CHAPTER ONE

Introduction and overview

1.1 Some history

Chaotic dynamics may be said to have started with the work of the French mathematician Henri Poincaré at about the turn of the century. Poincaré's motivation was partly provided by the problem of the orbits of three celestial bodies experiencing mutual gravational attraction (e.g., a star and two planets). By considering the behavior of orbits arising from *sets* of initial points (rather than focusing on *individual* orbits), Poincaré was able to show that very complicated (now called chaotic) orbits were possible. Subsequent noteworthy early mathematical work on chaotic dynamics includes that of G. Birkhoff in the 1920s, M. L. Cartwright and J. E. Littlewood in the 1940s, S. Smale in the 1960s, and Soviet mathematicians, notably A. N. Kolmogorov and his coworkers. In spite of this work, however, the possibility of chaos in real physical systems was not widely appreciated until relatively recently. The reasons for this were first that the mathematical papers are difficult to read for workers in other fields, and second that the theorems proven were often not strong enough to convince researchers in these other fields that this type of behavior would be important in their systems. The situation has now changed drastically, and much of the credit for this can be ascribed to the extensive numerical solution of dynamical systems on digital computers. Using such solutions, the chaotic character of the time evolutions in situations of practical importance has become dramatically clear. Furthermore, the complexity of the dynamics cannot be blamed on unknown extraneous experimental effects, as might be the case when dealing with an actual physical system.

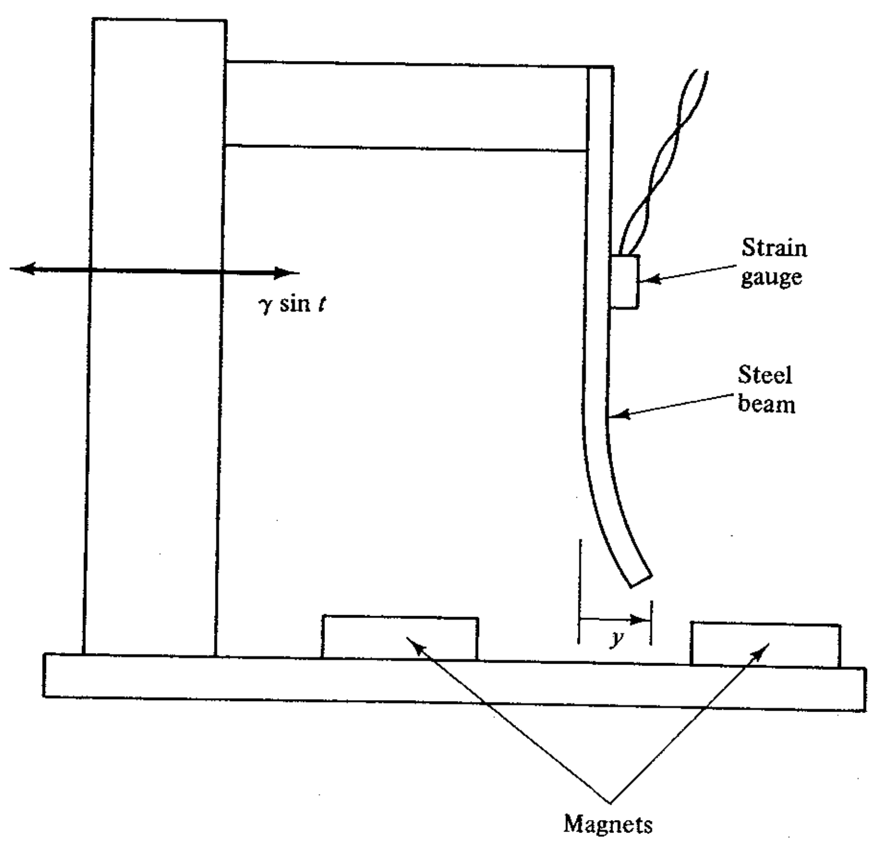
In this chapter, we shall provide some of the phenomenology of chaos and will introduce some of the more basic concepts. The aim is to provide a motivating overview[1] in preparation for the more detailed treatments to be pursued in the rest of this book.

1.2 Examples of chaotic behavior

Most students of science or engineering have seen examples of dynamical behavior which can be fully analyzed mathematically and in which the system eventually (after some transient period) settles either into periodic motion (a limit cycle) or into a steady state (i.e., a situation in which the system ceases its motion). When one relies on being able to specify an orbit analytically, these two cases will typically (and falsely) appear to be the only important motions. The point is that chaotic orbits are also very common but cannot be represented using standard analytical functions. Chaotic motions are neither steady nor periodic. Indeed, they appear to be very complex, and, when viewing such motions, adjectives like wild, turbulent, and random come to mind. In spite of the complexity of these motions, they commonly occur in systems which themselves are not complex and are even surprisingly simple. (In addition to steady state, periodic and chaotic motion, there is a fourth common type of motion, namely quasiperiodic motion. We defer our discussion of quasiperiodicity to Chapter 6.)

Before giving a definition of chaos we first present some examples and background material. As a first example of chaotic motion, we consider an

Figure 1.1 The apparatus of Moon and Holmes (1979).



experiment of Moon and Holmes (1979). The apparatus is shown in Figure 1.1. When the apparatus is at rest, the steel beam has two stable steady-state equilibria: either the tip of the beam is deflected toward the left magnet or toward the right magnet. In the experiment, the horizontal position of the apparatus was oscillated sinusoidally with time. Under certain conditions, when this was done, the tip of the steel beam was observed to oscillate in a very irregular manner. As an indication of this very irregular behavior, Figure 1.2(*a*) shows the output signal of a strain gauge attached to the beam (Figure 1.1). Although the apparatus appears to be very simple, one might attribute the observed complicated motion to complexities in the physical situation, such as the excitation of higher order vibrational modes in the beam, possible noise in the sinusoidal shaking device, etc. To show that it is not necessary to invoke such effects, Moon and Holmes considered a simple model for their experiment, namely, the forced Duffing equation in the following form,

$$\frac{d^2y}{dt^2} + \nu\frac{dy}{dt} + (y^3 - y) = g\sin t. \tag{1.1}$$

In Eq. (1.1), the first two terms represent the inertia of the beam and dissipative effects, while the third term represents the effects of the magnets and the elastic force. The sinusoidal term on the right-hand side represents the shaking of the apparatus. In the absence of shaking ($g = 0$), Eq. (1.1) possesses two stable steady states, $y = 1$ and $y = -1$, corresponding to the two previously mentioned stable steady states of the beam. (There is also an unstable steady state $y = 0$.) Figure 1.2(*b*) shows the results of a digital computer numerical solution of Eq. (1.1) for a particular choice of ν and g. We observe that the results of the physical experiment are qualitatively similar to those of the numerical solution. Thus, it is unnecessary to invoke complicated physical processes to explain the observed complicated motion.

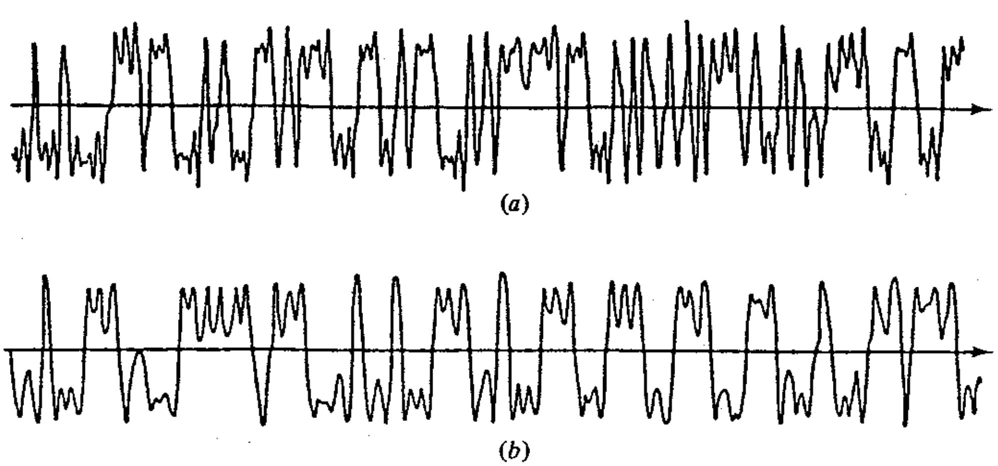
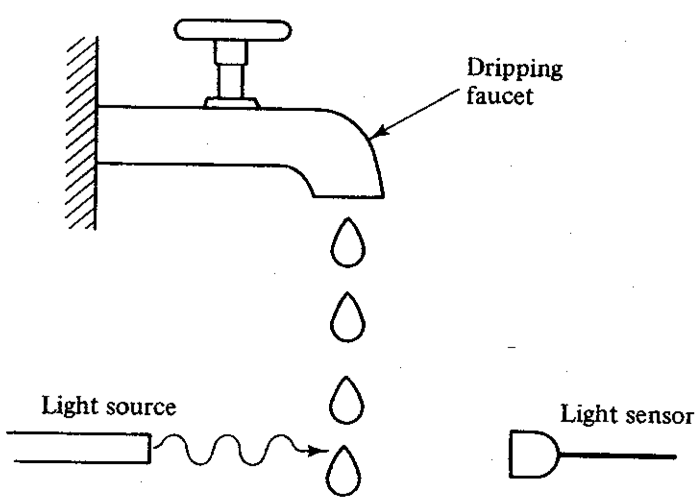


Figure 1.2(*a*) Signal from the strain gauge. (*b*) Numerical solution of Eq. (1.1) (Moon and Holmes 1979).

As a second example, we consider the experiment of Shaw (1984) illustrated schematically in Figure 1.3. In this experiment, a slow steady inflow of water to a 'faucet' was maintained. Water drops fall from the faucet, and the times at which successive drops pass a sensing device are recorded. Thus, the data consists of the discrete set of times $t_1, t_2, \ldots, t_n, \ldots$ at which drops were observed by the sensor. From these data, the time intervals between successive drops can be formed, $\Delta t_n \equiv t_{n+1} - t_n$. When the inflow rate to the faucet is sufficiently small, the time intervals Δt_n are all equal. As the inflow rate is increased, the time interval sequence becomes periodic with a short interval Δt_a followed by a longer interval Δt_b, so that the sequence of time intervals is of the form $\ldots, \Delta t_a, \Delta t_b, \Delta t_a, \Delta t_b, \Delta t_a, \ldots$. We call this a period two sequence since $\Delta t_n = \Delta t_{n+2}$. As the inflow rate is further increased, periodic sequences of longer and longer periods were observed, until, at sufficiently large inflow rate, the sequence $\Delta t_1, \Delta t_2, \Delta t_3, \ldots$ apparently has no regularity. This irregular sequence is argued to be due to chaotic dynamics (see Section 2.4.3).

As a third example, we consider the problem of chaotic Rayleigh–Benard convection, originally studied theoretically and computationally in the seminal paper of Lorenz (1963) and experimentally by, for example, Ahlers and Behringer (1978), Gollub and Benson (1980), Bergé *et al.* (1980) and Libchaber and Maurer (1980). In Rayleigh–Benard convection, one considers a fluid contained between two rigid plates and subjected to gravity, as shown in Figure 1.4. The bottom plate is maintained at a higher temperature $T_0 + \Delta T$ than the temperature T_0 of the top plate. As a result, the fluid near the warmer lower plate expands, and buoyancy creates a tendency for this fluid to rise. Similarly, the cooler more dense fluid near the top plate has a tendency to fall. While Lorenz's equations are too idealized a model to describe the experiments accurately, in the case where the experiments were done with vertical bounding

Figure 1.3 Schematic illustration of the experiment of Shaw (1984).



side-walls situated at a spacing of two to three times the distance between the horizontal walls, there was a degree of qualitative correspondence between the model and the experiments. In particular, in this case, for some range of values of the temperature difference ΔT, the experiments show that the fluid will execute a *steady* convective cellular flow, as shown in the figure. At a somewhat larger value of the temperature difference, the flow becomes time-dependent, and this time dependence is chaotic. This general behavior is also predicted by Lorenz's paper.

From these simple examples, it is clear that chaos should be expected to be a very common basic dynamical state in a wide variety of systems. Indeed, chaotic dynamics has by now been shown to be of potential importance in many different fields including fluids,[2] plasmas,[3] solid state devices,[4] circuits,[5] lasers,[6] mechanical devices,[7] biology,[8] chemistry,[9] acoustics,[10] celestial mechanics,[11] etc.

In both the dripping faucet example and the Rayleigh–Benard convection example, our discussions indicated a situation as shown schematically in Figure 1.5. Namely, there was a system parameter, labeled p in Figure 1.5, such that, at a value $p = p_1$, the motion is observed to be nonchaotic, and at another value $p = p_2$, the motion is chaotic. (For the faucet example, p is the inflow rate, while for the example of Rayleigh–Benard convection, p is the temperature difference ΔT.) The natural question raised by Figure 1.5 is *how does chaos come about as the parameter p is varied continuously from p_1 to p_2?* That is, how do the dynamical motions of the system evolve with continuous variation of p from p_1 to p_2? This question of the *routes to chaos*[12] will be considered in detail in Chapter 8.

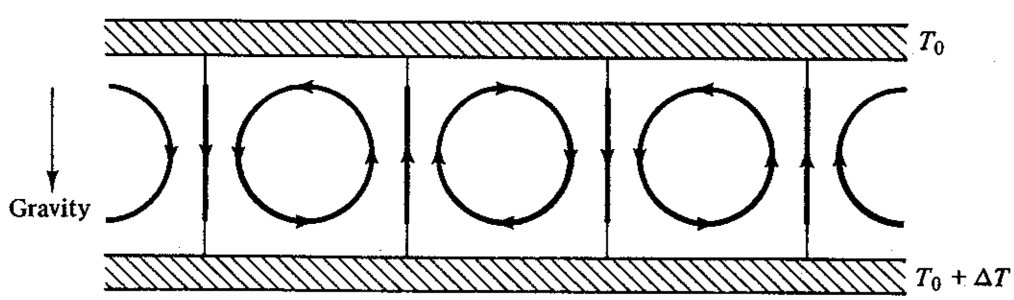


Figure 1.4 Rayleigh–Benard convection.

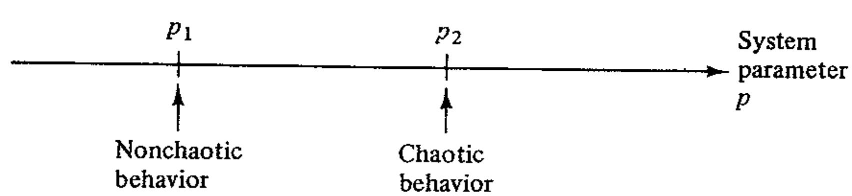


Figure 1.5 Schematic illustration of the question of the transition to chaos with variation of a system parameter.

1.3 Dynamical systems

A *dynamical system* may be defined as a deterministic mathematical prescription for evolving the state of a system forward in time. Time here either may be a continuous variable, or else it may be a discrete integer-valued variable. An example of a dynamical system in which time (denoted t) is a continuous variable is a system of N first-order, autonomous, ordinary differential equations,

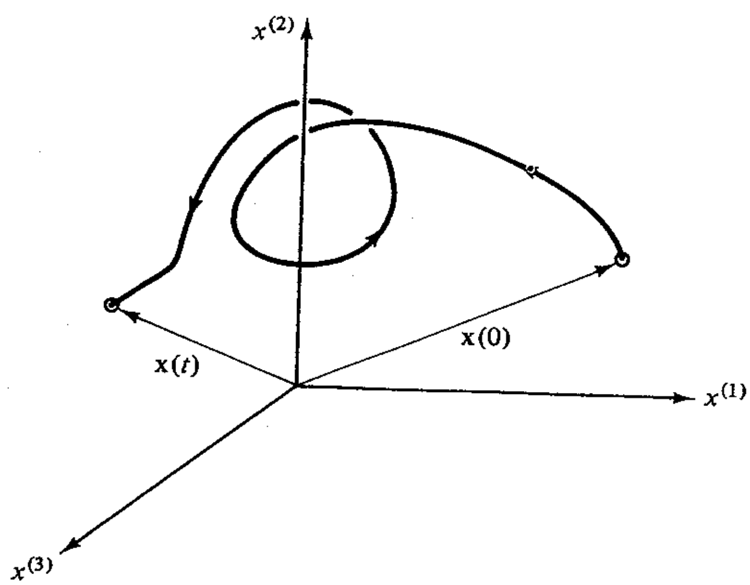
$$\left.\begin{aligned} \mathrm{d}x^{(1)}/\mathrm{d}t &= F_1(x^{(1)}, x^{(2)}, \ldots, x^{(N)}), \\ \mathrm{d}x^{(2)}/\mathrm{d}t &= F_2(x^{(1)}, x^{(2)}, \ldots, x^{(N)}), \\ &\vdots \\ \mathrm{d}x^{(N)}/\mathrm{d}t &= F_N(x^{(1)}, x^{(2)}, \ldots, x^{(N)}), \end{aligned}\right\} \tag{1.2}$$

which we shall often write in vector form as

$$\mathrm{d}\mathbf{x}(t)/\mathrm{d}t = \mathbf{F}[\mathbf{x}(t)], \tag{1.3}$$

where $\mathbf{x}$ is an N-dimensional vector. This is a dynamical system because, for any initial state of the system $\mathbf{x}(0)$, we can in principle solve the equations to obtain the future system state $\mathbf{x}(t)$ for $t > 0$. Figure 1.6 shows the path followed by the system state as it evolves with time in a case where $N = 3$. The space $(x^{(1)}, x^{(2)}, x^{(3)})$ in the figure is referred to as *phase space*, and the path in phase space followed by the system as it evolves with time is referred to as an *orbit* or *trajectory*. Also, it is common to refer to a continuous time dynamical system as a *flow*. (This latter terminology is apparently motivated by considering the trajectories generated by *all* the initial conditions in the phase space as roughly analogous to the paths followed by the particles of a flowing fluid.)

Figure 1.6 An orbit in a three-dimensional ($N = 3$) phase space.



In the case of discrete, integer-valued time (with n denoting the time variable, $n = 0, 1, 2, \ldots$), an example of a dynamical system is a map, which we write in vector form as

$$\mathbf{x}_{n+1} = \mathbf{M}(\mathbf{x}_n), \tag{1.4}$$

where $\mathbf{x}_n$ is N-dimensional, $\mathbf{x}_n = (x_n^{(1)}, x_n^{(2)}, \ldots, x_n^{(N)})$. Given an initial state $\mathbf{x}_0$, we obtain the state at time $n = 1$ by $\mathbf{x}_1 = \mathbf{M}(\mathbf{x}_0)$. Having determined $\mathbf{x}_1$, we can then determine the state at $n = 2$ by $\mathbf{x}_2 = \mathbf{M}(\mathbf{x}_1)$, and so on. Thus, given an initial condition $\mathbf{x}_0$, we generate an orbit (or trajectory) of the discrete time system: $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$. As we shall see, a continuous time system of dimensionality N can often profitably be reduced to a discrete time map of dimensionality $N - 1$ via the Poincaré surface of section technique.

It is reasonable to conjecture that the complexity of the possible structure of orbits can be greater for larger system dimensionality. Thus, a natural question is *how large does N have to be in order for chaos to be possible*? For the case of N first-order autonomous ordinary differential equations, the answer is that

$$N \geq 3 \tag{1.5}$$

is sufficient.[13] Thus, if one is given an autonomous first-order system with $N = 2$, chaos can be ruled out immediately.

Example: Consider the forced damped pendulum equation (cf. Figure 1.7)

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \nu\frac{\mathrm{d}\theta}{\mathrm{d}t} + \sin\theta = T\sin(2\pi f t), \tag{1.6a}$$

where the first term represents inertia, the second, friction at the pivot, the third, gravity, and the term on the right-hand side represents a sinusoidal torque applied at the pivot. (This equation also describes the behavior of a simple Josephson junction circuit.) We ask: is chaos ruled out for the

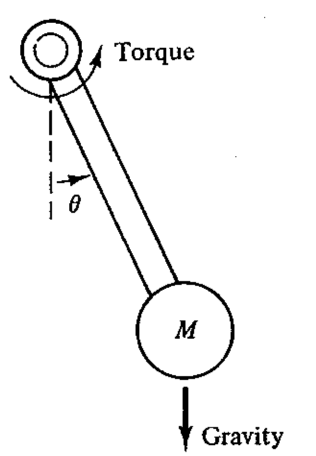


Figure 1.7 Forced, damped pendulum.

 $\nu \Rightarrow f \quad 2\pi f \Rightarrow \omega$

driven damped pendulum equation? To answer this question, we put the equation (which is second-order and nonautonomous) into first-order autonomous form by the substitution

$$x^1 = \mathrm{d}\theta/\mathrm{d}t,$$
$$x^{(2)} = \theta,$$
$$x^{(3)} = 2\pi ft.$$

(Note that, since both $x^{(2)}$ and $x^{(3)}$ appear in Eq. (1.6a) as the argument of a sine function, they can be regarded as angles and may, if desired, be defined to lie between 0 and 2π.) The driven damped pendulum equation then yields the following first-order autonomous system,

$$\left.\begin{aligned} \mathrm{d}x^{(1)}/\mathrm{d}t &= T\sin x^{(3)} - \sin x^{(2)} - \nu x^{(1)}, \\ \mathrm{d}x^{(2)}/\mathrm{d}t &= x^{(1)}, \\ \mathrm{d}x^{(3)}/\mathrm{d}t &= 2\pi f. \end{aligned}\right\} \tag{1.6b}$$

Since $N = 3$, chaos is not ruled out. Indeed, numerical solutions show that both chaotic and periodic solutions of the driven damped pendulum equation are possible depending on the particular choice of system parameters ν, T, and f.

We now consider the question of the required dimensionality for chaos for the case of maps. In this case, we must distinguish between invertible and noninvertible maps. We say the map $\mathbf{M}$ is invertible if, given $\mathbf{x}_{n+1}$, we can solve $\mathbf{x}_{n+1} = \mathbf{M}(\mathbf{x}_n)$ uniquely for $\mathbf{x}_n$. If this is so, we denote the solution for $\mathbf{x}_n$ as

$$\mathbf{x}_n = \mathbf{M}^{-1}(\mathbf{x}_{n+1}), \tag{1.7}$$

and we call $\mathbf{M}^{-1}$ the inverse of $\mathbf{M}$. For example, consider the one-dimensional ($N = 1$) map[14],

$$M(x) = rx(1 - x), \tag{1.8}$$

which is commonly called the 'logistic map.' As shown in Figure 1.8, this map is not invertible because for a given x_{n+1} there are two possible values of x_n from which it could have come. On the other hand, consider the two-dimensional map,

$$\begin{aligned} x_{n+1}^{(1)} &= f(x_n^{(1)}) - Jx_n^{(2)}, \\ x_{n+1}^{(2)} &= x_n^{(1)}. \end{aligned} \tag{1.9}$$

This map is clearly invertible as long as $J \neq 0$,

$$\begin{aligned} x_n^{(1)} &= x_{n+1}^{(2)}, \\ x_n^{(2)} &= J^{-1}[f(x_{n+1}^{(2)}) - x_{n+1}^{(1)}]. \end{aligned} \tag{1.10}$$

We can now state the dimensionality requirements on maps. If the map is invertible, then there can be no chaos unless

$$N \geq 2. \tag{1.11}$$

If the map is noninvertible, chaos is possible even in one-dimensional maps. Indeed, the logistic map Eq. (1.8) exhibits chaos for large enough r.

It is often useful to reduce a continuous time system (or 'flow') to a discrete time map by a technique called the Poincaré surface of section method. We consider N first-order autonomous ordinary differential equations (Eq. (1.2)). The 'Poincaré map' represents a reduction of the N-dimensional flow to an $(N-1)$-dimensional map. For illustrative purposes, we take $N = 3$ and illustrate the construction in Figure 1.9. Consider a solution of (1.2). Now, choose some appropriate $(N-1)$-dimensional surface (the 'surface of section') in the N-dimensional phase space, and observe the intersections of the orbit with the surface. In Figure 1.9, the surface of section is the plane $x^{(3)} = K$, but we emphasize that in general the choice of the surface can be tailored in a convenient way to the particular problem. Points A and B represent two successive crossings of the surface of section. Point A uniquely determines point B, because A can be used as an initial condition in (1.2) to determine B. Likewise, B uniquely determines A by reversing time in (1.2) and using B as the initial condition. Thus, the Poincaré map in this illustration represents an invertible two-dimensional map transforming the coordinates $(x_n^{(1)}, x_n^{(2)})$ of the nth piercing of the surface of section to the coordinates $(x_{n+1}^{(1)}, x_{n+1}^{(2)})$ at piercing $n + 1$. This equivalence of an N-dimensional flow with an $(N-1)$-dimensional invertible map shows that the requirement Eq. (1.11) for chaos in a map follows from Eq. (1.5) for chaos in a flow.

Another way to create a map from the flow generated by the system of autonomous differential equations (1.3) is to sample the flow at discrete times $t_n = t_0 + nT$ $(n = 0, 1, 2, \ldots)$, where the sampling interval T can be chosen on the basis of convenience. Thus, a continuous time trajectory $\mathbf{x}(t)$ yields a discrete time trajectory $\mathbf{x}_n \equiv \mathbf{x}(t_n)$. The quantity $\mathbf{x}_{n+1}$ is

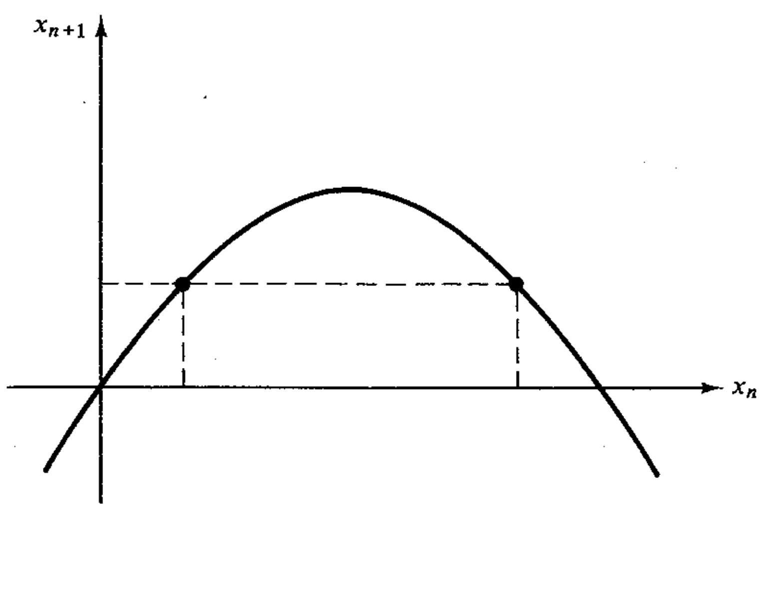


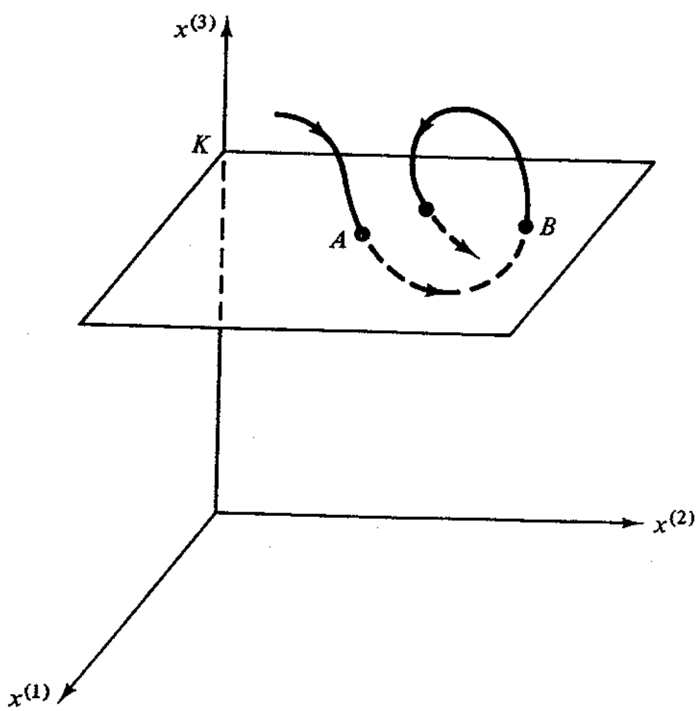
Figure 1.8 Noninvertibility of the logistic map.

uniquely determined from $\mathbf{x}_n$ since we can use $\mathbf{x}_n$ as an initial condition in Eqs. (1.3) and integrate the equations forward for an amount of time T to determine $\mathbf{x}_{n+1}$. Thus, in principle, we have a map $\mathbf{x}_{n+1} = \mathbf{M}(\mathbf{x}_n)$. We call this map the time T map. The time T map is invertible (like the Poincaré map), since the differential equations (1.3) can be integrated backward in time. Unlike the Poincaré map, the dimensionality of the time T map is the same as that of the flow.

1.4 Attractors

In Hamiltonian systems (cf. Chapter 7) such as arise in Newton's equations for the motion of particles without friction, there are choices of the phase space variables (e.g., the canonically conjugate position and momentum variables) such that phase space volumes are preserved under the time evolution. That is, if we choose an initial ($t = 0$) closed $(N - 1)$-dimensional surface S_0 in the N-dimensional $\mathbf{x}$-phase space, and then evolve each point on the surface S_0 forward in time by using them as initial conditions in Eq. (1.3), then the closed surface S_0 evolves to a closed surface S_t at some later time t, and the N-dimensional volumes $V(0)$ of the region enclosed by S_0 and $V(t)$ of the region enclosed by S_t are the same, $V(t) = V(0)$. We call such a volume preserving system *conservative*. On the other hand, if the flow does not preserve volumes, and cannot be made to do so by a change of variables, then we say that the system is

Figure 1.9 A Poincaré surface of section.



nonconservative. By the divergence theorem, we have that

$$\mathrm{d}V(t)/\mathrm{d}t = \int_{S_t} \nabla\cdot\mathbf{F}d^N x, \tag{1.12}$$

where $\int_{S_t}$ signifies the integral over the volume interior to the surface S_t, and $\nabla\cdot\mathbf{F} \equiv \Sigma_{i=1}^{N} \partial F_i(x^{(1)},\ldots,x^{(N)})/\partial x^{(i)}$. For example, for the forced damped pendulum equation written in first-order autonomous form, Eq. (1.6b), we have that $\nabla\cdot\mathbf{F} = -\nu$, which is independent of the phase space position $\mathbf{x}$ and is negative. From (1.12), we have $\mathrm{d}V(t)/\mathrm{d}t = -\nu V(t)$ so that V decreases exponentially with time, $V(t) = \exp(-\nu t)V(0)$. In general, $\nabla\cdot\mathbf{F}$ will be a function of the phase space position $\mathbf{x}$. If $\nabla\cdot\mathbf{F} < 0$ in some region of phase space (signifying volume contraction in that region), then we shall refer to the system as a *dissipative* system. It is an important concept in dynamics that dissipative systems typically are characterized by the presence of attracting sets or *attractors* in the phase space. These are bounded subsets to which regions of initial conditions of nonzero phase space volume asymptote as time increases. (Conservative dynamical systems do not have attractors; see the discussion of the Poincaré recurrence theorem in Chapter 7).

As an example of an attractor, consider the damped harmonic oscillator, $\mathrm{d}^2y/\mathrm{d}t^2 + \nu\,\mathrm{d}y/\mathrm{d}t + \omega^2 y = 0$. A typical trajectory in the phase space ($x^{(1)} = y$, $x^{(2)} = \mathrm{d}y/\mathrm{d}t$) is shown in Figure 1.10(*a*). We see that, as time goes on, the orbit spirals into the origin, and this is true for any initial condition. Thus, in this case the origin, $x^{(1)} = x^{(2)} = 0$, is said to be the 'attractor' of the dynamical system. As a second example, Figure 1.10(*b*) shows the case of a limit cycle (the dashed curve). The initial condition (labeled α) outside the limit cycle yields an orbit which, with time, spirals into the closed dashed curve on which it circulates in periodic motion in the $t \to +\infty$ limit. Similarly, the initial condition (labeled β) inside the limit cycle yields an orbit which spirals outward, asymptotically

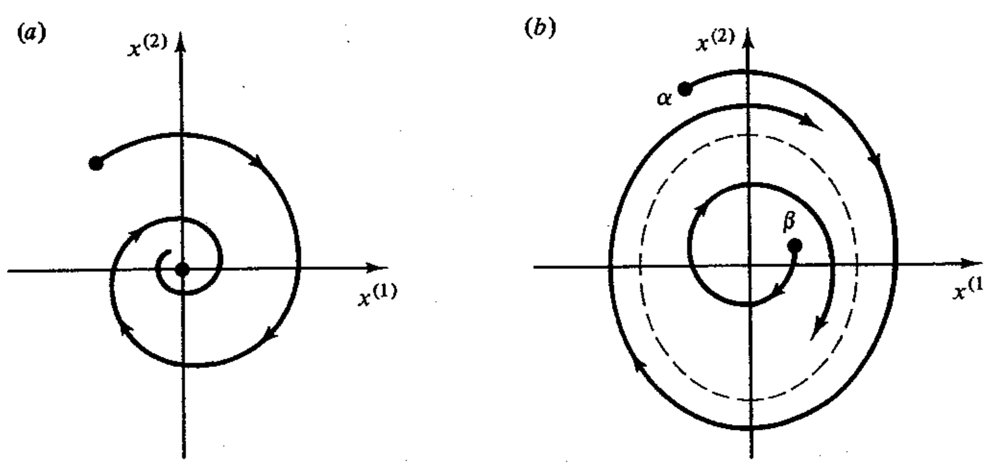


Figure 1.10(*a*) The attractor is the point at the origin. (*b*) The attractor is the closed dashed curve.

approaching the dashed curve. Thus, in this case, the dashed closed curve is the attractor. An example of an equation displaying a limit cycle attractor as illustrated in Figure 1.10(*b*) is the van der Pol equation,

$$\frac{d^2y}{dt^2} + (y^2 - \eta)\frac{dy}{dt} + \omega^2 y = 0. \tag{1.13}$$

This equation was introduced in the 1920s as a model for a simple vacuum tube oscillator circuit.

One can speak of conservative and dissipative *maps*. A conservative N-dimensional map is one which preserves N-dimensional phase space volumes on each iterate (or else can be made to do so by a suitable change of variables). A map is volume preserving if the magnitude of the determinant of its Jacobian matrix of partial derivatives is one,

$$J(\mathbf{x}) \equiv |\det[\partial \mathbf{M}(\mathbf{x})/\partial \mathbf{x}]| = 1.$$

For example, for a continuous time Hamiltonian system, a surface of section formed by setting one of the N canonically conjugate variables equal to a constant can be shown to yield a volume preserving map in the remaining $N - 1$ canonically conjugate variables (Chapter 7). On the other hand, if $J(\mathbf{x}) < 1$ in some regions, then we say the map is dissipative, and, as for flows, typically it can have attractors. For example, Figure 1.11 illustrates the Poincaré surface of section map for a three-dimensional

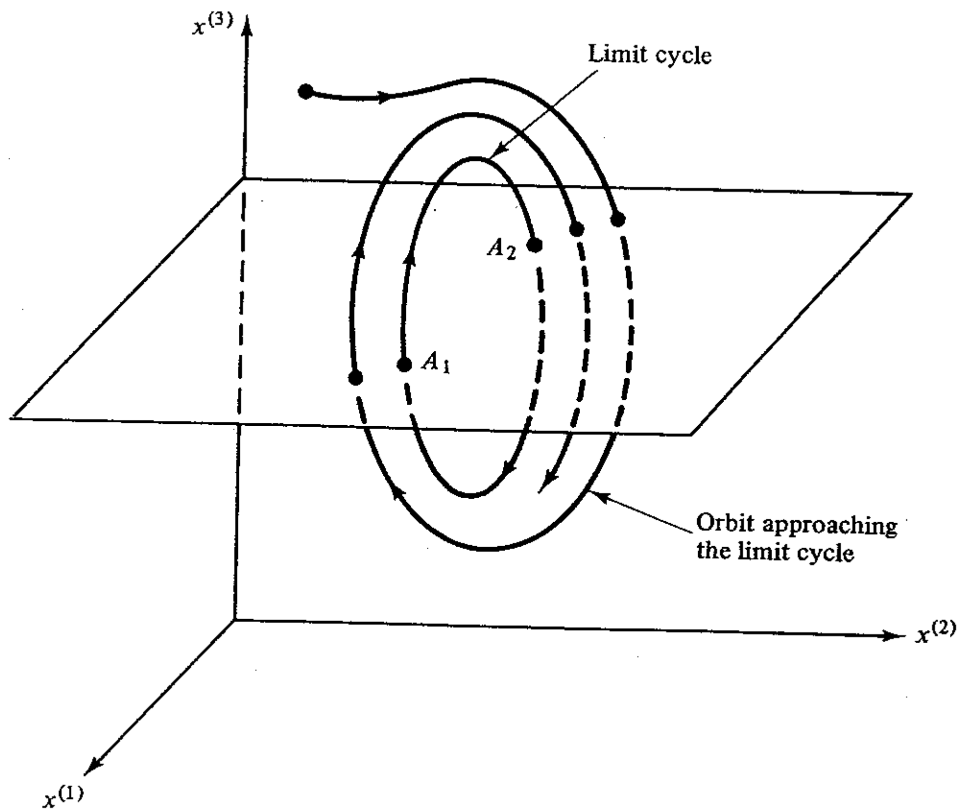


Figure 1.11 Surface of section for a three-dimensional flow with a limit cycle.

flow with a limit cycle. We see that *for the map*, the two points A_1 and A_2 together constitute the attractor. That is, the orbit of the two-dimensional surface of section map $\mathbf{x}_{n+1} = \mathbf{M}(\mathbf{x}_n)$ yields a sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ which converges to the set consisting of the two points A_1 and A_2, between which the map orbit sequentially alternates in the limit $n \to +\infty$.

In Figure 1.10, we have two examples, one in which the attractor of a continuous time system is a set of dimension zero (a single point) and one in which the attractor is a set of dimension one (a closed curve). In Figure 1.11, the attractor of the map has dimension zero (it is the two points, A_1 and A_2). It is a characteristic of chaotic dynamics that the resulting attractors often have a much more intricate geometrical structure in the phase space than do the examples of attractors cited above. In fact, according to a standard definition of dimension (Section 3.1), these attractors commonly have a value for this dimension which is not an integer. In the terminology of Mandelbrot, such geometrical objects are *fractals*. When an attractor is fractal, it is called a *strange attractor*.

As an example of a strange attractor, consider the attractor obtained for the two-dimensional Hénon map,

$$\left.\begin{aligned} x_{n+1}^{(1)} &= A - (x_n^{(1)})^2 + Bx_n^{(2)}, \\ x_{n+1}^{(2)} &= x_n^{(1)}, \end{aligned}\right\} \tag{1.14}$$

for $A = 1.4$ and $B = 0.3$. See Hénon (1976). (Note that Eq. (1.14) is in the form of Eq. (1.9).) Figure 1.12(a) shows the results of plotting 10^4 successive points obtained by iterating Eqs. (1.14) (with the initial transient before the orbit settles into the attractor deleted). The result is essentially a picture of the attractor. Figure 1.12(b) shows that a blow-up of the rectangle in Figure 1.12(a) reveals that the attractor apparently has a local small-scale structure consisting of a number of parallel lines. A blow-up of the rectangle in Figure 1.12(b) is shown in Figure 1.12(c) and reveals more lines. Continuation of this blow-up procedure would show that the attractor has similar structure on *arbitrarily small scale*. In fact, roughly speaking, we can regard the attractor in Figure 1.12(b) as consisting of an *uncountable* infinity of lines. Numerical computations show that the fractal dimension D_0 of the attractor in Figure 1.12 is a number between one and two, $D_0 \simeq 1.26$. Hence, this appears to be an example of a strange attractor.

As another example of a strange attractor, consider the forced damped pendulum (Eqs. (1.6) and Figure 1.7) with $\nu = 0.22$, $T = 2.7$, and $f = 1/2\pi$. Treating $x^{(3)}$ as an angle in phase space, we define

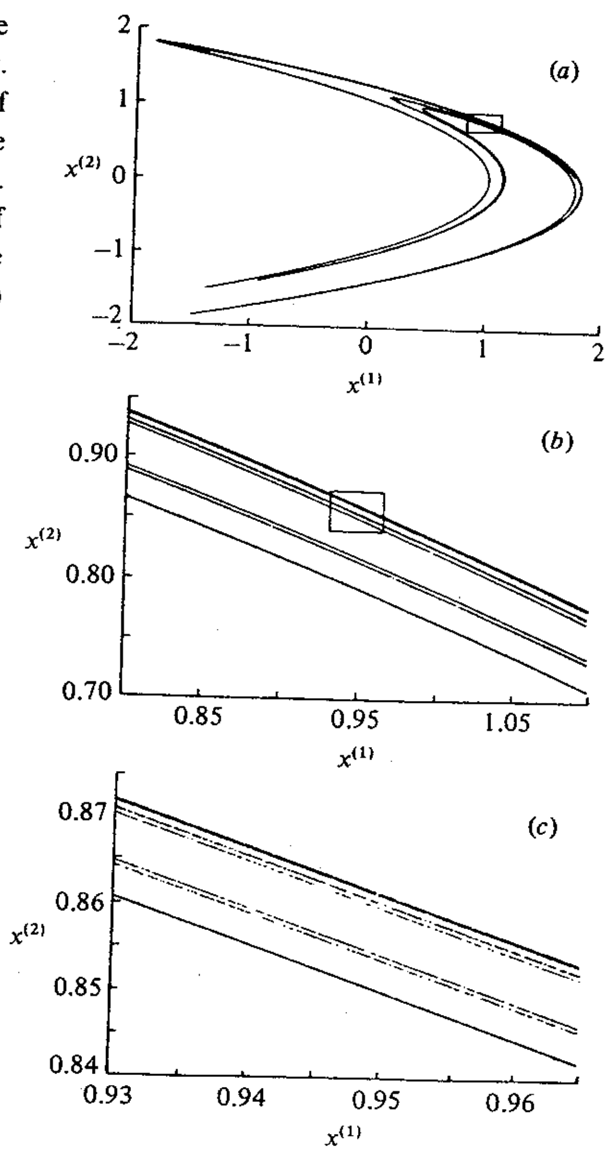
$$\bar{x}^{(3)} = x^{(3)} \text{ modulo } 2\pi$$

and choose a surface of section $\bar{x}^{(3)} = 0$. The modulo operation is defined as

$$y \text{ modulo } K \equiv y + pK.$$

where p is a positive or negative integer chosen to make $0 \leq y + pK < K$. The surface of section $\bar{x}^{(3)} = 0$ is crossed at the times $t = 0, 2\pi, 4\pi, 6\pi, \ldots$. (This type of surface of section for a periodically forced system is often referred to as a *stroboscopic* surface of section, since it shows the system state at successive 'snapshots' of the system at evenly spaced time intervals.) As seen in Figure 1.13(*a*) and in the blow-up of the rectangle (Figure 1.13(*b*)), the attractor again apparently consists of a number of parallel curves. The fractal dimension of the intersection of the attractor with the surface of section in this case is approximately 1.38. Correspondingly, if one considers the attracting set in the full three-dimensional phase space, it has a dimension 2.38 (i.e., one greater than its intersection with the surface of section).

Figure 1.12(*a*) The Hénon attractor. (*b*) Enlargement of region defined by the rectangle in (*a*). (*c*) Enlargement of region defined by the rectangle in (*b*) (Grebogi *et al.* 1987d).



1.5 Sensitive dependence on initial conditions

A defining attribute of an attractor on which the dynamics is *chaotic* is that it displays exponentially sensitive dependence on initial conditions. Consider two nearby initial conditions $\mathbf{x}_1(0)$ and $\mathbf{x}_2(0) = \mathbf{x}_0 + \Delta(0)$, and imagine that they are evolved forward in time by a continuous time dynamical system yielding orbits $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ as shown in Figure 1.14. At time t, the separation between the two orbits is $\Delta(t) = \mathbf{x}_2(t) - \mathbf{x}_1(t)$. If, in the limit $|\Delta(0)| \to 0$, and large t, orbits remain bounded and the difference between the solutions $|\Delta(t)|$ grows exponentially for typical orientation of the vector $\Delta(0)$ (i.e., $|\Delta(t)|/|\Delta(0)| \sim \exp(ht)$, $h > 0$), then we say that the system displays sensitive dependence on initial conditions and

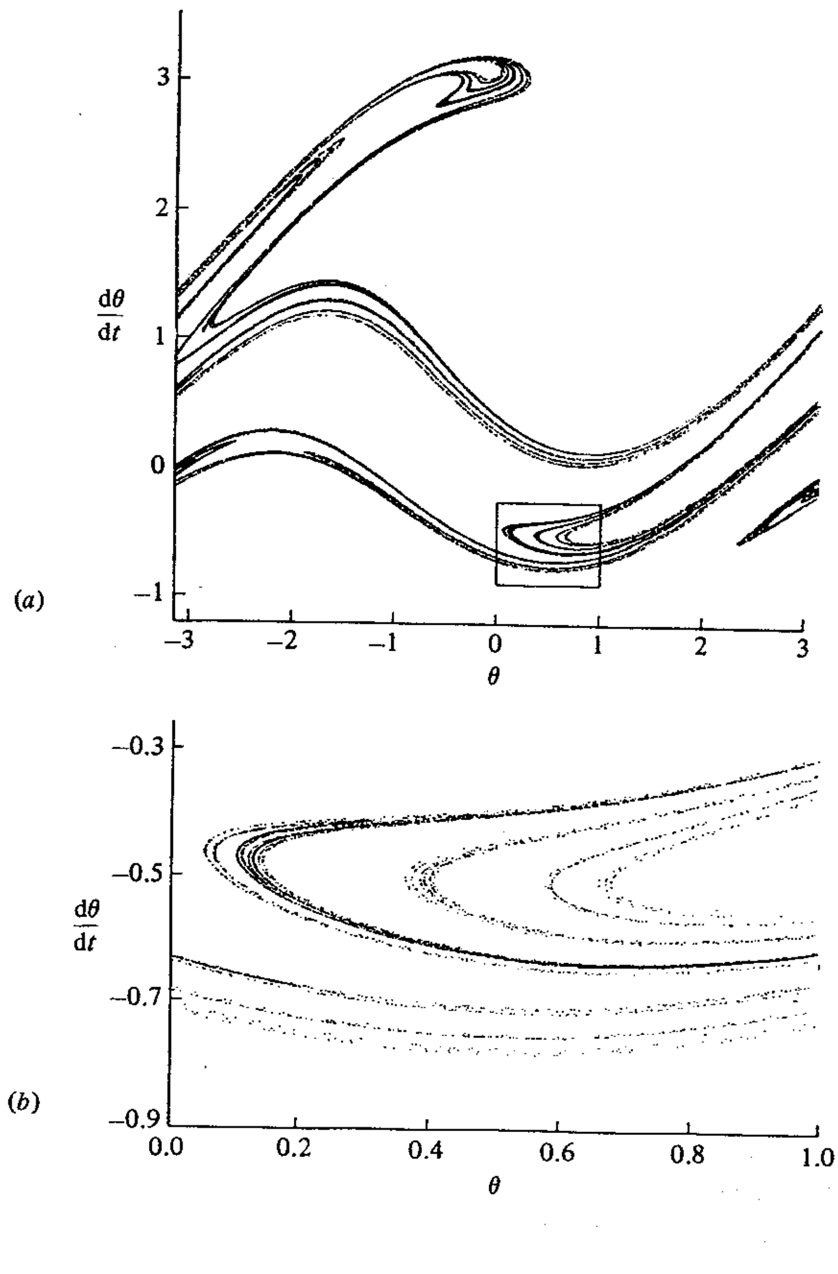
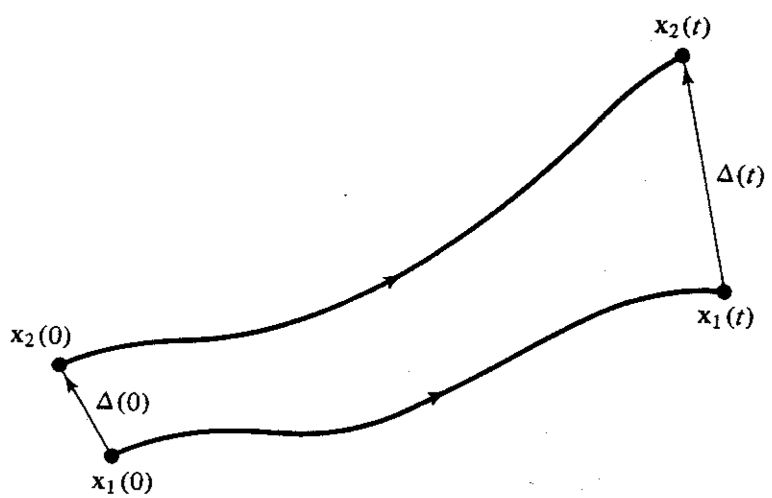


Figure 1.13 The attractor of the forced damped pendulum equation in the surface of section $x^{(3)}$ modulo $2\pi = 0$ (Grebogi *et al.* 1987d).

is chaotic. By bounded solutions, we mean that there is some ball in phase space, $|\mathbf{x}| < R < \infty$, which solutions never leave.[15] (Thus, if the motion is on an attractor, then the attractor lies in $|\mathbf{x}| < R$.) The reason we have imposed the restriction that orbits remain bounded is that, if orbits go to infinity, it is relatively simple for their distances to diverge exponentially. An example is the single, autonomous, linear, first-order differential equation $\mathrm{d}x/\mathrm{d}t = x$. This yields $\mathrm{d}[x_2(t) - x_1(t)]/\mathrm{d}t = [x_2(t) - x_1(t)]$ and hence $\Delta(t) \sim \exp(t)$. Our requirement of bounded solutions eliminates such trivial cases.[16] For the case of the driven damped pendulum equation, we defined three phase space variables, one of which was $x^{(3)} = 2\pi f t$. As defined, $x^{(3)}$ is unbounded since it is proportional to t. The reason we can speak of the driven damped pendulum as being chaotic is that, as previously mentioned, $x^{(3)}$ only occurs as the argument of a sine, and hence it (as well as $x^{(2)} = \theta$) can be regarded as an angle. Thus, the phase space coordinates can be taken as $x^{(1)}, \tilde{x}^{(2)}, \tilde{x}^{(3)}$, where $\tilde{x}^{(2,3)} \equiv x^{(2,3)}$ modulo 2π. Since the variables $\tilde{x}^{(2)}$ and $\tilde{x}^{(3)}$ lie between 0 and 2π, they are necessarily bounded.

The exponential sensitivity of chaotic solutions means that, as time goes on, small errors in the solution can grow very rapidly (i.e., exponentially) with time. Hence, after some time, effects such as noise and computer roundoff can totally change the solution from what it would be in the absence of these effects. As an illustration of this, Figure 1.15 shows the results of a computer experiment on the Hénon map, Eq. (1.14), with $A = 1.4$ and $B = 0.3$. In this figure, we show a picture of the attractor (as in Figure 1.12(*a*)) superposed on which are two computations of iterate numbers 32–36 of an orbit originating from the single initial condition $(x_0^{(1)}, x_0^{(2)}) = (0, 0)$ (labeled as an asterisk in the figure). The two computations of the orbits are done identically, but one uses single precision and

Figure 1.14 Evolution of two nearby orbits in phase space.



the other double precision. The roundoff error in the single precision computation is about 10^{-14}. The orbit computed using single precision is shown as open diamonds, while the orbit using double precision is shown as asterisks. A straight line joins the two orbit locations at each iterate. We see that the difference in the two computations has become as large as the variables themselves. Thus, we cannot meaningfully compute the orbit on the Hénon attractor using a computer with 10^{-14} roundoff for more than of the order of 30–40 iterates. Hence, given the state of a chaotic system, its future becomes difficult to predict after a certain point. Returning to the Hénon map example, we note that, after the first iterate, the two solutions differ by of the order of 10^{-14} (the roundoff). If the subsequent computations were made *without error*, and the error doubled on each iterate (i.e., an exponential increase of $2^n = \exp(n \ln 2)$), then the orbits would be separated by an amount of the order of the attractor size at a time roughly determined by $2^n 10^{-14} \sim 1$ or $n \sim 45$. If errors double on

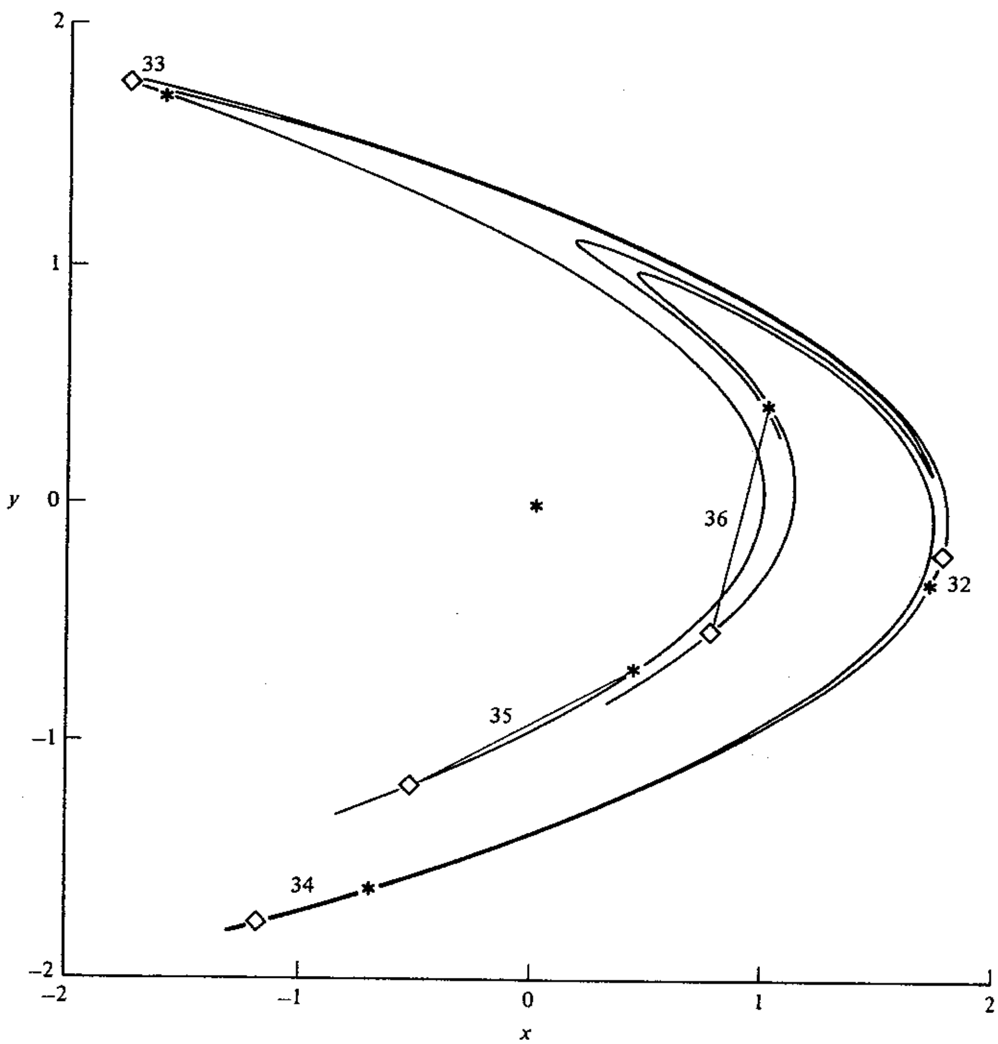


Figure 1.15 After a relatively small number of iterates, two trajectories, one computed using single precision, the other computed using double precision, both originating from the same initial condition, are far apart (This figure courtesy of Y. Du).

each iterate, it becomes almost impossible to improve prediction. Say, we can compute exactly, but our initial measurement of the system state is only accurate to within 10^{-14}. The above shows that we cannot predict the state of the system past $n \sim 45$. Suppose that we wish to predict to a longer time, say, twice as long, i.e., to $n \sim 90$. Then we must improve the accuracy of our initial measurement from 10^{-14} to 10^{-28}. That is, we must improve our accuracy by a tremendous amount, namely, 14 orders of magnitude! In any practical situation, this is likely to be impossible. Thus, the relatively modest goal of an improvement of prediction time by a factor of two is not feasible.

The fact that chaos may make prediction past a certain time difficult, and essentially impossible in a practical sense, has important consequences. Indeed, the work of Lorenz was motivated by the problem of weather prediction. Lorenz was concerned with whether it is possible to do long-range prediction of atmospheric conditions. His demonstration that thermally driven convection could result in chaos raises the possibility that the atmosphere is chaotic. Thus, even the smallest perturbation, such as a butterfly flapping its wings, *eventually* has a large effect. Long-term prediction becomes impossible.

Given the difficulty of accurate computation, illustrated in Figure 1.15, one might question the validity of pictures such as Figures 1.12 and 1.13 which show thousands of iterates of the Hénon map. Is the figure real, or is it merely an artifact of chaos-amplified computer roundoff? A partial answer to this question comes from rigorous mathematical proofs of the *shadowing* property for certain chaotic systems. Although a numerical trajectory diverges exponentially from the true trajectory with the same initial condition, there exists a true (i.e. errorless) trajectory with a slightly

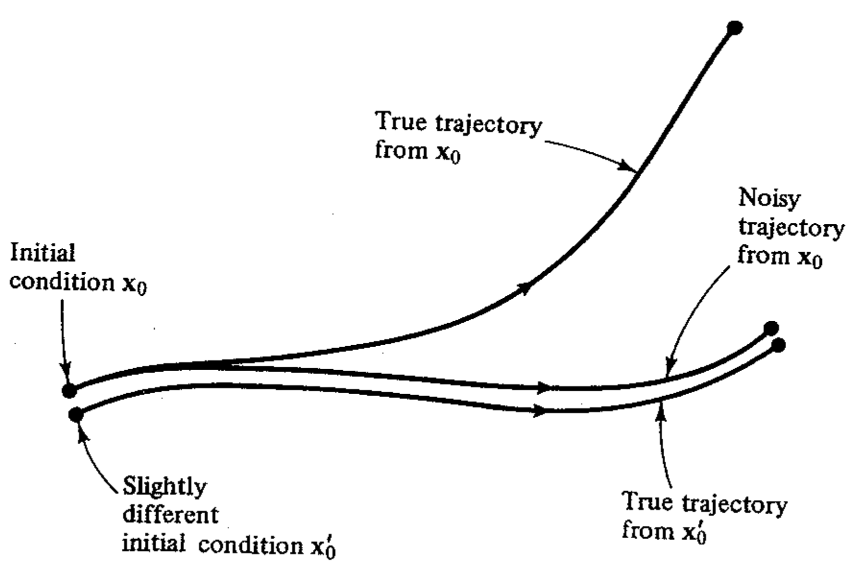


Figure 1.16 Given a noisy trajectory from the initial condition x_0, it is possible to find a slightly different initial condition x_0', such that the true (i.e., noiseless) trajectory from x_0' shadows the noisy trajectory from x_0.

different initial condition (Fig. 1.16) that stays near (shadows) the numerical trajectory (Anosov, 1967; Bowen, 1970; Hammel, Yorke and Grebogi, 1987). Thus, there is good reason to believe that the apparent fractal structure seen in pictures like Figures 1.12 and 1.13 is real.

We emphasize that the nonchaotic cases, shown in Figure 1.10(a) and 1.10(b), do not yield long-term exponential divergence of solutions. For the damped harmonic oscillator example (Figure 1.10(a)), two initially nearby points approach the point attractor and their energies decrease exponentially to zero with time. Hence, orbits *converge* exponentially for large time. For the case of a limit cycle (Figure 1.10(b)), orbits initially separated by an amount $\Delta(0)$ typically eventually wind up on the limit cycle attractor separated by an amount of order $|\Delta(0)|$ and maintain a separation of this order forever. Thus, a small initial error leads to small errors *for all time*. As another example, consider the motion of a particle in a one-dimensional anharmonic potential well in the absence of friction (a conservative system). The total particle energy (potential energy plus kinetic energy) is constant with time on an orbit. Each orbit is periodic and the period depends on the particle energy. Two nearby initial conditions, in general, will have slightly different energies and hence slightly different orbit frequencies. This leads to divergence of these orbits, but the divergence is only linear with time rather than exponential; $|\Delta(t)| \sim (\Delta\omega)t$, where $\Delta\omega$ is the difference of the orbital frequencies. Thus, if $|\Delta(0)|$ is reduced by a factor of two (reducing $\Delta\omega$ by a factor of two), then t can be doubled, and the same error will be produced. This is in contrast with our chaotic example above where errors doubled on each iterate. In that case, to increase the time by a factor of two, $|\Delta(0)|$ had to be reduced by a factor of order 10^{14}.

The dynamics on an attractor is said to be chaotic if there is exponential sensitivity to initial conditions. We will say that an attractor is strange if it is fractal (this definition of strange is often used but is not universally accepted). Thus, chaos describes the dynamics on the attractor, while 'strange' refers to the geometry of the attractor. It is possible for chaotic attractors not to be strange (typically the case for one-dimensional maps (see the next chapter)), and it is also possible for attractors to be strange but not chaotic (Grebogi *et al.*, 1984; Romeiras and Ott, 1987). For most cases involving differential equations, strangeness and chaos commonly occur together.

1.6 Delay coordinates

In experiments one cannot always measure all the components of the vector $\mathbf{x}(t)$ giving the state of the system. Let us suppose that we can only measure one component, or, more generally, one scalar function of the state vector,

$$g(t) = G(\mathbf{x}(t)). \tag{1.15}$$

Given such a situation, can we obtain phase space information on the geometry of the attractor? For example, can we somehow make a surface of section revealing fractal structure as in Figures 1.12 and 1.13? The answer is yes. To see that this is so, define the so-called delay coordinate vector (Takens, 1980), $\mathbf{y} = (y^{(1)}, y^{(2)}, \ldots, y^{(M)})$, by

$$\left.\begin{aligned} y^{(1)}(t) &= g(t), \\ y^{(2)}(t) &= g(t-\tau), \\ y^{(3)}(t) &= g(t-2\tau), \\ &\vdots \\ y^{(M)}(t) &= g[t-(M-1)\tau], \end{aligned}\right\} \tag{1.16}$$

where τ is some fixed time interval, which should be chosen to be of the order of the characteristic time over which $g(t)$ varies. Given $\mathbf{x}$ at a specific time t_0, one could, in principle, obtain $\mathbf{x}(t_0 - m\tau)$ by integrating Eq. (1.3) backwards in time by an amount $m\tau$. Thus, $\mathbf{x}(t_0 - m\tau)$ is uniquely determined by $\mathbf{x}(t_0)$ and can hence be regarded as a function of $\mathbf{x}(t_0)$,

$$\mathbf{x}(t - m\tau) = \mathbf{L}_m(\mathbf{x}(t)).$$

Hence, $g(t - m\tau) = G(\mathbf{L}_m(\mathbf{x}(t)))$, and we may thus regard the vector $\mathbf{y}(t)$ as a function of $\mathbf{x}(t)$

$$\mathbf{y} = \mathbf{H}(\mathbf{x}).$$

We can now imagine making a surface of section in the $\mathbf{y}$-space. It can be shown (Section 3.8) that, if the number of delays M is sufficiently large, then we will typically see a qualitatively similar structure as would be seen had we made our surface of section in the original phase space $\mathbf{x}$.

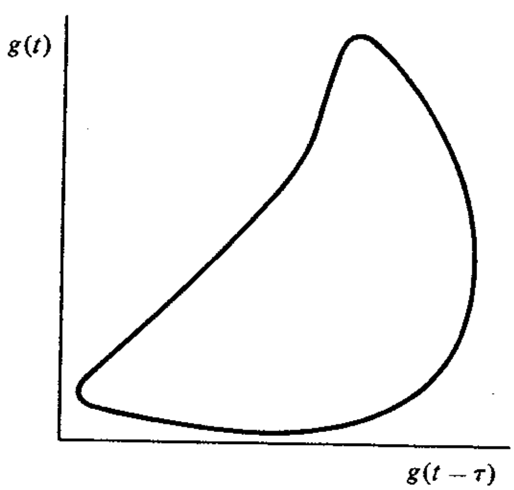


Figure 1.17 Experimental delay coordinate plot showing a closed curve corresponding to a limit cycle attractor.

Alternatively, we might simply examine the continuous time trajectory in y. For example, Figure 1.17 shows a result for an experiment involving chemical reactions (cf. Section 2.4.3). The vertical axis is the measured concentration $g(t)$ of one chemical constituant at time t and the horizontal axis is the same quantity evaluated at $t - (8.8 \text{ seconds})$. We see that the delay coordinates $\mathbf{y} = (g(t), g(t - 8.8))$ traces out a closed curve indicating a limit cycle.

Problems

1. Consider the following systems and specify (i) whether chaos can or cannot be ruled out for these systems, and (ii) whether the system is conservative or dissipative. Justify your answer

   (a) $\theta_{n+1} = [\theta_n + \Omega + 1.5 \sin \theta_n]$ modulo 2π,

   (b) $\theta_{n+1} = [\theta_n + \Omega + 0.5 \sin \theta_n]$ modulo 2π,

   (c) $x_{n+1} = [2x_n - x_{n-1} + k \sin x_n]$ modulo 2π,

   (d) $x_{n+1} = x_n + k(x_n - y_n)^2$, $y_{n+1} = y_n + k(x_n - y_n)^2$,

   (e) $\mathrm{d}x/\mathrm{d}t = v$, $\mathrm{d}v/\mathrm{d}t = -\alpha v + C \sin(\omega t - kx)$,

   (f) $\mathrm{d}x/\mathrm{d}t = B \cos y + C \sin z$

   $\mathrm{d}y/\mathrm{d}t = C \cos z + A \sin x$

   $\mathrm{d}z/\mathrm{d}t = A \cos x + B \sin y$.

2. Consider the one-dimensional motion of a free particle which bounces elastically between a stationary wall located at $x = 0$ and a wall whose position oscillates with time and is given by $x = L + \Delta \sin(\omega t)$. Derive a map relating the times T_n of the nth bounce off the oscillating wall and the particle speed v_n between the nth bounce and the $(n + 1)$th bounce off the oscillating wall to T_{n+1} and v_{n+1}. Assume that $L \gg \Delta$ so that $v_n(T_{n+1} - T_n) \approx 2L$. Is the map relating (T_n, v_n) to (T_{n+1}, v_{n+1}) conservative? Show that a new variable can be introduced in place of T_n, such that the new variable is bounded and results in a map which yields the same v_n as for the original map for all n.

3. Write a computer program to take iterates of the Hénon map. Considering the case $A = 1.4$, $B = 0.3$ and starting from an initial condition $(x_0, y_0) = (0, 0)$ iterate the map 20 times and then plot the next 1000 iterates to get a picture of the attractor.

4. Plot the first 25 iterates of the map given by Eq. (1.8) starting from $x_0 = 1/2$; (a) for $r = 3.8$ (chaotic attractor), (b) for $r = 2.5$ (period one attractor), and (c) for $r = 3.1$ (period two attractor).

5. For the map (1.8) with $r = 3.8$ plot the iterates of the two orbits orginating from the initial conditions $x_0 = 0.2$ and $x_0 = 0.2 + 10^{-5}$ versus iterate number. When does the separation between the two orbits first exceed 0.2?

Notes

1. Some review articles giving compact overview of chaotic dynamics are those of Helleman (1980), Ott (1981), Shaw (1981), and Grebogi *et al.* (1987d).
2. Some experiments on chaos in fluids are those of Libchaber and Maurer (1980), Gollub and Benson (1980), Bergé *et al.* (1980), Brandstater *et al.* (1983), and Sreenivasan (1986).
3. Applications of chaos to plasmas as well as many other topics are dealt with in the book by Sagdeev, Usikov and Zaslavsky (1990).
4. For example, Bryant and Jefferies (1984), Iansiti *et al.* (1985), Carroll *et al.* (1987), Roukes and Alerhand (1990), and Ditto *et al.* (1990b).
5. For example, Linsay (1981), Testa *et al.* (1982), and Rollins and Hunt (1984).
6. For example, Arecchi *et al.* (1982), Gioggia and Abraham (1984), and Mork *et al.* (1990).
7. The book by Moon (1987) on chaos contains outlines of results from a number of mechanical applications.
8. Chaotic phenomena and nonlinear dynamics in biology are dealt with in the book by Glass and Mackey (1988).
9. For example, Rössler (1976), Roux *et al.* (1980), Hudson and Mankin (1981), and Simoyi *et al.* (1982).
10. For example, Lauterborn (1981).
11. For example, Wisdom (1987) and Petit and Hénon (1986).
12. A review on the topic of routes to chaos is that of Eckmann (1981).
13. For example, according to the Poincaré–Bendixon theorem (e.g., see Hirsch and Smale (1974)), the only possible attracting solutions of (1.3) for $\mathbf{x}$ a two-dimensional vector in the plane are periodic solutions, steady states and solutions in which the orbit approaches a figure 8 or one of its lobes. In all these cases, the solution is not chaotic.
14. See May (1976) for an early discussion of the dynamics of this map.
15. Note that, if we take $|\Delta(0)|$ to be a small constant value (rather than examining $|\Delta(t)|/|\Delta(0)|$ in the limit that $|\Delta(0)| \to 0$), then the growth of $|\Delta(t)|$ cannot be exponential forever. In particular, $|\Delta(t)| < 2R$, and hence exponential growth must cease when $|\Delta(t)|$ becomes of the order of the attractor size. Thus later on (in Chapters 2 and 4) we shall be defining sensitive dependence on initial conditions in terms of the exponential growth of *differential* separations between orbits.
16. The definition of chaos given here is for chaotic *attractors*. When dealing with *nonattracting* chaotic sets (treated in Chapter 5) a more general definition of chaos is called for. Such a more general definition, which seems suitable very broadly, equates chaos with the condition of positive topological entropy. Topological entropy is defined in Chapter 4.