Summary

Not just any nonlinear equation can lead to chaos. For example, iterating a simple power law ($y = ax^b$) or exponential law ($y = ac^{bx}$) won't necessarily produce chaos. For a nonlinear equation to produce chaos, proposed requirements in the equation are a switchback or hump, noninvertibility, and at least one unstable fixed point. A suitable choice of parameter values in a general nonlinear equation can often fulfil one or more such requirements and lead to chaos. The literature contains many examples of nonlinear equations that can be iterated to yield chaos.

PART IV

CHARACTERISTICS OF CHAOS

I've mentioned some characteristics of chaos in earlier chapters; others come in later chapters. To get a more complete picture, here's a list of chaos's main characteristics:

1. Chaos results from a deterministic process.
2. It happens only in nonlinear systems.
3. The motion or pattern for the most part looks disorganized and erratic, although sustained. In fact, it can usually pass all statistical tests for randomness.
4. It happens in feedback systems—systems in which past events affect today's events, and today's events affect the future.
5. It can result from relatively simple systems. With discrete time, chaos can take place in a system that has only one variable. With continuous time, it can happen in systems with as few as three variables.
6. For given conditions or control parameters, it's entirely self-generated. In other words, changes in other (i.e. external) variables or parameters aren't necessary.
7. It isn't the result of data inaccuracies, such as sampling error or measurement error. Any particular value of x_t (right or wrong), as long as the control parameter is within an appropriate range, can lead to chaos.
8. In spite of its disjointed appearance, it includes one or more types of order or structure.
9. The ranges of the variables have finite bounds. The bounds restrict the attractor to a certain finite region in phase space.
10. Details of the chaotic behavior are hypersensitive to changes in initial conditions (minor changes in the starting values of the variables).
11. Forecasts of long-term behavior are meaningless. The reasons are sensitivity to initial conditions and the impossibility of measuring a variable to infinite accuracy.
12. Short-term predictions, however, can be relatively accurate.
13. Information about initial conditions is irretrievably lost. In the mathematician's jargon, the equation is "noninvertible." In other words, we can't determine a chaotic system's prior history.

14. The Fourier spectrum is "broad" (mostly uncorrelated noise) but with some periodicities sticking up here and there.
15. The phase space trajectory may have **fractal** properties. (We'll discuss fractals in Ch. 17.)
16. As a control parameter increases systematically, an initially nonchaotic system follows one of a select few typical scenarios, called routes, to chaos.

The next few chapters give a closer look at some of the more important of these characteristics. We'll discuss other items in the list later in the book.

Chapter 14

Sensitive dependence on initial conditions

This chapter looks more closely at an important feature that can help identify chaos.

Trajectories in both the nonchaotic and chaotic regimes of nonlinear, dissipative systems have at least one thing in common: once we choose the control parameter, all trajectories (no matter what value we use for x_0 to start iterating) go to an attractor. There's also an important difference in trajectories of the two regimes. For nonchaotic circumstances, trajectories for two different values of starting conditions (x_0) get closer together (in most cases) or remain equidistant (on tori), for the same value of the control parameter. In the chaotic domain, however, it's just the opposite: trajectories diverge.

The trajectory divergence in the chaotic regime means that two neighboring values of x_0, differing only by the minutest of magnitudes, such as in the fourth or fifth decimal place, can evolve to different trajectories under the same value of the control parameter. A tiny difference or error, compounded over many iterations (a long time), grows into an enormous difference or error. Eventually, there's no relation at all between the two trajectories. A loose analogy (not necessarily a chaotic process) is the shuffling of a deck of cards. Two cards that start next to one another may end up quite far apart after many shuffles. Another analogy is a pair of water molecules in a river or a pair of air molecules in the breeze. The molecules may be close at one time and far apart some time later. In general, the slightest change in a variable's first value or the system's state at one time leads ultimately to very different evolutionary paths. Chaologists refer to such a trait as **sensitive dependence on initial conditions**, or simply sensitivity to initial conditions. ("Initial" in this sense means any time at which we begin comparing the pair of neighboring trajectories.)

Some chaologists use "sensitivity to initial conditions" in a more restrictive sense. They use the expression only when the rate of divergence of two nearby trajectories is *exponential* with time. (We'll discuss exponential divergence in Ch. 25.)

From tiny to huge differences

Let's look at a quick example of long-term effects of minor differences in initial conditions. We'll use the logistic equation, take a k value (3.75) that leads to chaos, and see what happens to two trajectories whose starting points differ by only a tiny amount, say 0.0001. Specifically, we'll compare the trajectory (list of iterates) for $x_0 = 0.4100$ to the one for $x_0 = 0.4101$. Here are the first five iterations, with those for $x_0 = 0.4100$ listed first and those for $x_0 = 0.4101$ in parentheses: 0.9071 (0.9072), 0.3159 (0.3157), 0.8104 (0.8102), 0.5761 (0.5767), and 0.9158 (0.9154). Short-term predictions (the early iterates) don't vary significantly despite the minor difference in x_0. This implies that short-term predictions in a chaotic system can be reasonably reliable. (In contrast, we can't make reliable predictions for uncorrelated noise, even for one time-step into the future.)

At iterations 20–24, on the other hand, the values are 0.6135 (0.6366), 0.8892 (0.8675), 0.3694 (0.4309), 0.8736 (0.9196), and 0.4142 (0.2773). Even after only these few iterations, some computed values diverge by quite a bit (e.g. 0.4142 versus 0.2773). With other equations or examples—that is, depending on noise, the system, and other features—sensitivity to initial conditions might become manifest over fewer iterations or less time. In fact, within the noise range, large differences might be possible after just one iteration. At the other extreme, differences might not become significant until many more iterations or measuring events than those associated with the example I gave here. The important point is that differences sooner or later do become significant (even though I didn't rigorously prove that here). Any original separation at all, even infinitesimal, sooner or later can lead to important separations.

The difference between the two starting values of 0.4100 and 0.4101 is only 0.024 per cent. Differences or errors somewhat larger than that are typical of most data (and certainly of mine). Figure 14.1 shows the effects of errors of 0.1 per cent and 2 per cent in the starting input values, again using the logistic equation and k = 3.75. A 2 per cent error in the input value (Fig. 14.1b) leads to major discrepancies after fewer than ten iterations. By about the 15th iteration, there's no apparent relation at all between the two trajectories. Therefore, predictions for the long term (and 15 iterations isn't long) are highly sensitive to the initial condition (values of k and x_0).

The key feature about extreme sensitivity to initial conditions, then, is that a seemingly tiny difference in input conditions becomes amplified or compounded over time and eventually can lead to a large difference in the predicted output. (Again, to invoke that type of behavior, the control parameter has to be large enough to put the system into the chaotic domain.)

Minor differences or errors in input values are absolutely unavoidable, both in the practical world of measurement and in the mathematician's world of computer calculations. In practice, nobody ever measures anything with infinite accuracy. Any measurement inevitably includes some uncontrollable and indeterminate error. Such error might be due to human mistakes, human frailties, or instrument

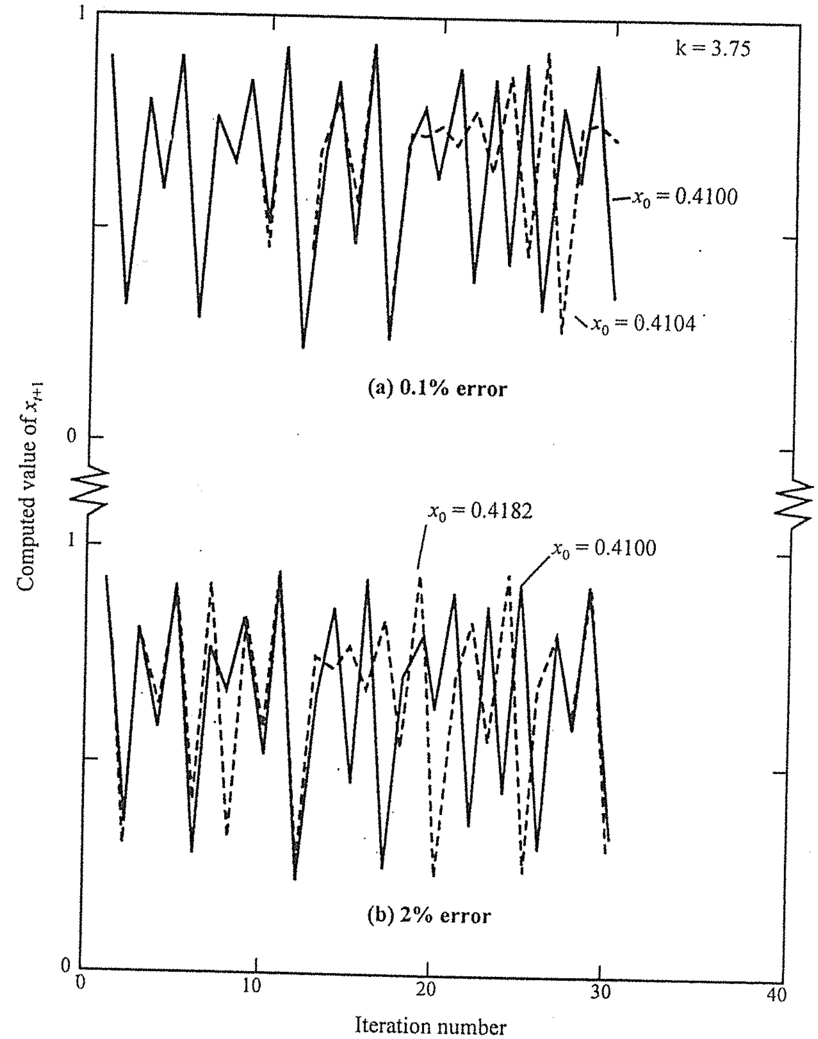

Figure 14.1 Variation in computed iterations of the logistic equation owing to minor differences (error) in initial input value x_0. (a) Error of 0.1 per cent. (b) Error of 2.0 per cent.

imprecision. Even if infinite accuracy were possible, we'd have to round off the value to a reasonable number of decimal places to make practical computations, so there'd be some roundoff error. In addition, there are background fluctuations or noise. Because of these considerations, a "datum point" on a graph is really a "datum region," owing to the finite uncertainty associated with each value.

Such inaccuracies may not affect short-term predictions. Long-term predictions, on the other hand, may be greatly affected and therefore meaningless. Given the inaccuracy of the input value, the value of a computed iterate at a far-off future time can be anywhere within a wide range. In fact, it can be anywhere on the attractor. The size of the attractor in phase space therefore represents an irreducible margin of error for any long-term prediction for a chaotic system. (At the same time, it's also a bound on the error.)

In conclusion, iterating an equation gives specific predictions for any given input values, but original input values are never completely accurate. It's an extension of the garbage-in, garbage-out theme: inaccuracies in, greater inaccuracies out.

Computer inaccuracy and imprecision

It's tempting to assume that describing a measurement more accurately, as computers enable us to do, might resolve the problem of sensitivity to initial conditions. However, greater accuracy just extends the time or "grace period" until the uncertainty expands to the same magnitude as the attractor. Furthermore, computers are a perfect example of inescapable sensitivity to initial conditions. The next few paragraphs explain why.

Ranked by increasing power of the machine, computers are of three types. (The classification is only approximate, as there can be some overlap between types.)

- *Microcomputers or personal computers ("PCs")* These are the friendly little things that sit on our desks or on our laps. Typical users of these computers include individuals, small and large businesses, and schools.
- *Minicomputers* Minicomputers usually have more than one user and are often found at colleges and universities. Some fit on a desktop, others are floor-mounted.
- *Mainframes, including "supercomputers"* A mainframe is a large multi-user system used by major establishments. It's usually located in a large air-conditioned room.

All of these computers operate on a so-called **binary system**. A computer's binary system reduces everything to some combination of zeroes and ones. The computer uses binary principles to represent numbers, store and manipulate data, and so on. There are many possible ways to use the binary system to achieve those purposes. All of them have one thing in common—they are imperfect. One reason is that, because of its binary system, a computer doesn't interpret an input number exactly. Another reason is that a computer only has a finite number of discrete states. For instance, a computer's memory capacity is always a finite limit on the "accuracy" of any input value. Also, a computer doesn't carry out mathematical operations exactly. Instead, it does so only approximately. (The approximation, of course, is plenty good enough for most purposes.) In other words, to some usually minor degree, computers are inherently inaccurate.

As a simple example of computer inaccuracy, let's take the logistic equation $x_{t+1} = kx_t(1-x_t)$, assign k = 3.8 and insert 0.4 for x_t. Using pencil and paper, you and I can calculate the first iterate (x_{t+1}) to be 3.8 (0.4) (1−0.4) = *exactly* 0.912. Suppose we want the answer expressed to 15 decimal places. The correct value then is 0.912 000 000 000 000. A computer, interestingly, won't give that result! Instead, it reports an answer such as 0.911 999 940 872 192 or 0.912 000 000 476 837. The error is, of course, tiny and probably negligible for most purposes. However, under iteration the error can snowball and become significant, as shown below.

The amount of inaccuracy varies with computer brand and model (IBM PC, Apple MacIntosh, IBM mainframe, Cray, etc.), computer operating system (MS-DOS, UNIX, Primos, VMS, etc.), type of computation (e.g. simple addition, iteration), compiler, software instructions, and probably other features. Thus, although we may think a computer calculates with accuracy to a certain number of decimal places, the real accuracy is something different from that, and it depends on all of the factors just mentioned.

Instructing the computer to calculate to a specified "precision" (e.g. single versus double "precision") or to a specified number of significant **digits** doesn't help. Results are still different because of the factors just mentioned. For instance, Rothman (1991: 58) describes computer simulations that Tomio Petrosky ran at the University of Texas. In the experiments, Petrosky routed a comet from outside the galaxy between the planet Jupiter and the Sun. The idea was to calculate how many times the comet would orbit the sun before sailing off again into outer space. With the computer set for six-digit accuracy, the model predicted 757 orbits; setting the computer to seven digits produced a prediction of 38 orbits; at eight digits, it gave 236; at nine, 44; at ten, 12; at 11, 157; and so on. The answer varied wildly, depending on the prescribed accuracy of the computations. Peitgen et al. (1992: 531–5) give an interesting discussion of the general problem.

The problem may be insignificant for some purposes but critical for others. It's insignificant, for instance, in the nonchaotic domain of the logistic equation. Within that domain, differences between successive iterates for conditions other than period-doubling become smaller (nearly zero) as iteration continues, regardless of the prescribed "precision." The same is true for any member of a period-doubling sequence. That's because iterates in the nonchaotic domain gravitate toward a single point or a repeated sequence of points.

The logistic equation's chaotic domain, on the other hand, is quite sensitive to initial conditions. Minor differences, such as those associated with computer "precision," are systematically amplified (Jackson 1991: 212). A good example is iterations done at single versus double precision on a Prime minicomputer with k = 3.8. Those iterations show a difference only in the eighth decimal place for the first iteration, but later differences build on that discrepancy and increase (Fig. 14.2). After only 35 or 40 iterations the absolute differences, attributable solely to the prescribed "precision," are as large as 0.8. That means, for instance, that the computations at one "precision" might generate a value of $x_{t+1} = 0.9$, whereas those at the other might report $x_{t+1} = 0.1$ for the same iteration number. Considering that the

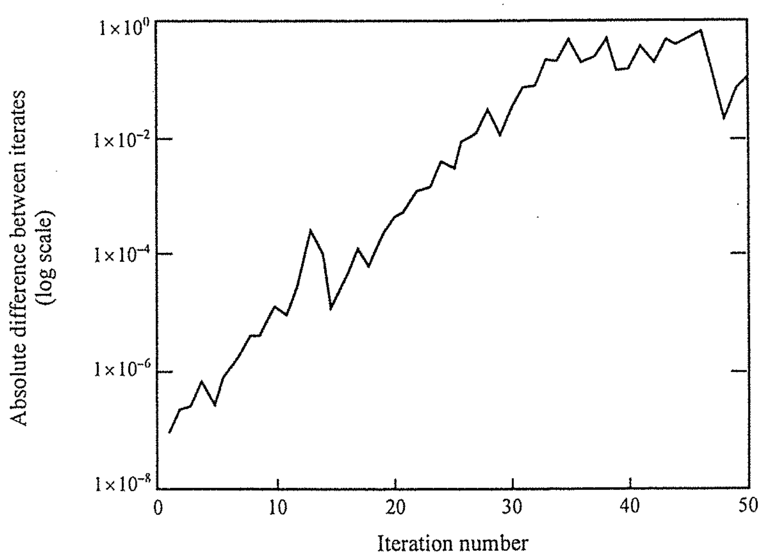

Figure 14.2 Temporal growth of difference between iterates obtained with single and double precision on a Prime minicomputer, using the logistic equation with k = 3.8.

values can range only from zero to 1, that maximum possible difference of 0.8 is immense. Differences remain large but fluctuate irregularly once the early stage of systematic increase is over. Any other type of computer gives similar results.

The above examples involved comparisons done on one computer. Different results arose because of different original conditions, in this case the number of digits to the right of the decimal in the computations. Such differences are a matter of accuracy (correctness or exactness). Besides that problem, iterates differ according to the particular computer. Differences attributable to brand or model of computer are a matter of precision or reproducibility rather than accuracy. There's no way to control imprecision. We're stuck with it regardless of the quality of our measurements.

As a quick example of imprecision, let's look at iterations of the logistic equation done on two different machines—a Data General Aviion Work Station (here called "DG") and a Prime minicomputer ("Prime"). I used Fortran language, "single precision," and $x_0 = 0.4$ on both computers. In both the nonchaotic and chaotic domains, results are similar to those described above. For instance, differences in the nonchaotic domain quickly become constant and negligible. In contrast, differences in results in the chaotic domain (k = 3.8) systematically increase (Fig. 14.3). After 35–40 iterations, x_{t+1} as given by one computer differed by as much as 0.8 from that of the other. For further iterates, differences in x_{t+1} between the two computers fluctuate irregularly.

Out of curiosity, I made similar chaotic-domain computations on my Hewlett–

Packard HP-97 desk calculator and compared the results to those of the other two machines. Each of the three machines gave a different series of iterates. Differences between iterates for any two of the three machines are similar and behave in the same way, soon becoming quite significant (Fig. 14.3).

The problems just described have led to an active research area: the interaction or relationships between finite computer arithmetic (such as the orbits a computer generates) on the one hand and the true dynamics of the theoretical system on the other.

Because of the inaccuracy and imprecision inherent in computers, we must treat iterations with care and suspicion. The inaccuracy and imprecision won't be significant for a whole range of parameter values (0 to about 3.56 for the logistic equation) and hence for a wide range of problems. The chaotic domain, on the other hand, shows sensitivity to initial conditions. Here differences between any two iteration sequences can be substantial, depending on many computer-related factors. However, ratios of iterates, quantitative measures based on iterates, and trends in iterates (or in their differences) probably are reliable. In other words, even though iterates may differ, they nonetheless fall on the same attractor. The same quantitative measures of the attractor's properties also emerge.

Where iterations simulate real data, carry the values only to a realistic number of digits. All digits that the computer produces beyond that significant number are garbage. Drop those numbers after each iteration; they can (in the chaotic domain) affect the results because of sensitive dependence on initial conditions.

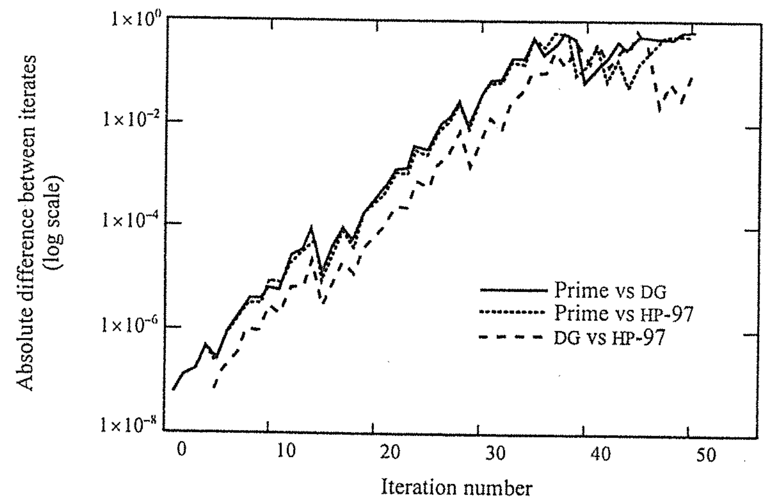

Figure 14.3 Temporal growth of differences between iterates done on different computers, using the logistic equation with k = 3.8. "Prime" = Prime minicomputer, "DG" = Data General Aviion Work Station, and "HP-97" = Hewlett–Packard-97 desk calculator.

Interpretations

Unpredictability is the most common interpretation of sensitive dependence on initial conditions. "Unpredictable" here refers to the system's future state, condition, or behavior.[1] Stated more formally, extreme sensitivity to initial conditions, combined with the inevitable measuring errors, roundoff errors, and computer precision, imposes limits on how accurately we can predict the long-term temporal behavior of any chaotic process. Beyond a certain time, as discussed in Chapters 15 and 25, long-term behavior looks random, is indeterminable, and cannot be reliably predicted.

This has important practical implications. Measurements of any variable over time may not help in accurately forecasting its future behavior beyond a brief time range. That is, the behavior may be chaotic, so that the long-term future evolution can't be predicted except to within certain broad limits (the size of the attractor). That could explain why weather forecasters often can't correctly predict the weather beyond a day or two (sometimes hours or minutes). Similarly, knowing the record of large floods over the past century or two may be useless in reliably predicting the frequency of future floods.

A second way of interpreting sensitive dependence on initial conditions is that a tiny historical accident can bring about a radically different eventual outcome. In other words, an infinitely small change in any external force can lead to radically different and unexpected long-term behavior. Or, an apparently steady pattern can be highly unstable to small disturbances. For example, continuing the flood analogy, seemingly minor changes in climate might lead to a major change in future flood frequencies. For species population, two iterations of the model, using the same parameters and almost the same first populations, might lead to vastly different predicted eventual populations. The timeworn and farfetched example from the chaos literature is the so-called "**butterfly effect**." It says that a butterfly flapping its wings in, say Brazil, might create different "initial conditions" and trigger a later tornado elsewhere (e.g. Texas), at least theoretically.

Sensitivity to initial conditions is a main feature of chaos. Some chaologists say it's the fundamental cause of chaos. Rediscovering and emphasizing that sensitivity is one of chaos's chief contributions to science. However, extreme sensitivity to first conditions doesn't necessarily, by itself, lead to chaos. For instance, even if we're just playing with numbers at our desk we still need an equation that can lead to chaos. Also, sensitivity to initial conditions can occur with random data (Wegman 1988). That's especially true if we interpret sensitivity loosely as any separation at all over time. How commonly it occurs with random data hasn't yet been resolved.

Sensitivity to initial conditions also points out the folly of placing too much emphasis on the idea of determinism. On the one hand, natural phenomena are subject to rigid physical laws, and some equations (at least theoretical ones) may be exact. In practice, however, noise, perturbations, and measurement limitations are always present and can reduce determinism to a physically meaningless concept, regardless of any underlying mathematical validity. Many physical systems behave so erratically that they are indistinguishable from a random process, even though they are strictly determinate in a mathematical sense.

1. We can predict numbers, of course. It's just that the numbers are meaningless. For example, we just saw that the same person using two different computers can come up with widely disparate long-term predictions.

Summary

Sensitive dependence on initial conditions means that a seemingly insignificant difference in the starting value of a variable can, over time, lead to vast differences in output. (Some authors prefer a more limited definition, namely that the differences increase exponentially.) Measurement error, noise, or roundoff in the data values can cause such tiny differences in input. In that sense, sensitivity to initial conditions implies sensitivity to measurement error and to noise. Computers, in particular, are susceptible to sensitivity to initial conditions because of their binary system, operating system, software details, the way they are built, and so on. Sensitivity to initial conditions is a very important characteristic of chaos, but it's not a foolproof indicator; random data can also show it. One practical effect of sensitivity to initial conditions is that, even though a system may be deterministic and the governing laws known, long-term predictions are meaningless.

Chapter 15
The chaotic (strange) attractor

Nonchaotic attractors generally are points, cycles, or smooth surfaces (corresponding to static, periodic, and multifrequency systems, respectively). Their geometry is regular. Small initial errors or minor perturbations generally don't have significant long-term effects. (We saw this in following the routes of various trajectories as they went to a point- or limit-cycle attractor.) Also, neighboring trajectories stay close to one another. Predictions of a trajectory's motion on nonchaotic attractors therefore are fairly meaningful and useful, in spite of errors or differences in starting conditions. Now, just the opposite characteristics describe chaotic attractors, which I'll define as attractors within the chaotic regime (but see next paragraph). As of yet, there isn't any universal agreement on a definition of a chaotic attractor.

Because of the unexpected and quite different features just mentioned, you'll often see the term strange attractor in place of "chaotic attractor." The attractor is strange (in the sense of unfamiliar, poorly understood, or unknown) in terms of both dynamics and geometry. However, Grebogi et al. (1984) and some other authors distinguish between "chaotic" and "strange" attractors. To them, "chaotic" refers to the trajectory dynamics on the attractor. In particular, a chaotic attractor in their view is one on which two nearby trajectories diverge *exponentially* with time. Such chaotic attractors also have fractal properties (Ch. 17). In contrast, those authors use "strange" in connection with the attractor's geometrical structure. That is, they define a strange attractor as one having fractal (Cantor set) structure (Ch. 17) or certain other special geometrical properties. "Strange attractors" in their definition don't show sensitivity to initial conditions (exponential divergence). In other words, their chaotic attractors are strange (have fractal properties), but their strange attractors aren't chaotic (don't show exponential divergence of neighboring trajectories with time). In many cases it's either impossible or unnecessary to confirm one or both of sensitivity and strangeness. I'll use the terms "chaotic attractor" and "strange attractor" synonymously.

Similarities with other attractors

Before examining the characteristics of a chaotic attractor, let's see why it still qualifies as an attractor. Here are several features it has in common with all attractors.

- It's still the set of points (but in this case an infinite number of points) that the system settles down to in phase space.
- It occupies only certain zones (and is therefore still a shape) within the bounded phase space. All data points are confined to that shape. That is, all possible trajectories still arrive at and stay "on" the attractor. (As with non-chaotic attractors, a trajectory technically never gets completely onto a chaotic attractor but only approaches it asymptotically.) In that sense, a chaotic attractor is a unit made up of all chaotic trajectories. Figure 15.1 shows a two-dimensional view of the well known Rössler (1976) attractor, a three-dimensional strange attractor designed as a simplification of the Lorenz attractor.
- A chaotic attractor shows zones of recurrent behavior in the form of orderly periodicity, as explained below.
- It's quite reproducible.
- It has an invariant probability distribution, as explained in the following section.

Invariant probability distribution of attractor

First of all, don't be intimidated by the formidable-sounding name of this section. This item is no big deal—a sheep in wolf's clothing. All it does is indicate where, in phase space, a trajectory likes to spend its time while cruising around on an attractor. If the system has only one variable, a simple histogram of the time-series data gives a crude idea of those locations.

For a given value of the control parameter, a trajectory goes to its attractor and stays there forever. The attractor consists of phase space cells, each of which has its own unique address. The trajectory visits those cells as the system evolves. It might visit some cells more often than others. That preferential visiting is the feature we'll look at right now.

A periodic trajectory repeats the same circuitous phase space route. It therefore visits only a relatively few phase space cells or addresses. For example, the logistic equation with k = 3.4 iterates to an attractor of two fixed x^* values, specifically 0.452 and 0.842. Suppose we let that trajectory get to its attractor, then sample it for a long time and compile a frequency distribution of its wanderings. As usual, we'll assume relative frequency equals probability—the probability of finding the system in some particular phase space location at any given time. A graph of the relative frequency distribution or probability distribution for this example is simply a plot of probability versus x value for the possible range of values (here the

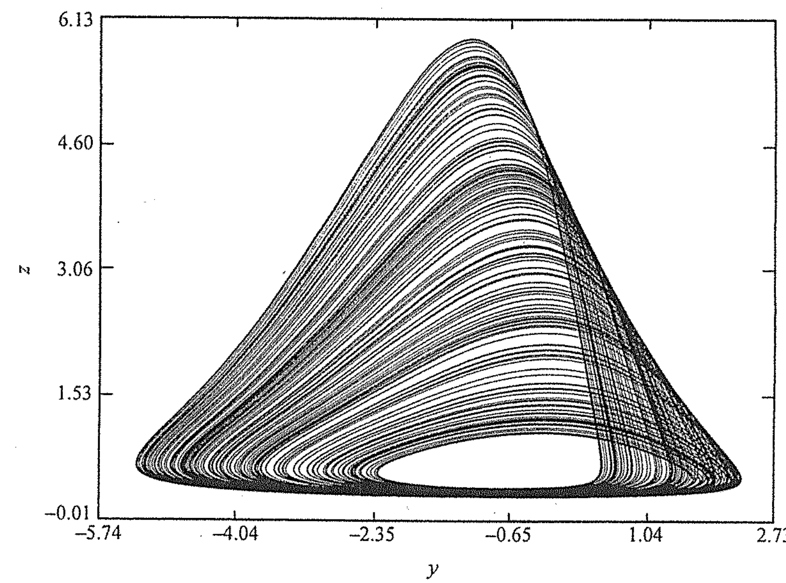

Figure 15.1 Two-dimensional projection of the Rössler strange attractor. Computer-generated graphics by Sebastian Kuzminsky.

interval $0 \leq x \leq 1$). The system only visits two phase space sites (here values of x), and those in regular alternation. The frequency distribution therefore consists simply of 50 per cent of the visits at $x = 0.452$ and the other 50 per cent at $x = 0.842$. A plot of that distribution (Fig. 15.2a) consists of two equal-height spikes (one at $x = 0.452$, the other at $x = 0.842$). Each spike extends upward to an ordinate (probability) value of 0.50.

That's all there is to it. We have just built a probability distribution for this particular attractor or variable, for the given parameter value. Furthermore, that distribution is time invariant (doesn't change with time). That's because, assuming a thorough sampling of the attractor (i.e. a long time series) and a fixed binning scheme, we get the same distribution, no matter where we start in the time sequence. Hence, "**invariant probability distribution** of attractor." You might also see it called an **invariant measure**, **natural probability measure**, or simply a probability distribution. All such terms mean the relative frequency with which a trajectory visits the different phase space regions of an attractor.

The sampling duration ideally ought to be long enough to characterize the system's long-term average behavior adequately. In truth, the distribution usually obtained for a chaotic attractor isn't exactly invariant. Rather, it becomes asymptotic toward some limiting distribution as time or number of observations goes to

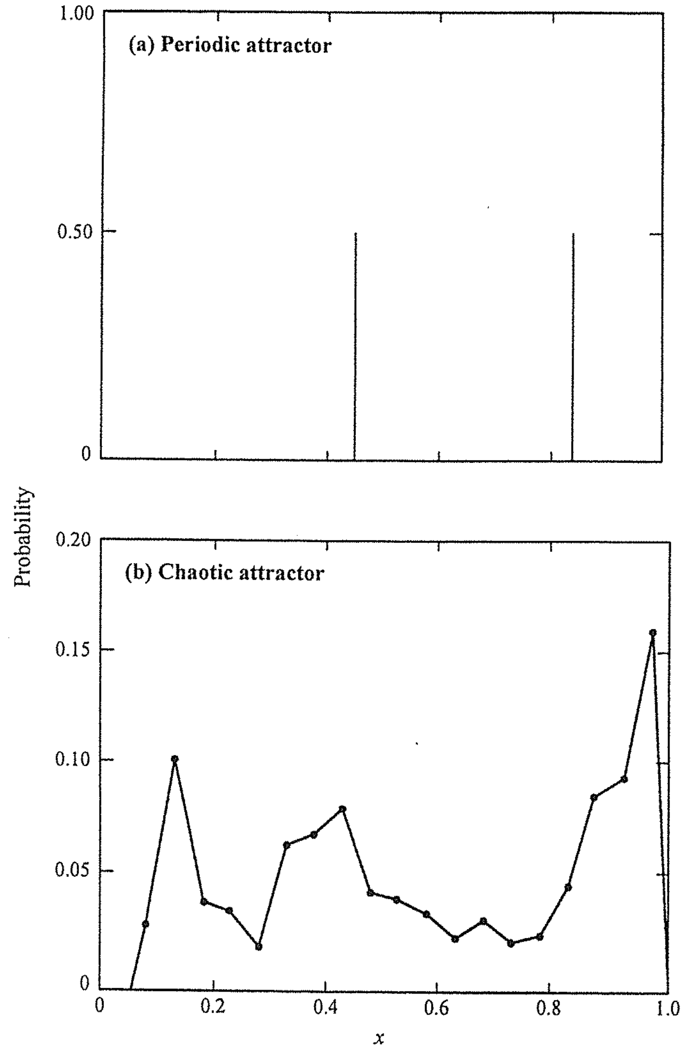

Figure 15.2 Invariant probability distributions based on the logistic equation, using k = 3.4 (upper) and k = 3.9 (lower).

infinity. As a result, an invariant probability distribution really is the distribution that the data approach as time goes to infinity.

The nonchaotic regime by definition includes only fixed points, limit cycles (regular periodicities), and quasiperiodic motion. The attractor's probability distribution for a given parameter value within the nonchaotic regime shows one or

more pronounced spikes. For example, the "trajectory" of a fixed point is always at the same phase space address. The probability distribution for a fixed point therefore is a spike located at the point's value along the abscissa and extending vertically to a probability of 1.0. If the trajectory instead alternates between two phase space addresses, the plot shows two spikes (Fig. 15.2a). In general, the number of spikes corresponds to the periodicity (two spikes indicating period two, etc.). For quasiperiodic motion the spikes may not be as well defined.

The logistic equation's *chaotic* trajectory, in contrast, goes to many locations within phase space, that is, to relatively many sites on its attractor. True, it shows some favoritism for certain x values, for a given parameter value. True, also, it becomes periodic at some parameter values (Ch. 16). In general, however, the trajectory eventually gets to most x values within its possible range. The frequency distribution for a chaotic trajectory (Fig. 15.2b) doesn't have the pronounced spikes that typify the nonchaotic trajectory. However, we still compiled it in the same way, namely by partitioning the interval into classes or bins and counting the number of points that fall into each bin.

The frequency distribution for a given dataset varies with the number of classes (Fig. 6.3). We want to designate the number of classes (class width) such that the maximum number (narrowest width) of the bins gives a bin width greater than the noise level (accuracy) of the data. For example, if the data values range from 0 to 1 and are only accurate to within 0.05, it's senseless to set up boxes narrower than 0.05, such as with widths of 0.001. Furthermore, dividing a dataset into many classes tends to produce fewer observations within any one class. Some classes may not get a representative number of observations; others may get a number small enough to be statistically insignificant. (A relatively large dataset, say many thousands of observations, helps combat the problem of few observations in a class.) At the other extreme, a small number of classes (say fewer than about five) tends to obscure important details of the distribution. There isn't any simple formula to specify the appropriate number of classes to use for a given case. It's an arbitrary, subjective judgement. Figure 15.2 uses 20 classes.

A frequency distribution can be compiled for a system that involves virtually any number of variables. When more than one variable is involved, it's a *joint probability distribution* (Ch. 6). The idea is always the same, namely to define compartments by prescribing various class limits for each compartment, no matter how many variables the system has. Graphically, a system of two variables has two axes, such as x and y; the compartments then are rectangles that together make up a grid on the graph. For three variables (the most we can draw on a graph) the compartments are cubes. For four or more variables the compartments are called **hypercubes**. We have to use our imaginations to figuratively cover the attractor with hypercubes of a certain size; analytically, it's a straightforward matter to compile joint frequencies by having the computer count the frequencies in hypercube A, hypercube B, and so on.

Besides data consisting of one variable or of several variables, a third type of data for which chaologists estimate an invariant probability distribution (in this case a joint probability distribution) is lagged data of one variable. In that case the frequency distribution reflects x_t and subsets based on specified lags. For instance, in two dimensions (x_t and x_{t+m}) with x ranging from 0 to 1 we might set up bins having intervals of 0.1 for both "variables." The first bin might consist of x_t ranging from 0 to 0.1 and x_{t+m} ranging from 0 to 0.1, and so on. Then we count the number of observations falling into each bin, similar to the scheme of Figure 6.6. That kind of invariant probability distribution is popular in chaos analyses. The distributions for all three types of data show important global features of the frequency with which orbits visit various regions of the attractor.

In theory, then, a plot of the attractor's long-term probability distribution shows whether a system is in a nonchaotic or chaotic regime. In practice, that doesn't necessarily work. First, it may not be possible to collect enough data to estimate the probability distribution reliably. Secondly, a broad band of frequencies without pronounced spikes carries the same identification problem that a Fourier analysis does. That problem is that such a pattern *might* indicate chaos, but it might instead indicate a random process (Glass & Mackey 1988: 42–7). In other words, the broad band on a Fourier analysis is a necessary but not sufficient criterion of chaos.

Distinctive features of chaotic attractors

The distinctive features of chaotic attractors are:

- A trajectory within the chaotic regime (e.g. Fig. 15.1) is usually more complex than just a simple, regular loop. At some values of the control parameter, it supposedly never repeats itself (never stabilizes). (A trajectory that never repeats itself is an **aperiodic** or **nonperiodic** trajectory.) Many other parameter values seem to bring periodic orbits of long periodicity. (For example, iterating the logistic equation within the chaotic regime on my computer can produce periodicities that range from a few hundred to several thousand iterations, depending on k.) In practice it might be very difficult to distinguish cycles of long periodicity from aperiodic trajectories. Other parameter values (called windows) within the chaotic regime have special cascading (period-doubling) periodicities, as discussed in the next chapter.
- Trajectories on a chaotic attractor do not cross. If they did, then the system could behave in very different ways whenever the conditions at the crossing point recur.
- Two trajectories that at one time are quite close together diverge and eventually follow very different paths. That's because of the sensitivity to initial conditions that characterizes the chaotic regime. Topologists look upon such phase space divergence of neighboring trajectories as a stretching, as mentioned in Chapter 10.
- The phase space path of a chaotic trajectory also does a folding maneuver. That occurs when the trajectory reaches its phase space boundary (at limiting values of one or more variables) and rebounds or deflects back in its plotted pattern. Stretching and folding are really just imaginary topological actions but are helpful notions nonetheless. In fact, Abraham & Shaw (1983: 107) write that "the basic dynamical feature of chaotic attractors is bounded expansion, or divergence and folding together of trajectories within a bounded space."
- A chaotic attractor has a complex, many-layered internal structure. The reason is that "folding" happens over and over again. That internal structure is usually (but not always) fractal. (A fractal [Ch. 17] is an elegant geometric pattern that looks the same regardless of any change in scale.)
- The external appearance is elaborate and variable compared to the loops or smooth-surfaced tori of the nonchaotic attractor. To date, many chaotic attractors have been found. Many more probably will be discovered.

  It's instructive to look at chaotic attractors from different angles in phase space. Modern computer-graphics packages enable us to look at attractors and other graphed objects from virtually any direction. That's a big advantage, especially for an attractor in three-dimensional phase space.
- Its dimension doesn't have to be an integer, such as 2 or 3. (We'll discuss dimensions in later chapters.) The noninteger and usually fractal nature of chaotic attractors led Mandelbrot (1983: 197) to recommend calling them fractal attractors rather than chaotic or strange attractors.

Definitions of a chaotic attractor

Based on the above features, chaologists proposed several definitions of a chaotic (strange) attractor. Two of the best are:

- A chaotic attractor is a complex phase space surface to which the trajectory is asymptotic in time and on which it wanders chaotically (Grebogi et al. 1982).
- A chaotic attractor is an attractor that shows extreme sensitivity to initial conditions (Eckmann & Ruelle 1985, Holden & Muhammad 1986).

According to Ralph Abraham (Fisher 1985: 31): "The chaotic attractor emerged in mathematical theory in 1932 or so, then came into view in science with Lorenz in 1971. And now it's cresting in a wave of fanatical popularity in all the sciences."

Summary

As with all attractors, chaotic (strange) attractors:

- are the set of conditions that a system can take on
- occupy certain zones (or have a certain geometric shape) in phase space

- have zones that are more popular, sometimes with periodic visits
- are reproducible
- have an invariant probability distribution.

The invariant probability distribution of an attractor is simply a histogram or relative frequency distribution representing the long-term relative frequency with which the system visits its various possible phase space locations, for a given value of the control parameter. In the nonchaotic regime, that distribution shows spikes. The number of spikes equals the periodicity. In the chaotic regime, the distribution tends to lack pronounced spikes. Random data also lack pronounced spikes, so an attractor's invariant probability distribution won't distinguish chaotic data from random data. The main characteristics of a chaotic or strange attractor are:

- an irregular, erratic, trajectory that can be periodic or nonperiodic
- a total absence of crossing trajectories
- stretching (divergence) of trajectories that originally were close together (sensitivity to initial conditions)
- folding, owing to confinement within the bounded phase space
- complex, many-layered, usually fractal internal structure
- elaborate or unusual outer geometry
- noninteger dimension.

Two definitions of a chaotic attractor are that it is an attractor that shows extreme sensitivity to initial conditions or a complex phase space surface to which the trajectory is asymptotic in time and on which it wanders chaotically.

Chapter 16

Order within chaos

Chaos *looks* erratic. However, it isn't just a vast sea of disorder. Against a general backdrop of apparent randomness are some curious features. For example, continued increases of a control parameter, such as k in the logistic equation, don't necessarily bring increased degrees of chaos. Each systematic increase in the control parameter, no matter how slight, does bring about a different trajectory. (And hence, even within chaos, the parameter is king!) However, within the chaotic domain such an increase often leads to some type of regularity, a kind of order. All of those regularities are quite reproducible, for a given value of the control parameter. Curiously, the order usually lasts only over brief ranges of the control parameter.

The orderly features as a group are striking enough that some people emphasize them more than the chaos. At the very least, chaos consists of various types of order camouflaged under random-like behavior. In that sense, order within chaos is the rule rather than the exception.

When chaos comes from iterating an equation, typical forms of order within chaos are windows (regions of periodicity), various routes to chaos (e.g. period-doubling), the chaotic attractor itself, zones of popularity that the trajectory prefers to visit more frequently, and fractal structure in a chaotic attractor.

Windows

One orderly peculiarity within chaos is **windows**—zones of k values for which iterations from any x_0 produce our friend, the periodic attractor, instead of a chaotic attractor. In other words, scattered throughout the chaotic realm, at known k values, are stable conditions where there actually isn't any chaos. Instead, there's a nonchaotic attractor consisting of a fixed periodicity.

Grossman & Thomae (1977) describe the chaotic regime as a "mixed state where periodic and chaotic time development are mingled with each other" or as a "superposition of a periodic motion and a chaotic motion in state space." Wolf (1983) calls it a complex sequence of periodic and chaotic trajectories. Chaotic trajectories can be difficult to identify; we may not be able to tell whether they are indeed chaotic or just periodic but with a very long period.

As with many other features of chaos, windows were discovered and studied only by iterating an equation, naturally using a computer. The apparent number of windows within the chaotic domain depends on computer precision. Iterations on powerful computers suggest that the chaotic regime of the logistic equation (Fig. 16.1), and probably of other iterated equations, includes an infinite number of windows of periodic behavior (Wolf 1986: 276; Grebogi et al. 1987; Peitgen et al. 1992: 635).

The most prominent window generated on the average personal computer using the logistic equation is at k ≈ 3.83, where a period-three cycle turns up. On Figure 16.1 that window is a broad, vertical white band; the three black lines crossing the band are the only values the system takes on, within that interval of k. For the quadratic map (Eq. 10.2), the most prominent window within chaos is at k ≈ 1.76 (Grebogi et al. 1987), again with a period-three cycle (Fig. 16.2).[1] Olsen & Degn (1985) mention that there are windows for all uneven periodicities (3, 5, 7, etc.) from 3 to infinity.

Any regular periodicity within the chaotic regime usually lasts for just a brief range of k values. Most windows, for example, are so narrow that they are invisible on a graph except under very high magnification. Also, the periodicity can be difficult to detect. For one thing, the periodicity usually is high. For another, the

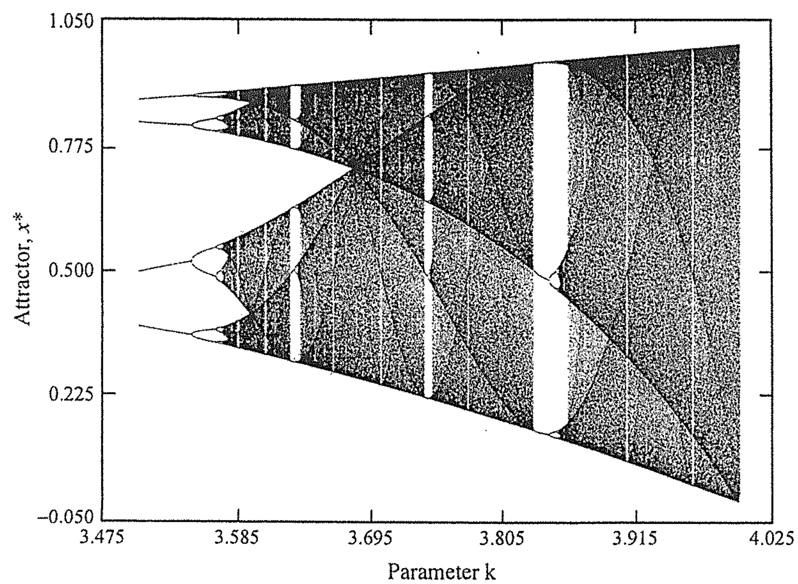

Figure 16.1 Chaotic domain of logistic equation. Computer-generated graphics by Sebastian Kuzminsky.

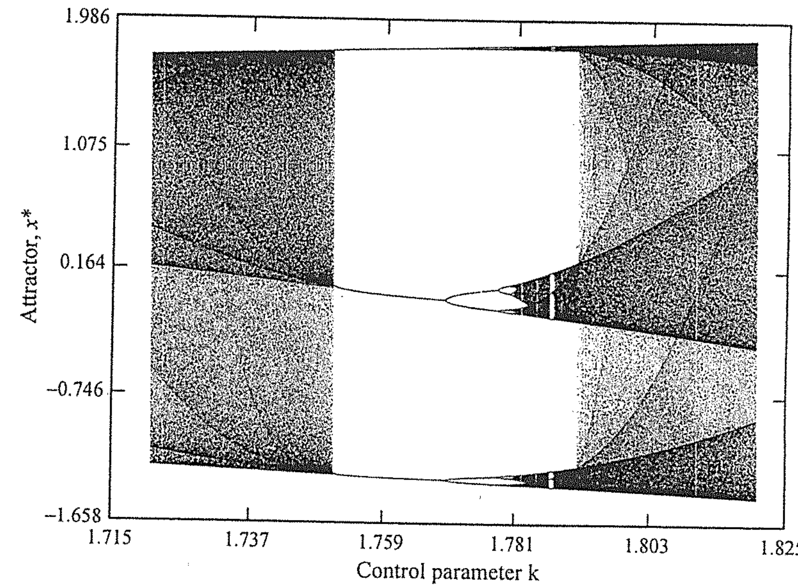

Figure 16.2 Window consisting of a period-three cycle, occurring at about k = 1.76 for the quadratic equation (after Grebogi et al. 1987). Computer-generated graphics by Sebastian Kuzminsky.

trajectory often looks chaotic (Hofstadter 1981, 1985). Even so, windows typically have a global scaling structure (Yorke et al. 1985; Peitgen et al. 1992: 636).

Real-world data typically defy any attempts to study such detailed phenomena. Windows aren't of much use in detecting chaos from such data because of noise and because they occur only over narrow and special conditions (Schaffer et al. 1986).

Routes between windows and chaos

A second kind of order within chaos is the routes to chaos, such as period-doubling and intermittency. The period-doubling goes on in the same cascading way that it does in the pre-chaotic domain. With the period-three attractor at k = 3.83 in the

1. Figures 16.1 and 16.2 were generated by starting with the lowest k value shown, then:
   1. choosing x_0
   2. iterating the equation thousands of times
   3. throwing out the first few hundred iterates as possibly atypical
   4. plotting the remaining points on the graph
   5. increasing k by a tiny amount and repeating steps 1–4 with that new k
   6. repeating step 5 until the desired range of k values had been covered.

logistic equation, for example, slight increases in k bring bifurcations to cycles of period 6, 12, 24, and so on (Jensen 1987).

Within chaotic-regime windows, period-doubling sequences themselves have orderly (in fact universal) features (besides the cyclical cascades). The logistic equation and quadratic map (Eqs 10.1 and 10.2) (Yorke et al. 1985) are examples. Here the width of any window bears an approximately constant relation to the distance from the start of the window to the first bifurcation. The general relation is

$$\lim_{n\to\infty} \frac{k_\infty - k_0}{k_2 - k_0} \to 2.25 \qquad (16.1)$$

where n here is the periodicity of the attractor, k_0 is the k value at which the window first appears, k_2 is the k value at which the starting periodicity undergoes its first doubling, and k_∞ is the k value at which chaos becomes re-established. As with the other universal numbers associated with period-doubling (Ch. 12), the ratio in Equation 16.1 comes closer and closer to the limiting number (2.25) as the periodicity gets higher, as the limit term suggests. The calculation is only approximate for windows that start at a low periodicity. For instance, for the period-three window of Figure 16.2, the window begins at about k = 1.75, bifurcates into period six at about k = 1.768, and ends its period-doubling in chaos at about k = 1.79. Plugging these values into Equation 16.1 gives (1.79–1.75) / (1.768–1.75) = 2.222.

Intermittency within chaos works in a way opposite from within the pre-chaotic domain, as the control parameter increases. That is, rather than moving from order to chaos, the scenario now goes from chaos to order. At a k value just below that associated with a window, periodicity alternates with chaos as iterations continue. As k increases, the periodic parts become longer (i.e. range over larger intervals of x) relative to the chaotic bursts. At a critical value of k, the duration (interval) of the periodic behavior becomes infinitely long. In other words, the periodic attractor then becomes fixed or firm (Grebogi et al. 1987).

A special route to chaos, known as a **crisis** (Grebogi et al. 1982), shows up only within the chaotic domain. A crisis is an abrupt change in a chaotic attractor, such that with increase in the control parameter the attractor suddenly enlarges greatly or gets destroyed. We don't yet know how common such events are. They occur in nonlinear circuits and in lasers (Grebogi et al. 1983).

According to Grebogi et al. (1983), transitions between order and chaos are reversible, at least within the chaotic domain. That is, within critical zones of the control parameter (edges of windows), chaos can be created or destroyed, depending on increases or decreases in the control parameter.

The chaotic attractor

A third structural aspect or underlying framework of chaos is the chaotic attractor. During a chaotic system's long-term evolution, the variables don't haphazardly

occupy every possible location within the phase space. Instead, the same chaotic attractor eventually takes shape, regardless of where the iterations began (value of x_0). Although different values of x_0 do lead to different trajectories on the attractor, the trajectories never leave the attractor, and the attractor occupies only part of the phase space.

Zones of popularity on the attractor

The fourth orderly structure of chaos is zones of relatively greater popularity on each chaotic attractor. These are zones a chaotic system is more likely to visit during its evolution. On the graphs of Figures 16.1 and 16.2, areas the trajectory visits the most are proportionally darker. The darkest zones are most likely to be visited, for a given k. In contrast, the trajectory never goes to white zones at all. The zones of greater popularity could provide some guidance in devising a statistical theory for accurately predicting the likelihood of x_t taking on a particular value (Jensen 1987). A fifth orderly feature of chaos is the fractal structure, discussed in the next chapter.

All this regularity indicates that order is a basic ingredient in chaos. That's why some people look upon chaos as order disguised as disorder. Some authors believe that pure chaos sets in only when the control parameter is at a maximum (such as k = 4 in the logistic equation).

Self-organization, complexity, and emergent systems

Judging from the logistic equation, most orderly features of chaos depend strictly on the control parameter. For a given value of that parameter, the order develops spontaneously, that is, without external cause. Such spontaneous development seems to be a special class of an interesting process called **self-organization**. Self-organization is the act whereby a self-propagating system, without outside influence, takes itself from seeming irregularity into some sort of order. It seems to reflect a tendency for a dynamical system to organize itself into more complex structures. The structure can be spatial, temporal, or operational (functional). The time over which the structure lasts varies from one case to another.

Examples of self-organization are the organizing of birds into an orderly flock, of fish into a clearly arranged school, of sand particles into ripple marks, of weather elements (wind, moisture, etc.) into hurricanes, of water molecules into laminar flow, of stars into the spiral arms of a galaxy, and of the demand for goods, services, labor, salaries, and so on, into economic markets. Briggs & Peat (1989) cite, as other examples:

- the lattice of hexagonal cells that form after the onset of chaos by heating a pan of liquid from below

- successions of ordered or oscillatory regimes that follow chaos in various chemical reactions
- termite nests resulting from random-like termite activity
- rush-hour traffic patterns following less-busy, earlier random-like traffic
- organized amoeba (slime mold) migrations appearing after random-like aggregation.

Self-organization is a main feature of a kind of behavior called **complexity**. In this specialized sense, complexity (Lewin 1992, Waldrop 1992) is a type of dynamic behavior that never reaches equilibrium and in which many independent particle-like units or "agents" perpetually interact and seek mutual accommodation in any of many possible ways. The units or agents spontaneously organize and re-organize themselves in the process into ever larger and more involved structures over time. "Complex" dynamic behavior has at least six ingredients:

- A large number of somewhat similar but independent items, particles, members, components or agents.
- Dynamism—the particles' persistent movement and readjustment. Each agent continually acts on and responds to its fellow agents in perpetually novel ways.
- Adaptiveness: the system conforms or adjusts to new situations so as to insure survival or to bring about some advantageous realignment.
- Self-organization, whereby some order inevitably and spontaneously forms.
- Local rules that govern each cell or agent.
- Hierarchical progression in the evolution of rules and structures. As evolution goes on, the rules become more efficient and sophisticated, and the structure becomes more complex and larger. For instance, atoms form molecules, molecules form cells, cells form people, and thence to families, cities, nations, and so forth.

Because of those characteristics, complex adaptive systems are called **emergent** systems. Their chief characteristic is the *emergence* of new, more complex levels of order over time. Complexity, like chaos, implies that we can't necessarily understand a system by isolating its components and analyzing each component individually. Instead, looking at the system as a whole might provide greater—or at least equally helpful—insight.

Authors used to categorize dynamical behavior either as orderly (e.g. having a fixed-point or periodic attractor) or random (Crutchfield & Young 1989). Chaos and complexity are important additional types. So, as of today, dynamical behavior might be classified into order, complexity, chaos, and randomness. In fact, those four types might form a progression or hierarchy, in the order just listed. Also, the four types aren't mutually exclusive. Orderly, chaotic, and random-like behavior, for instance, all have elements of determinism. Similarly, complexity and chaos both contain order and randomness.

To the extent that there is order within the chaotic regime, the word chaos (since its normal usage implies utter confusion or total disorder) is a misnomer. But it's too late now.

Summary

In spite of its name and appearance, chaos has a great deal of order or regularity. Examples are windows (regions of periodicity within the chaotic domain), various routes to chaos from within a window, the unique geometric shape (chaotic attractor) within the phase space, zones of greater popularity on the attractor, and the attractor's fractal structure. Because of those orderly features, you might see chaos mentioned in the context of two newly emerging concepts. One is self-organization—the process whereby a self-propagating system, without any apparent external influence, takes itself from seeming irregularity or uniformity into a pattern or structure. The other is complexity—dynamic behavior characterized by continuous give-and-take among many independent agents, resulting in an adaptive, self-organizing system that forms larger and more complex structures over time.

Chapter 17
Fractal structure

A fifth geometric characteristic or type of order within many chaotic domains is *fractal structure*. I'll discuss this within a general treatment of fractals and their characteristics.

Definitions

A fractal is a pattern that repeats the same design and detail or definition over a broad range of scale. Any piece of a fractal appears the same as we repeatedly magnify it. For instance, a twig and its appendages from the edge of some species of trees form a pattern that repeats the design of the trunk and main branches of the tree. Such repetition of detail, or recurrence of statistically identical geometrical patterns as we look at smaller-upon-smaller parts of the original object, is the unifying theme of fractals. Fractal patterns don't have any characteristic size.

The definition just given is general. Experts don't agree on a more explicit or mathematical definition.[1]

Fractals are all around us. Examples are:

- branching configurations (e.g. in bronchial tubes, blood vessels, coral reefs, mineral deposits, river networks, trees, ferns)

---

1. The word "fractal" comes from a Latin word *fractus*, which loosely means "to break into irregular fragments." "Fractal" was coined in 1975 by the acknowledged "father of fractals," **Benoit B. Mandelbrot** (1924–) of International Business Machines, Yorktown Heights, New York. Mandelbrot was born in Warsaw, educated mostly in Paris, and has held many university positions in Europe and the USA. As an applied mathematician, he has contributed to such varied subjects as statistics, economics, engineering, physiology, geomorphology, languages, astronomy, and physics. He is a self-designated "nomad-by-choice" in regard to the topics he works on (Gleick 1987: 90). Many mathematicians and theoretical physicists at first disregarded his work on fractals, but fractals now seem to be here to stay. (As a separately defined concept, fractals are only a couple of decades old, although their mathematical antecedents go back many decades.) Mandelbrot is now a Professor of Mathematical Sciences at Yale University and a Watson Fellow at the IBM Thomas J. Watson research center in Yorktown Heights, New York.

- rough surfaces in general (landscapes, mountains, outcrops of rock)
- fracture networks and cracks in a surface
- objects that undergo fragmentation (coal and rock, soil, asteroids and meteorites, volcanic ejecta)
- objects that result from aggregation or disorderly growth (soot, dust)
- viscous fingering (e.g. injection of water or air into a subsurface oilfield)
- earthquake features (distribution of their magnitudes, spatial distribution, frequency distribution of aftershocks)
- flow patterns (turbulence, eddies, wakes, jets, clouds, smoke, other mists, streamlines of water particles)
- the irregular trace of a water/land interface (edge of a mud puddle, shoreline of a continent)
- galaxy distributions, size distribution of craters, rings of Saturn, fluctuations in interplanetary magnetic fields
- lightning
- changes in stock prices
- incomes of rich people.

Figure 17.1 shows a few photographs of fractals.

Fractals and chaos

Fractals deal with geometric patterns and quantitative ways of characterizing those patterns. Chaos, in contrast, deals with time evolution and its underlying or distinguishing characteristics. Fractals are a class of geometric forms; chaos is a class of dynamical behavior.

Fractals and chaos are closely intertwined and often occur together. For instance, most chaotic attractors have a fractal striated texture. Points on such attractors plot as a set of layers that look the same over a wide range of scales. The layered structure can be difficult to see because of noise and the small number of points in some datasets. The points then seem to plot on a single curve. Figure 17.2 is a chaotic attractor (Eq. 13.11) on which the same amount of detail appears as we repeatedly zoom in on smaller and smaller sections (and enlarge and refocus them for appropriate viewing). In general, "the chaotic attractors of flows or invertible maps typically are fractals; the chaotic attractors of noninvertible maps may or may not be fractals. The chaotic attractors of the logistic equation, for example, are not fractals" (Eubank & Farmer 1990). Other chaos-related geometric objects, such as the boundary between periodic and chaotic motions in phase space, also may have fractal properties (Moon 1992: 325). Because of those close relationships, fractals can help detect chaos.

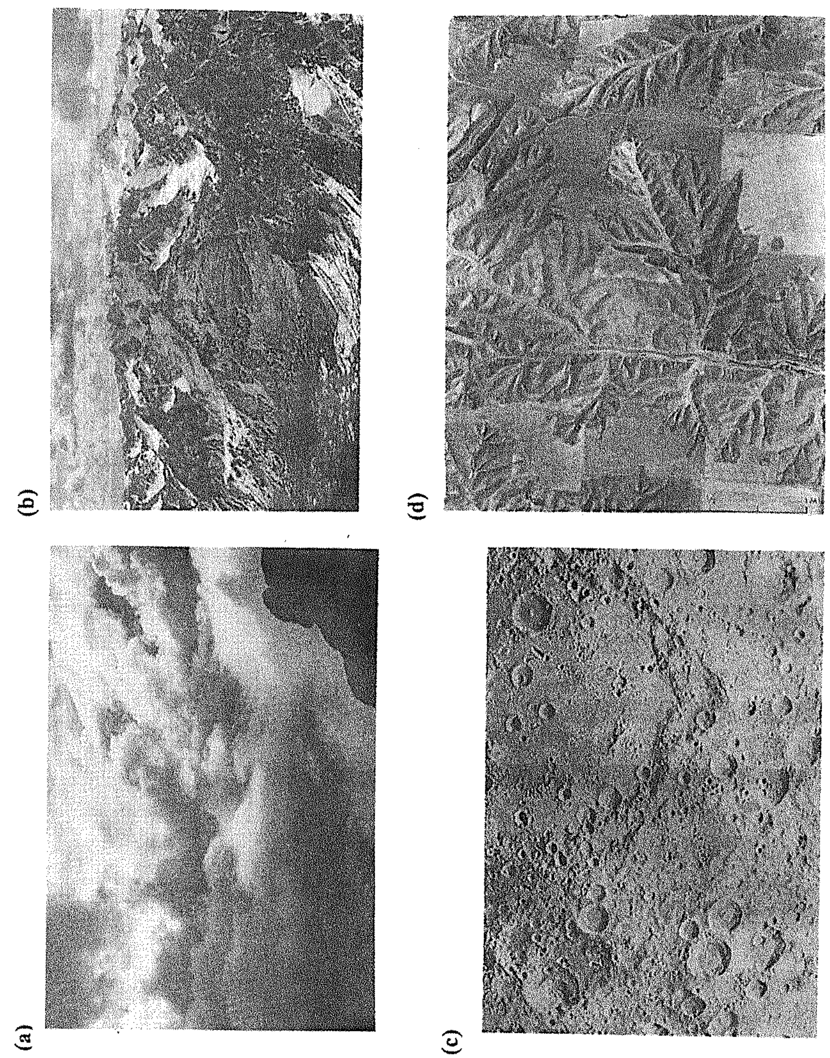

Figure 17.1 Photographs of natural features that have given rise to fractal patterns. (a) Clouds (photo: Nolan Doesken). (b) Earth surface topography (photo: Bob Broshears). (c) Asteroid impact craters, here on the Moon (photo: NASA/USGS, Lunar Orbiter I, 38M). (d) Gullies (photo: James Brice).

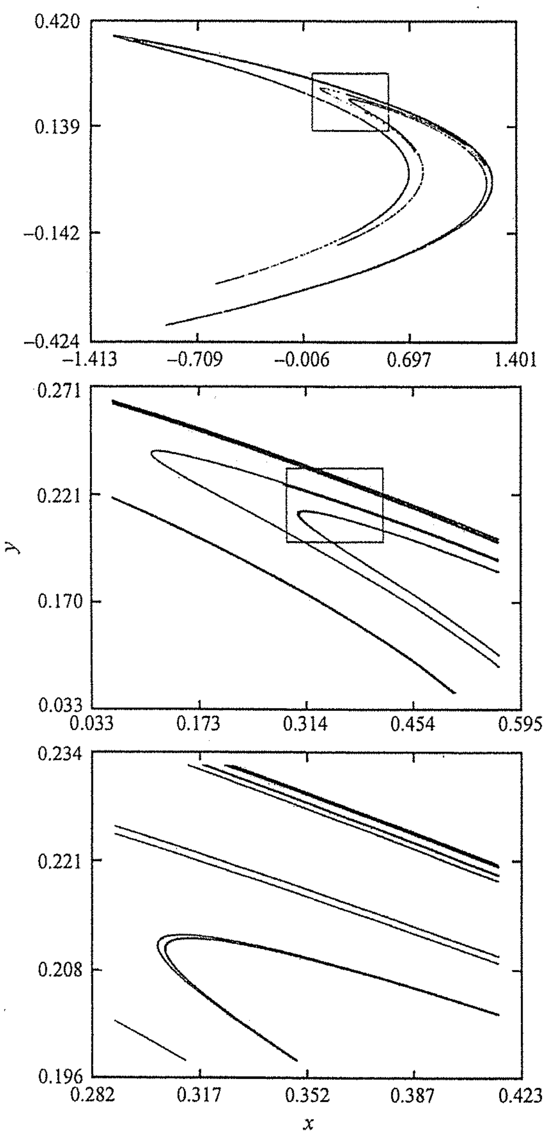

Figure 17.2 Three different magnifications of part of the Hénon attractor. Entire middle diagram shows area outlined by the rectangle within upper diagram; entire lower diagram is region outlined by the rectangle within middle diagram. Computer-generated graphics by Sebastian Kuzminsky.

Characteristics

All fractals have four common characteristics.

- They have a statistical geometric regularity. The amount of detail, or the geometric structure, look the same at one length scale as at another length scale. The recurrence of the same pattern over a range of scales is called **self-similarity** or **scale invariance**. People define these terms in slightly different ways, but the two terms mean the same thing.

  Self-similarity means that any part of the object, enlarged and refocused to whatever extent necessary for comparison, looks like the whole. The central theme is the preservation of detail upon magnification. Objects that aren't self-similar don't keep the same degree of detail upon magnification. Theoretically at least, a self-similar object has infinite detail.

  Scale invariance means that the object lacks a preferred or characteristic size (length, radius, etc.). Something scale-invariant is independent of size (looks pretty much the same at all sizes). Many geologic patterns, for example, are scale-invariant. (And that's why geologists place something of known scale in photographs of rock outcrops. Geology has few characteristic lengths.)

  Fractals don't have to be exactly self-similar. Some are only approximately or statistically self-similar, as explained below. Others—called **self-affine**—are only self-similar if some rotation, stretching, squashing, or other distortion (i.e. a direction-dependent rescaling) accompanies the magnification. (For instance, in two dimensions, one dimension might be stretched more than the other. In three dimensions, the two dimensions in the horizontal might retain the same relative scaling, but the vertical dimension would be scaled differently.)
- A characteristic number, in the form of a noninteger dimension (discussed in later chapters), quantifies the scaling of their pattern or complexity over a range of scales.
- Fractals usually are generated by many repetitions of a given operation. The operation might be natural, such as the disintegration of rock fragments, or it might be a mathematical exercise, as in solving the same equation over and over (iteration). If done mathematically, the many repeated calculations ordinarily require a computer.
- Fractals aren't smooth. They generally look rough, broken, disorganized, jagged, bumpy, or shaggy. Iteration, for example, doesn't produce a smooth object. Instead, it produces a line or surface that has detail over a wide range of scale. Consequently, there's no simple algebraic equation that can specify a particular spot on a fractal object.

An object can have one of the above features and still not be a fractal. For instance, a straight line, square, and cube are all self-similar but aren't fractals. Also, iteration doesn't necessarily yield a fractal.

Fractals are of two types (Saupe 1988: 72), as follows:

- **Deterministic** (ordered or exact) **fractals** These fractals look exactly the same (even in minute detail) at all levels of magnification. Theoretically at least, their range in terms of size is unlimited or infinite; they range over all scales, with no upper and lower limits. They are more common in mathematical and theoretical worlds. Popular examples in the literature are the von **Koch snowflake, Cantor set**, and Sierpinski gasket.
- **Natural** (approximate, statistical, or stochastic) **fractals** These are the fractals found in nature. We can compose them according to the same rules as deterministic fractals but with the additional element of randomness or noise included. The random element is critical for reproducing natural features, such as landscapes, branches, and coastlines. Such features, when scaled down or up, never look completely alike in detail; they only look generally or statistically alike. (All differences between the original and scaled versions are attributable to chance.) A second important distinction between natural and deterministic fractals is that natural fractals, since they simulate natural objects, are limited to a range of sizes. They exist only over a finite range of sizes. A rock particle, for example, can't be smaller than an atom and can only be so large, even though we don't know exactly how large. Because of that size limitation, there's an upper and lower limit beyond which self-similarity no longer holds for a natural fractal.

Fractals first became established in a spatial context (in analyzing static shapes, such as trees). Further work (West 1990, Plotnick & Prestegaard 1993) has begun to make a case for considering them in statistical and dynamical (temporal) contexts as well; certain diagnostic statistics recur at all levels of magnification. In a fractal dynamical sense, a time series looks the same over a wide range of timescales.

Value of fractals

Like chaos, fractals generate controversy within the mathematics community (Pool 1990a). Some mathematicians charge that fractals lack real mathematical content, theorems, and proofs and that fractals emphasize pretty pictures or computer-generated designs. In addition, fractals, like chaos, don't provide physical explanations. Defenders rejoin that fractals are a new and efficient way of mathematically describing natural objects (e.g. Burrough 1984). That is, fractals are a mathematical way of describing the variability of irregular features over a range of scales.

Fractals have geometric features that make them useful as idealized models for our natural world. Examples include the human circulatory system (Barcellos 1984), soil pollution, cave systems, the seeping of oil through porous rock (Peterson 1988), drainage networks, and coral reefs. The general principle in such modeling usually is to add the same structure at progressively smaller scales.

Scientists see fractals partly as a vehicle for opening up a whole new realm of questions about the physical world. Why is an object fractal? What physical processes produce fractals? Why is the same scaling valid over a wide range of scales? Mathematicians can use fractals to develop an intuition for certain mathematical problems. Such intuition in turn brings new conjectures and new approaches to solving important theorems.

Society in general (especially the television and motion picture industries) has taken a keen interest in fractals. Fractals, used with computer graphics, provide beautiful and realistic forgeries of many natural shapes, such as rocks, mountains, landscapes, stars, planets, and the like. The Star Trek television programs and motion pictures are a notable example of applying fractals in that way. Furthermore, fractals are easy and fun to generate, at least for someone who likes personal computers. Commercially available computer software programs show fractals in fascinating color and detail.

These and other important and practical applications came about only as unexpected by-products, after the concept of fractals had crystallized and been developed. That sort of thing happens quite often in the world of research. The applications show that basic research—research that may not promise any obvious benefit when first undertaken—can lead to practical and useful ends.

Summary

A fractal is a line, surface, or pattern that looks the same over a wide range of scales. In other words, a fractal is self-similar (any part of the object, suitably magnified and refocused, looks like the whole), and scale-invariant (lacks a characteristic size). Other typical features of fractals are that a single number (a noninteger dimension) describes their scaling, they can be formed by many repetitions of the same operation, and they are never smooth (being instead rough, jagged, etc.). A chaotic attractor usually is a fractal. The two main types of fractals are deterministic and natural fractals. Deterministic fractals are generated by iterating an equation or by following a specified numerical recipe. Therefore, they are exact. Natural fractals include noise and therefore are only approximately similar under change of scale. Natural fractals simulate natural objects, such as rocks, surface textures, and landscapes.