orbit, which then dictates the behavior of the corresponding solution.

Given these two simplifications, it then becomes much easier to comprehend the complicated chaotic behavior that often arises for systems of differential equations. While the study of discrete dynamical systems is a topic that could easily fill this entire book, we will restrict attention here primarily to the portion of this theory that helps us understand chaotic behavior in one dimension. In the following chapter we will extend these ideas to higher dimensions.

15.1 Introduction to Discrete Dynamical Systems

Throughout this chapter we will work with real functions $f: \mathbb{R} \to \mathbb{R}$. As usual, we assume throughout that f is C^∞, although there will be several special examples where this is not the case.

Let f^n denote the nth iterate of f. That is, f^n is the n-fold composition of f with itself. Given $x_0 \in \mathbb{R}$, the *orbit* of x_0 is the sequence

$$x_0,\ x_1 = f(x_0),\ x_2 = f^2(x_0), \ldots,\ x_n = f^n(x_0), \ldots.$$

The point x_0 is called the *seed* of the orbit.

Example. Let $f(x) = x^2 + 1$. Then the orbit of the seed 0 is the sequence

$$\begin{aligned}
x_0 &= 0 \\
x_1 &= 1 \\
x_2 &= 2 \\
x_3 &= 5 \\
x_4 &= 26 \\
&\vdots \\
x_n &= \text{big} \\
x_{n+1} &= \text{bigger} \\
&\vdots
\end{aligned}$$

and so forth, so we see that this orbit tends to ∞ as $n \to \infty$. ■

In analogy with equilibrium solutions of systems of differential equations, *fixed points* play a central role in discrete dynamical systems. A point x_0 is called a fixed point if $f(x_0) = x_0$. Obviously, the orbit of a fixed point is the constant sequence $x_0, x_0, x_0, \ldots$.

The analog of closed orbits for differential equations is given by *periodic points of period n*. These are seeds x_0 for which $f^n(x_0) = x_0$ for some $n > 0$. As a consequence, like a closed orbit, a periodic orbit repeats itself:

$$x_0, x_1, \ldots, x_{n-1}, x_0, x_1, \ldots, x_{n-1}, x_0 \ldots.$$

Periodic orbits of period n are also called *n-cycles*. We say that the periodic point x_0 has *minimal period n* if n is the least positive integer for which $f^n(x_0) = x_0$.

Example. The function $f(x) = x^3$ has fixed points at $x = 0, \pm 1$. The function $g(x) = -x^3$ has a fixed point at 0 and a periodic point of period 2 at $x = \pm 1$, since $g(1) = -1$ and $g(-1) = 1$, so $g^2(\pm 1) = \pm 1$. The function

$$h(x) = (2 - x)(3x + 1)/2$$

has a 3-cycle given by $x_0 = 0, x_1 = 1, x_2 = 2, x_3 = x_0 = 0 \ldots$. ■

A useful way to visualize orbits of one-dimensional discrete dynamical systems is via *graphical iteration*. In this picture, we superimpose the curve $y = f(x)$ and the diagonal line $y = x$ on the same graph. We display the orbit of x_0 as follows: Begin at the point (x_0, x_0) on the diagonal and draw a vertical line to the graph of f, reaching the graph at $(x_0, f(x_0)) = (x_0, x_1)$. Then draw a horizontal line back to the diagonal, ending at (x_1, x_1). This procedure moves us from a point on the diagonal directly over the seed x_0 to a point directly over the next point on the orbit, x_1. Then we continue from (x_1, x_1): First go vertically to the graph to the point (x_1, x_2), then horizontally back to the diagonal at (x_2, x_2). On the x-axis this moves us from x_1 to the next point on the orbit, x_2. Continuing, we produce a sequence of pairs of lines, each of which terminates on the diagonal at a point of the form (x_n, x_n).

In Figure 15.1a, graphical iteration shows that the orbit of x_0 tends to the fixed point z_0 under iteration of f. In Figure 15.1b, the orbit of x_0 under g lies on a 3-cycle: $x_0, x_1, x_2, x_0, x_1, \ldots$.

As in the case of equilibrium points of differential equations, there are different types of fixed points for a discrete dynamical system. Suppose that x_0 is a fixed point for f. We say that x_0 is a *sink* or an *attracting fixed point* for f if there is a neighborhood $\mathcal{U}$ of x_0 in $\mathbb{R}$ having the property that, if $y_0 \in \mathcal{U}$, then

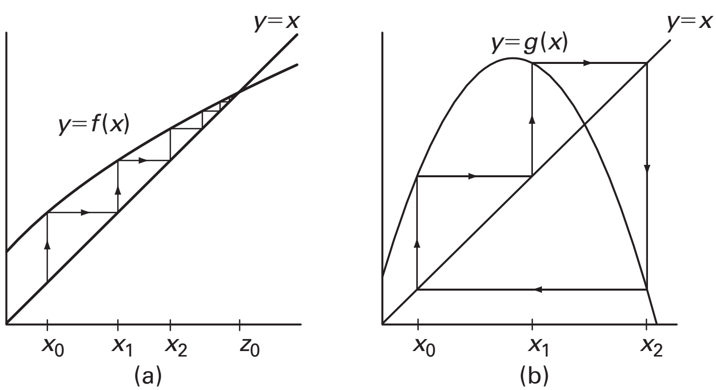


Figure 15.1 (a) The orbit of x_0 tends to the fixed point at z_0 under iteration of f, while (b) the orbit of x_0 lies on a 3-cycle under iteration of g.

$f^n(y_0) \in \mathcal{U}$ for all n and, moreover, $f^n(y_0) \to x_0$ as $n \to \infty$. Similarly, x_0 is a *source* or a *repelling fixed point* if all orbits (except x_0) leave $\mathcal{U}$ under iteration of f. A fixed point is called *neutral* or *indifferent* if it is neither attracting nor repelling.

For differential equations, we saw that it was the derivative of the vector field at an equilibrium point that determined the type of the equilibrium point. This is also true for fixed points, although the numbers change a bit.

Proposition. *Suppose f has a fixed point at x_0. Then*

1. *x_0 is a sink if $|f'(x_0)| < 1$;*
2. *x_0 is a source if $|f'(x_0)| > 1$;*
3. *we get no information about the type of x_0 if $f'(x_0) = \pm 1$.*

Proof: We first prove case (1). Suppose $|f'(x_0)| = \nu < 1$. Choose K with $\nu < K < 1$. Since f' is continuous, we may find $\delta > 0$ so that $|f'(x)| < K$ for all x in the interval $I = [x_0 - \delta, x_0 + \delta]$. We now invoke the mean value theorem. Given any $x \in I$, we have

$$\frac{f(x) - x_0}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} = f'(c)$$

for some c between x and x_0. Hence we have

$$|f(x) - x_0| < K|x - x_0|.$$

It follows that $f(x)$ is closer to x_0 than x and so $f(x) \in I$. Applying this result again, we have

$$|f^2(x) - x_0| < K|f(x) - x_0| < K^2|x - x_0|,$$

and, continuing, we find

$$|f^n(x) - x_0| < K^n|x - x_0|,$$

so that $f^n(x) \to x_0$ in I as required, since $0 < K < 1$.

The proof of case (2) follows similarly. In case (3), we note that each of the functions

1. $f(x) = x + x^3$;
2. $g(x) = x - x^3$;
3. $h(x) = x + x^2$

has a fixed point at 0 with $f'(0) = 1$. But graphical iteration (Figure 15.2) shows that f has a source at 0; g has a sink at 0; and 0 is attracting from one side and repelling from the other for the function h. ∎

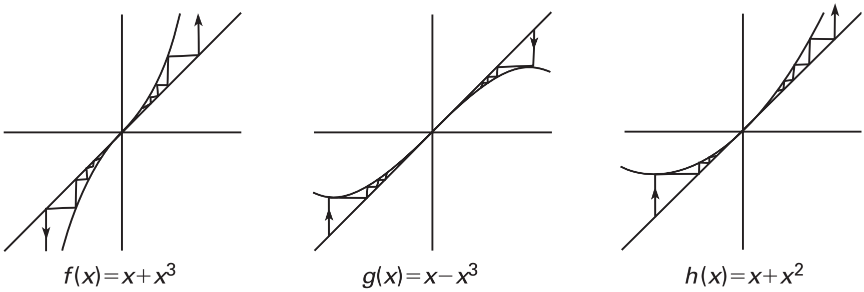


Figure 15.2 In each case, the derivative at 0 is 1, but f has a source at 0; g has a sink; and h has neither.

Note that, at a fixed point x_0 for which $f'(x_0) < 0$, the orbits of nearby points jump from one side of the fixed point to the other at each iteration. See Figure 15.3. This is the reason why the output of graphical iteration is often called a *web diagram.*

Since a periodic point x_0 of period n for f is a fixed point of f^n, we may classify these points as sinks or sources depending on whether $|(f^n)'(x_0)| < 1$ or $|(f^n)'(x_0)| > 1$. One may check that $(f^n)'(x_0) = (f^n)'(x_j)$ for any other point x_j on the periodic orbit, so this definition makes sense (see Exercise 6 at the end of this chapter).

Example. The function $f(x) = x^2 - 1$ has a 2-cycle given by 0 and -1. One checks easily that $(f^2)'(0) = 0 = (f^2)'(-1)$, so this cycle is a sink. In Figure 15.4, we show a graphical iteration of f with the graph of f^2 superimposed. Note that 0 and -1 are attracting fixed points for f^2. ■

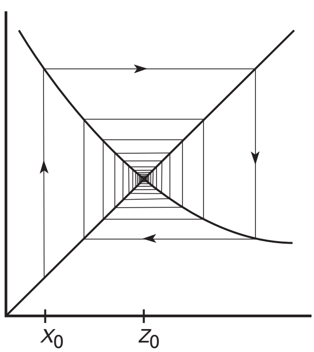


Figure 15.3 Since $-1 < f'(z_0) < 0$, the orbit of x_0 "spirals" toward the attracting fixed point at z_0.

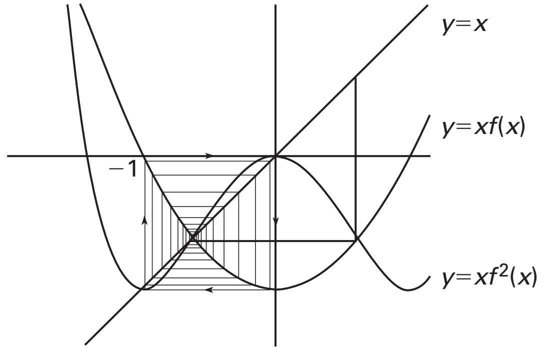


Figure 15.4 The graphs of $f(x) = x^2 - 1$ and f^2 showing that 0 and -1 lie on an attracting 2-cycle for f.

15.2 Bifurcations

Discrete dynamical systems undergo bifurcations when parameters are varied just as differential equations do. We deal in this section with several types of bifurcations that occur for one-dimensional systems.

Example. Let $f_c(x) = x^2 + c$ where c is a parameter. The fixed points for this family are given by solving the equation $x^2 + c = x$, which yields

$$p_\pm = \frac{1}{2} \pm \frac{\sqrt{1-4c}}{2}.$$

Hence there are no fixed points if $c > 1/4$; a single fixed point at $x = 1/2$ when $c = 1/4$; and a pair of fixed points at $p_\pm$ when $c < 1/4$. Graphical iteration shows that all orbits of f_c tend to ∞ if $c > 1/4$. When $c = 1/4$, the fixed point at $x = 1/2$ is neutral, as is easily seen by graphical iteration. See Figure 15.5. When $c < 1/4$, we have $f_c'(p_+) = 1 + \sqrt{1-4c} > 1$, so p_+ is always repelling. A straightforward computation also shows that $-1 < f_c'(p_-) < 1$ provided $-3/4 < c < 1/4$. For these c-values, p_- is attracting. When $-3/4 < c < 1/4$, all orbits in the interval $(-p_+, p_+)$ tend to p_- (though, technically, the orbit of $-p_-$ is *eventually fixed*, since it maps directly onto p_-, as do the orbits of certain other points in this interval when $c < 0$). Thus as c decreases through the bifurcation value $c = 1/4$, we see the birth of a single neutral fixed point, which then immediately splits into two fixed points, one attracting and one repelling. This is an example of a *saddle-node* or *tangent bifurcation*. Graphically, this bifurcation is essentially the same as its namesake for first-order differential equations as described in Chapter 8. See Figure 15.5. ■

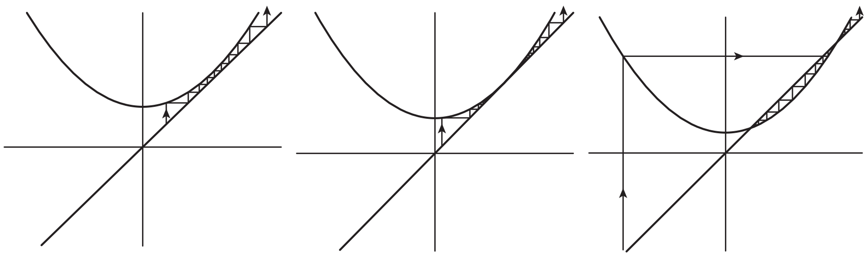


Figure 15.5 The saddle-node bifurcation for $f_{c(x)} = x^2 + c$ at $c = 1/4$.

Note that, in this example, at the bifurcation point, the derivative at the fixed point equals 1. This is no accident, for we have:

Theorem. (The Bifurcation Criterion) *Let f_λ be a family of functions depending smoothly on the parameter λ. Suppose that $f_{\lambda_0}(x_0) = x_0$ and $f'_{\lambda_0}(x_0) \neq 1$. Then there are intervals I about x_0 and J about λ_0 and a smooth function $p\colon J \to I$ such that $p(\lambda_0) = x_0$ and $f_\lambda(p(\lambda)) = p(\lambda)$. Moreover, f_λ has no other fixed points in I.*

Proof: Consider the function defined by $G(x, \lambda) = f_\lambda(x) - x$. By hypothesis, $G(x_0, \lambda_0) = 0$ and

$$\frac{\partial G}{\partial x}(x_0, \lambda_0) = f'_{\lambda_0} x_0) - 1 \neq 0.$$

By the implicit function theorem, there are intervals I about x_0 and J about λ_0, and a smooth function $p\colon J \to I$ such that $p(\lambda_0) = x_0$ and $G(p(\lambda), \lambda) \equiv 0$ for all $\lambda \in J$. Moreover, $G(x, \lambda) \neq 0$ unless $x = p(\lambda)$. This concludes the proof. ■

As a consequence of this result, f_λ may undergo a bifurcation involving a change in the number of fixed points only if f_λ has a fixed point with derivative equal to 1. The typical bifurcation that occurs at such parameter values is the saddle-node bifurcation (see Exercises 18 and 19). However, many other types of bifurcations of fixed points may occur.

Example. Let $f_\lambda(x) = \lambda x(1 - x)$. Note that $f_\lambda(0) = 0$ for all λ. We have $f'_\lambda(0) = \lambda$, so we have a possible bifurcation at $\lambda = 1$. There is a second fixed point for f_λ at $x_\lambda = (\lambda - 1)/\lambda$. When $\lambda < 1$, x_λ is negative, and when $\lambda > 1$, x_λ is positive. When $\lambda = 1$, x_λ coalesces with the fixed point at 0 so there is a single fixed point for f_1. A computation shows that 0 is repelling and x_λ is attracting if $\lambda > 1$ (and $\lambda < 3$), while the reverse is true if $\lambda < 1$. For this reason, this type of bifurcation is known as an *exchange bifurcation*. ■

Example. Consider the family of functions $f_\mu(x) = \mu x + x^3$. When $\mu = 1$ we have $f_1(0) = 0$ and $f_1'(0) = 1$ so we have the possibility for a bifurcation. The fixed points are 0 and $\pm\sqrt{1-\mu}$, so we have three fixed points when $\mu < 1$ but only one fixed point when $\mu \geq 1$, so a bifurcation does indeed occur as μ passes through 1. ■

The only other possible bifurcation value for a one-dimensional discrete system occurs when the derivative at the fixed (or periodic) point is equal to -1, since at these values the fixed point may change from a sink to a source or from a source to a sink. At all other values of the derivative, the fixed point simply remains a sink or source and there are no other periodic orbits nearby. Certain portions of a periodic orbit may come close to a source, but the entire orbit cannot lie close by (see Exercise 7). In the case of derivative -1 at the fixed point, the typical bifurcation is a *period doubling* bifurcation.

Example. As a simple example of this type of bifurcation, consider the family $f_\lambda(x) = \lambda x$ near $\lambda_0 = -1$. There is a fixed point at 0 for all λ. When $-1 < \lambda < 1$, 0 is an attracting fixed point and all orbits tend to 0. When $|\lambda| > 1$, 0 is repelling and all nonzero orbits tend to $\pm\infty$. When $\lambda = -1$, 0 is a neutral fixed point and all nonzero points lie on 2-cycles. As λ passes through -1, the type of the fixed point changes from attracting to repelling; meanwhile, a family of 2-cycles appears. ■

Generally, when a period doubling bifurcation occurs, the 2-cycles do not all exist for a single parameter value. A more typical example of this bifurcation is provided next.

Example. Again consider $f_c(x) = x^2 + c$, this time with c near $c = -3/4$. There is a fixed point at

$$p_- = \frac{1}{2} - \frac{\sqrt{1-4c}}{2}.$$

We have seen that $f'_{-3/4}(p_-) = -1$ and that p_- is attracting when c is slightly larger than $-3/4$ and repelling when c is less than $-3/4$. Graphical iteration shows that more happens as c descends through $-3/4$: We see the birth of an (attracting) 2-cycle as well. This is the period doubling bifurcation. See Figure 15.6. Indeed, one can easily solve for the period two points and check that they are attracting (for $-5/4 < c < -3/4$; see Exercise 8). ■

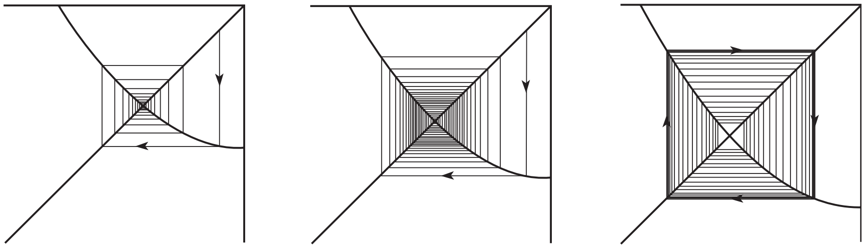


Figure 15.6 The period doubling bifurcation for $f_{c(x)} = x^2 + c$ at $c = -3/4$. The fixed point is attracting for $c \geq -0.75$ and repelling for $c < -0.75$.

15.3 The Discrete Logistic Model

In Chapter 1 we introduced one of the simplest nonlinear first-order differential equations, the logistic model for population growth

$$x' = ax(1 - x).$$

In this model we took into account the fact that there is a carrying capacity for a typical population, and we saw that the resulting solutions behaved quite simply: All nonzero solutions tended to the "ideal" population. Now something about this model may have bothered you way back then: Populations generally are not continuous functions of time! A more natural type of model would measure populations at specific times, say, every year or every generation. Here we introduce just such a model, the *discrete logistic model* for population growth.

Suppose we consider a population whose members are counted each year (or at other specified times). Let x_n denote the population at the end of year n. If we assume that no overcrowding can occur, then one such population model is the *exponential growth* model where we assume that

$$x_{n+1} = kx_n$$

for some constant $k > 0$. That is, the next year's population is directly proportional to this year's. Thus we have

$$\begin{aligned} x_1 &= kx_0 \\ x_2 &= kx_1 = k^2x_0 \\ x_3 &= kx_2 = k^3x_0 \\ &\vdots \end{aligned}$$

Clearly, $x_n = k^n x_0$ so we conclude that the population explodes if $k > 1$, becomes extinct if $0 \le k < 1$, or remains constant if $k = 1$.

This is an example of a first-order *difference equation*. This is an equation that determines x_n based on the value of x_{n-1}. A second-order difference equation would give x_n based on x_{n-1} and x_{n-2}. From our point of view, the successive populations are given by simply iterating the function $f_k(x) = kx$ with the seed x_0.

A more realistic assumption about population growth is that there is a maximal population M such that, if the population exceeds this amount, then all resources are used up and the entire population dies out in the next year.

One such model that reflects these assumptions is the *discrete logistic population model*. Here we assume that the populations obey the rule

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

where k and M are positive parameters. Note that, if $x_n \ge M$, then $x_{n+1} \le 0$, so the population does indeed die out in the ensuing year.

Rather than deal with actual population numbers, we will instead let x_n denote the fraction of the maximal population, so that $0 \le x_n \le 1$. The logistic difference equation then becomes

$$x_{n+1} = \lambda x_n (1 - x_n)$$

where $\lambda > 0$ is a parameter. We may therefore predict the fate of the initial population x_0 by simply iterating the quadratic function $f_\lambda(x) = \lambda x(1 - x)$ (also called the *logistic map*). Sounds easy, right? Well, suffice it to say that this simple quadratic iteration was only completely understood in the late 1990s, thanks to the work of hundreds of mathematicians. We will see why the discrete logistic model is so much more complicated than its cousin, the logistic differential equation, in a moment, but first let's do some simple cases.

We consider only the logistic map on the unit interval I. We have $f_\lambda(0) = 0$, so 0 is a fixed point. The fixed point is attracting in I for $0 < \lambda \le 1$, and repelling thereafter. The point 1 is eventually fixed, since $f_\lambda(1) = 0$. There is a second fixed point $x_\lambda = (\lambda - 1)/\lambda$ in I for $\lambda > 1$. The fixed point x_λ is attracting for $1 < \lambda \le 3$ and repelling for $\lambda > 3$. At $\lambda = 3$ a period doubling bifurcation occurs (see Exercise 4). For λ-values between 3 and approximately 3.4, the only periodic points present are the two fixed points and the 2-cycle.

When $\lambda = 4$, the situation is much more complicated. Note that $f_\lambda'(1/2) = 0$ and that 1/2 is the only critical point for f_λ for each λ. When $\lambda = 4$, we have $f_4(1/2) = 1$, so $f_4^2(1/2) = 0$. Therefore f_4 maps each of the half-intervals $[0, 1/2]$ and $[1/2, 1]$ onto the entire interval I. Consequently, there exist points $y_0 \in [0, 1/2]$ and $y_1 \in [1/2, 1]$ such that $f_4(y_j) = 1/2$ and hence $f_4^2(y_j) = 1$.

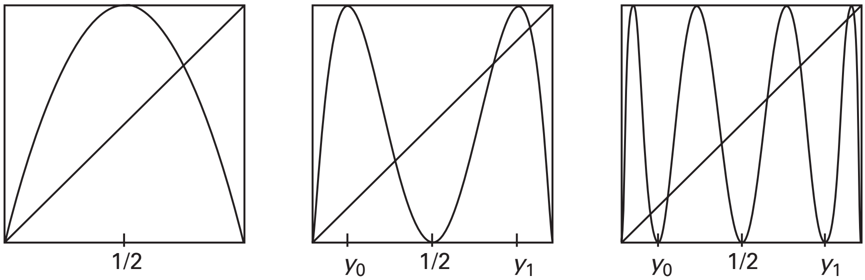


Figure 15.7 The graphs of the logistic function $f_{\lambda}(x) - \lambda\, x(1 - x)$ as well as f_{λ}^{2} and f_{λ}^{3} over the interval I.

Therefore we have

$$f_4^2[0, y_0] = f_4^2[y_0, 1/2] = I$$

and

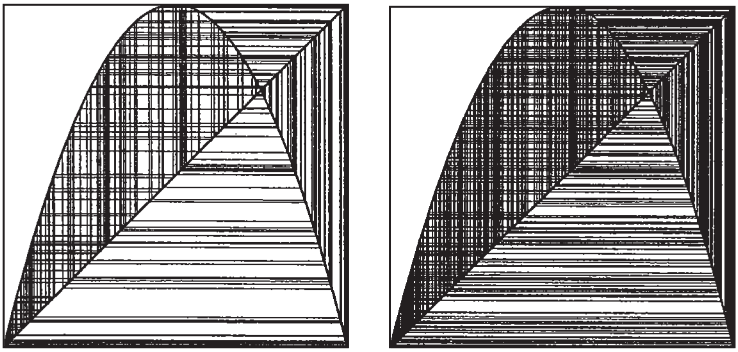
$$f_4^2[1/2, y_1] = f_4^2[y_1, 1] = I.$$

Since the function f_4^2 is a quartic, it follows that the graph of f_4^2 is as depicted in Figure 15.7. Continuing in this fashion, we find 2^3 subintervals of I that are mapped onto I by f_4^3, 2^4 subintervals mapped onto I by f_4^4, and so forth. We therefore see that f_4 has two fixed points in I; f_4^2 has four fixed points in I; f_4^3 has 2^3 fixed points in I; and, inductively, f_4^n has 2^n fixed points in I. The fixed points for f_4 occur at 0 and 3/4. The four fixed points for f_4^2 include these two fixed points plus a pair of periodic points of period 2. Of the eight fixed points for f_4^3, two must be the fixed points and the other six must lie on a pair of 3-cycles. Among the 16 fixed points for f_4^4 are two fixed points, two periodic points of period 2, and twelve periodic points of period 4. Clearly, a lot has changed as λ varies from 3. 4 to 4.

On the other hand, if we choose a random seed in the interval I and plot the orbit of this seed under iteration of f_4 using graphical iteration, we rarely see any of these cycles. In Figure 15.8 we have plotted the orbit of 0. 123 under iteration of f_4 using 200 and 500 iterations. Presumably, there is something "chaotic" going on.

15.4 Chaos

In this section we introduce several quintessential examples of chaotic one-dimensional discrete dynamical systems. Recall that a subset $U \subset W$ is said



(a) (b)

Figure 15.8 The orbit of the seed 0.123 under f_4 using (a) 200 iterations and (b) 500 iterations.

to be *dense* in W if there are points in U arbitrarily close to any point in the larger set W. As in the Lorenz model, we say that a map f, which takes an interval $I = [\alpha, \beta]$ to itself, is *chaotic* if

1. Periodic points of f are dense in I;
2. f is transitive on I; that is, given any two subintervals U_1 and U_2 in I, there is a point $x_0 \in U_1$ and an $n > 0$ such that $f^n(x_0) \in U_2$;
3. f has sensitive dependence in I; that is, there is a *sensitivity constant* β such that, for any $x_0 \in I$ and any open interval U about x_0, there is some seed $y_0 \in U$ and $n > 0$ such that

$$|f^n(x_0) - f^n(y_0)| > \beta.$$

It is known that the transitivity condition is equivalent to the existence of an orbit that is dense in I. Clearly, a dense orbit implies transitivity, for such an orbit repeatedly visits any open subinterval in I. The other direction relies on the Baire category theorem from analysis, so we will not prove this here.

Curiously, for maps of an interval, condition 3 in the definition of chaos is redundant [8]. This is somewhat surprising, since the first two conditions in the definition are topological in nature, while the third is a metric property (it depends on the notion of distance).

We now discuss several classical examples of chaotic one-dimensional maps.

Example. (The Doubling Map) Define the discontinuous function $D\colon [0,1) \to [0,1)$ by $D(x) = 2x \bmod 1$. That is,

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

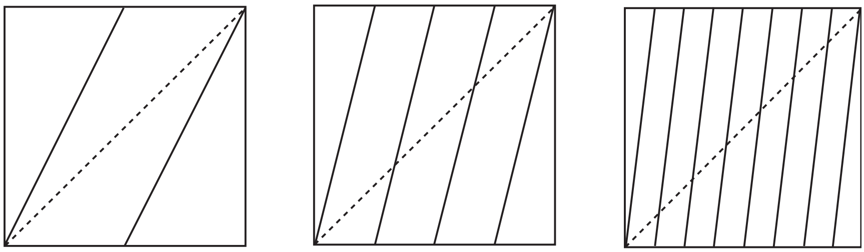


Figure 15.9 The graph of the doubling map D and its higher iterates D^2 and D^3 on [0, 1).

An easy computation shows that $D^n(x) = 2^n x \bmod 1$, so that the graph of D^n consists of 2^n straight lines with slope 2^n, each extending over the entire interval $[0, 1)$. See Figure 15.9.

To see that the doubling function is chaotic on $[0, 1)$, note that D^n maps any interval of the form $[k/2^n, (k+1)/2^n)$ for $k = 0, 1, \ldots 2^n - 2$ onto the interval $[0, 1)$. Hence the graph of D^n crosses the diagonal $y = x$ at some point in this interval, and so there is a periodic point in any such interval. Since the lengths of these intervals are $1/2^n$, it follows that periodic points are dense in $[0, 1)$. Transitivity also follows, since, given any open interval J, we may always find an interval of the form $[k/2^n, (k+1)/2^n)$ inside J for sufficiently large n. Hence D^n maps J onto all of $[0, 1)$. This also proves sensitivity, where we choose the sensitivity constant 1/2.

We remark that it is possible to write down all of the periodic points for D explicitly (see Exercise 5a). It is also interesting to note that, if you use a computer to iterate the doubling function, then it appears that all orbits are eventually fixed at 0, which, of course, is false! See Exercise 5c for an explanation of this phenomenon. ∎

Example. (The Tent Map) Now consider a continuous cousin of the doubling map given by

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

T is called the *tent map*. See Figure 15.10. The fact that T is chaotic on $[0, 1]$ follows exactly as in the case of the doubling function, using the graphs of T^n (see Exercise 15).

Looking at the graphs of the tent function T and the logistic function $f_4(x) = 4x(1 - x)$ that we discussed in Section 15.3, it appears that they should share many of the same properties under iteration. Indeed, this is the case. To understand this, we need to reintroduce the notion of conjugacy, this time for discrete systems.

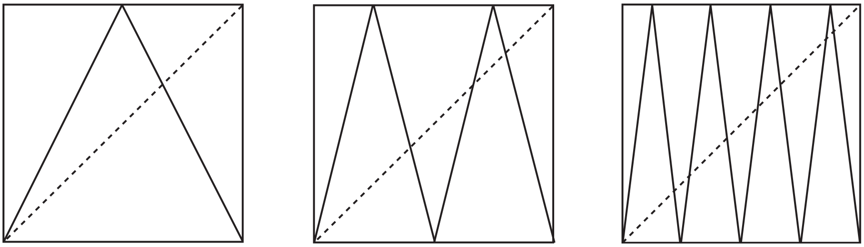


Figure 15.10 The graph of the tent map T and its higher iterates T^2 and T^3 on [0, 1].

Suppose I and J are intervals and $f : I \to I$ and $g : J \to J$. We say that f and g are *conjugate* if there is a homeomorphism $h : I \to J$ such that h satisfies the *conjugacy equation* $h \circ f = g \circ h$. Just as in the case of flows, a conjugacy takes orbits of f to orbits of g. This follows since we have $h(f^n(x)) = g^n(h(x))$ for all $x \in I$, so h takes the nth point on the orbit of x under f to the nth point on the orbit of $h(x)$ under g. Similarly, h^{-1} takes orbits of g to orbits of f. ■

Example. Consider the logistic function $f_4 : [0, 1] \to [0, 1]$ and the quadratic function $g : [-2, 2] \to [-2, 2]$ given by $g(x) = x^2 - 2$. Let $h(x) = -4x + 2$ and note that h takes $[0, 1]$ to $[-2, 2]$. Moreover, we have $h(4x(1 - x)) = (h(x))^2 - 2$, so h satisfies the conjugacy equation and f_4 and g are conjugate. ■

From the point of view of chaotic systems, conjugacies are important since they map one chaotic system to another.

Proposition. *Suppose $f : I \to I$ and $g : J \to J$ are conjugate via h, where both I and J are closed intervals in $\mathbb{R}$ of finite length. If f is chaotic on I, then g is chaotic on J.*

Proof: Let U be an open subinterval of J and consider $h^{-1}(U) \subset I$. Since periodic points of f are dense in I, there is a periodic point $x \in h^{-1}(U)$ for f. Say x has period n. Then

$$g^n(h(x)) = h(f^n(x)) = h(x)$$

by the conjugacy equation. This gives a periodic point $h(x)$ for g in U and shows that periodic points of g are dense in J.

If U and V are open subintervals of J, then $h^{-1}(U)$ and $h^{-1}(V)$ are open intervals in I. By transitivity of f, there exists $x_1 \in h^{-1}(U)$ such that

$f^m(x_1) \in h^{-1}(V)$ for some m. But then $h(x_1) \in U$ and we have $g^m(h(x_1)) = h(f^m(x_1)) \in V$, so g is transitive also.

For sensitivity, suppose that f has sensitivity constant β. Let $I = [\alpha_0, \alpha_1]$. We may assume that $\beta < \alpha_1 - \alpha_0$. For any $x \in [\alpha_0, \alpha_1 - \beta]$, consider the function $|h(x + \beta) - h(x)|$. This is a continuous function on $[\alpha_0, \alpha_1 - \beta]$ that is positive. Hence it has a minimum value β'. It follows that h takes intervals of length β in I to intervals of length at least β' in J. Then it is easy to check that β' is a sensitivity constant for g. This completes the proof. ∎

It is not always possible to find conjugacies between functions with equivalent dynamics. However, we can relax the requirement that the conjugacy be one to one and still salvage the preceding proposition. A continuous function h that is at most n to one and that satisfies the conjugacy equation $f \circ h = h \circ g$ is called a *semiconjugacy* between g and f. It is easy to check that a semiconjugacy also preserves chaotic behavior on intervals of finite length (see Exercise 12). A semiconjugacy need not preserve the minimal periods of cycles, but it does map cycles to cycles.

Example. The tent function T and the logistic function f_4 are semiconjugate on the unit interval. To see this, let

$$h(x) = \frac{1}{2}\left(1 - \cos 2\pi x\right).$$

Then h maps the interval [0, 1] in two-to-one fashion over itself, except at 1/2, which is the only point mapped to 1. Then we compute

$$\begin{aligned}
h(T(x)) &= \frac{1}{2}\left(1 - \cos 4\pi x\right) \\
&= \frac{1}{2} - \frac{1}{2}\left(2\cos^2 2\pi x - 1\right) \\
&= 1 - \cos^2 2\pi x \\
&= 4\left(\frac{1}{2} - \frac{1}{2}\cos 2\pi x\right)\left(\frac{1}{2} + \frac{1}{2}\cos 2\pi x\right) \\
&= f_4(h(x)).
\end{aligned}$$

Thus h is a semiconjugacy between T and f_4. As a remark, recall that we may find arbitrarily small subintervals mapped onto all of [0, 1] by T. Hence f_4 maps the images of these intervals under h onto all of [0, 1]. Since h is continuous, the images of these intervals may be chosen arbitrarily small. Hence we may choose 1/2 as a sensitivity constant for f_4 as well. We have proven the following proposition: ■

Proposition. *The logistic function $f_4(x) = 4x(1 - x)$ is chaotic on the unit interval.* ■

15.5 Symbolic Dynamics

We turn now to one of the most useful tools for analyzing chaotic systems, *symbolic dynamics*. We give just one example of how to use symbolic dynamics here; several more are included in the next chapter.

Consider the logistic map $f_\lambda(x) = \lambda x(1 - x)$ where $\lambda > 4$. Graphical iteration seems to imply that almost all orbits tend to $-\infty$. See Figure 15.11. Of course, this is not true, because we have fixed points and other periodic points for this function. In fact, there is an unexpectedly "large" set called a Cantor set that is filled with chaotic behavior for this function, as we shall see.

Unlike the case $\lambda \leq 4$, the interval $I = [0, 1]$ is no longer invariant when $\lambda > 4$: Certain orbits escape from I and then tend to $-\infty$. Our goal is to understand the behavior of the nonescaping orbits. Let Λ denote the set of points in I whose orbits never leave I. As shown in Figure 15.12a, there is an open interval A_0 on which $f_\lambda > 1$. Hence $f_\lambda^2(x) < 0$ for any $x \in A_0$ and, as a consequence, the orbits of all points in A_0 tend to $-\infty$. Note that any orbit that leaves I must first enter A_0 before departing toward $-\infty$. Also, the

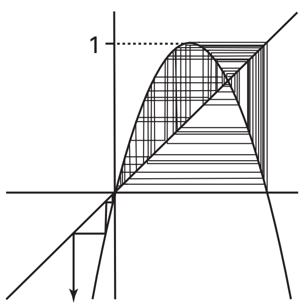


Figure 15.11 Typical orbits for the logistic function f_λ with $\lambda > 4$ seem to tend to $-\infty$ after wandering around the unit interval for a while.

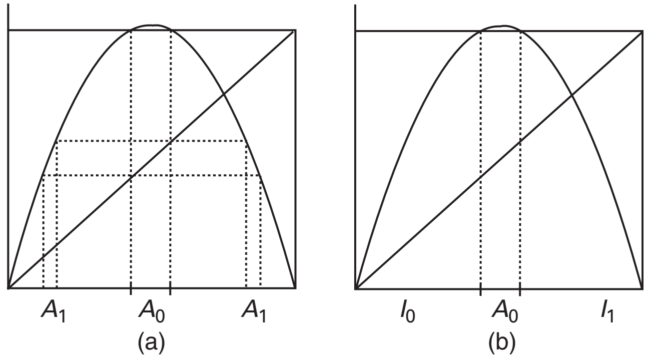


Figure 15.12 (a) The exit set in I consists of a collection of disjoint open intervals. (b) The intervals I_0 and I_1 lie to the left and right of A_0.

orbits of the endpoints of A_0 are eventually fixed at 0, so these endpoints are contained in Λ. Now let A_1 denote the preimage of A_0 in I: A_1 consists of two open intervals in I, one on each side of A_0. All points in A_1 are mapped into A_0 by f_λ, and hence their orbits also tend to $-\infty$. Again, the endpoints of A_1 are eventual fixed points. Continuing, we see that each of the two open intervals in A_1 has as a preimage a pair of disjoint intervals, so there are four open intervals that consist of points whose first iteration lies in A_1, the second in A_0, and so, again, all of these points have orbits that tend to $-\infty$. Call these four intervals A_2. In general, let A_n denote the set of points in I whose nth iterate lies in A_0. A_n consists of set 2^n disjoint open intervals in I. Any point whose orbit leaves I must lie in one of the A_n. Hence we see that

$$\Lambda = I - \bigcup_{n=0}^{\infty} A_n.$$

To understand the dynamics of f_λ on I, we introduce *symbolic dynamics.* Toward that end, let I_0 and I_1 denote the left and right closed interval respectively in $I - A_0$. See Figure 15.12b. Given $x_0 \in \Lambda$, the entire orbit of x_0 lies in $I_0 \cup I_1$. Hence we may associate an infinite sequence $S(x_0) = (s_0 s_1 s_2 \ldots)$ consisting of 0's and 1's to the point x_0 via the rule

$$s_j = k \text{ if and only if } f_\lambda^j(x_0) \in I_k.$$

That is, we simply watch how $f_\lambda^j(x_0)$ bounces around I_0 and I_1, assigning a 0 or 1 at the jth stage depending on which interval $f_\lambda^j(x_0)$ lies in. The sequence $S(x_0)$ is called the *itinerary* of x_0.

Example. The fixed point 0 has itinerary $S(0) = (000\ldots)$. The fixed point x_λ in I_1 has itinerary $S(x_\lambda) = (111\ldots)$. The point $x_0 = 1$ is eventually fixed and has itinerary $S(1) = (1000\ldots)$. A 2-cycle that hops back and forth between I_0 and I_1 has itinerary $(\overline{01}\ldots)$ or $(\overline{10}\ldots)$ where $\overline{01}$ denotes the infinitely repeating sequence consisting of repeated blocks 01.

Let Σ denote the set of all possible sequences of 0's and 1's. A "point" in the space Σ is therefore an infinite sequence of the form $s = (s_0 s_1 s_2 \ldots)$. To visualize Σ, we need to tell how far apart different points in Σ are. To do this, let $s = (s_0 s_1 s_2 \ldots)$ and $t = (t_0 t_1 t_2 \ldots)$ be points in Σ. A *distance function* or *metric* on Σ is a function $d = d(s,t)$ that satisfies

1. $d(s,t) \geq 0$ and $d(s,t) = 0$ if and only if $s = t$;
2. $d(s,t) = d(t,s)$;
3. the triangle inequality: $d(s,u) \leq d(s,t) + d(t,u)$.

Since Σ is not naturally a subset of a Euclidean space, we do not have a Euclidean distance to use on Σ. Hence we must concoct one of our own. Here is the distance function we choose:

$$d(s,t) = \sum_{i=0}^{\infty} \frac{|s_i - t_i|}{2^i}.$$

Note that this infinite series converges: The numerators in this series are always either 0 or 1, so this series converges by comparison to the geometric series:

$$d(s,t) \leq \sum_{i=0}^{\infty} \frac{1}{2^i} = \frac{1}{1 - 1/2} = 2.$$

It is straightforward to check that this choice of d satisfies the three requirements to be a distance function (see Exercise 13). While this distance function may look a little complicated at first, it is often easy to compute. ■

Example.

$$(1)\quad d\left((\overline{0}),(\overline{1})\right) = \sum_{i=0}^{\infty} \frac{|0-1|}{2^i} = \sum_{i=0}^{\infty} \frac{1}{2^i} = 2$$

$$(2)\quad d\left((\overline{01}),(\overline{10})\right) = \sum_{i=0}^{\infty} \frac{1}{2^i} = 2$$

$$(3)\quad d\left((\overline{01}),(\overline{1})\right) = \sum_{i=0}^{\infty} \frac{1}{4^i} = \frac{1}{1 - 1/4} = \frac{4}{3}.$$

■

The importance of having a distance function on Σ is that we now know when points are close together or far apart. In particular, we have

Proposition. *Suppose* $s = (s_0 s_1 s_2 \ldots)$ *and* $t = (t_0 t_1 t_2 \ldots) \in \Sigma$.

1. *If* $s_j = t_j$ *for* $j = 0, \ldots, n$, *then* $d(s,t) \leq 1/2^n$;
2. *Conversely, if* $d(s,t) < 1/2^n$, *then* $s_j = t_j$ *for* $j = 0, \ldots, n$.

Proof: In case (1), we have

$$\begin{aligned} d(s,t) &= \sum_{i=0}^{n} \frac{|s_i - s_i|}{2^i} + \sum_{i=n+1}^{\infty} \frac{|s_i - t_i|}{2^i} \\ &\leq 0 + \frac{1}{2^{n+1}} \sum_{i=0}^{\infty} \frac{1}{2^i} \\ &= \frac{1}{2^n}. \end{aligned}$$

If, on the other hand, $d(s,t) < 1/2^n$, then we must have $s_j = t_j$ for any $j \leq n$, because otherwise $d(s,t) \geq |s_j - t_j|/2^j = 1/2^j \geq 1/2^n$. ■

Now that we have a notion of closeness in Σ, we are ready to prove the main theorem of this chapter:

Theorem. *The itinerary function* $S \colon \Lambda \to \Sigma$ *is a homeomorphism provided* $\lambda > 4$.

Proof: Actually, we will only prove this for the case in which λ is sufficiently large that $|f_\lambda'(x)| > K > 1$ for some K and for all $x \in I_0 \cup I_1$. The reader may check that $\lambda > 2 + \sqrt{5}$ suffices for this. For the more complicated proof in the case where $4 < \lambda \leq 2 + \sqrt{5}$, see [25].

We first show that S is one to one. Let $x, y \in \Lambda$ and suppose $S(x) = S(y)$. Then, for each n, $f_\lambda^n(x)$ and $f_\lambda^n(y)$ lie on the same side of 1/2. This implies that f_λ is monotone on the interval between $f_\lambda^n(x)$ and $f_\lambda^n(y)$. Consequently, all points in this interval remain in $I_0 \cup I_1$ when we apply f_λ. Now $|f_\lambda'| > K > 1$ at all points in this interval, so, as in Section 15.1, each iteration of f_λ expands this interval by a factor of K. Hence the distance between $f_\lambda^n(x)$ and $f_\lambda^n(y)$ grows without bound, so these two points must eventually lie on opposite sides of A_0. This contradicts the fact that they have the same itinerary.

To see that S is onto, we first introduce the following notation. Let $J \subset I$ be a closed interval. Let

$$f_\lambda^{-n}(J) = \{x \in I \mid f_\lambda^n(x) \in J\},$$

so that $f_\lambda^{-n}(J)$ is the preimage of J under f_λ^n. A glance at the graph of f_λ when $\lambda > 4$ shows that, if $J \subset I$ is a closed interval, then $f_\lambda^{-1}(J)$ consists of two closed subintervals, one in I_0 and one in I_1.

Now let $s = (s_0 s_1 s_2 \ldots)$. We must produce $x \in \Lambda$ with $S(x) = s$. To that end we define

$$\begin{aligned} I_{s_0 s_1 \ldots s_n} &= \{x \in I \mid x \in I_{s_0}, f_\lambda(x) \in I_{s_1}, \ldots, f_\lambda^n(x) \in I_{s_n}\} \\ &= I_{s_0} \cap f_\lambda^{-1}(I_{s_1}) \cap \ldots \cap f_\lambda^{-n}(I_{s_n}). \end{aligned}$$

We claim that the $I_{s_0 \ldots s_n}$ form a nested sequence of nonempty closed intervals. Note that

$$I_{s_0 s_1 \ldots s_n} = I_{s_0} \cap f_\lambda^{-1}(I_{s_1 \ldots s_n}).$$

By induction, we may assume that $I_{s_1 \ldots s_n}$ is a nonempty subinterval, so that, by the observation above, $f_\lambda^{-1}(I_{s_1 \ldots s_n})$ consists of two closed intervals, one in I_0 and one in I_1. Hence $I_{s_0} \cap f_\lambda^{-1}(I_{s_1 \ldots s_n})$ is a single closed interval. These intervals are nested because

$$I_{s_0 \ldots s_n} = I_{s_0 \ldots s_{n-1}} \cap f_\lambda^{-n}(I_{s_n}) \subset I_{s_0 \ldots s_{n-1}}.$$

Therefore we conclude that

$$\bigcap_{n \geq 0}^{\infty} I_{s_0 s_1 \ldots s_n}$$

is nonempty. Note that if $x \in \cap_{n \geq 0} I_{s_0 s_1 \ldots s_n}$, then $x \in I_{s_0}, f_\lambda(x) \in I_{s_1}$, etc. Hence $S(x) = (s_0 s_1 \ldots)$. This proves that S is onto.

Observe that $\cap_{n \geq 0} I_{s_0 s_1 \ldots s_n}$ consists of a unique point. This follows immediately from the fact that S is one to one. In particular, we have that the diameter of $I_{s_0 s_1 \ldots s_n}$ tends to 0 as $n \to \infty$.

To prove continuity of S, we choose $x \in \Lambda$ and suppose that $S(x) = (s_0 s_1 s_2 \ldots)$. Let $\epsilon > 0$. Pick n so that $1/2^n < \epsilon$. Consider the closed subintervals $I_{t_0 t_1 \ldots t_n}$ defined above for all possible combinations $t_0 t_1 \ldots t_n$. These subintervals are all disjoint, and Λ is contained in their union. There are 2^{n+1} such subintervals, and $I_{s_0 s_1 \ldots s_n}$ is one of them. Hence we may choose δ such that $|x - y| < \delta$ and $y \in \Lambda$ implies that $y \in I_{s_0 s_1 \ldots s_n}$. Therefore, $S(y)$ agrees with $S(x)$ in the first $n + 1$ terms. So, by the previous proposition, we have

$$d(S(x), S(y)) \leq \frac{1}{2^n} < \epsilon.$$

This proves the continuity of S. It is easy to check that S^{-1} is also continuous. Thus, S is a homeomorphism. ∎

15.6 The Shift Map

We now construct a map on $\sigma : \Sigma \to \Sigma$ with the following properties:

1. σ is chaotic;
2. σ is conjugate to f_λ on Λ;
3. σ is completely understandable from a dynamical systems point of view.

The meaning of this last item will become clear as we proceed.

We define the *shift map* $\sigma : \Sigma \to \Sigma$ by

$$\sigma(s_0 s_1 s_2 \ldots) = (s_1 s_2 s_3 \ldots).$$

That is, the shift map simply drops the first digit in each sequence in Σ. Note that σ is a two-to-one map onto Σ. This follows since, if $(s_0 s_1 s_2 \ldots) \in \Sigma$, then we have

$$\sigma(0 s_0 s_1 s_2 \ldots) = \sigma(1 s_0 s_1 s_2 \ldots) = (s_0 s_1 s_2 \ldots).$$

Proposition. *The shift map $\sigma : \Sigma \to \Sigma$ is continuous.*

Proof: Let $s = (s_0 s_1 s_2 \ldots) \in \Sigma$, and let $\epsilon > 0$. Choose n so that $1/2^n < \epsilon$. Let $\delta = 1/2^{n+1}$. Suppose that $d(s, t) < \delta$, where $t = (t_0 t_1 t_2 \ldots)$. Then we have $s_i = t_i$ for $i = 0, \ldots, n+1$.

Now $\sigma(t) = (s_1 s_2 \ldots s_n t_{n+1} t_{n+2} \ldots)$ so that $d(\sigma(s), \sigma(t)) \leq 1/2^n < \epsilon$. This proves that σ is continuous. ∎

Note that we can easily write down all of the periodic points of any period for the shift map. Indeed, the fixed points are $(\overline{0})$ and $(\overline{1})$. The 2 cycles are $(\overline{01})$ and $(\overline{10})$. In general, the periodic points of period n are given by repeating sequences that consist of repeated blocks of length n: $(\overline{s_0 \ldots s_{n-1}})$. Note how much nicer σ is compared to f_λ: Just try to write down explicitly all of the periodic points of period n for f_λ someday! They are there and we know roughly where they are, because we have:

Theorem. *The itinerary function $S : \Lambda \to \Sigma$ provides a conjugacy between f_λ and the shift map σ.*

Proof: In the previous section we showed that S is a homeomorphism. So it suffices to show that $S \circ f_\lambda = \sigma \circ S$. To that end, let $x_0 \in \Lambda$ and suppose that $S(x_0) = (s_0 s_1 s_2 \ldots)$. Then we have $x_0 \in I_{s_0}$, $f_\lambda(x_0) \in I_{s_1}$, $f_\lambda^2(x_0) \in I_{s_2}$, and so forth. But then the fact that $f_\lambda(x_0) \in I_{s_1}$, $f_\lambda^2(x_0) \in I_{s_2}$, etc., says that $S(f_\lambda(x_0)) = (s_1 s_2 s_3 \ldots)$, so $S(f_\lambda(x_0)) = \sigma(S(x_0))$, which is what we wanted to prove. ■

Now, not only can we write down all periodic points for σ, but we can in fact write down explicitly a point in Σ whose orbit is dense. Here is such a point:

$$s^* = (\underbrace{0\ 1}_{\text{1 blocks}}\ |\underbrace{00\ 01\ 10\ 11}_{\text{2 blocks}}|\underbrace{000\ 001\cdots}_{\text{3 blocks}}|\ \underbrace{\cdots}_{\text{4 blocks}}\ ).$$

The sequence s^* is constructed by successively listing all possible blocks of 0's and 1's of length 1, length 2, length 3, and so forth. Clearly, some iterate of σ applied to s^* yields a sequence that agrees with any given sequence in an arbitrarily large number of initial places. That is, given $t = (t_0 t_1 t_2 \ldots) \in \Sigma$, we may find k so that the sequence $\sigma^k(s^*)$ begins

$$(t_0 \ldots t_n\ s_{n+1}\ s_{n+2} \ldots)$$

so that

$$d(\sigma^k(s^*), t) \leq 1/2^n.$$

Hence the orbit of s^* comes arbitrarily close to every point in Σ. This proves that the orbit of s^* under σ is dense in Σ and so σ is transitive. Note that we may construct a multitude of other points with dense orbits in Σ by just rearranging the blocks in the sequence s^*. Again, think about how difficult it would be to identify a seed whose orbit under a quadratic function like f_4 is dense in $[0, 1]$. This is what we meant when we said earlier that the dynamics of σ are "completely understandable."

The shift map also has sensitive dependence. Indeed, we may choose the sensitivity constant to be 2, which is the largest possible distance between two points in Σ. The reason for this is, if $s = (s_0 s_1 s_2 \ldots) \in \Sigma$ and $\hat{s}_j$ denotes "not s_j" (that is, if $s_j = 0$, then $\hat{s}_j = 1$, or if $s_j = 1$ then $\hat{s}_j = 0$), then the point $s' = (s_0 s_1 \ldots s_n \hat{s}_{n+1} \hat{s}_{n+2} \ldots)$ satisfies:

1. $d(s, s') = 1/2^n$, but
2. $d(\sigma^n(s), \sigma^n(s')) = 2$.

As a consequence, we have proved the following:

Theorem. *The shift map σ is chaotic on Σ, and so by the conjugacy in the previous theorem, the logistic map f_λ is chaotic on Λ when $\lambda > 4$.* ■

Thus symbolic dynamics provides us with a computable model for the dynamics of f_λ on the set Λ, despite the fact that f_λ is chaotic on Λ.

15.7 The Cantor Middle-Thirds Set

We mentioned earlier that Λ was an example of a Cantor set. Here we describe the simplest example of such a set, the *Cantor middle-thirds set C*. As we shall see, this set has some unexpectedly interesting properties.

To define C, we begin with the closed unit interval $I = [0, 1]$. The rule is, each time we see a closed interval, we remove its open middle third. Hence, at the first stage, we remove $(1/3, 2/3)$, leaving us with two closed intervals, $[0, 1/3]$ and $[2/3, 1]$. We now repeat this step by removing the middle thirds of these two intervals. We are left with four closed intervals $[0, 1/9], [2/9, 1/3], [2/3, 7/9]$, and $[8/9, 1]$. Removing the open middle thirds of these intervals leaves us with 2^3 closed intervals, each of length $1/3^3$. Continuing in this fashion, at the nth stage we are left with 2^n closed intervals each of length $1/3^n$. The Cantor middle-thirds set C is what is left when we take this process to the limit as $n \to \infty$. Note how similar this construction is to that of Λ in Section 15.5. In fact, it can be proved that Λ is homeomorphic to C (see Exercises 16 and 17).

What points in I are left in C after removing all of these open intervals? Certainly 0 and 1 remain in C, as do the endpoints 1/3 and 2/3 of the first removed interval. Indeed, each endpoint of a removed open interval lies in C because such a point never lies in an open middle-third subinterval. At first glance, it appears that these are the only points in the Cantor set, but in fact, that is far from the truth. Indeed, most points in C are **not** endpoints!

To see this, we attach an address to each point in C. The address will be an infinite string of L's or R's determined as follows. At each stage of the construction, our point lies in one of two small closed intervals, one to the left of the removed open interval or one to its right. So at the nth stage we may assign an L or R to the point depending on its location left or right of the interval removed at that stage. For example, we associate $LLL\ldots$ to 0 and $RRR\ldots$ to 1. The endpoints 1/3 and 2/3 have addresses $LRRR\ldots$ and $RLLL\ldots$, respectively. At the next stage, 1/9 has address $LLRRR\ldots$ since 1/9 lies in $[0, 1/3]$ and $[0, 1/9]$ at the first two stages, but then always lies in the right-hand interval. Similarly, 2/9 has address $LRLLL\ldots$, while 7/9 and 8/9 have addresses $RLRRR\ldots$ and $RRLLL\ldots$.

Notice what happens at each endpoint of C. As the above examples indicate, the address of an endpoint always ends in an infinite string of all L's or all R's. But there are plenty of other possible addresses for points in C. For example, there is a point with address $LRLRLR\ldots$. This point lies in

$$[0, 1/3] \cap [2/9, 1/3] \cap [2/9, 7/27] \cap [20/81, 7/27] \ldots.$$

Note that this point lies in the nested intersection of closed intervals of length $1/3^n$ for each n, and it is the unique such point that does so. This shows that most points in C are not endpoints, for the typical address will not end in all L's or all R's.

We can actually say quite a bit more: The Cantor middle-thirds set contains uncountably many points. Recall that an infinite set is *countable* if it can be put in one-to-one correspondence with the natural numbers; otherwise, the set is *uncountable*.

Proposition. *The Cantor middle-thirds set is uncountable.*

Proof: Suppose that C is countable. This means that we can pair each point in C with a natural number in some fashion, say as

$$\begin{array}{ccl} 1 & : & LLLLL\ldots \\ 2 & : & RRRR\ldots \\ 3 & : & LRLR\ldots \\ 4 & : & RLRL\ldots \\ 5 & : & LRRLRR\ldots \end{array}$$

and so forth. But now consider the address whose first entry is the opposite of the first entry of sequence 1, whose second entry is the opposite of the second entry of sequence 2; and so forth. This is a new sequence of L's and R's (which, in the example above, begins with $RLRRL\ldots$). Thus we have created a sequence of L's and R's that disagrees in the nth spot with the nth sequence on our list. Hence this sequence is not on our list and so we have failed in our construction of a one-to-one correspondence with the natural numbers. This contradiction establishes the result. ■

We can actually determine the points in the Cantor middle-thirds set in a more familiar way. To do this we change the address of a point in C from a sequence of L's and R's to a sequence of 0's and 2's; that is, we replace each L with a 0 and each R with a 2. To determine the numerical value of a point $x \in C$ we approach x from below by starting at 0 and moving $s_n/3^n$ units to the right for each n, where $s_n = 0$ or 2 depending on the nth digit in the address for $n = 1, 2, 3 \ldots$.

For example, 1 has address $RRR\ldots$ or $222\ldots$, so 1 is given by

$$\frac{2}{3}+\frac{2}{3^2}+\frac{2}{3^3}+\cdots=\frac{2}{3}\sum_{n=0}^{\infty}\frac{1}{3^n}=\frac{2}{3}\left(\frac{1}{1-1/3}\right)=1.$$

Similarly, 1/3 has address $LRRR\ldots$ or $0222\ldots$, which yields

$$\frac{0}{3}+\frac{2}{3^2}+\frac{2}{3^3}+\cdots=\frac{2}{9}\sum_{n=0}^{\infty}\frac{1}{3^n}=\frac{2}{9}\cdot\frac{3}{2}=\frac{1}{3}.$$

Finally, the point with address $LRLRLR\ldots$ or $020202\ldots$ is

$$\frac{0}{3}+\frac{2}{3^2}+\frac{0}{3^3}+\frac{2}{3^4}+\cdots=\frac{2}{9}\sum_{n=0}^{\infty}\frac{1}{9^n}=\frac{2}{9}\left(\frac{1}{1-1/9}\right)=\frac{1}{4}.$$

Note that this is one of the non-endpoints in C referred to earlier.

The astute reader will recognize that the address of a point x in C with 0's and 2's gives the *ternary expansion* of x. A point $x \in I$ has ternary expansion $a_1a_2a_3\ldots$ if

$$x=\sum_{i=1}^{\infty}\frac{a_i}{3^i}$$

where each a_i is either 0, 1, or 2. Thus we see that points in the Cantor middle-thirds set have ternary expansions that may be written with no 1's among the digits.

We should be a little careful here. The ternary expansion of 1/3 is $1000\ldots$. But 1/3 also has ternary expansion $0222\ldots$ as we saw above. So 1/3 may be written in ternary form in a way that contains no 1's. In fact, every endpoint in C has a similar pair of ternary representations, one of which contains no 1's.

We have shown that C contains uncountably many points, but we can say even more:

Proposition. *The Cantor middle-thirds set contains as many points as the interval* $[0, 1]$.

Proof: C consists of all points whose ternary expansion $a_0a_1a_2\ldots$ contains only 0's or 2's. Take this expansion and change each 2 to a 1 and then think of this string as a binary expansion. We get every possible binary expansion in this manner. We have therefore made a correspondence (at most two to one) between the points in C and the points in $[0, 1]$, since every such point has a binary expansion. ∎

Finally, we note that

Proposition. *The Cantor middle-thirds set has length 0.*

Proof: We compute the "length" of C by adding up the lengths of the intervals removed at each stage to determine the length of the complement of C. These removed intervals have successive lengths 1/3, 2/9, 4/27… and so the length of $I - C$ is

$$\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \cdots = \frac{1}{3}\sum_{n=0}^{\infty}\left(\frac{2}{3}\right)^n = 1. \qquad \blacksquare$$

This fact may come as no surprise since C consists of a "scatter" of points. But now consider the Cantor middle-fifths set, obtained by removing the open middle-fifth of each closed interval in similar fashion to the construction of C. The length of this set is nonzero, yet it is homeomorphic to C. These Cantor sets have, as we said earlier, unexpectedly interesting properties! And remember, the set Λ on which f_4 is chaotic is just this kind of object.

15.8 Exploration: Cubic Chaos

In this exploration, you will investigate the behavior of the discrete dynamical system given by the family of cubic functions $f_\lambda(x) = \lambda x - x^3$. You should attempt to prove rigorously everything outlined below.

1. Describe the dynamics of this family of functions for all $\lambda < -1$.
2. Describe the bifurcation that occurs at $\lambda = -1$. *Hint:* Note that f_λ is an odd function. In particular, what happens when the graph of f_λ crosses the line $y = -x$?
3. Describe the dynamics of f_λ when $-1 < \lambda < 1$.
4. Describe the bifurcation that occurs at $\lambda = 1$.
5. Find a λ-value, λ^*, for which f_{λ^*} has a pair of invariant intervals $[0, \pm x^*]$ on each of which the behavior of f_λ mimics that of the logistic function $4x(1 - x)$.
6. Describe the change in dynamics that occurs when λ increases through λ^*.
7. Describe the dynamics of f_λ when λ is very large. Describe the set of points Λ_λ whose orbits do not escape to $\pm\infty$ in this case.
8. Use symbolic dynamics to set up a sequence space and a corresponding shift map for λ large. Prove that f_λ is chaotic on Λ_λ.
9. Find the parameter value $\lambda' > \lambda^*$ above, which the results of the previous two investigations hold true.
10. Describe the bifurcation that occurs as λ increases through λ'.

15.9 Exploration: The Orbit Diagram

Unlike the previous exploration, this exploration is primarily experimental. It is designed to acquaint you with the rich dynamics of the logistic family as the parameter increases from 0 to 4. Using a computer and whatever software seems appropriate, construct the *orbit diagram* for the logistic family $f_\lambda(x) = \lambda x(1-x)$ as follows: Choose N equally spaced λ-values $\lambda_1, \lambda_2, \ldots, \lambda_N$ in the interval $0 \le \lambda_j \le 4$. For example, let $N = 800$ and set $\lambda_j = 0.005j$. For each λ_j, compute the orbit of 0.5 under f_{λ_j} and plot this orbit as follows.

Let the horizontal axis be the λ-axis and let the vertical axis be the x-axis. Over each λ_j, plot the points $(\lambda_j, f^k_{\lambda_j}(0.5))$ for, say, $50 \le k \le 250$. That is, compute the first 250 points on the orbit of 0.5 under f_{λ_j}, but display only the last 200 points on the vertical line over $\lambda = \lambda_j$. Effectively, you are displaying the "fate" of the orbit of 0.5 in this way.

You will need to magnify certain portions of this diagram; one such magnification is displayed in Figure 15.13, where we have displayed only that portion of the orbit diagram for λ in the interval $3 \le \lambda \le 4$.

1. The region bounded by $0 \le \lambda < 3.57\ldots$ is called the *period 1 window*. Describe what you see as λ increases in this window. What type of bifurcations occur?
2. Near the bifurcations in the previous question, you sometimes see a smear of points. What causes this?
3. Observe the *period 3 window* bounded approximately by $3.828\ldots < \lambda < 3.857\ldots$. Investigate the bifurcation that gives rise to this window as λ increases.

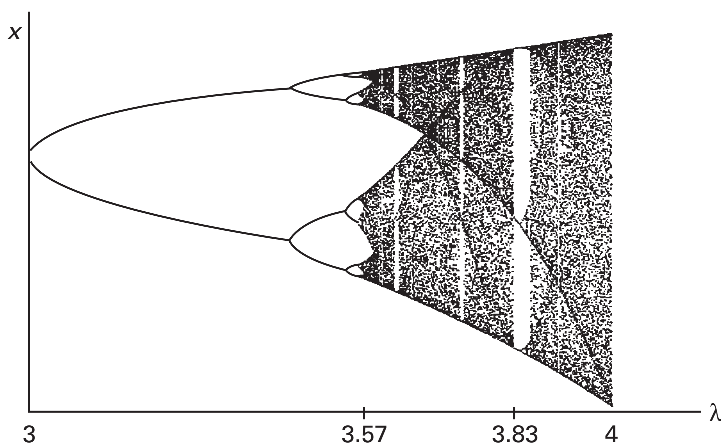


Figure 15.13 The orbit diagram for the logistic family with $3 \le \lambda \le 4$.

4. There are many other period n windows (named for the least period of the cycle in that window). Discuss any pattern you can find in how these windows are arranged as λ increases. In particular, if you magnify portions between the period 1 and period 3 windows, how are the larger windows in each successive enlargement arranged?
5. You observe "darker" curves in this orbit diagram. What are these? Why does this happen?

EXERCISES

1. Find all periodic points for each of the following maps and classify them as attracting, repelling, or neither.

(a) $Q(x) = x - x^2$	(b) $Q(x) = 2(x - x^2)$
(c) $C(x) = x^3 - \frac{1}{9}x$	(d) $C(x) = x^3 - x$
(e) $S(x) = \frac{1}{2}\sin(x)$	(f) $S(x) = \sin(x)$
(g) $E(x) = e^{x-1}$	(h) $E(x) = e^x$
(i) $A(x) = \arctan x$	(j) $A(x) = -\frac{\pi}{4}\arctan x$

2. Discuss the bifurcations that occur in the following families of maps at the indicated parameter value

(a) $S_\lambda(x) = \lambda \sin x, \quad \lambda = 1$

(b) $C_\mu(x) = x^3 + \mu x, \quad \mu = -1$ (*Hint:* Exploit the fact that C_μ is an odd function.)

(c) $G_\nu(x) = x + \sin x + \nu, \quad \nu = 1$

(d) $E_\lambda(x) = \lambda e^x, \quad \lambda = 1/e$

(e) $E_\lambda(x) = \lambda e^x, \quad \lambda = -e$

(f) $A_\lambda(x) = \lambda \arctan x, \quad \lambda = 1$

(g) $A_\lambda(x) = \lambda \arctan x, \quad \lambda = -1$

3. Consider the linear maps $f_k(x) = kx$. Show that there are four open sets of parameters for which the behavior of orbits of f_k is similar. Describe what happens in the exceptional cases.

4. For the function $f_\lambda(x) = \lambda x(1 - x)$ defined on $\mathbb{R}$:

(a) Describe the bifurcations that occur at $\lambda = -1$ and $\lambda = 3$.

(b) Find all period 2 points.

(c) Describe the bifurcation that occurs at $\lambda = -1.75$.

5. For the doubling map D on $[0, 1)$:

(a) List all periodic points explicitly.

(b) List all points whose orbits end up landing on 0 and are thereby eventually fixed.

(c) Let $x \in [0, 1)$ and suppose that x is given in binary form as $a_0 a_1 a_2 \ldots$ where each a_j is either 0 or 1. First give a formula for the binary representation of $D(x)$. Then explain why this causes orbits of D generated by a computer to end up eventually fixed at 0.

6. Show that, if x_0 lies on a cycle of period n, then

$$(f^n)'(x_0) = \prod_{i=0}^{n-1} f'(x_i).$$

Conclude that

$$(f^n)'(x_0) = (f^n)'(x_j)$$

for $j = 1, \ldots, n-1$.

7. Prove that if f_{λ_0} has a fixed point at x_0 with $|f'_{\lambda_0}(x_0)| > 1$, then there is an interval I about x_0 and an interval J about λ_0 such that, if $\lambda \in J$, then f_λ has a unique fixed source in I and no other orbits that lie entirely in I.

8. Verify that the family $f_c(x) = x^2 + c$ undergoes a period doubling bifurcation at $c = -3/4$ by

(a) Computing explicitly the period two orbit.

(b) Showing that this orbit is attracting for $-5/4 < c < -3/4$.

9. Show that the family $f_c(x) = x^2 + c$ undergoes a second period doubling bifurcation at $c = -5/4$ by using the graphs of f_c^2 and f_c^4.

10. Find an example of a bifurcation in which more than three fixed points are born.

11. Prove that $f_3(x) = 3x(1-x)$ on I is conjugate to $f(x) = x^2 - 3/4$ on a certain interval in $\mathbb{R}$. Determine this interval.

12. Suppose $f, g\colon [0, 1] \to [0, 1]$ and that there is a semiconjugacy from f to g. Suppose that f is chaotic. Prove that g is also chaotic on $[0, 1]$.

13. Prove that the function $d(s, t)$ on Σ satisfies the three properties required for d to be a distance function or metric.

14. Identify the points in the Cantor middle-thirds set C whose addresses are

(a) $LLRLLRLLR\ldots$

(b) $LRRLLRRLLRRL\ldots$

15. Consider the tent map

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

Prove that T is chaotic on $[0, 1]$.

16. Consider a different "tent function" defined on all of $\mathbb{R}$ by

$$T(x) = \begin{cases} 3x & \text{if } x \leq 1/2 \\ -3x + 3 & \text{if } 1/2 \leq x. \end{cases}$$

Identify the set of points Λ whose orbits do not go to $-\infty$. What can you say about the dynamics of this set?

17. Use the results of the previous exercise to show that the set Λ in Section 15.5 is homeomorphic to the Cantor middle-thirds set.

18. Prove the following saddle-node bifurcation theorem: Suppose that f_λ depends smoothly on the parameter λ and satisfies:

(a) $f_{\lambda_0}(x_0) = x_0$

(b) $f'_{\lambda_0}(x_0) = 1$

(c) $f''_{\lambda_0}(x_0) \neq 0$

(d) $\left.\dfrac{\partial f_\lambda}{\partial \lambda}\right|_{\lambda=\lambda_0} (x_0) \neq 0$

Then there is an interval I about x_0 and a smooth function $\mu \colon I \to \mathbb{R}$ satisfying $\mu(x_0) = \lambda_0$ and such that

$$f_{\mu(x)}(x) = x.$$

Moreover, $\mu'(x_0) = 0$ and $\mu''(x_0) \neq 0$. *Hint:* Apply the implicit function theorem to $G(x, \lambda) = f_\lambda(x) - x$ at (x_0, λ_0).

19. Discuss why the saddle-node bifurcation is the "typical" bifurcation involving only fixed points.

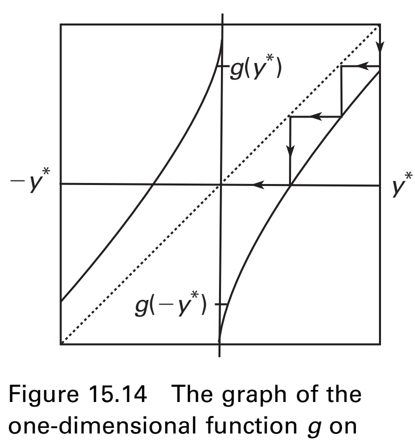


Figure 15.14 The graph of the one-dimensional function g on $[-y^*, y^*]$.

20. Recall that comprehending the behavior of the Lorenz system in Chapter 14 could be reduced to understanding the dynamics of a certain one-dimensional function g on an interval $[-y^*, y^*]$ whose graph is shown in Figure 15.14. Recall also $|g'(y)| > 1$ for all $y \neq 0$ and that g is undefined at 0. Suppose now that $g^3(y^*) = 0$ as displayed in this graph. By symmetry, we also have $g^3(-y^*) = 0$. Let $I_0 = [-y^*, 0)$ and $I_1 = (0, y^*]$ and define the usual itinerary map on $[-y^*, y^*]$.

(a) Describe the set of possible itineraries under g.

(b) What are the possible periodic points for g?

(c) Prove that g is chaotic on $[-y^*, y^*]$.