

Chapter 2

Logistic Map

The most important aspects of chaotic behavior should appear in systems of lowest dimension. Thus, we would like in a first step to reduce as much as possible the dimension of state space. However, this quickly conflicts with the requirement of invertibility. On the one hand, it can be shown that maps based on a one-dimensional homeomorphism can only display stationary or periodic regimes, and hence cannot be chaotic. On the other hand, if we sacrifice invertibility temporarily, thereby introducing *singularities*, one-dimensional chaotic systems can easily be found, as illustrated by the celebrated logistic map. Indeed, this simple system will be seen to display many of the essential features of deterministic chaos.

It is, in fact, no coincidence that chaotic behavior appears in its simplest form in a noninvertible system. As emphasized in this book, singularities and noninvertibility are intimately linked to the mixing processes (stretching and squeezing) associated with chaos.

Because of the latter, a dissipative invertible chaotic map becomes formally noninvertible when infinitely iterated (i.e., when phase space has been infinitely squeezed). Thus, the dynamics is, in fact, organized by an underlying singular map of lower dimension, as can be shown easily in model systems such as the horseshoe map. A classical example of this is the Hénon map, a diffeomorphism of the plane into itself that is known to have the logistic map as a backbone.

A noninvertible one-dimensional map has at least one point where its derivative vanishes. The simplest such maps are quadratic polynomials, which can always be brought to the form $f(x) = a - x^2$ under a suitable change of variables. The logistic map¹

$$x_{n+1} = a - x_n^2 \tag{2.1}$$

which depends on a single parameter a , is thus the simplest one-dimensional map displaying a singularity. As can be seen from its graph [Fig. 2.1(a)], the most important consequence of the singularity located at the critical point $x = 0$ is that each value in the range of the map f has exactly two preimages, which

¹A popular variant is $x_{n+1} = \lambda x_n(1 - x_n)$, with parameter λ .

will prove to be a key ingredient to generate chaos. Maps with a single critical point are called *unimodal*. It will be seen later that all unimodal maps display very similar dynamical behavior.

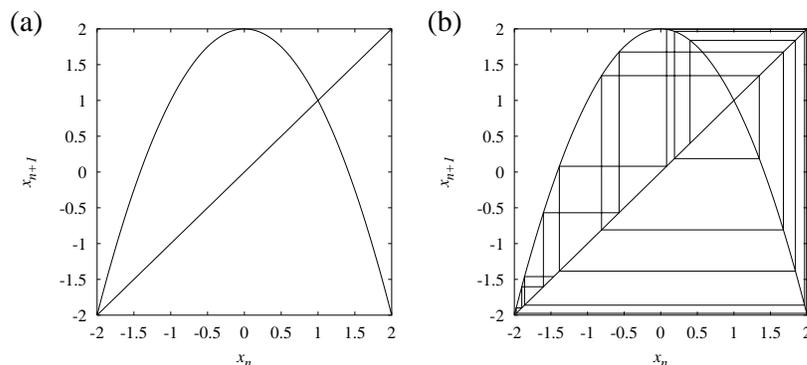


Figure 2.1: (a) Graph of the logistic map for $a = 2$. (b) Graphical representation of the iteration of (2.1).

As is often the case in dynamical systems theory, the action of the logistic map can not only be represented algebraically, as in Eq. 2.1, but also geometrically. Given a point x_n , the graph of the logistic map provides $y = f(x_n)$. To use y as the starting point of the next iteration, we must find the corresponding location in the x space, which is done simply by drawing the line from the point $[x_n, f(x_n)]$ to the diagonal $y = x$. This simple construction is then repeated ad libitum, as illustrated in Fig. 2.1(b).

The various behaviors displayed by the logistic map are easily explored, as this map depends on a single parameter a . As illustrated in Fig. 2.2, one finds quickly that two main types of dynamical regimes can be observed: stationary or periodic regimes on the one hand, and “chaotic” regimes on the other hand. In the first case, iterations eventually visit only a finite set of different values that are forever repeated in a fixed order. In the latter case, the state of the system never repeats itself exactly and seemingly evolves in a disordered way, as in Fig. 2.1(b). Both types of behaviors have been observed in the experiment discussed in Chapter 1.

What makes the study of the logistic map so important is not only that the organization in parameter space of these periodic and chaotic regimes can be completely understood with simple tools, but that despite of its simplicity it displays the most important features of low-dimensional chaotic behavior. By studying how periodic and chaotic behavior are interlaced, we will learn much about the mechanisms responsible for the appearance of chaotic behavior. Moreover, the logistic map is not only a paradigmatic system: One-dimensional maps will later prove also to be a fundamental tool for understanding the topological structure of flows.

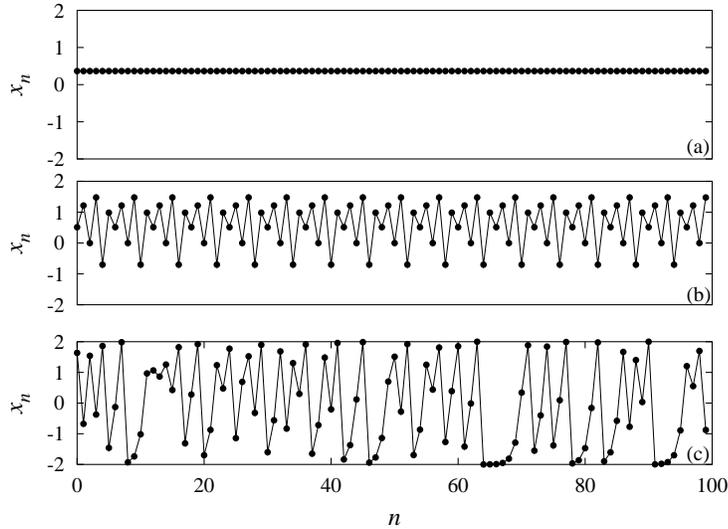


Figure 2.2: Different dynamical behaviors observed in the logistic map system are represented by plotting successive iterates: (a) stationary regime, $a = 0.5$; (b) periodic regime of period 5, $a = 1.476$; (c) chaotic regime, $a = 2.0$.

2.1 Bifurcation Diagrams

A first step in classifying the dynamical regimes of the logistic map is to obtain a global representation of the various regimes that are encountered as the control parameter a is varied. This can be done with the help of bifurcation diagrams, which are tools commonly used in nonlinear dynamics. Bifurcation diagrams display some characteristic property of the asymptotic solution of a dynamical system as a function of a control parameter, allowing one to see at a glance where qualitative changes in the asymptotic solution occur. Such changes are termed *bifurcations*.

In the case of the logistic map that has a single dynamical variable, the bifurcation diagram is readily obtained by plotting a sample set of values of the sequence (x_n) as a function of the parameter a , as shown in Fig. 2.3.

For $a < a_0 = -\frac{1}{4}$, iterations of the logistic map escape to infinity from all initial conditions. For $a > a_R = 2$ almost all initial conditions escape to infinity. The bifurcation diagram is thus limited to the range $a_0 < a < a_R$, where bounded solutions can be observed.

Between $a_0 = -\frac{1}{4}$ and $a_1 = \frac{3}{4}$, the limit set consists of a single value. This corresponds to a stationary regime, but one that should be considered in this context as a period-1 *periodic orbit*. At $a = a_1$, a bifurcation occurs, giving birth to a period-2 *periodic orbit*: Iterations oscillate between two values. As detailed in Section ??, this is an example of a *period-doubling bifurcation*. At $a = a_2 = \frac{5}{4}$, there is another period-doubling bifurcation where the period-2

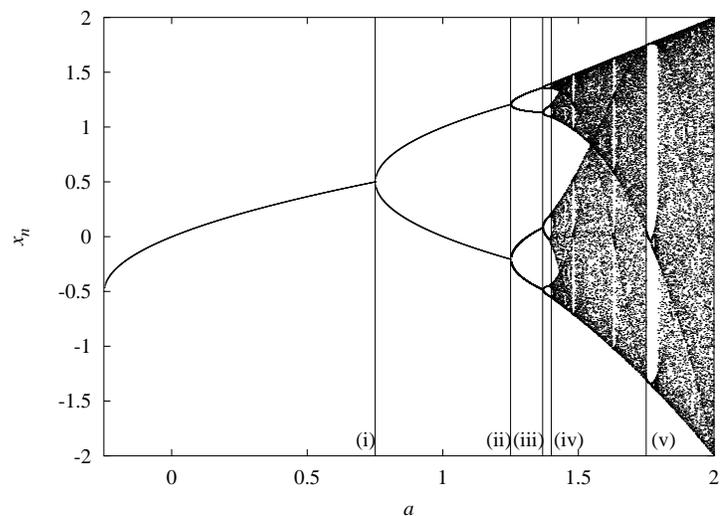


Figure 2.3: Bifurcation diagram of the logistic map. For a number of parameter values between $a = -0.25$ and $a = 2.0$, 50 successive iterates of the logistic map are plotted after transients have died out. From left to right, the vertical lines mark the creations of (i) a period-2 orbit; (ii) a period-4 orbit; (iii) a period-8 orbit, and (iv) the accumulation point of the period-doubling cascade; (v) the starting point of a period-3 window.

orbit gives place to a period-4 orbit.

The period-doubling bifurcations occurring at $a = a_1$ and $a = a_2$ are the first two members of an infinite series, known as the *period-doubling cascade*, in which an orbit of period 2^n is created for every integer n . The bifurcation at $a = a_3$ leading to a period-8 orbit is easily seen in the bifurcation diagram of Fig. 2.3, the one at $a = a_4$ is hardly visible, and the following ones are completely indiscernible to the naked eye.

This is because the parameter values a_n at which the period- 2^n orbit is created converge geometrically to the accumulation point $a_\infty = 1.401155189\dots$ with a convergence ratio substantially larger than 1:

$$\lim_{n \rightarrow \infty} \frac{a_n - a_{n-1}}{a_{n+1} - a_n} = \delta \sim 4.6692016091\dots \quad (2.2)$$

The constant δ appearing in 2.2 was discovered by Feigenbaum [?, ?] and is named after him. This distinction is justified by a remarkable property: Period-doubling cascades observed in an extremely large class of systems (experimental or theoretical, defined by maps or by differential equations...) have a convergence rate given by δ .

At the accumulation point a_∞ , the period of the solution has become infinite. Right of this point, the system can be found in chaotic regimes, as can be guessed from the abundance of dark regions in this part of the bifurcation diagram, which indicate that the system visits many different states. The period-doubling cascade is one of the best-known *routes to chaos and can be observed in many low-dimensional systems [?]*. *It has many universal properties that are in no way restricted to the case of the logistic map.*

However, the structure of the bifurcation diagram is more complex than a simple division between periodic and chaotic regions on both sides of the accumulation point of the period-doubling cascade. For example, a relatively large periodic window, which corresponds to the domain of stability of a period-3 orbit, is clearly seen to begin at $a = \frac{7}{4}$, well inside the chaotic zone. In fact, periodic windows and chaotic regions are arbitrarily finely interlaced as illustrated by Fig. 2.4. As will be shown later, there are infinitely many periodic windows between any two periodic windows. To interpret Fig. 2.4, it should be noted that periodic windows are visible to the naked eye only for very low periods. For higher periods, (1) the periodic window is too narrow compared to the scale of the plot, and (2) the number of samples is sufficiently large that the window cannot be distinguished from the chaotic regimes.

Ideally, we would like to determine for each periodic solution the range of parameter values over which it is stable. In Section 2.2 we will perform this analysis for the simple cases of the period-1 and period-2 orbits, so that we get a better understanding of the two types of bifurcation that are encountered in the logistic map. This is motivated by the fact that these are the two bifurcations that are generically observed in low-dimensional dynamical systems (omitting the Hopf bifurcation, which we discuss later).

However, we will not attempt to go much further in this direction. First, the complexity of Figs. 2.3 and 2.4 shows that this task is out of reach. Moreover,

which has two solutions:

$$x_-(a) = \frac{-1 - \sqrt{1 + 4a}}{2} \quad x_+(a) = \frac{-1 + \sqrt{1 + 4a}}{2} \quad (2.4)$$

The fixed points of a one-dimensional map can also be located geometrically, since they correspond to the intersections of its graph with the diagonal (Fig. 2.1).

Although a single period-1 regime is observed in the bifurcation diagram, there are actually two period-1 orbits. Later we will see why. Expressions 2.4 are real-valued only for $a > a_0 = -\frac{1}{4}$. Below this value, all orbits escape to infinity. Thus, the point at infinity, which we denote x_∞ in the following, can formally be considered as another fixed point of the system, albeit unphysical.

The important qualitative change that occurs at $a = a_0$ is our first example of a ubiquitous phenomenon of low-dimensional nonlinear dynamics, a *tangent*, or *saddle-node*, bifurcation: The two fixed points 2.4 become simultaneously real and are degenerate: $x_-(a_0) = x_+(a_0) = -\frac{1}{2}$. The two designations point to two different (but related) properties of this bifurcation.

The saddle-node qualifier is related to the fact that the two bifurcating fixed points have different stability properties. For a slightly above a_0 , it is found that orbits located near x_+ converge to it, whereas those starting in the neighborhood of x_- leave it to either converge to x_+ or escape to infinity, depending on whether they are located right or left of x_- . Thus, the fixed point x_+ (and obviously also x_∞) is said to be *stable* while x_- is *unstable*. They are called the *node* and the *saddle*, respectively.

Since trajectories in their respective neighborhoods converge to them, x_+ and x_∞ are *attracting sets*, or *attractors*. The sets of points whose orbits converge to an attractor of a system is called the *basin of attraction* of this point. From Fig. 2.5 we see that the unstable fixed point x_- is on the boundary between the *basins of attraction* of the two stable fixed points x_+ and x_∞ . The other boundary point is the preimage $f^{-1}(x_-)$ of x_- (Fig. 2.5).

It is easily seen that the stability of a fixed point depends on the derivative of the map at the fixed point. Indeed, if we perturb a fixed point $x_* = f(x_*)$ by a small quantity δx_n , the perturbation δx_{n+1} at the next iteration is given by

$$\delta x_{n+1} = f(x_* + \delta x_n) - x_* = \left. \frac{df(x)}{dx} \right|_{x_*} \delta x_n + O(\delta x_n^2) \quad (2.5)$$

If we start with an infinitesimally small δx_0 , the perturbation after n iterations is thus $\delta x_n \approx (\mu_*)^n \delta x_0$, where μ_* , the *multiplier* of the fixed point, is given by the map derivative at $x = x_*$.

A fixed point is thus stable (resp., unstable) when the absolute value of its multiplier is smaller (resp., greater) than unity. Here the multipliers μ_\pm of the two fixed points of the logistic map are given by

$$\mu_- = \left. \frac{df(x)}{dx} \right|_{x_-} = -2x_- = 1 + \sqrt{1 + 4a} \quad (2.6)$$

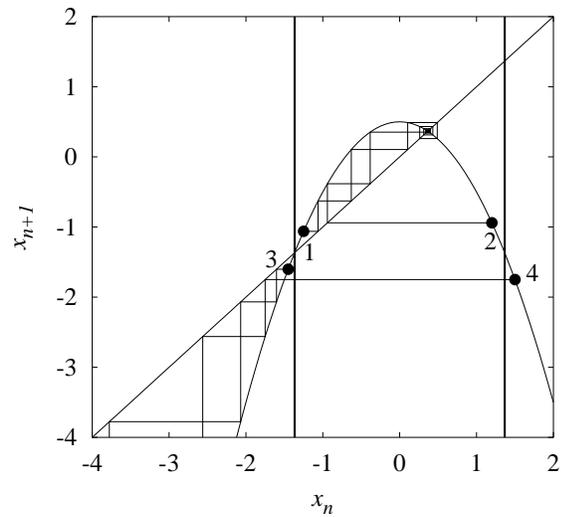


Figure 2.5: The basin of attraction of the x_+ fixed point is located between the left fixed point x_- and its preimage, indicated by two vertical lines. The orbits labeled 1 and 2 are inside the basin and converge towards x_+ . The orbits labeled 3 and 4 are outside the basin and escape to infinity (i.e., converge to the point at infinity x_∞ .)

$$\mu_+ = \left. \frac{df(x)}{dx} \right|_{x_+} = -2x_+ = 1 - \sqrt{1+4a} \quad (2.7)$$

Equation 2.6 shows that x_- is unconditionally unstable on its entire domain of existence, and hence is generically not observed as a stationary regime, whereas x_+ is stable for parameters a just above $a_0 = -\frac{1}{4}$, as mentioned above. This is why only x_+ can be observed on the bifurcation diagram shown in Fig. 2.3.

More precisely, x_+ is stable for $a \in [a_0, a_1]$, where $a_1 = \frac{3}{4}$ is such that $\mu_+ = -1$. This is consistent with the bifurcation diagram of Fig. 2.3. Note that at $a = 0 \in [a_0, a_1]$, the multiplier $\mu_+ = 0$ and thus perturbations are damped out faster than exponentially: The fixed point is then said to be *superstable*.

At the saddle-node bifurcation, both fixed points are degenerate and their multiplier is $+1$. This fundamental property is linked to the fact that at the bifurcation point, the graph of the logistic map is tangent to the diagonal (see Fig. 2.6), which is why this bifurcation is also known as the *tangent bifurcation*. Tangency of two smooth curves (here, the graph of f and the diagonal) is generic at a multiple intersection point. This is an example of a structurally unstable situation: An arbitrarily small perturbation of f leads to two distinct intersections or no intersection at all (alternatively, to two real or to two complex roots).

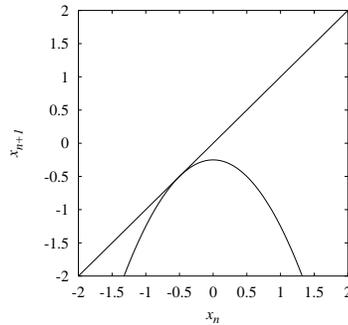


Figure 2.6: Graph of the logistic map at the initial saddle-node bifurcation.

It is instructive to formulate the intersection problem in algebraic terms. The fixed points of the logistic equations are zeros of the equation $G(x, a) = f(x; a) - x = 0$. This equation defines implicit functions $x_+(a)$ and $x_-(a)$ of the parameter a . In structurally stable situations, these functions can be extended to neighboring parameter values by use of the implicit function theorem.

Assume that $x_*(a)$ satisfies $G(x_*(a), a) = 0$ and that we shift the parameter a by an infinitesimal quantity δa . Provided that $\partial G(x_*(a), a)/\partial x \neq 0$, the corresponding variation δx_* in x_* is given by

$$G(x, a) = G(x_* + \delta x_*, a + \delta a) = G(x, a) + \frac{\partial G}{\partial x} \delta x_* + \frac{\partial G}{\partial a} \delta a = 0 \quad (2.8)$$

which yields:

$$\delta x_* = -\frac{\partial G/\partial a}{\partial G/\partial x} \delta a \quad (2.9)$$

showing that $x_*(a)$ is well defined on both sides of a if and only if $\partial G/\partial x \neq 0$. The condition

$$\frac{\partial G(x_*(a), a)}{\partial x} = 0 \quad (2.10)$$

is thus the signature of a bifurcation point. In this case, the Taylor series 2.8 has to be extended to higher orders of δx_* . If $\partial^2 G(x_*(a), a)/\partial x^2 \neq 0$, the variation δx_* in the neighborhood of the bifurcation is given by

$$(\delta x_*)^2 = -2 \frac{\partial G/\partial a}{\partial^2 G/\partial x^2} \delta a \quad (2.11)$$

From 2.11, we recover the fact that there is a twofold degeneracy at the bifurcation point, two solutions on one side of the bifurcation and none on the other side. The stability of the two bifurcating fixed points can also be analyzed: Since $G(x, a) = f(x; a) - x$, their multipliers are given by $\mu_* = 1 + \partial G(x_*, a)/\partial x$ and are thus equal to 1 at the bifurcation.

Just above the bifurcation point, it is easy to show that the multipliers of the two fixed points x_+ and x_- are given to leading order by $\mu_{\pm} = 1 \mp \alpha \sqrt{|\delta a|}$, where the factor α depends on the derivatives of G at the bifurcation point. It is thus generic that one bifurcating fixed point is stable while the other one is unstable. In fact, this is a trivial consequence of the fact that the two nondegenerate zeros of $G(x, a)$ must have derivatives $\partial G/\partial x$ with opposite signs.

This is linked to a fundamental theorem, which we state below in the one-dimensional case but which can be generalized to arbitrary dimensions by replacing derivatives with Jacobian determinants. Define the degree of a map f as

$$\deg f = \sum_{f(x_i)=y} \text{sign} \frac{df}{dx}(x_i) \quad (2.12)$$

where the sum extends over all the preimages of the arbitrary point y , and $\text{sign } z = +1$ (resp., -1) if $z > 0$ (resp., $z < 0$). It can be shown that $\deg f$ does not depend on the choice of y provided that it is a regular value (the derivatives at its preimages x_i are not zero) and that it is invariant by homotopy. Let us apply this to $G(x, a)$ for $y = 0$. Obviously, $\deg G = 0$ when there are no fixed points, but also for any a since the effect of varying a is a homotopy. We thus see that fixed points must appear in pairs having opposite contributions to $\deg G$. As discussed above, these opposite contributions correspond to different stability properties at the bifurcation.

The discussion above shows that although we have introduced the tangent bifurcation in the context of the logistic map, much of the analysis can be carried to higher dimensions. In an n -dimensional state space, the fixed points are determined by an n -dimensional vector function \mathbf{G} . In a structurally stable situation,

the Jacobian $\partial\mathbf{G}/\partial\mathbf{X}$ has rank n . As one control parameter is varied, bifurcations will be encountered at parameter values where $\partial\mathbf{G}/\partial\mathbf{X}$ is of lower rank. If the Jacobian has rank $n - 1$, it has a single null eigenvector, which defines the direction along which the bifurcation takes place. This explains why the essential features of tangent bifurcations can be understood from a one-dimensional analysis.

The theory of bifurcations is in fact a subset of a larger field of mathematics, the theory of singularities [?, ?], which includes catastrophe theory [?, ?] as a special important case. The tangent bifurcation is an example of the simplest type of singularity: The fold singularity, which typically corresponds to twofold degeneracies.

In the next section we see an example of a threefold degeneracy, the cusp singularity, in the form of the period-doubling bifurcation.

As shown in Section 2.2.1, the fixed point x_+ is stable only for $a \in [a_0, a_1]$, with $\mu_+ = 1$ at $a = a_1 = -\frac{1}{4}$ and $\mu_+ = -1$ at $a = a_1 = \frac{3}{4}$. For $a > a_1$, both fixed points 2.4 are unstable, which precludes a period-1 regime. Just above the bifurcation, what is observed instead is that successive iterates oscillate between two distinct values (see Fig. 2.3), which comprise a period-2 orbit. This could have been expected from the fact that at $a = a_1$, $\mu_+ = -1$ indicates that perturbations are reproduced every other period. The qualitative change that occurs at $a = a_1$ (a fixed point becomes unstable and gives birth to an orbit of twice the period) is another important example of bifurcation: The period-doubling bifurcation, which is represented schematically in Fig. 2.7. Saddle-node and period-doubling bifurcations are the only two types of local bifurcation that are observed for the logistic map. With the Hopf bifurcation, they are also the only bifurcations that occur generically in one-parameter paths in parameter space, and consequently, in low-dimensional systems.

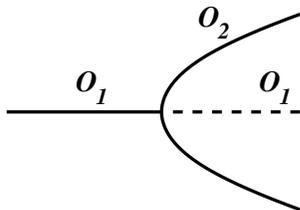


Figure 2.7: Period-doubling bifurcation. The orbit O_1 becomes unstable in giving birth to an orbit O_2 , whose period is twice that of O_1 .

Before we carry out the stability analysis for the period-2 orbit created at $a = a_1$, an important remark has to be made. Expression 2.4 shows that the period-1 orbit x_+ exists for every $a > a_0$: Hence it does not disappear at the period-doubling bifurcation but merely becomes unstable. It is thus present in all the dynamical regimes observed after its loss of stability, including in the chaotic regimes of the right part of the bifurcation diagram of Fig. 2.3. In fact,

this holds for all the periodic solutions of the logistic map. As an example, the logistic map at the transition to chaos ($a = a_\infty$) has an infinity of (unstable) periodic orbits of periods 2^n for any n , as Fig. 2.8 shows.

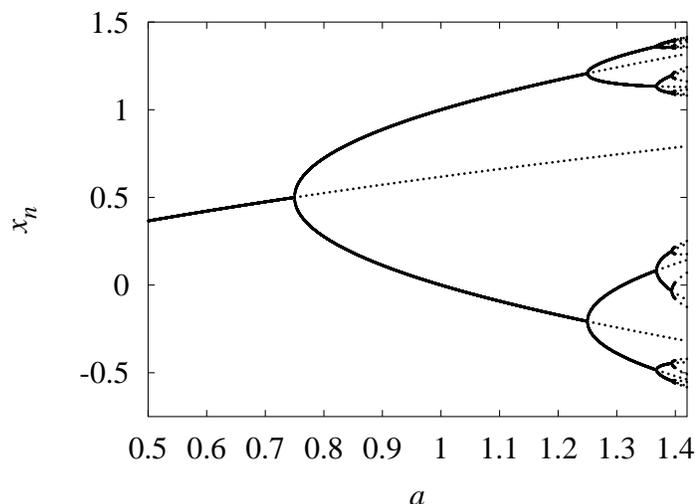


Figure 2.8: Orbits of period up to 16 of the period-doubling cascade. Stable (resp., unstable) periodic orbits are drawn with solid (resp., dashed) lines.

We thus expect periodic orbits to play an important role in the dynamics even after they have become unstable. We will see later that this is indeed the case and that much can be learned about a chaotic system from its set of periodic orbits, both stable and unstable.

Since the period-2 orbit can be viewed as a fixed point of the second iterate of the logistic map, we can proceed as above to determine its range of stability. The two periodic points $\{x_1, x_2\}$ are solutions of the quartic equation

$$x = f(f(x)) = a - (a - x^2)^2 \quad (2.13)$$

To solve for x_1 and x_2 , we take advantage of the fact that the fixed points x_+ and x_- are obviously solutions of Eq. (2.13). Hence, we just have to solve the quadratic equation

$$p(x) = \frac{f(f(x)) - x}{f(x) - x} = 1 - a - x + x^2 = 0 \quad (2.14)$$

whose solutions are

$$x_1 = \frac{1 - \sqrt{-3 + 4a}}{2} \quad x_2 = \frac{1 + \sqrt{-3 + 4a}}{2} \quad (2.15)$$

We recover the fact that the period-2 orbit (x_1, x_2) appears at $a = a_1 = \frac{3}{4}$, and exists for every $a > a_1$. By using the chain rule for derivatives, we obtain the multiplier of the fixed point x_1 of f^2 as

$$\mu_{1,2} = \left. \frac{df^2(x)}{dx} \right|_{x_1} = \left. \frac{df(x)}{dx} \right|_{x_2} \times \left. \frac{df(x)}{dx} \right|_{x_1} = 4x_1x_2 = 4(1-a) \quad (2.16)$$

Note that x_1 and x_2 viewed as fixed points of f^2 have the same multiplier, which is defined to be the multiplier of the orbit (x_1, x_2) . At the bifurcation point $a = a_1$, we have $\mu_{1,2} = 1$, a signature of the two periodic points x_1 and x_2 being degenerate at the period-doubling bifurcation.

However, the structure of the bifurcation is not completely similar to that of the tangent bifurcation discussed earlier. Indeed, the two periodic points x_1 and x_2 are also degenerate with the fixed point x_+ . The period-doubling bifurcation of the fixed point x_+ is thus a situation where the second iterate f^2 has three degenerate fixed points. If we define $G_2(x, a) = f^2(x; a) - x$, the signature of this threefold degeneracy is $G_2 = \partial G_2 / \partial x = \partial^2 G_2 / \partial x^2 = 0$, which corresponds to a higher-order singularity than the fold singularity encountered in our discussion of the tangent bifurcation. This is, in fact, our first example of the cusp singularity. Note that x_+ has a multiplier of -1 as a fixed point of f at the bifurcation and hence exists on both sides of the bifurcation: It merely becomes unstable at $a = a_1$. On the contrary, x_1 and x_2 have multiplier 1 for the lowest iterate of f of which they are fixed points, and thus exist only on one side of the bifurcation.

We also may want to verify that $\deg f^2 = 0$ on both sides of the bifurcation. Let us denote $d(x_*)$ as the contribution of the fixed point x_* to $\deg f^2$. We do not consider x_- , which is not involved in the bifurcation. Before the bifurcation, we have $d(x_+) = -1$ [$df^2/dx(x_+) < 1$]. After the bifurcation, $d(x_+) = 1$ but $d(x_1) = d(x_2) = -1$, so that the sum is conserved.

The period-2 orbit is stable only on a finite parameter range. The other end of the stability domain is at $a = a_2 = \frac{5}{4}$ where $\mu_{1,2} = -1$. At this parameter value, a new period-doubling bifurcation takes place, where the period-2 orbit loses its stability and gives birth to a period-4 orbit. As shown in Figs. 2.3 and 2.8, period-doubling occurs repeatedly until an orbit of infinite period is created.

Although one might in principle repeat the analysis above for the successive bifurcations of the period-doubling cascade, the algebra involved quickly becomes intractable. Anyhow, the sequence of parameters a_n at which a solution of period 2^n emerges converges so quickly to the accumulation point a_∞ that this would be of little use, except perhaps to determine the exact value of a_∞ , after which the first chaotic regimes are encountered.

A fascinating property of the period-doubling cascade is that we do not need to analyze directly the orbit of period 2^∞ to determine very accurately a_∞ . Indeed, it can be remarked that the orbit of period 2^∞ is formally its own period-doubled orbit. This indicates some kind of scale invariance. Accordingly, it was

recognized by Feigenbaum that the transition to chaos in the period-doubling cascade can be analyzed by means of renormalization group techniques [?, ?].

In this section we have analyzed how the periodic solutions of the logistic map are created. After discussing changes of coordinate systems in the next section, we shall take a closer look at the chaotic regimes appearing in the bifurcation diagram of Fig. 2.3. We will then be in position to introduce more sophisticated techniques to analyze the logistic map, namely *symbolic dynamics*, and to gain a complete understanding of the bifurcation diagram of a large class of maps of the interval.

2.3 Fully Developed Chaos in the Logistic Map

The first chaotic regime that we study in the logistic map is the one observed at the right end of the bifurcation diagram, namely at $a = 2$. At this point, the logistic map is surjective on the interval $I = [-2, 2]$: Every point $y \in I$ is the image of two different points, $x_1, x_2 \in I$. Moreover, I is then an invariant set since $f(I) = I$.

It turns out that the dynamical behavior of the surjective logistic map can be analyzed in a particularly simple way by using a suitable change of coordinates, namely $x = 2 \sin(\pi x')/4$. This is a one-to-one transformation between I and itself, which is a diffeomorphism everywhere except at the endpoints $x = \pm 2$, where the inverse function $x'(x)$ is not differentiable. With the help of a few trigonometric identities, the action of the logistic map in the x' space can be written as

$$x'_{n+1} = g(x'_n) = 2 - 2|x'_n| \quad (2.17)$$

a piecewise linear map known as the *tent map*.

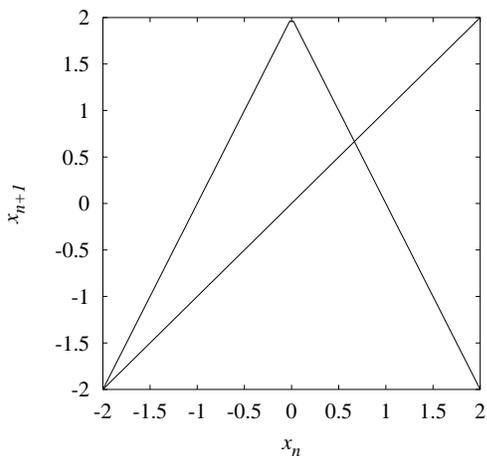


Figure 2.9: Graph of the tent map 2.17.

Figure 2.9 shows that the graph of the tent map is extremely similar to that of the logistic map (Fig. 2.1). In both cases, the interval I is decomposed into two subintervals: $I = I_0 \cup I_1$, such that each restriction $f_k : I_k \rightarrow f(I_k)$ of f is a homeomorphism, with $f(I_0) = f(I_1) = I$. Furthermore, f_0 (resp., f_1) is orientation-preserving (resp., orientation-reversing).

In fact, these topological properties suffice to determine the dynamics completely and are characteristic features of what is often called a *topological horseshoe*. In the remainder of Section 2.3, we review a few fundamental properties of chaotic behavior that can be shown to be direct consequences of these properties.

2.3.1 Iterates of the Tent Map

The advantage of the tent map over the logistic map is that calculations are simplified dramatically. In particular, higher-order iterates of the tent map, which are involved in the study of the asymptotic dynamics, are themselves piecewise-linear maps and are easy to compute. For illustration, the graphs of the second iterate g^2 and of the fourth iterate g^4 are shown in Fig. 2.10. Their structure is seen to be directly related to that of the tent map.

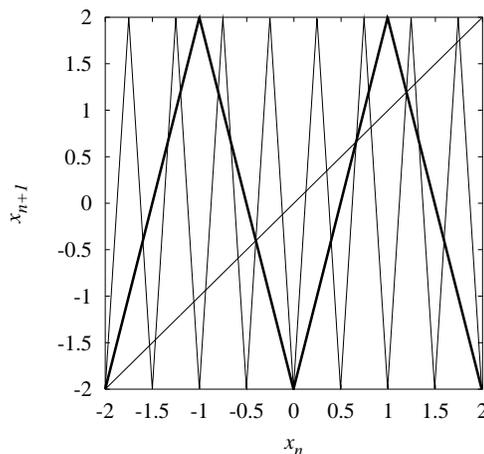


Figure 2.10: Graphs of the second (heavy line) and fourth (light line) iterates of the tent map 2.17.

Much of the structure of the iterates g^n can be understood from the fact that g maps linearly each of the two subintervals I_0 and I_1 to the whole interval I . Thus, the graph of the restriction of g^2 to each of the two components I_k reproduces the graph of g on I . This explains the two-“hump” structure of g^2 . Similarly, the trivial relation

$$\forall x \in I_k, \quad g^n(x) = g(g^{n-1}(x)) = g^{n-1}(g(x)) \quad (2.18)$$

shows that the graph of g^n consists of two copies of that of g^{n-1} . Indeed, Eq. 2.18 can be viewed as a semiconjugacy between g^n and g^{n-1} via the 2-to-1 transformation $x' = g(x)$.

By recursion, the graph of g^n shows 2^{n-1} scaled copies of the graph of g , each contained in a subinterval $I_k^n = [X_k - \epsilon_n, X_k + \epsilon_n]$ ($0 \leq k < 2^{n-1}$), where $\epsilon_n = 1/2^{n-2}$ and $X_k = -2 + (2k + 1)\epsilon_n$. The expression of g^n can thus be obtained from that of g by

$$\forall x \in I_k^n = [X_k - \epsilon_n, X_k + \epsilon_n] \quad g^n(x) = g(\epsilon_n(x - X_k)) \quad (2.19)$$

An important consequence of 2.19 is that each subinterval I_k^n is mapped to the whole interval I in no more than n iterations of g :

$$\forall k = 0 \dots 2^{n-1} \quad g^n(I_k^n) = g(I) = I \quad (2.20)$$

More precisely, one has $g(I_k^n) = I_{k'}^{n-1}$, where $k' = k$ (resp., $k' = 2^{n-1} - k$) if $k < 2^{n-2}$ (resp., $k \geq 2^{n-2}$). Note also that each I_k^n can itself be split into two intervals $I_{k,i}^n$ on which g^n is monotonic and such that $g^n(I_{k,i}^n) = I$.

Because the diameter of I_k^n is $|I_k^n| = 2^{3-n}$ and can be made arbitrarily small if n is chosen sufficiently large, this implies that an arbitrary subinterval $J \subset I$, however small, contains at least one interval I_k^n :

$$\forall J \subset I \quad \exists N_0 \quad n > N_0 \Rightarrow \exists k \quad I_k^n \subset J \quad (2.21)$$

Thus, how the iterates g^n act on the intervals I_k^n can help us to understand how they act on an arbitrary interval, as we will see later. In general, chaotic dynamics is better characterized by studying how sets of points are globally mapped rather than by focusing on individual orbits.

2.3.2 Lyapunov Exponents

An important feature of the tent map 2.17 is that the slope $|dg(x)/dx| = 2$ is constant on the whole interval $I = [-2, 2]$. This simplifies significantly the study of the stability of solutions of 2.17. From 2.5, an infinitesimal perturbation δx_0 from a reference state will grow after n iterations to $|\delta x_n| = 2^n |\delta x_0|$. Thus, any two distinct states, however close they may be, will eventually be separated by a macroscopic distance. This shows clearly that no periodic orbit can be stable (see Section 2.2.1) and thus that the asymptotic motion of 2.17 is aperiodic.

This exponential divergence of neighboring trajectories, or sensitivity to initial conditions, can be characterized quantitatively by Lyapunov exponents, which correspond to the average separation rate. For a one-dimensional map, there is only one Lyapunov exponent, defined by

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \log \frac{|\delta x_{n+1}|}{|\delta x_n|} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \log \left| \frac{df}{dx}(f^i(x_0)) \right| \quad (2.22)$$

which is a geometric average of the stretching rates experienced at each iteration. It can be shown that Lyapunov exponents are independent of the initial condition x_0 , except perhaps for a set of measure zero [?].

Since the distance between infinitesimally close states grows exponentially as $\delta x_n \sim e^{n\lambda} \delta x_0$, sensitivity to initial conditions is associated with a strictly positive Lyapunov exponent. It is easy to see that the Lyapunov exponent of surjective tent map 2.17 is $\lambda = \ln 2$.

2.3.3 Sensitivity to Initial Conditions and Mixing

Sensitivity to initial conditions can also be expressed in a way that is more topological, without using distances. The key property we use here is that any subinterval $J \subset I$ is eventually mapped to the whole I :

$$\forall J \subset I, \quad \exists N_0, \quad n > N_0 \Rightarrow g^n(J) = I \quad (2.23)$$

This follows directly from the fact that J contains one of the basis intervals I_k^n , and that these expand to I under the action of g [see 2.20 and 2.21].

We say that a map is expansive if it satisfies property 2.23. In plain words, the iterates of points in any subinterval can take every possible value in I after a sufficient number of iterations. Assume that J represents the uncertainty in the location of an initial condition x_0 : We merely know that $x_0 \in J$, but not its precise position. Then 2.23 shows that chaotic dynamics is, although deterministic, asymptotically unpredictable: After a certain amount of time, the system can be anywhere in the state space. Note that the time after which all the information about the initial condition has been lost depends only logarithmically on the diameter $|J|$ of J . Roughly, 2.20 indicates that $N_0 \simeq -\ln |J| / \ln 2 \sim -\ln |J| / \lambda$.

In the following, we use property 2.23 as a topological definition of chaos in one-dimensional noninvertible maps. To illustrate it, we recall the definitions of various properties that have been associated with chaotic behavior [?] and which can be shown to follow from 2.23.

A map $f : I \rightarrow I$:

- Has *sensitivity to initial conditions* if $\exists \delta > 0$ such that for all $x \in I$ and any interval $J \ni x$, there is a $y \in J$ and an $n > 0$ such that $|f^n(x) - f^n(y)| > \delta$.
- Is *topologically transitive* if for each pair of open sets $A, B \subset I$, there exists n such that $f^n(A) \cap B \neq \emptyset$.
- Is *mixing* if for each pair of open sets $A, B \subset I$, there exists $N_0 > 0$ such that $n > N_0 \Rightarrow f^n(A) \cap B \neq \emptyset$. A *mixing map* is obviously topologically transitive.

Sensitivity to initial conditions trivially follows from 2.23, since any neighborhood of $x \in I$ is eventually mapped to I . The mixing property, and hence transitivity, is also a consequence of expansiveness because the N_0 in the definition can be chosen so that $f^{N_0}(A) = I$ intersects any $B \subset I$. It can be shown that a topologically transitive map has at least a dense orbit (i.e., an orbit that passes arbitrarily close to any point of the invariant set).

Note that 2.23 precludes the existence of an invariant subinterval $J \subset I$ other than I itself: We would have simultaneously $f(J) = J$ and $f^{N_0}(J) = I$ for some N_0 . Thus, invariant sets contained in I necessarily consist of isolated points; these are the periodic orbits discussed in the next section.

2.3.4 Chaos and Density of (Unstable) Periodic Orbits

It has been proposed by Devaney [?] to say that a map f is chaotic if it:

- Displays sensitivity to initial conditions
- Is topologically transitive
- Has a set of periodic orbits that is dense in the invariant set

The first two properties were established in Section 2.3.3. It remains to be proved that 2.23 implies the third. When studying the bifurcation diagram of the logistic map (Section ??), we have noted that chaotic regimes contain many (unstable) periodic orbits. We are now in a position to make this observation more precise. We begin by showing that the tent map $x' = g(x)$ has infinitely many periodic orbits.

Number of Periodic Orbits of the Tent Map

A periodic orbit of g of period p is a fixed point of the p th iterate g^p . Thus, it satisfies $g^p(x) = x$ and is associated with an intersection of the graph of g^p with the diagonal. Since g itself has exactly two such intersections (corresponding to period-1 orbits), 2.19 shows that g^p has

$$N_f(p) = 2^p \tag{2.24}$$

fixed points (see Fig. 2.10 for an illustration).

Some of these intersections might actually be orbits of lower period: For example, the four fixed points of g^2 consist of two period-1 orbits and of two points constituting a period-2 orbit. As another example, note on Fig. 2.10 that fixed points of g^2 are also fixed points of g^4 . The number of periodic orbits of lowest period p is thus

$$N(p) = \frac{N_f(p) - \sum_q qN(q)}{p} \tag{2.25}$$

where the q are the divisors of p . Note that this is a recursive definition of $N(p)$. As an example, $N(6) = [N_f(6) - 3N(3) - 2N(2) - N(1)]/6 = (2^6 - 3 \times 2 - 2 \times 1 - 2)/6 = 9$ with the computation of $N(3)$, $N(2)$, and $N(1)$ being left to the reader. As detailed in Section 2.4.5, one of these nine orbits appears in a period doubling and the eight others are created by pairs in saddle-node bifurcations. Because $N_f(p)$ increases exponentially with p , $N(p)$ is well approximated for large p by $N(p) \simeq N_f(p)/p$.

We thus have the important property that there are an infinite number of periodic points and that the number $N(p)$ of periodic orbits of period p increases exponentially with the period. The corresponding growth rate,

$$h_P = \lim_{p \rightarrow \infty} \frac{1}{p} \ln N(p) = \lim_{p \rightarrow \infty} \frac{1}{p} \ln \frac{N_f(p)}{p} = \ln 2 \quad (2.26)$$

provides an accurate estimate of a central measure of chaos, the *topological entropy* h_T . In many cases it can be proven rigorously that $h_P = h_T$. Topological entropy itself can be defined in several different but equivalent ways.

Expansiveness Implies Infinitely Many Periodic Orbits

We now prove that if a continuous map $f : I \rightarrow I$ is expansive, its unstable periodic orbits are dense in I : Any point $x \in I$ has periodic points arbitrarily close to it. Equivalently, any subinterval $J \subset I$ contains periodic points.

We first note that if $J \subset f(J)$ (this is a particular case of a topological covering), then J contains a fixed point of f as a direct consequence of the intermediate value theorem.² Similarly, J contains at least one periodic point of period p if $J \subset f^p(J)$.

Now, if 2.23 is satisfied, every interval $J \subset I$ is eventually mapped to I : $f^n(J) = I$ [and thus $f^n(J) \subset J$], for $n > N_0(J)$. Using the remark above, we deduce that J contains periodic points of period p for any $p > N_0(J)$, but also possibly for smaller p . Therefore, *any interval contains an infinity of periodic points with arbitrarily high periods. A graphical illustration is provided by Fig. 2.10: Each intersection of a graph with the diagonal corresponds to a periodic point.*

Thus, the expansiveness property 2.23 implies that unstable periodic points are dense. We showed earlier that it also implies topological transitivity and sensitivity to initial conditions. Therefore, any map satisfying 2.23 is chaotic according to the definition given at the beginning of this section.

It is quite fascinating that sensitivity to initial conditions, which makes the dynamics unpredictable, and unstable periodic orbits, which correspond to perfectly ordered motion are so deeply linked: In a chaotic regime, order and disorder are intimately entangled.

Unstable periodic orbits will prove to be a powerful tool to analyze chaos. They form a skeleton around which the dynamics is organized. Although they can be characterized in a finite time, they provide invaluable information on the asymptotic dynamics because of the density property: The dynamics in the neighborhood of an unstable periodic orbit is governed largely by that orbit.

²Denote $x_a, x_b \in J$ the points such that $f(J) = [f(x_a), f(x_b)]$. If $J \subset f(J)$, one has $f(x_a) \leq x_a$ and $x_b \leq f(x_b)$. Thus, the function $F(x) = f(x) - x$ has opposite signs in x_a and x_b . If f is continuous, F must take all the values between $F(x_a)$ and $F(x_b)$. Thus, there exists $x_* \in J$ such that $F(x_*) = f(x_*) - x_* = 0$: x_* is a fixed point of f .

2.3.5 Symbolic Coding of Trajectories: First Approach

We showed above that because of sensitivity to initial conditions, the dynamics of the surjective tent map is asymptotically unpredictable (Section 2.3.3). However, we would like to have a better understanding of how irregular, or random, typical orbits can be. We also learned that there is a dense set of unstable periodic orbits embedded in the invariant set I , and that this set has a well-defined structure. What about the other orbits, which are aperiodic?

In this section we introduce a powerful approach to chaotic dynamics that answers these questions: symbolic dynamics. To do so as simply as possible, let us consider a dynamical system extremely similar to the surjective tent map, defined by the map

$$x_{n+1} = 2x_n \pmod{1} \quad (2.27)$$

It only differs from the tent map in that the two branches of its graph are both orientation-preserving (Fig. 2.11). As with the tent map, the interval $[0, 1]$ is decomposed in two subintervals I_k such that the restrictions $f_k : I_k \rightarrow f_k(I_k)$ are homeomorphisms.

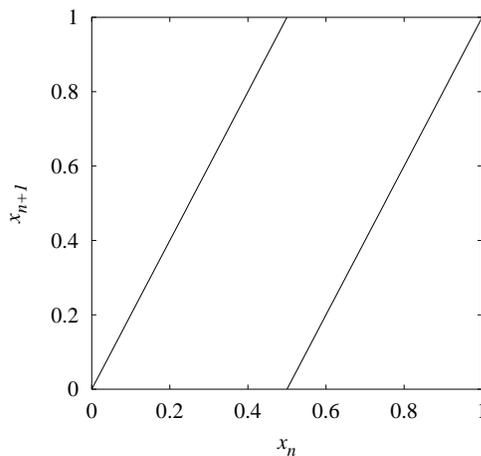


Figure 2.11: Graph of map 2.27.

The key step is to recognize that because the slope of the graph is 2 everywhere, the action of 2.27 is trivial if the coordinates $x \in [0, 1]$ are represented in base 2. Let x_n have the binary expansion $x_n = 0.d_0d_1 \dots d_k \dots$, with $d_k \in \{0, 1\}$. It is easy to see that the next iterate will be

$$x_{n+1} = (d_0.d_1d_2 \dots d_k \dots) \pmod{1} = 0.d_1d_2 \dots d_k \dots \quad (2.28)$$

Thus, the base-2 expansion of x_{n+1} is obtained by dropping the leading digit in the expansion of x_n . This leading digit indicates whether x is greater than

or equal to $\frac{1}{2} = 0.1\bar{0}$ (\bar{s} represents an infinite repetition of the string s), thus which interval $I_0 = [0, 0.5)$ or $I_1 = [0.5, 1]$ the point belongs to. Note that in the present case $0.1\bar{0}$ and $0.0\bar{1}$, which usually represent the same number $\frac{1}{2}$, correspond here to different trajectories because of the discontinuity. The former is located at $(\frac{1}{2})^+$ and remains on the fixed point $x = 1$ forever, while the latter is associated with $(\frac{1}{2})^-$ and converges to the fixed point $x = 0$.

Thus, there is a 1:1 correspondence between orbits of the dynamical system 2.27 (parameterized by their initial condition x) and infinite digit sequences $(d_k) \in \{0, 1\}^N$. Furthermore, the action of the map in the latter space has a particularly simple form. This correspondence allows one to establish extremely easily all the properties derived for the tent map in previous sections.

- **Sensitivity to initial conditions.** Whether the n th iterate of x falls in I_0 or I_1 is determined by the n th digit of the binary expansion of x . A small error in the initial condition (e.g., the n th digit is false) becomes macroscopic after a sufficient amount of time (i.e., after n iterations).
- **Existence of a dense orbit.** Construct an infinite binary sequence such that it contains all possible finite sequences. For example, concatenate all sequences of length $1, 2, \dots, n$ for arbitrarily large n . The iterates of the associated point $x = 0.0|1|00|01|10|11|000|001\dots$ will pass arbitrarily close to any point of the interval. The existence of a dense orbit implies topological transitivity.
- **Density of periodic orbits.** Each periodic point of 2.27 obviously corresponds to a periodic binary sequence. It is known that a periodic or eventually periodic digit expansion is a characteristic property of rational numbers. Since it is a classical result that rational numbers are dense in $[0, 1]$, we deduce immediately that periodic or eventually periodic points are dense in the interval $[0, 1]$. Alternatively, each point x can be approximated arbitrarily well by a sequence of periodic points $x_*(n)$ whose sequences consist of the infinitely repeated n first digits of x , with $n \rightarrow \infty$.

This analysis can easily be transposed to the case of the surjective tent map. Since its right branch is orientation-reversing, the action of this map on the binary expansion of a point x located in this branch differs slightly from that of (2.28). Assuming that the tent map is defined on $[0, 1]$, its expression at the right (resp., left) of the critical point is $x' = 2(1 - x)$ (resp., $x' = 2x$). Consequently, we have the additional rule that if the leading digit is $d_0 = 1$, all the digits d_i , $i \in N$, should be replaced by $\tilde{d}_i = 1 - d_i$ before dropping the leading digit d_0 as with the left branch (in fact, the two operations can be carried out in any order). The operation $d_i \rightarrow \tilde{d}_i$ is known as *complement to one*.

Example: Under the tent map, $0.01001011 \rightarrow 0.1001011 \rightarrow 0.110100$. For the first transition, since $0.0100100 < \frac{1}{2}$, we simply remove the decimal one digit to the right. In the second transition, since $x = 0.1001011 > \frac{1}{2}$, we first complement x and obtain $x' = (1 - x) = 0.0110100$, then multiply by 2: $2x' = 0.1101100$.

Except for this minor difference in the coding of trajectories, the arguments used above to show the existence of chaos in the map $x' = 2x \pmod{1}$ can be followed without modification. The binary coding we have used is thus a powerful method to prove that the tent map displays chaotic behavior.

The results of this section naturally highlight two important properties of chaotic dynamics:

- A series of coarse-grained measurements of the state of a system can suffice to estimate it with arbitrary accuracy if carried out over a sufficiently long time. By merely noting which branch is visited (one-bit digitizer) at each iteration of the map (2.27), all the digits of an initial condition can be extracted.
- Although a system such as (2.27) is perfectly deterministic, its asymptotic dynamics is as random as coin flipping (all sequences of 0 and 1 can be observed).

However, the coding used in these two examples (n -ary expansion) is too naive to be extended to maps that do not have a constant slope equal to an integer. In the next section we discuss the general theory of symbolic dynamics for one-dimensional maps. This topological approach will prove to be an extremely powerful tool to characterize the dynamics of the logistic map, not only in the surjective case but for any value of the parameter a .

2.4 One-Dimensional Symbolic Dynamics

2.4.1 Partitions

Consider a continuous map $f : I \rightarrow I$ that is singular. We would like to extend the symbolic dynamical approach introduced in Section 2.3.5 in order to analyze its dynamics. To this end, we have to construct a coding associating each orbit of the map with a symbol sequence.

We note that in the previous examples, each digit of the binary expansion of a point x indicates whether x belongs to the left or right branch of the map. Accordingly, we decompose the interval I in N disjoint intervals I_α , $\alpha = 0 \dots N-1$ (numbered from left to right), such that

- $I = I_0 \cup I_1 \cup \dots \cup I_{N-1}$
- In each interval I_α , the restriction $f|_{I_\alpha} : I_\alpha \rightarrow f(I_\alpha)$, which we denote f_α , is a homeomorphism.

For one-dimensional maps, such a partition can easily be constructed by choosing the critical points of the map as endpoints of the intervals I_α , as Fig. 2.12 illustrates. At each iteration, we record the symbol $\alpha \in \mathcal{A} = \{0, \dots, N-1\}$ that identifies the interval to which the current point belongs. The alphabet \mathcal{A} consists of the N values that the symbol can assume.

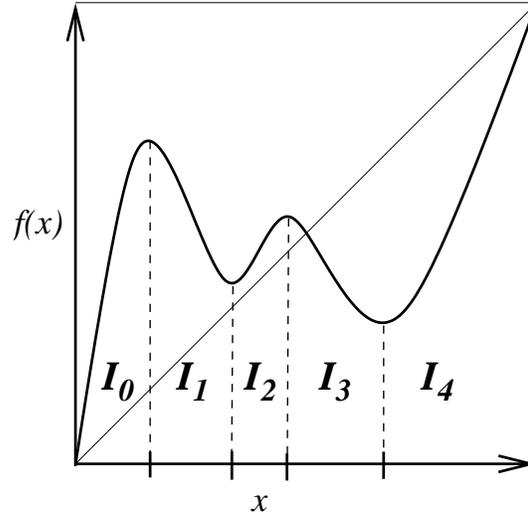


Figure 2.12: Decomposition of the domain of a map f into intervals I_α such that the restrictions $f : I_\alpha \rightarrow f(I_\alpha)$ to the intervals I_α are homeomorphisms.

We denote by $s(x)$ the corresponding coding function:

$$s(x) = \alpha \iff x \in I_\alpha \quad (2.29)$$

Any orbit $\{x, f(x), f^2(x), \dots, f^i(x), \dots\}$ of initial condition x can then be associated with the infinite sequence of symbols indicating the intervals visited successively by the orbit:

$$\Sigma(x) = \{s(x), s(f(x)), s(f^2(x)), \dots, s(f^i(x)), \dots\} \quad (2.30)$$

The sequence $\Sigma(x)$ is called the *itinerary* of x . We will also use the compact notation $\Sigma = s_0 s_1 s_2 \dots s_i \dots$ with the s_i being the successive symbols of the sequence (e.g., $\Sigma = 01101001\dots$). The set of all possible sequences in the alphabet \mathcal{A} is denoted $\mathcal{A}^{\mathbb{N}}$ and $\Sigma(I) \subset \mathcal{A}^{\mathbb{N}}$ represents the set of sequences actually associated with a point of I :

$$\Sigma(I) = \{\Sigma(x); x \in I\} \quad (2.31)$$

The finite sequence made of the n leading digits of $\Sigma(x)$ will later be useful. We denote it $\Sigma_n(x)$. For example, if $\Sigma(x) = 10110\dots$, then $\Sigma_3(x) = 101$. Accordingly, the set of finite sequences of length n involved in the dynamics is $\Sigma_n(I)$.

An important property of the symbolic representation 2.30 is that the expression of the time-one map becomes particularly simple. Indeed, if we compare the symbolic sequence of $f(x)$:

$$\Sigma(f(x)) = \{s(f(x)), s(f^2(x)), s(f^3(x)), \dots, s(f^{i+1}(x)), \dots\} \quad (2.32)$$

with that of x given in 2.30, we observe that the former can be obtained from the latter by dropping the leading symbol and shifting the remaining symbols to the left. Accordingly, we define the *shift operator* σ by

$$\Sigma = \{s_0, s_1, s_2, \dots, s_i, \dots\} \xrightarrow{\sigma} \{s_1, s_2, \dots, s_i, \dots\} = \sigma\Sigma \quad (2.33)$$

Applying f to a point $x \in I$ is equivalent to applying the shift operator σ on its symbolic sequence $\Sigma(x) \in \Sigma(I)$:

$$\Sigma(f(x)) = \sigma\Sigma(x) \quad (2.34)$$

which corresponds to the commutative diagram

$$x @> f >> f(x) @VV \Sigma V @VV \Sigma V \{s_i\}_{i \in N} @> \sigma >> \{s_{i+1}\}_{i \in N} \quad (2.35)$$

Note that because only forward orbits $\{f^n(x)\}_{n \geq 0}$ can be computed with a noninvertible map, the associated symbolic sequences are one-sided and extend to infinity only in the direction of forward time. This makes the operator σ noninvertible, as f is itself. Formally, we can define several “inverse” operators σ_α^{-1} acting on a sequence by inserting the symbol α at its head:

$$\Sigma = \{s_0, s_1, s_2, \dots, s_i, \dots\} \xrightarrow{\sigma_\alpha^{-1}} \{\alpha, s_0, s_1, \dots, s_{i-1}, \dots\} = \sigma_\alpha^{-1}\Sigma \quad (2.36)$$

However, note that $\sigma \circ \sigma_\alpha^{-1} = \text{Id} \neq \sigma_\alpha^{-1} \circ \sigma$.

Periodic sequences $\Sigma = \{s_i\}$ with $s_i = s_{i+p}$ for all $i \in N$ will be of particular importance in the following. Indeed, they satisfy $\sigma^p \Sigma = \Sigma$, which translates into $f^p(x) = x$ for the associated point that is thus periodic. Infinite periodic sequences will be represented by overlining the base pattern (e.g., $\overline{01011} = 010110101101011\dots$). When there is no ambiguity, the base pattern will be used as the name of the corresponding periodic orbit (e.g., the orbit $\overline{01011}$ has sequence $\overline{01011}$).

2.4.2 Symbolic Dynamics of Expansive Maps

To justify the relevance of symbolic coding, we now show that it is a faithful representation. Namely, the correspondence $x \in I \leftrightarrow \Sigma(x) \in \Sigma(I)$ defined by (2.29) and 2.30 can under appropriate conditions be made a bijection, that is,

$$x_1 \neq x_2 \iff \Sigma(x_1) \neq \Sigma(x_2) \quad (2.37)$$

We might additionally require some form of continuity so that sequences that are close according to some metric are associated with points that are close in space.

In plain words, the symbolic sequence associated to a given point is sufficient to distinguish it from any other point in the interval I . The two dynamical systems (I, f) and $(\Sigma(I), \sigma)$ can then be considered as equivalent, with $\Sigma(x)$

playing the role of a change of coordinate. Partitions of state space that satisfy 2.37 are said to be *generating*.

In Section 2.3.5, we have seen two particular examples of one-to-one correspondence between orbits and symbolic sequences. Here we show that such a bijection holds if the following two conditions are true: (1) the restriction of the map to each member of the partition is a homeomorphism (Section 2.4.1); (2) the map satisfies the expansiveness property 2.23. This will illustrate the intimate connection between symbolic dynamics and chaotic behavior.

In the tent map example, it is obvious how the successive digits of the binary expansion of a point x specify the position of x with increasing accuracy. As we show below, this is also true for general symbolic sequences under appropriate conditions.

As a simple example, assume that a point x has a symbol sequence $\Sigma(x) = 101\dots$. From the leading symbol we extract the top-level information about the position of x , namely that $x \in I_1$. Since the second symbol is 0, we deduce that $f(x) \in I_0$ [i.e., $x \in f^{-1}(I_0)$]. This second-level information combined with the first-level information indicates that $x \in I_1 \cap f^{-1}(I_0) \equiv I_{10}$. Using the first three symbols, we obtain $x \in I_{101} = I_1 \cap f^{-1}(I_0) \cap f^{-2}(I_1) = I_1 \cap f^{-1}(I_0 \cap f^{-1}(I_1))$. We note that longer symbol sequences localize the point with higher accuracy: $I_1 \supset I_{10} \supset I_{101}$.

More generally, define the interval $I_\Lambda = I_{s_0 s_1 \dots s_{n-1}}$ as the set of points whose symbolic sequence begins by the finite sequence $\Lambda = s_0 s_1 \dots s_{n-1}$, the remaining part of the sequence being arbitrary:

$$\begin{aligned} I_\Lambda = I_{s_0 s_1 \dots s_{n-1}} &= \{x; \Sigma_n(x) = s_0 s_1 \dots s_{n-1}\} \\ &= \{x; s(f^i(x)) = s_i, i < n\} \\ &= \{x; f^i(x) \in I_{s_i}, i < n\} \end{aligned} \quad (2.38)$$

Such sets are usually termed *cylinders*, with an n -cylinder being defined by a sequence of length n . We now show that cylinders can be expressed simply using inverse branches of the function f . We first define

$$\forall J \subset I \quad f_\alpha^{-1}(J) = I_\alpha \cap f^{-1}(J) \quad (2.39)$$

This is a slight abuse of notation, since we have only that $f_\alpha(f_\alpha^{-1}(J)) \subset J$ without the equality being always satisfied, but it makes the notation more compact. With this convention, the base intervals I_α can be written as

$$I_\alpha = \{x; s(x) = \alpha\} = f_\alpha^{-1}(I) \quad \forall \alpha \in \mathcal{A} \quad (2.40)$$

To generate the whole set of cylinders, this expression can be generalized to longer sequences by noting that

$$I_{\alpha\Lambda} = f_\alpha^{-1}(I_\Lambda) \quad (2.41)$$

which follows directly from definitions (2.38) and (2.39). Alternatively, (2.41) can be seen merely to express that $\alpha\Lambda = \sigma_\alpha^{-1}\Lambda$. By applying (2.41) recursively,

one obtains

$$I_\Lambda = f_\Lambda^{-n}(I) \quad (2.42)$$

where f_Λ^{-n} is defined by

$$f_\Lambda^{-n} = f_{s_0 s_1 \dots s_{n-1}}^{-n} = f_{s_0}^{-1} \circ f_{s_1}^{-1} \circ \dots \circ f_{s_{n-1}}^{-1} \quad (2.43)$$

Just as the restriction of f to any interval I_α is a homeomorphism $f_\alpha : I_\alpha \rightarrow f(I_\alpha)$, the restriction of f^n to any set I_Λ with Λ of length n is a homeomorphism³ $f_\Lambda^n : I_\Lambda \rightarrow f^n(I_\Lambda)$. The function f_Λ^{-n} defined by 2.43 is the inverse of this homeomorphism, which explains the notation. For a graphical illustration, see Fig. 2.10: Each interval of monotonic behavior of the graph of g^2 (resp., g^4) corresponds to a different interval I_Λ , with Λ of length 2 (resp., 4).

Note that because I is connected and the f_α^{-1} are homeomorphisms, all the I_Λ are connected sets, hence are intervals in the one-dimensional context. This follows directly from 2.41 and the fact that the image of a connected set by a homeomorphism is a connected set. This property will be important in the following.

To illustrate relation 2.42, we apply it to the case $\Lambda = 101$ considered in the example above:

$$\begin{aligned} I_{101} = (f_1^{-1} \circ f_0^{-1} \circ f_1^{-1})(I) &= (f_1^{-1} \circ f_0^{-1})(I_1) \\ &= f_1^{-1}(I_0 \cap f^{-1}(I_1)) \\ &= I_1 \cap f^{-1}(I_0 \cap f^{-1}(I_1)) \end{aligned}$$

and verify that it reproduces the expression obtained previously.

The discussion above shows that the set of n -cylinders $\mathcal{C}_n = \{I_\Lambda; \Lambda \in \Sigma_n(I)\}$ is a partition of I :

$$I = \bigcup_{\Lambda \in \Sigma_n(I)} I_\Lambda \quad I_\Lambda \cap I_{\Lambda'} = \emptyset \quad (2.44)$$

with \mathcal{C}_n being a refinement of \mathcal{C}_{n-1} (i.e., each member of \mathcal{C}_n is a subset of a member of \mathcal{C}_{n-1}). As $n \rightarrow \infty$, the partition \mathcal{C}_n becomes finer and finer (again, see Fig. 2.10). What we want to show is that the partition is arbitrarily fine in this limit, with the size of *each* interval of the partition converging to zero.

Consider an arbitrary symbolic sequence $\Sigma(x)$, with $\Sigma_n(x)$ listing its n leading symbols. Since by definition $I_{\Sigma_{n+1}(x)} \subset I_{\Sigma_n(x)}$, the sequence $(I_{\Sigma_n(x)})_{n \in \mathbb{N}}$ is decreasing, hence it converges to a limit $I_{\Sigma(x)}$. All the points in $I_{\Sigma(x)}$ share the same infinite symbolic sequence.

As the limit of a sequence of connected sets, $I_{\Sigma(x)}$ is itself a connected set, hence is an interval or an isolated point. Assume that $I_{\Sigma(x)}$ is an interval. Then, because of the expansiveness property 2.23, there is N_0 such that $f^{N_0}(I_{\Sigma(x)}) = I$. This implies that for points $x \in I_{\Sigma(x)}$, the symbol $s_{N_0} = s(f^{N_0})(x)$ can take any value $\alpha \in \mathcal{A}$, in direct contradiction with $I_{\Sigma(x)}$ corresponding to a unique sequence. Thus, the only possible solution is that the limit $I_{\Sigma(x)}$ is an isolated

³Note that f^n is a homeomorphism only on set I_Λ defined by symbolic strings Λ of length $p \geq n$ (e.g., f^2 has a singularity in the middle of I_0).

point, showing that the correspondence between points and symbolic sequences is one-to-one. Thus, the symbolic dynamical representation of the dynamics is faithful.

This demonstration assumes a partition of I constructed so that each interval of monotonicity corresponds to a different symbol (see Fig. 2.12). This guarantees that all the preimages of a given point are associated with different symbols since they belong to different intervals.

It is easy to see that partitions not respecting this rule cannot be generating. Assume that two points x_0 and x_1 have the same image $f(x_0) = f(x_1) = y$ and that they are coded with the same symbol $s(x_0) = s(x_1) = \alpha_k$ (Fig. 2.13). They are then necessarily associated with the same symbolic sequence, consisting of the common symbol α_k concatenated with the symbolic sequence of their common image: $\Sigma(x_0) = \Sigma(x_1) = \alpha_k \Sigma(y)$. In other words, associating x_0 and x_1 with different symbols is the only chance to distinguish them because they have exactly the same future.

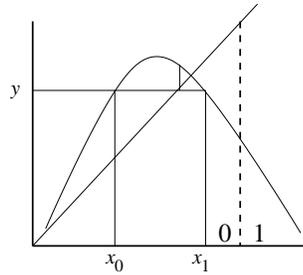


Figure 2.13: The dashed line indicates the border of a partition such that the preimages x_0 and x_1 of the same point y are coded with the same symbol (0). As a consequence, the symbolic sequences associated to x_0 and x_1 are identical.

In one-dimensional maps, two preimages of a given point are always separated by a critical point. Hence, the simplest generating partition is obtained by merely dividing the base interval I into intervals connecting two adjacent critical points, and associating each with a different symbol.

Remark 1: This is no longer true for higher-dimensional noninvertible maps, which introduces some ambiguity in the symbolic coding of trajectories.

Remark 2: Invertible chaotic maps do not have singularities, hence the construction of generating partitions is more involved. The forthcoming examples of the horseshoe and of the Hénon map will help us to understand how the rules established in the present section can be generalized.

2.4.3 Grammar of Chaos: First Approach

Symbolic dynamics provides a simple but faithful representation of a chaotic dynamical system (Section 2.4.2). It has allowed us to understand the structure

of the chaotic and periodic orbits of the surjective tent map, and hence of the surjective logistic map (Section 2.3.5). But there is more.

As a control parameter of a one-dimensional map varies, the structure of its invariant set and of its orbits changes (perestroika). Symbolic dynamics is a powerful tool to analyze these modifications: As orbits are created or destroyed, symbolic sequences appear or disappear from the associated symbolic dynamics. Thus, a regime can be characterized by a description of its set of forbidden sequences. We refer to such a description as the *grammar of chaos*. Changes in the structure of a map are characterized by changes in this grammar.

As we illustrate below with simple examples, which sequences are allowed and which are not can be determined entirely geometrically. In particular, the orbit of the critical point plays a crucial role. The complete theory, namely *kneading theory*, is detailed in Section 2.4.4.

Interval Arithmetics and Invariant Interval

We begin by determining the smallest invariant interval I [i.e., such that $f(I) = I$]. This is where the asymptotic dynamics will take place. Let us first show how to compute the image of an arbitrary interval $J = [x_l, x_h]$. If J is located entirely to the left or right of the critical point x_c , one merely needs to take into account that the logistic map is orientation-preserving (resp., orientation-reversing) at the left (resp., right) of x_c . Conversely, if $x_c \in J$, J can be decomposed as $[x_l, x_h] = [x_l, x_c] \cup [x_c, x_h]$. This gives

$$f([x_l, x_h]) = \begin{cases} [f(x_l), f(x_h)] & \text{if } x_l, x_h \leq x_c \\ [f(x_h), f(x_l)] & \text{if } x_c \leq x_l, x_h \\ [\min\{f(x_l), f(x_h)\}, f(x_c)] & \text{if } x_l \leq x_c \leq x_h \end{cases} \quad (2.45)$$

As was noted by Poincaré, the apparent complexity of chaotic dynamics is such that it makes little sense to follow individual orbits: What is relevant is how regions of the state space are mapped between each other. One-dimensional maps are no exception, and in fact many properties of the logistic map can be extracted from the interval arithmetics defined by 2.45. Here, we use them to show in a simple way that some symbolic sequences are forbidden.

Let us now determine $I = [x_{\min}, x_{\max}]$ such that $f(I) = I$. We are interested only in situations where this interval contains the critical point, so that the dynamics is nontrivial. Note that this implies that $x_c \leq f(x_c)$ because we must have $x_c = f(y) \leq f(x_c)$ (hence the top of the parabola must be above the diagonal). We use the third case of Eq. 2.45 to obtain the equation

$$[\min\{f(x_{\min}), f(x_{\max})\}, f(x_c)] = [x_{\min}, x_{\max}] \quad (2.46)$$

The upper bound is thus the image of the critical point: $x_{\max} = f(x_c)$. The lower bound x_{\min} satisfies the equation

$$x_{\min} = \min\{f(x_{\min}), f(x_{\max})\} = \min\{f(x_{\min}), f^2(x_c)\} \quad (2.47)$$

An obvious solution is $x_{\min} = f^2(x_c)$, which is valid provided that $f(f^2(x_c)) > f^2(x_c)$. This is always the case between the parameter value where the period-1 orbit is superstable and the one where bounded solutions cease to exist. The other possible solution is the fixed point $x_- = f(x_-)$. In the parameter region of interest, however, one has $x_- < f^2(x_c)$, and thus the smallest invariant interval is given by

$$I = [f^2(x_c), f(x_c)] \quad (2.48)$$

That it depends only on the orbit of the critical point x_c is remarkable. However, this merely prefigures Section 2.4.4 where we shall see that this orbit determines the dynamics completely. Note that $I \neq \emptyset$ as soon as $f(x_c) > x_c$, which is the only interesting parameter region from a dynamical point of view.

Existence of Forbidden Sequences

As shown previously, the set I_Λ of points whose symbolic sequence begins by the finite string Λ is given by $I_\Lambda = f_\Lambda^{-n}(I)$, where f_Λ^{-n} is defined by (2.43) and 2.39. It is easy to see that if $I_\Lambda = \emptyset$, the finite symbol sequence Λ is forbidden.

From the discussion above, the base intervals are

$$I_0 = [f^2(x_c), x_c] \quad I_1 = [x_c, f(x_c)] \quad (2.49)$$

which are nonempty for $f(x_c) > x_c$. The existence of symbolic sequences of length two is determined by the intervals

$$I_{00} = I_0 \cap f_0^{-1}(I_0) = [f^2(x_c), f_0^{-1}(x_c)] \quad (2.50)$$

$$I_{01} = I_0 \cap f_0^{-1}(I_1) = [f_0^{-1}(x_c), x_c] \quad (2.51)$$

$$I_{10} = I_1 \cap f_1^{-1}(I_0) = [f_1^{-1}(x_c), f(x_c)] \quad (2.52)$$

$$I_{11} = I_1 \cap f_1^{-1}(I_1) = [x_c, f_1^{-1}(x_c)] \quad (2.53)$$

which are computed by means of the interval arithmetics 2.45 but can also be obtained graphically (Fig. 2.14). The last three do not provide useful information: They are nonempty whenever $f(x_c) > x_c$, [i.e., as soon as I given by 2.48 is well-defined].

On the contrary, 2.50 yields a nontrivial condition for I_{00} to be nonempty, namely that $f^2(x_c) < f_0^{-1}(x_c)$. This interval has zero width when its two bounds are equal, thus the string “00” becomes allowed when the critical point belongs to a period-3 orbit:

$$f^2(x_c) = f_0^{-1}(x_c) \Rightarrow f^3(x_c) = x_c \quad (2.54)$$

which is then superstable since the derivative of f is zero at the critical point. For the logistic map 2.1, this occurs precisely at $a = a_{00} = 1.75487766\dots$, inside the unique period-3 window that can be seen in the bifurcation diagram of Fig. 2.3.

Since $I_{00} = \emptyset$ for $a < a_{00}$, we conclude that the symbolic string “00” never appears in the symbolic dynamics of regimes located at the left of the period-3 window. Thus, the presence or absence of this string suffices to distinguish regimes located before and after this window.

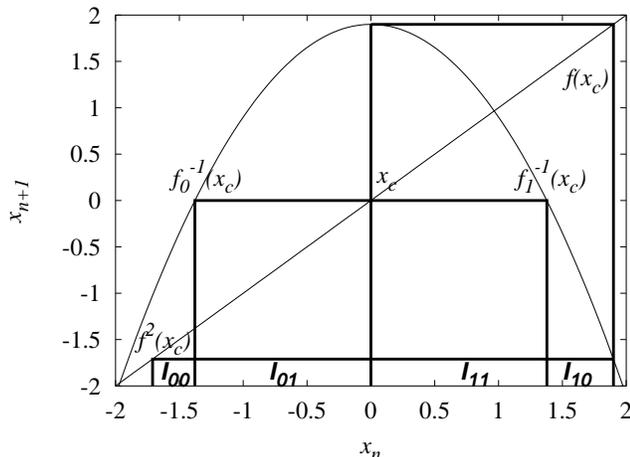


Figure 2.14: Intervals I_{00} , I_{01} , I_{11} , and I_{10} defined in 2.4.3. In each interval, the itineraries have the same leading two symbols.

In particular, this has consequences on the order of the appearance of periodic orbits. The first periodic orbit carrying the “00” string is the period-3 orbit $\overline{001}$. Therefore, all other periodic orbits whose names contain “00,” for example $\overline{001011}$, must appear after $\overline{001}$. This shows that the geometrical structure of the map has a deep influence on the order of appearance of periodic orbits, as detailed later in a more systematic way.

Reproducing the calculation above for longer symbolic strings, we would find that new symbolic sequences always appear when the critical point is part of a periodic orbit (i.e., at the parameter inside the periodic window where the orbit is superstable). This is not surprising if we note that the bounds of all the intervals I_Λ can be expressed in terms of the images and preimages of the critical point x_c .

As a result, the condition of zero width of these intervals can always be rewritten as an equation of the type $f_\Lambda^{-n}(x_c) = x_c$, expressing that x belongs to a periodic orbit of period n and of symbolic sequence Λ . For example, 2.54 corresponds to $f_{100}^{-3}(x_c) = x_c$. However, we will not proceed in this direction. The observation that the grammar of the symbolic dynamics is governed completely by the orbit of the critical point will lead us to a much more efficient framework for classifying symbolic sequences of orbits.

We conclude this section with the important remark that the symbolic dynamics of a chaotic dynamical system is in general *intimately related to its geometrical structure*. In the case of unimodal maps, the structure of the forbidden sequences depends only on the position of the image of the critical point organizing the dynamics. Thus, given an arbitrary symbolic sequence, it is in principle possible to determine whether it has been generated by a one-dimensional map.

More generally, extracting the structure of a map from the grammar of the symbolic dynamics it generates is a fascinating problem. It has been much less explored for two-dimensional invertible maps than for maps of the interval, and even less for noninvertible maps of dimension 2 and higher.

2.4.4 Kneading Theory

Rather than solve algebraic equations such as 2.54 to determine forbidden sequences, it would be preferable to work completely in the space of symbolic sequences. Since the orbit of the critical point plays a crucial role in understanding which symbolic sequences are forbidden, it is natural to study more closely the distinguished symbolic sequence associated with the critical point.

Since the first symbol of this sequence does not carry any information (the critical point x_c is the border between intervals I_0 and I_1), we accordingly define the *kneading sequence* $K(f)$ as the itinerary of the image of x_c :

$$K(f) = \Sigma(f(x_c)) = \{s(f(x_c)), s(f^2(x_c)), \dots\} \quad (2.55)$$

Note that the first two symbols are constant inside the parameter region where $f^2(x_c) < x_c < f(x_c)$: $f^2(x_c)$ and $f(x_c)$ are the left and right ends of the invariant interval I defined in 2.48 and are thus associated with the symbols 0 and 1, respectively. Since the value of the third symbol depends on whether $f^3(x_c)$ is located at the left or at the right of the critical point, it changes when $f^3(x_c) = x_c$ (i.e., when the string “00” becomes allowed), and thus

$$\begin{aligned} a < a_{00} &\Rightarrow K(f) = \{1, 0, 1, \dots\} \\ a > a_{00} &\Rightarrow K(f) = \{1, 0, 0, \dots\} \end{aligned} \quad (2.56)$$

This confirms the importance of the kneading sequence 2.55: The appearance of the symbolic string “00” in the symbolic dynamics of the logistic map coincides with its appearance in the kneading sequence.

To go beyond this observation, we need to be able to determine from $K(f)$ alone which sequences are allowed and which are not. The distinctive property of $f(x_c)$ is that it is the rightmost point of the invariant interval 2.48. To see that there is indeed a similar property for the kneading sequence, we first show that an order on itineraries can be defined.

Ordering of Itineraries

In the example of the $x_{n+1} = 2x_n \pmod{1}$ map (Section 2.3.5), the itinerary of a point (i.e., its binary expansion) not only identifies it uniquely, it also contains information about its position relative to the other points. In that case, the lexicographic order on symbolic sequences reflects exactly the order of the associated points on the interval. More generally, we would like to define for an arbitrary map an order relation \prec on itineraries so that

$$\Sigma(x) \prec \Sigma(x') \iff x < x' \quad (2.57)$$

Ordering two itineraries is easy when their leading symbols differ. If the base intervals I_α are numbered sequentially from left to right as in Fig. 2.12, the itinerary with the smallest leading symbol is associated with the leftmost point and should be considered “smaller” than the other.

If the two itineraries have a common leading substring, one has to take into account the fact that the map f can be orientation-reversing on some intervals I_α . For example, the two-symbol cylinders $I_{\alpha\alpha'}$ given by Eq. (2.4.3) and shown in Fig. 2.14 appear left to right in the order I_{00} , I_{01} , I_{11} , and I_{10} .

Thus, $11 \prec 10$ for the logistic map, which markedly differs from the lexicographic order. This is because both strings have a leading 1, which is associated with the orientation-reversing branch f_1 . Indeed, assume that $x_{11} \in I_{11}$, $x_{10} \in I_{10}$. From the second symbol, we know that $f(x_{10}) < f(x_{11})$ because $0 < 1$. However, since f is orientation-reversing in I_1 , this implies that $x_{11} < x_{10}$, hence $11 \prec 10$. With this point in mind, two arbitrary itineraries Σ, Σ' can be ordered as follows.

Assume that the two sequences $\Sigma = \Lambda s_m \dots$ and $\Sigma' = \Lambda s'_m \dots$ have a common leading symbolic string Λ of length m , and first differ in symbols s_m and s'_m . Thus, the corresponding points x and x' are such that $f^m(x)$ and $f^m(x')$ belong to different intervals I_α , hence can be ordered. As in the example above, it then suffices to determine whether the restriction f_Λ^m of f^m to the interval I_Λ is orientation-preserving or orientation-reversing (has a positive or a negative slope, respectively) to obtain the ordering of x and x' , and thus that of Σ and Σ' . Define the branch parity

$$\epsilon(\alpha) = \begin{cases} +1 & \text{if } f_\alpha : I_\alpha \rightarrow I \text{ is orientation-preserving} \\ -1 & \text{if } f_\alpha : I_\alpha \rightarrow I \text{ is orientation-reversing} \end{cases} \quad (2.58)$$

The parity of the finite sequence $\Lambda = s_0 s_1 \dots s_{m-1}$ is then given by

$$\epsilon(\Lambda) = \epsilon(s_0) \times \epsilon(s_1) \times \dots \times \epsilon(s_{m-1}). \quad (2.59)$$

If the map $f_\Lambda^m = f_{s_{m-1}} \circ \dots \circ f_{s_1} \circ f_{s_0}$ (i.e., the restriction of f^m to the interval I_Λ) is orientation-preserving (resp., orientation-reversing), then $\epsilon(\Lambda) = +1$ (resp., -1). In the case of unimodal maps, $\epsilon(\Lambda) = +1$ if there is an even number of “1” (or of the symbol associated with the orientation-reversing branch), and -1 otherwise.

We can now define the order

$$\Sigma = \Lambda s \dots \prec \Sigma' = \Lambda s' \dots \iff \begin{cases} s < s' & \text{and } \epsilon(\Lambda) = +1 \\ \text{or} & \\ s > s' & \text{and } \epsilon(\Lambda) = -1 \end{cases} \quad (2.60)$$

This order satisfies condition (2.57). Let us illustrate these rules with the example of period-4 orbit $\overline{0111}$ of the logistic map. The relative order of the four periodic orbits is

$$\overline{0111} \prec \overline{1101} \prec \overline{1110} \prec \overline{1011} \quad (2.61)$$

as detailed in Fig. 2.15.

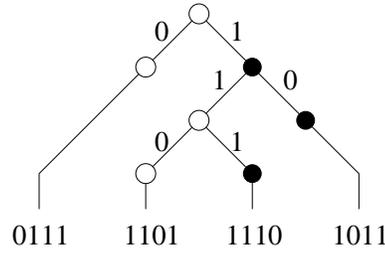


Figure 2.15: Determination of the relative order of symbolic sequences. White (resp., black) nodes correspond to positive (resp., negative) parity. The topmost node corresponds to the empty sequence, and sequences are formed by following edges carrying the symbols 0 or 1. When an edge “1” is followed, the parity of the node changes. A white node has an edge “0” on its left and an edge “1” on its right (this is the lexicographic order). At a black node, these two edges are in the opposite order because of the negative parity. To order a set of symbolic sequences, one follows the edges corresponding to the successive symbols of the sequence until no other sequence remains in the branch. The ordered sequences can then be read from left to right.

Another common technique for ordering symbolic sequences is to use *invariant coordinates*. Given a sequence $\Sigma = s_0 s_1 s_2 \dots s_k \dots \in \{0, \dots, N - 1\}^N$, we define its invariant coordinate $\theta(\Sigma)$ by

$$\theta(\Sigma) = \sum_{i=0}^{\infty} \frac{t_i}{N^{i+1}} \quad t_i = \begin{cases} s_i & \text{if } \epsilon(s_0 \dots s_{i-1}) = +1 \\ (N - 1) - s_i & \text{if } \epsilon(s_0 \dots s_{i-1}) = -1 \end{cases} \quad (2.62)$$

so that $0 \leq \theta(\Sigma) \leq 1$. By inspecting 2.60 and 2.62, one easily verifies that two symbolic sequences can be ordered by comparing their invariant coordinates:

$$\Sigma \prec \Sigma' \iff \theta(\Sigma) < \theta(\Sigma') \quad (2.63)$$

As an example, the invariant coordinate of the periodic point $\overline{1011}$ of the logistic map is

$$\begin{aligned} \theta(\overline{1011}) &= \left(\frac{1}{2^1} + \frac{\mathbf{1}}{2^2} + \frac{\mathbf{0}}{2^3} + \frac{1}{2^4} + \frac{\mathbf{0}}{2^5} + \frac{0}{2^6} + \frac{1}{2^7} + \frac{\mathbf{0}}{2^8} \right) \\ &\quad \times \left(1 + \frac{1}{2^8} + \frac{1}{2^{16}} + \dots \right) = \frac{105}{128} \times \frac{256}{255} = \frac{14}{17} \end{aligned} \quad (2.64)$$

where the digits in bold are those that have been inverted with respect to the original sequence. Because $\overline{1011}$ has negative parity, the binary digit sequence of $\theta(\Sigma)$ has period 8 instead of 4. The first factor in 2.64 corresponds to the basic pattern “11010010,” and the second term comes from the infinite repetition of this pattern. Note that the fraction obtained is the position of the corresponding

periodic point of the tent map defined on $[0, 1]$: The reader may verify as an exercise that $2 \times |1 - \theta(\overline{1011})| = \theta(\overline{0111})$ and that $2 \times \theta(\overline{0111}) = \theta(\overline{1110})$.

Admissible Sequences

We showed earlier that each point x inside the invariant interval (2.48) satisfies $f^2(x_c) < x < f(x_c)$. Using (2.57), we can now translate this ordering relation between points into a ordering relation between symbolic sequences:

$$\forall x \in I \quad \sigma K(f) \prec \Sigma(x) \prec K(f) \quad (2.65)$$

since $K(f) = \Sigma(f(x_c))$, by definition. Moreover, the orbit of a point $x \in I$ is forever contained in I , by definition. A necessary condition for a sequence Σ to be the itinerary $\Sigma(x)$ of a point $x \in I$ is thus that (2.65) holds for any $\Sigma(f^n(x))$ and thus that

$$\forall n \geq 0 \quad \sigma^n \Sigma \preceq K(f) \quad (2.66)$$

One of the fundamental results of one-dimensional symbolic dynamics is that this is also a sufficient condition: Condition (2.66) completely determines whether a sequence occurs as the itinerary of a point [?, ?]. A sequence satisfying it is said to be *admissible* [equivalently, one can test whether $\sigma K(f) < \sigma^n \Sigma$ for all n].

Therefore, all the information about the symbolic dynamics of a map is contained in its kneading sequence $K(f)$. As a matter of fact, it can be shown that if two unimodal maps have the same kneading sequence, and that if this sequence is dense (i.e., the orbit of x_c is aperiodic), the two maps are topologically conjugate.

Condition (2.66) is particularly simple to test when the symbolic sequence Σ is periodic, since the shifts $\sigma^n \Sigma$ are finite in number. For example, let us assume that $K(f) = 1001001 \dots$, and that we want to know whether the periodic sequences $\overline{01101101}$ and $\overline{00101}$ are admissible. We first determine the rightmost periodic points (for which $\sigma^n \Sigma$ is maximal) of the two orbits: These are $\overline{10110110}$ and $\overline{10010}$. We then compare them to the kneading sequence $K(f)$ and find that

$$\overline{10110110} \prec K(f) = 1001001 \dots \prec \overline{10010}$$

Thus, the period-8 sequence $\overline{10110110}$ is admissible, whereas the period-5 sequence $\overline{10010}$ is not. This indicates that the periodic orbit associated with the latter sequence does not exist in maps with the given $K(f)$. We also see that every map that has the second periodic orbit also has the first. Therefore, the order of appearance of periodic orbits is fixed, and the structure of the bifurcation diagram of Fig. 2.3 is universal for unimodal maps. We investigate this universality in the next section.

2.4.5 Bifurcation Diagram of the Logistic Map Revisited

We are now in a position to understand the structure of the bifurcation diagram shown in Fig. 2.3 using the tools of symbolic dynamics introduced in the previous

sections. This bifurcation diagram displays two types of bifurcations: saddle-node and period-doubling bifurcations. Each saddle-node bifurcation creates a pair of periodic orbits of period p , one unstable (the saddle) and the other stable (the node). The latter is the germ of a period-doubling cascade with orbits of periods $p \times 2^n$.

As discussed in Section 2.4.4, the kneading sequence governs which symbolic sequences are admissible and which are forbidden, hence the order in which new sequences appear. Therefore, there must be a simple relation between the symbolic names of the orbits involved in a saddle-node or in a period-doubling bifurcation. Furthermore, the different saddle-node bifurcations and their associated period-doubling cascades must be organized rigidly.

Saddle-Node Bifurcations

At a saddle-node bifurcation, the two newly born periodic orbits of period p are indistinguishable. Thus, they have formally the same symbolic name. This is not in contradiction with the one-to-one correspondence between orbits and itineraries that was shown to hold for chaotic regimes: At the bifurcation, the node is stable and there is no sensitivity to initial conditions.

When the two orbits are born, they have a multiplier of $+1$ (Section 2.2.1), which implies that f^p is orientation-preserving in the neighborhood of the orbit. Consequently, the common symbolic itinerary of the two orbits must contain an even number of symbols “1” (in the general case, an even number of symbols with negative parity).

A symbolic itinerary can change only if one of the periodic points crosses the critical point, which is the border of the partition. This happens to the stable node when it becomes superstable, changing its parity on its way to the period-doubling bifurcation where its multiplier crosses -1 . Thus, its parity must be negative, and its final symbolic name (i.e., the one in the unstable regime) must differ from that of the saddle by a single symbol.

One can proceed as follows to see which symbol differs. Each periodic point is associated with a cyclic permutation of the symbolic name. For example, the orbit 01111 has periodic points with sequences $\overline{01111}$, $\overline{11110}$, $\overline{11101}$, $\overline{11011}$, and $\overline{10111}$. These periodic points can be ordered using the kneading order 2.60, here

$$\overline{01111} \prec \overline{11110} \prec \overline{11011} \prec \overline{11101} \prec \overline{10111}$$

The symbol that is flipped at the superstable parameter value is obviously associated with the point that is then degenerate with the critical point. The image of this point is thus the rightmost periodic point, corresponding to the highest sequence in the kneading order (Fig. 2.16). Consequently, a simple rule to obtain the symbolic name of the saddle-node partner of an orbit of given name is to flip the last symbol of the highest itinerary (i.e., of the itinerary of the rightmost point). Alternatively, one can flip the second to last symbol of the leftmost itinerary.

In the example above, the saddle-node partner of the 01111 orbit is thus 01101. Other examples of saddle-node pairs include $0_1^0 1$, $00_1^0 1$, and 01101_1^0 .

Universal Sequence

Consider two periodic itineraries $\Sigma \prec \Sigma'$. For some parameter a , the kneading sequence $K(f_a)$ is such that $\Sigma \prec K(f_a) \prec \Sigma'$, so that Σ satisfies the admissibility condition 2.66 but not Σ' . Thus, the periodic orbit associated to Σ must be created before the one associated to Σ' .

This observation suffices to build a complete list of the successive bifurcations occurring in the bifurcation diagram of Fig. 2.3. Using the rules derived in previous sections, we can classify all the symbolic names according to which series of bifurcations they belong.

To this end, all periodic itineraries up to a given period p are sorted according to the kneading order, with saddle-node pairs and orbits of the same period-doubling cascade grouped together. We denote the i th bifurcation creating period- P orbits as P_i , with the node being called $P_i f$ (for flip) and the saddle $P_i r$ (for regular). This is illustrated in Table 2.1, which lists the symbolic names of all periodic orbits of period up to 8 of the logistic map. These names are sorted by order of appearance, and the bifurcation in which they appear is indicated.

Table 2.1: Sequence of bifurcations in the logistic map up to period 8 (from top and to bottom and left to right)^a

Name	Bifurcation	Name	Bifurcation	Name	Bifurcation
0_1	$1_1[s_1]$	00101_1^0	$7_3[s_7^2]$	0001_1^0	$6_4[s_6^3]$
01	$2_1[s_1 \times 2^1]$	001010_1^0	$8_5[s_8^4]$	000111_1^0	$8_{11}[s_8^9]$
0111	$4_1[s_1 \times 2^2]$	001_1^0	$5_2[s_5^2]$	00011_1^0	$7_7[s_7^7]$
01010111	$8_1[s_1 \times 2^3]$	001110_1^0	$8_6[s_8^5]$	000110_1^0	$8_{12}[s_8^{10}]$
0111_1^0	$6_1[s_6^1]$	00111_1^0	$7_4[s_7^4]$	000_1^0	$5_3[s_5^3]$
011111_1^0	$8_2[s_8^1]$	001111_1^0	$8_7[s_8^6]$	000010_1^0	$8_{13}[s_8^{11}]$
01111_1^0	$7_1[s_7^1]$	0011_1^0	$6_3[s_6^2]$	00001_1^0	$7_8[s_7^8]$
011_1^0	$5_1[s_5^1]$	001101_1^0	$8_8[s_8^7]$	000011_1^0	$8_{14}[s_8^{12}]$
01101_1^0	$7_2[s_7^2]$	00110_1^0	$7_5[s_7^2]$	0000_1^0	$6_5[s_6^4]$
011011_1^0	$8_3[s_8^2]$	00_1^0	$4_2[s_4^1]$	000001_1^0	$8_{15}[s_8^{13}]$
0_1^0	$3_1[s_3]$	00010011	$8_9[s_4^1 \times 2^1]$	00000_1^0	$7_9[s_7^9]$
001011	$6_2[s_3 \times 2^1]$	00010_1^0	$7_6[s_7^6]$	000000_1^0	$8_{16}[s_8^{14}]$
001011_1^0	$8_4[s_8^3]$	000101_1^0	$8_{10}[s_8^8]$		

^aThe notation P_i refers to the i th bifurcation of period P . We also give inside brackets an alternative classification that distinguishes between saddle-node and period-doubling bifurcations. In this scheme, the i th saddle-node bifurcation of period P is denoted s_P^i , and $s_P^i \times 2^k$ is the orbit of period $P \times 2^k$ belonging to the period-doubling cascade originating from s_P^i .

This sequence of symbolic names, often referred to as the *universal sequence*, was discovered by Metropolis, Stein, and Stein [?]. It is universal in that it depends only on the kneading order 2.60: The bifurcation diagram of any unimodal map will display exactly the same bifurcations in exactly the same order.

Note, however, that this holds only for one-dimensional maps. If a two-dimensional map is sufficiently dissipative so that its return map can be well approximated by a one-dimensional map, most of the bifurcation sequences will occur in the order predicted by the universal sequence. However, there will be a few discrepancies, and the order of many bifurcations will be reversed as one decreases dissipation [?].

Self-Similar Structure of the Bifurcation Diagram

In this section we mention briefly another surprising property of the bifurcation diagram of the logistic map that is unveiled by symbolic dynamics. Look at the period-3 window beginning at $a = 1.75$ in Fig. 2.4. There is a whole parameter range where the attractor is contained in three disconnected pieces, before it expands suddenly. These pieces are visited successively in a fixed order. We call this parameter region the generalized period-3 window. Look more closely at, say, the middle branch: This is a complete copy of the whole bifurcation diagram! In particular, there is a period-9 window which is to the period-3 window what the period-3 is itself to the whole diagram.

To understand this, we note that the base symbols 0 and 1 can be viewed as the names of the period-1 orbits organizing the global dynamics. Similarly, let us denote by $X = 101$ and $Y = 100$ the symbolic names of the two period-3 orbits born in the saddle-node bifurcation initiating the period-3 window. All periodic orbits appearing in the generalized period-3 window can be written as words in the letters X and Y .

Indeed, since the attractor is split into three pieces visited successively, the dynamics can be simplified by considering the third iterate f^3 . Each of the three pieces is a different attractor of f^3 . The return map for each attractor is a unimodal map, with two “period-1” orbits that are in fact periodic points of the two period-3 orbits $\overline{100}$ and $\overline{101}$. Any pair of symbols X' and Y' such that the sequences $\overline{X'}$ and $\overline{Y'}$ correspond to periodic points that are degenerate at the period-3 saddle-node bifurcation can thus be used to code orbits of this map. Because we chose X and Y above to be higher in the kneading order than all their cyclic permutations, they satisfy this condition as well as any pair $\sigma^k X, \sigma^k Y$.

Since the two words $X = 101$ and $Y = 100$ are such that (1) $\overline{X} \prec \overline{Y}$ and (2) they have parities $\epsilon(X) = +1$ and $\epsilon(Y) = -1$, it is easy to see that the ordering of two sequences $W_1(X, Y)$ and $W_2(X, Y)$ will be exactly the same as for the corresponding sequences $W_1(0, 1)$ and $W_2(0, 1)$. For example,

$$\overline{YXYYXYYX} \prec \overline{YXXYX} \iff \overline{10110110} \prec \overline{10010}$$

This explains why the bifurcation diagram in the generalized period-3 window has exactly the same structure as the whole diagram. Using the names of the standard period-doubling cascade given in 2.67, we find that the orbits involved in the period-doubling cascade of this window are $Y, XY, XYYY, XYYYXYXY$. The first orbits to appear in the window are the X and Y

orbits (naming them after their sequences in the unstable regime), the last is the YX^∞ orbit.

In fact, the results of this section could have been foreseen: They are a consequence of the qualitative universality of bifurcations in unimodal maps. Inside the period-3 window, the third return map is a unimodal map and therefore displays the same series of bifurcations as the first return map.

2.5 Shift Dynamical Systems, Markov Partitions, and Entropy

In Section 2.4, we have seen how a chaotic system can be analyzed with the tools of symbolic dynamics. In particular, each regime of the logistic map is characterized by a different grammar (i.e., a set of forbidden symbolic sequences). Moreover, symbolic dynamics can be shown in some cases to provide a complete description of a dynamical system; for example, it is known that chaotic unimodal maps are conjugate if they have the same kneading sequence.

It is thus natural to study systems whose evolution laws are defined directly in a symbolic space by rules specifying which sequences are admissible. Such systems are usually referred to as *symbolic dynamical systems*, or as *shift dynamical systems* when they are based on the shift map [?]. Tools developed to characterize these systems can then be applied to any physical system for which a symbolic dynamical description has been obtained. This is illustrated by computations of entropy, an important measure of chaotic dynamics.

Here, we limit ourselves to *shifts of finite type*, which are characterized by a finite set of forbidden sequences. The interest of finite shifts is twofold. First, there are dynamical systems, those for which a *Markov partition* exists, that can be shown to be equivalent to a finite shift. Second, systems whose grammar cannot be specified by a finite set of rules can always be approximated with increasing accuracy by a sequence of finite shifts of increasing order.

2.5.1 Shifts of Finite Type and Topological Markov Chains

The natural phase space of a symbolic dynamical system is the set of infinite or bi-infinite sequences of symbols from an alphabet \mathcal{A} . Here we assume that the alphabet is finite and choose $\mathcal{A} = \{0, \dots, N-1\}$, where N is the number of symbols. The systems we consider here share the same time-one map: the shift operator σ , which shifts symbols one place to the left (Section 2.4.1).

In the case of the logistic map, the symbolic space consisted of one-sided symbolic sequences. We noted in Section 2.4.1 that this makes σ noninvertible, since memory of the leading symbol is lost after each time step. If the shift operator has to be invertible, its action must not discard information. Thus, sequences must be bi-infinite (two-sided), for example,

$$\Sigma = \dots s_{-3}s_{-2}s_{-1}.s_0s_1s_2 \quad (2.68)$$

with the dot separating the *forward sequence* $\Sigma_+ = s_0s_1\dots$ from the *backward sequence* $\Sigma_- = s_{-1}s_{-2}\dots$. These two sequences describe the future and the past of the point, respectively. The action of the shift operator on a sequence is then given by

$$\sigma(\dots s_{-1}.s_0s_1\dots) = \dots s_{-1}s_0.s_1\dots$$

The dot is merely moved to the right, which obviously preserves the information contained in the sequence. This is illustrated with the horseshoe map in Section ??.

The distinction between invertible and noninvertible dynamics is not made by most methods developed for characterizing symbolic dynamical systems. As we see below, they usually involve determining which finite blocks of symbols can appear in a typical sequence and which cannot. Thus, whether sequences are one- or two-sided is not relevant.

Full shifts are the simplest shift symbolic dynamical systems: Any sequence made of letters of the alphabet is allowed. Thus the symbolic space is $\mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$, depending on the invertibility of the dynamics. Two full shifts can differ only by the number of symbols of their alphabets. A full shift on r symbols is termed an r -shift.

In a general shift dynamical system, not all sequences are allowed: It is then called a *subshift*. We are interested only in subshifts whose set of allowed sequences \mathcal{S} is shift-invariant (i.e., $\sigma\mathcal{S} = \mathcal{S}$). This implies that whether a finite symbol string can be found in a sequence $\Sigma \in \mathcal{S}$ does not depend on its position in the sequence but only in the content of the string. Finite symbol strings are also often referred to as *blocks*, with an n -block containing n symbols, or as *words*.

One could therefore provide a complete description of a dynamical system (\mathcal{S}, σ) by specifying its *language* (i.e., the list of all finite strings of symbols that can be extracted from infinite sequences). It is usually much more convenient to specify its set \mathcal{F} of *irreducible forbidden words* (IFWs). An IFW never appears in a sequence of \mathcal{S} and does not contain any other forbidden word. For example, assume that $\mathcal{F} = \{00\}$. The word 001 is a forbidden word but is not irreducible because it contains 00. More generally, any word of the form $u00v$, where u and v are arbitrary words is not irreducible. By construction, any forbidden word of the language must have one of the elements of \mathcal{F} as a substring, or is itself an IFW. IFWs are of length $l \geq 2$, since length-1 forbidden words can be removed by reducing the alphabet.

Of particular importance are the *shifts of finite type* (SFTs), which are described by a finite number of IFWs. Indeed, they can be specified with a finite amount of information, and invariant quantities such as the entropies described later can then be computed exactly. If the longest IFW of a shift of finite type is of length $L + 1$, the *order* of the shift is L .

Shifts of finite type of order 1 are also called *topological Markov chains*. Since their set of IFWs contains only 2-blocks, the structure of this set can be

described by a transition matrix M such that

$$M_{s_1, s_0} = \begin{cases} 0 & \text{if } s_0 s_1 \text{ is forbidden} \\ 1 & \text{if } s_0 s_1 \text{ is allowed} \end{cases} \quad (2.69)$$

That is, M_{s_1, s_0} is nonzero if s_1 is allowed to follow s_0 in a sequence (i.e., there is a transition $s_0 \rightarrow s_1$). A simple example is the transition matrix of the SFT with $\mathcal{F} = \{00\}$:

$$M = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad (2.70)$$

which characterizes the symbolic dynamics of the logistic map immediately before the period-3 window (see Section 2.4.3).

Markov chains are all the more important as any shift of finite type of order L can be reformulated as a Markov chain by recoding sequences appropriately. Assume that there are N' allowed L -blocks, and denote \mathcal{A}_L the alphabet made of these N' symbols. Any sequence $s_0 s_1 \dots s_k \dots$ can then be recoded as $S_0 S_1 \dots S_k \dots$, where the new symbol $S_k \in \mathcal{A}_L$ is the L -block starting at position k : $S_k = s_k s_{k+1} \dots s_{L-1+k}$. For example, $S_0 = s_0 s_1 \dots s_{L-1}$ and $S_1 = s_1 s_2 \dots s_L$.

Now, assume that $S = s_0 s_1 \dots s_{L-1}$ and $S' = s'_0 s'_1 \dots s'_{L-1}$ are two symbols of \mathcal{A}_L . The element $M_{S', S}$ of the new transition matrix is 1 if:

- The head of S coincides with the tail of S' : $s'_0 s'_1 \dots s'_{L-2} = s_1 \dots s_{L-1}$
- $s_0 s_1 \dots s_{L-2} s'_{L-1}$ is an allowed L -block of the original shift

and is 0 otherwise. For example, assume that $\mathcal{F} = \{00, 0110\}$ and sequences are recoded using blocks of four symbols. Then one has $M_{101, 011} = 0$ because 11 (tail of S) differs from 10 (head of S'), $M_{110, 011} = 0$ because 0110 is an IFW, but $M_{111, 011} = 1$ because 0111 does not contain any IFW.

Therefore, we see that any shift of finite type can be described completely by a transition matrix M . In the following section we show how to extract from this matrix information about the spectrum of periodic orbits of the dynamical system and an important measure of chaos, topological entropy.

2.5.2 Periodic Orbits and Topological Entropy of a Markov Chain

As discussed previously, periodic points of a symbolic dynamical system correspond to periodic sequences Σ satisfying $\sigma^p \Sigma = \Sigma$. Since periodic orbits play a crucial role, we want to be able to compute the number of periodic sequences of period p of an arbitrary Markov chain, given its transition matrix M . This problem is part of a more general one, which is to determine the number of allowed symbol strings of length n . An important measure of chaos, *topological entropy*, characterizes how this number increases when $n \rightarrow \infty$.

A transition matrix is conveniently represented by a directed graph. To each symbol corresponds a node, which can be viewed as a state. When the transition

from “state” i to “state” j is allowed (i.e., $M_{ji} \neq 0$), there is a directed edge going from node i to node j (Fig. 2.17). The problems stated above can be reformulated as follows: How many distinct paths of length n does the graph have? How many of these paths are closed?

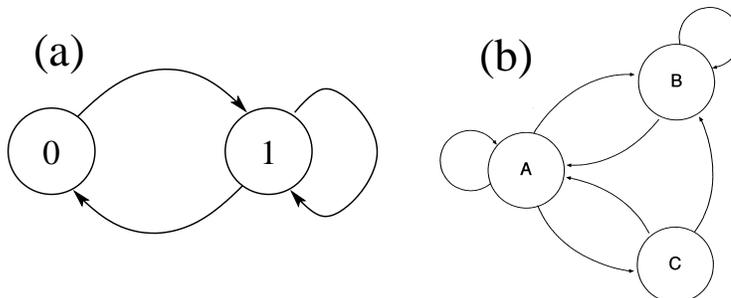


Figure 2.17: A transition matrix can be represented by a directed graph. Nodes correspond to possible states and edges indicate whether transition from one state to another is possible. (a) This graph corresponds to the transition matrix 2.70. (b) This graph describes a grammar in which C must be preceded by A : The sequences BC and CC cannot occur.

We first compute the number P_{ji}^n of paths connecting node i to node j in exactly n steps. This can be done inductively. Obviously, one can go from i to j in one step only if there is an edge between the two sites, thus $P_{ji}^1 = M_{ji}$. Then we note that each itinerary linking i to j in n steps goes from i to some site k in $n - 1$ steps, then follows the edge from k to j in one step. By summing over all possible intermediary sites k , one obtains

$$P_{ji}^n = \sum_{k=0}^{N-1} M_{jk} P_{ki}^{n-1}$$

which is immediately seen to be the rule for matrix multiplication. Since $P_{ji}^1 = M_{ji}$, it follows that

$$P_{ji}^n = (M^n)_{ji} \quad (2.71)$$

Hence all the relevant information is contained in the successive powers of the transition matrix.

Periodic sequences of period p correspond to paths of length p that begin and end at the same node. Thus, the number $N_f(p)$ of periodic points of period p is

$$N_f(p) = \sum_{k=0}^{N-1} M_{kk}^p = \text{tr} M^p \quad (2.72)$$

Similarly, the number $N_s(p)$ of p -symbol strings equals the total number of paths of length $p - 1$ and thus is given by the sum of the elements of M^{p-1} . This can

be formalized as follows. Let V^p be the vector whose i th component V_i^p is the number of length- p symbol strings beginning by symbol i . It is easy to see that the components of V^1 are all 1, and that $V^p = M^{p-1}V^1$. Thus,

$$N_s(p) = (V^1)^T V^p = (V^1)^T M^{p-1} V^1 \quad (2.73)$$

Expressions 2.72 and 2.73 show that $N_f(p)$ and $N_s(p)$ have the same asymptotic behavior. Indeed, the action of M^p for large p is determined by its largest eigenvalue λ_{\max} and the associated eigenvector. It is easily shown that if $\lambda_{\max} > 1$, then

$$\lim_{p \rightarrow \infty} \frac{\log N_f(p)}{p} = \lim_{p \rightarrow \infty} \frac{\log N_s(p)}{p} = \log \lambda_{\max} \quad (2.74)$$

The growth rate of the number $N_s(p)$ of p -blocks

$$h_T = \lim_{p \rightarrow \infty} \frac{\log N_s(p)}{p} = \log \lambda_{\max} \quad (2.75)$$

is called the *topological entropy*. It measures the average amount of information that is extracted by reading one symbol of a typical sequence. Equation 2.74 shows that the topological entropy of a Markov chain depends in a very simple way on the transition matrix. It also illustrates the fact that in general the growth rate of the number of periodic points is equal to the topological entropy, as noted in Section 2.3.4. More sophisticated techniques to compute topological entropy are presented in Section 2.5.6. Let us consider two examples:

- The eigenvalues of the transition matrix 2.70 are $(1 \pm \sqrt{5})/2$. The largest one is $\lambda_{\max} \sim 1.6180339$ and is known as the *golden mean*. This Markov chain, accordingly called the *golden mean shift*, has topological entropy $h_T \sim 0.4812118$.
- The transition matrix of the full N -shift is filled with 1's. Its largest eigenvalue is N , and the topological entropy is $h_T = \ln N$. In particular, the 2-shift has $h_T = 0.6931471$, which is greater than for the golden mean shift: Topological entropy increases as chaos becomes more developed.

2.5.3 Markov Partitions

Markov chains are not only interesting as model dynamical systems but also because there are some classical dynamical systems whose symbolic dynamics can be represented exactly by a topological Markov chain. We have already encountered a few examples of such systems. The simplest ones are the tent and logistic maps associated with a full 2-shift: Every sequence of “0” and “1” is associated with a physical orbit. The logistic map at the beginning of the period-3 window is another example of a finite shift; the only forbidden sequence is the string “00.” A natural question then is: Under which conditions is a dynamical system described faithfully by a Markov chain?

The symbolic coding of a dynamical system relies on the existence of a partition $\mathcal{P} = \{\mathcal{P}_0, \mathcal{P}_1, \dots, \mathcal{P}_{N-1}\}$ of phase space into N disjoint regions \mathcal{P}_i . At each time step, the current system state is coded with the symbol i associated with the region \mathcal{P}_i to which it belongs. Because the partition and the time-one map determine completely the symbolic dynamics, it is not surprising that the condition for being describable by a Markov chain involves the partition and the images of the members of the partition.

We now state this condition without a proof. Assume that there exists a partition $\mathcal{P} = \{\mathcal{P}_i\}$ such that the intersection of any member with the image of another is either itself or is empty:

$$\forall i, j \quad \mathcal{P}_i \cap f(\mathcal{P}_j) = \begin{cases} \mathcal{P}_i \\ \emptyset \end{cases} \quad (2.76)$$

The structure of such a partition can be described concisely but faithfully by a transition matrix $M^{\mathcal{P}}$ defined by

$$M_{ij}^{\mathcal{P}} = \begin{cases} 1 & \text{if } \mathcal{P}_i \cap f(\mathcal{P}_j) \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \quad (2.77)$$

It can be shown that the dynamical system coded by the partition \mathcal{P} is then completely equivalent to the Markov chain of transition matrix $M^{\mathcal{P}}$. In particular, the topological entropy of the original dynamical system is equal to that of the Markov chain, and both systems have the same spectrum of periodic orbits. Accordingly, a partition \mathcal{P} satisfying (2.76) is called a *Markov partition*. Examples of Markov partitions in the logistic map at special values of the parameter a are given in Section 2.6.2.

Note that even when there is a generating partition (such as the one based on critical points of a one-dimensional map) and it is not of Markov type, the existence of a Markov partition is not precluded. If the system is equivalent to a shift of finite type, an analysis of the symbolic dynamics obtained with the generating partition should reveal that there are a finite number of irreducible forbidden words. As described in Section 2.5.1, a Markov chain can then be obtained with a suitable recoding and the topological entropy computed using the associated transition matrix.

2.5.4 Approximation by Markov Chains

In fact, only a small fraction of the regimes of the logistic map can be represented by a Markov chain exactly. Indeed, there is only a countable number of finite matrices of 0 and 1, whereas these regimes are indexed by the parameter a , which is a real number [?]. Moreover, chaotic regimes are associated with kneading sequences that are not eventually periodic, which makes it generally impossible to describe the symbolic dynamics by a finite number of IFWs.

However, this does not make shifts of finite type irrelevant. Indeed, it is not possible to analyze arbitrarily long symbol sequences. In practice, there is an upper bound on the length of the longest symbolic sequence that can

be obtained in a reasonable time. This limits the search for forbidden symbol blocks to a maximal length. Otherwise, longer symbol blocks may be classified incorrectly as forbidden only because their probability of occurrence is too small. For example, assume that an orbit of 1 million ($\sim 2^{20}$) points has been recorded and coded on two symbols. It is certainly pointless to determine forbidden blocks longer than 20 symbols, since the least probable one will occur at most once in the best case, where all blocks are equiprobable.

Therefore, Markov chains are still relevant to characterize dynamical systems that are not conjugate to a shift of finite type, provided that a generating partition is known and a long symbolic sequence has been recorded. If the list of forbidden words has been determined for word lengths up to L , this gives a natural approximation of the system under study by a shift of finite type of order L , and hence by a Markov chain after a suitable recoding. Note that this systematically overestimates topological entropy estimates because higher-order forbidden sequences are neglected. The expression $h_T = \ln \lambda_{\max}$ for a Markov chain assumes that the number $N_s(p)$ of p -blocks can be determined for arbitrary p from the transition matrix. If there is more “pruning” than described by this transition matrix, the actual number of sequences will be lower, as well as the topological entropy. Carrying out this computation for increasing block lengths and comparing the results may help to estimate its accuracy.

2.5.5 Zeta Function

As shown in Section 2.5.2, the number of periodic points of period n , P_n , can be computed as $P_n = \text{tr } M^n$, where M is the Markov transition matrix. The information contained in the transition matrix can be transformed to a generating function for P_n by defining

$$\zeta(t) = \exp\left(\sum_{n=1}^{\infty} \frac{P_n}{n} t^n\right) \quad (2.78)$$

With a little bit of algebraic calisthenics, it is possible to show that

$$\zeta(M, t) = \frac{1}{\det(Id - tM)} \quad (2.79)$$

We illustrate one use of the zeta function in the example below.

Example: The spectrum of orbits forced by the period-three orbit 3_1 of the logistic map is computed using the Markov transition matrix M (cf. (2.70)). If the matrices $Id = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $M = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ have been defined previously, as well as the positive integer N ($= 12$ below), the generating function $\sum_n P_n/n t^n$ is given up to degree N by the simple Maple call

$$\begin{aligned} > \text{taylor}(-\log(\det(Id - t * M)), t = 0, N + 1); \\ & t + \frac{3}{2} t^2 + \frac{4}{3} t^3 + \frac{7}{4} t^4 + \frac{11}{5} t^5 + \frac{18}{6} t^6 + \frac{29}{7} t^7 + \frac{47}{8} t^8 + \frac{76}{9} t^9 \\ & + \frac{123}{10} t^{10} + \frac{199}{11} t^{11} + \frac{322}{12} t^{12} + \mathcal{O}(t^{13}) \end{aligned} \quad (2.80)$$

Table 2.2: Number of orbits up to period $p = 12$ forced by 3_1 computed using the zeta function based on the golden mean matrix (2.70)

p	$N(p)$	Lower-Period Orbits	Period- Doubled	Number of Saddle- Node Pairs
1	1	1_1		
2	3	$(1_1+$	$2_1)$	
3	4	$1_1 + 3_1$		
4	7	$(1_1 + 2_1+$	$4_1)$	0
5	11	1_1		1
6	18	$1_1 + 2_1 + 3_1$		1
7	29	1_1		2
8	47	$(1_1 + 2_1 + 4_1$	$8_1)$	2
9	76	$1_1 + 3_1$		4
10	123	$1_1 + 2_1 + (5_1+$	$10_4)$	5
11	199	1_1		9
12	322	$(1_1 + 2_1 + 4_1) + 3_1 + (6_1+$	$12_2)$	12

We read these results as follows:

1. There is one period-1 point 1_1 .
2. There are three period-2 points. One is the period-1 point 1_1 . The other two belong to the single period-2 orbit 2_1 .
3. There are four period-3 points. One is 1_1 . The other three belong to the *degenerate* saddle-node pair 3_1 (001 and 011).
4. There are seven period-4 points, which belong to the orbits 1_1 , 2_1 , and 4_1 of the initial period-doubling cascade.
5. There are 11 period-5 points. One belongs to 1_1 . The remaining 10 belong to two period-5 orbits, which comprise the saddle-node pair 01111 and 01101.

Continuing in this way, we construct the remainder of the results. These are summarized in Table 2.2. It is a simple matter to verify that the results of this table are consistent with the results of Table 2.1 up to period 8.

2.5.6 Dealing with Grammars

At the saddle-node bifurcation of the period-three orbit, the adjacency matrix (Markov transition matrix) is given by Eq. (2.70). This matrix tells us that the symbol 0 must be followed by the symbol 1, and the symbol 1 can be followed by either of the symbols 0 or 1.

Simple Grammars

It is useful to introduce an alternative representation for the dynamics. This involves introducing two symbols (words) $A = 01$ and $B = 1$. These two words have length 2 and 1, respectively. In this representation, A can be followed by either A or B , as is true also for B . The transition matrix is full:

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad (2.81)$$

and the grammar is simple (no transitions are forbidden).

The periodic orbits in the dynamics can be constructed as follows. Replace the nonzero elements in the first row of M by the symbol A , those in the second row by the symbol B (A and B do not commute).

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} A & A \\ B & B \end{bmatrix} = \mathcal{M} \quad (2.82)$$

Then the complete set of periodic orbits is constructed by computing $\text{tr } \mathcal{M}^n$, $n = 1, 2, 3, \dots$

Example: For $n = 1, 2$ and 3 we find

n	$\text{tr } \mathcal{M}^n$	Orbits
1	$A + B$	$01 + 1$
2	$2A^2 + AB + B^2$	$2(01)^2 + 011 + (1)^2$
3	$2A^3 + 2ABA + 2A^2B + AB^2 + B^3$	$2(01)^3 + 4(01101) + 0111 + (1)^3$

Reduction to simple grammars is often useful in analyzing experimental data. For example, chaotic data generated by the Belousov–Zhabotinskii reaction (cf. Chapter 7) have been reduced to a symbolic code sequence involving the two symbols 0 and 1. The rules of grammar observed in the experimental data are:

1. The symbol 0 must be followed by the symbol 1.
2. The symbol 1 can be followed by 0 or 1.
3. Symbol sequences $(11 \cdots 1)$ of length p can occur for $p = 1, 2, 3, 4$ but not for $p > 4$.

It appears that the vocabulary of this dynamics consists of the four words 01, 011, 0111, and 01111. It also appears that any of these words can be followed by any other word. The grammar is simple and represented by a 4×4 Markov transition matrix whose 16 elements are 1. The periodic orbits are obtained as described above for the golden mean case.

The topological entropy for dynamics consisting of a finite number of words of varying length obeying a simple grammar (full shift) is easily determined as follows. Assume that there are $w(p)$ words of length p , $p = 1, 2, 3, \dots$. Then the number of ways, $N(T)$, of constructing a word of length T is determined by the difference equation

$$N(T) = w(1)N(T-1) + w(2)N(T-2) + \cdots = \sum_{p=1} w(p)N(T-p) \quad (2.83)$$

The number $N(T)$ behaves asymptotically like $N(T) \sim A(X_M)^T$, where A is some constant and X_M is the largest real root of the characteristic equation

$$X^T = \sum_{p=1} w(p)X^{T-p} \quad \text{or} \quad 1 = \sum_{p=1} \frac{w(p)}{X^p} \quad (2.84)$$

Example 1: For the full shift on two symbols 0 and 1,

$$1 = \frac{2}{X} \implies h_T = \log 2$$

Example 2: For golden mean dynamics, $w(1) = 1$, $w(2) = 1$, and

$$1 = \frac{1}{X} + \frac{1}{X^2} \implies X_M = \frac{1 + \sqrt{5}}{2} \quad \text{and} \quad h_T = 0.481212$$

Example 3: For the Belousov data described above, $w(p) = 1$ for $p = 2, 3, 4, 5$ and $w(p) = 0$ for $p = 1$ and $p > 5$, so that

$$1 = \frac{1}{X^2} + \frac{1}{X^3} + \frac{1}{X^4} + \frac{1}{X^5} \quad X_M = 1.534158 \quad \text{and} \quad h_T = 0.427982$$

Complicated Grammars

There are many cases in which the dynamics either consists of, or is well approximated by, a finite vocabulary with a nontrivial grammar. The two questions addressed above (description of periodic orbits, computation of entropy) are still of interest.

The spectrum of periodic orbits can be computed by following the algorithm used above.

1. Write out the Markov transition matrix M for the vocabulary.
2. Replace each nonzero matrix element 1 in row i by the noncommuting symbol w_i representing the i th word, effecting the transition $M \rightarrow \mathcal{M}$.
3. Compute $\text{tr } \mathcal{M}^n$ for $n = 1, 2, \dots$.
4. Replace each word sequence by the appropriate sequence of symbols from the original alphabet (i.e., 0 and 1).

The problem of computing the topological entropy for this dynamics is more subtle. It is isomorphic to the problem of computing the capacity of a transmission channel. The capacity of a transmission channel is (Shannon, [?, ?])

$$C = \lim_{T \rightarrow \infty} \frac{1}{T} \log N(T)$$

Here $N(T)$ is the number of allowed signals of duration T , and \log is to base e .

In many grammars, not all symbol sequences are allowed (qu is OK, qv is not). In such cases, assume that there are m states b_1, b_2, \dots, b_m . For each state only certain symbols from the set S_1, S_2, \dots, S_n can be transmitted (different subsets for different states). The transmission of symbol S_k from state b_i to state b_j (b_i may be the same as b_j) takes time $t_{ij}^{(k)}$, where k indexes all possible paths from b_i to b_j . This process is illustrated by a graph such as that shown in Fig. 2.17.

Theorem: The channel capacity C is $\log W_0$, where W_0 is the largest real root of the $m \times m$ determinantal equation

$$\left| \sum_k W^{-t_{ij}^{(k)}} - \delta_{ij} \right| = 0 \quad (2.85)$$

For our purposes, we can regard each state as a word and t_{ij} is the length of word i .

Example: Assume that the vocabulary has three words A, B , and C or w_1, w_2 , and w_3 of lengths p, q , and r and a grammar defined by Fig. 2.17(b). The Markov matrix is

$$M = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

The determinantal equation constructed from the Markov matrix and word lengths is

$$\begin{bmatrix} \frac{1}{W^p} - 1 & \frac{1}{W^p} & \frac{1}{W^p} \\ \frac{1}{W^q} & \frac{1}{W^q} - 1 & 0 \\ \frac{1}{W^r} & \frac{1}{W^r} & -1 \end{bmatrix} = 0$$

The characteristic equation for this dynamical system is

$$\frac{1}{W^p} + \frac{1}{W^q} + \frac{1}{W^{p+r}} = 1$$

The topological entropy is $h_T = \log W_0$, where W_0 is the largest real eigenvalue of this characteristic equation.

2.6 Fingerprints of Periodic Orbits and Orbit Forcing

2.6.1 Permutation of Periodic Points as a Topological Invariant

Using kneading theory, periodic points with given symbolic itineraries can be ordered along the interval. This is not only useful to determine the order in which periodic orbits appear, but also to identify periodic orbits.

Indeed, consider the period-5 orbit with symbolic name 01011 of the logistic map (Fig. 2.16). Its five periodic points are associated with the five cyclic permutations of the symbolic name, ordered by kneading theory as follows:

$$\overline{01101} \prec \overline{01011} \prec \overline{11010} \prec \overline{10101} \prec \overline{10110} \quad (2.86)$$

Label the sequences in 2.86 from left to right by Σ_i , $i = 1 \dots 5$, and express them in terms of the leftmost sequence $\Sigma_1 = 01101$ and of powers of the shift operator. We have

$$\Sigma_1 \prec \Sigma_2 = \sigma^3 \Sigma_1 \prec \Sigma_3 = \sigma \Sigma_1 \prec \Sigma_4 = \sigma^2 \Sigma_1 \prec \Sigma_5 = \sigma^4 \Sigma_1 \quad (2.87)$$

We observe that under the action of the shift operator σ , these sequences are permuted: The lowest sequence becomes third, the second one the last, and so on. The corresponding permutation

$$\pi(01011) = (\pi_i) = (3, 5, 4, 2, 1) \quad (2.88)$$

such that $\sigma \Sigma_i = \Sigma_{\pi_i}$ provides crucial information about the orbit. Its dynamical relevance owes much to two fundamental properties.

First, *the permutation can be extracted directly from the periodic orbit without using any symbolic encoding and without having the graph of the map.* Indeed, consider the period-5 orbit in Fig. 2.16. If we label the periodic points x_1, x_2, \dots, x_5 from left to right, we can determine as above a permutation π such that $f(x_i) = x_{\pi_i}$. Obviously, this permutation is identical to 2.88 (the image of the first point is the third, etc.), which can easily be checked in Fig. 2.16.

Second, the permutation 2.88 *remains identical on the entire domain of existence of the orbit in parameter space.* As a parameter is varied, the points x_i of a periodic orbit move along the interval, but they do so without ever becoming degenerate (otherwise, we would have one point with two images, in contradiction with the deterministic nature of the map). Thus the relative order of points is preserved, hence the corresponding permutation.

An important property of this invariance property is that the permutations associated with orbits interacting in a bifurcation will be strongly related. For example, saddle-node partners will have identical permutations since they are indistinguishable at the bifurcation. Similarly, the permutation associated with a period-doubled orbit can easily be obtained from that of its mother.

Since there is a definite relation between symbolic names and permutations on the one hand, and periodic orbits and permutations on the other hand, we see that the symbolic names of the orbits are more than a convenient labeling and that they carry topological information. To illustrate this, we now show that in one-dimensional maps, much about the symbolic name of an orbit can be recovered merely from the permutation. Consider the graphical representation in Fig. 2.18 of the permutation 2.88 extracted from Fig. 2.16.

The global shape of Fig. 2.18 is characteristic of unimodal permutations: The relative order of the leftmost points is preserved, while that of the rightmost points is reversed (in particular, the rightmost point is mapped into the

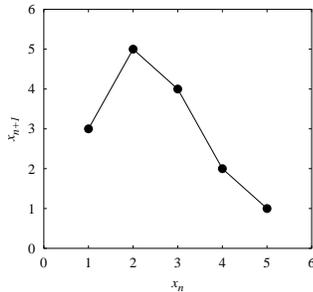


Figure 2.18: Graphical representation (i, π_i) of the permutation 2.88. Orientation-preserving and orientation-reversing parts are easily distinguished.

leftmost). This is a signature of the existence of branches with different parities in the underlying map (assumed unknown). In fact, Fig. 2.18 can be viewed as a topological representation of this map in a coordinate system where the points x_i are equidistant.

Let us note the orientation-preserving (resp., orientation-reversing) branch 0 (resp., 1). It can be seen from Fig. 2.18 that x_1 is necessarily on branch 0, while x_3 , x_4 , and x_5 must be on branch 1. The coding of x_2 is ambiguous: It can be on one branch or the other without modifying the permutation. Taking into account that the orbit of x_1 is $x_1 \rightarrow x_3 \rightarrow x_4 \rightarrow x_2 \rightarrow x_5$, the symbolic name of this orbit is thus $011 \begin{smallmatrix} 0 \\ 1 \end{smallmatrix} 1$ (i.e., any of the two saddle-node partners born in the 5_1 saddle-node bifurcation listed in Table 2.1). If additional information is available, such as parity (is the multiplier of the orbit positive or negative?), it is possible to distinguish between these two orbits.

By generalizing the example above, it is easy to see that every pair of saddle-node partners and every period-doubled orbit is associated with a different permutation. The permutation realized by a periodic orbit can thus be viewed as a *genuine fingerprint* of this orbit.

It is comforting to know that much of the discussion of this section is still relevant for orbits in three-dimensional flows. As discussed throughout this book, orbits in these systems will be associated with braids (a generalization of permutations deeply linked to knot theory). These braids will be characterized by topological invariants that do not depend on parameters and contain much information about the symbolic dynamics and the genealogy of periodic orbits. Just as the structure of unimodal maps governs that of the permutations (see Fig. 2.18), there is a systematic way to study the global organization of braids in three-dimensional systems.

2.6.2 Topological Entropy of a Periodic Orbit

The permutation associated with a periodic orbit not only provides qualitative information, it can also provide estimates of fundamental quantitative measures

of chaotic dynamics, as we show next. The key idea is that if the underlying map is continuous, the way in which the points of the orbit are mapped onto each other provides information on orbits in an extended neighborhood of the orbit.

For simplicity, let us consider a superstable periodic orbit such as the period-5 orbit of Fig. 2.16. As previously, the points are numbered from left to right. Since x_2 is the critical point x_c , the leftmost and rightmost points x_1 and x_5 correspond to the lower and upper bounds of the invariant interval $I = [f(x_c), f^2(x_c)]$ where the relevant dynamics is confined. Using the fact that the x_i are mapped exactly onto each other, this will allow us to build a topological model of the dynamics.

To this end, consider the following partition of the invariant interval I :

$$I = [x_1, x_5] = I_1 \cup I_2 \cup I_3 \cup I_4 = [x_1, x_2] \cup [x_2, x_3] \cup [x_3, x_4] \cup [x_4, x_5] \quad (2.89)$$

Using the interval arithmetics 2.45 and the permutation 2.88, we find easily that

$$f(I_1) = f([x_1, x_2]) = [x_3, x_5] = I_3 \cup I_4 \quad (2.90)$$

$$f(I_2) = f([x_2, x_3]) = [x_4, x_5] = I_4 \quad (2.91)$$

$$f(I_3) = f([x_3, x_4]) = [x_2, x_4] = I_2 \cup I_3 \quad (2.92)$$

$$f(I_4) = f([x_4, x_5]) = [x_1, x_2] = I_1 \quad (2.93)$$

The set of relations 2.6.2 is the analog of the relations $f(I_1) = f(I_2) = I_1 \cup I_2$ which characterize the surjective logistic and tent maps (note that in a complete description of the map, the branch parities should also be specified). The key observation here is that both sets of relations define Markov partitions. In the example above, the Markov transition matrix as defined in (2.77) reads

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \quad (2.94)$$

where the nonzero entries correspond to pairs (I_i, I_j) such that $I_i \cap f(I_j) = I_i$. Even though the regime under study corresponds to a superstable orbit, the matrix (2.94) is a signature of a chaotic dynamics. Its largest eigenvalue is $\lambda_{\max} \sim 1.512876398$, which yields a topological entropy of $h_T = \ln \lambda_{\max} \sim 0.4140127381$. Moreover, the matrix is transitive, which indicates that the associated Markov chain is topologically mixing.

In fact, the topological entropy as computed above characterizes the periodic orbit rather than the dynamical system to which it belongs: It is obtained from the permutation associated to the orbit,⁴ not from the global structure of the

⁴For the sake of simplicity, we have indicated how to compute the topological entropy of an orbit only at a special parameter value, where the orbit is superstable. However, the topological entropy of an orbit depends only on the permutation associated with it and should be considered as a topological invariant of the orbit, defined on its entire domain of existence.

system. However, the entire system cannot be less chaotic than implied by the periodic orbit. In particular, its topological entropy is necessarily greater than the entropy of the orbit. Thus, the observation of an orbit with a positive topological entropy (as obtained from its permutation), even in a window of stability, indicates the presence of chaos in the system under study, and in particular the existence of an infinity of periodic orbits. This is illustrated in the next section with the “period-3 implies chaos” theorem. A similar statement will be made later for flows: Some periodic orbits have knot types that can exist only in a chaotic system.

Note that the tools introduced here show why one-dimensional diffeomorphisms are not chaotic: Since they globally preserve or reverse the order of points, the associated transition matrices cannot have eigenvalues larger than 1. In fact, the reader may want to check that a one-dimensional diffeomorphism can only have a period-1 orbit, possibly a period-2 orbit if it globally reverses orientation.

2.6.3 Period-3 Implies Chaos and Sarkovskii’s Theorem

To illustrate the discussion above, we present now the famous statement “period-3 implies chaos,” according to which the presence in a map of the interval of a periodic orbit of period 3 forces the presence of orbits of any other period [?].

If we carry out the same calculation for the superstable $0 \frac{0}{1} 1$ orbit as for the $011 \frac{0}{1} 1$ orbit in Section 2.6.2, we find a partition $I = [f^2(x_c), x_c] \cup [x_c, f(x_c)]$, whose transition matrix is the golden mean matrix (2.70) that we have already encountered:

$$A_{\text{GM}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad (2.95)$$

It is easy to prove that $\forall n \geq 2$, $A_{ii}^n \neq 0$, hence there are fixed points for any period p . This is our first example of the general fact that the existence of some periodic orbits can force the existence of many other (here, an infinity) periodic orbits. This phenomenon is usually referred to as *orbit forcing*.

Note that the Markov partitions constructed from superstable orbits are refinements of the partition used for symbolic codings, since the critical point is one of the border points. As a result, the transition matrices contain all the information needed to determine whether a given symbolic name is admissible. The example of the period-3 orbit is particularly simple since the Markov partition coincides with the coding partition: The transition matrix 2.95 indicates that all itineraries are allowed except those containing the string “00” (showing again that orbits of all periods exist). One may check in Table 2.1 that this is indeed what distinguishes orbits born before the 3_1 saddle-node bifurcation from orbits created after.

That the existence of an orbit of period 3 implies the existence of orbits of any other period can in fact be viewed as a particular case of a more general theorem due to Sarkovskii [?] (see also [?]). Consider the following ordering of

the natural integers, written as the product $2^k \times (2n + 1)$ of a power of 2 by a prime number:

$$2^k \times 1 \quad 2^l \times 1 \quad (k < l) \quad (2.96)$$

$$2^k \times 1 \quad 2^l \times (2n + 1) \quad (n > 0) \quad (2.97)$$

$$2^l \times (2n + 1) \quad 2^k \times (2n + 1) \quad (k < l, n > 0) \quad (2.98)$$

$$2^k \times (2n + 1) \quad 2^k \times (2m + 1) \quad (m < n) \quad (2.99)$$

For example,

$$122^2 2^3 2^4 \times 72^4 \times 32^2 \times 52^2 \times 32 \times 3753$$

Sarkovskii's theorem states that if a continuous map of an interval into itself has an orbit of lowest period p and qp , it also has an orbit of lowest period q . It is easily seen that (1) if there are infinitely many different periods in a map, all the periods corresponding to the period-doubling cascade of the period-1 orbit must be present; (2) since 3 comes last in 2.6.3, the presence of a period-3 orbit forces orbits of all other periods as shown above. It can be checked that the succession of bifurcations given in Table 2.1 satisfies the Sarkovskii theorem. In particular, the first period-7 orbit is created before the first period-5 orbit, which itself is created before the period-3 orbit. Moreover, all even periods are present when the first odd period appears.

2.6.4 Period-3 Does Not Always Imply Chaos: Role of Phase-Space Topology

Since so many properties of unimodal maps hold at least approximately for higher-dimensional systems, it might be troubling that numerous apparent counterexamples to the “period-3 implies chaos” theorem can be found in physical systems. In the modulated CO₂ lasers described in Chapter 1 and in Section 7.5.2, for example, it is quite common to observe multistability, with a period-3 orbit coexisting with the initial period-1 orbit, no other periodic orbit being present. Thus, a period-3 orbit does not necessarily imply chaos.

The clue to this paradox is that the modulated CO₂ laser can be described by a three-dimensional flow, hence by a two-dimensional Poincaré map, while the Sarkovskii theorem holds for a map of a one-dimensional interval into itself. It turns out that phase-space topology has a dramatic influence on orbit forcing.

The key topological difference between the two geometries is that in the two-dimensional case, it is possible to connect each of the three periodic points of the period-3 orbit to any other without encountering the third one. In the one-dimensional case, the two extreme points are isolated by the middle one.

Therefore, the three periodic points must be considered as arranged along a topological circle (obviously, this also applies to maps of a circle into itself; see Section ??). The three points divide this circle into three intervals versus two in the one-dimensional case. When the map is applied, the three points are cyclically permuted and the three intervals accordingly exchanged. As a result, the associated transition matrix is a simple permutation matrix, with zero

topological entropy. The action of the map on the periodic orbit is equivalent to a pure rotation, which does not itself imply the existence of a chaotic dynamics. However, although it appears that orbit forcing can be modified dramatically when the topology of phase space is changed, it remains true in all cases that some orbits can force the existence of an infinity of other orbits.

2.6.5 Permutations and Orbit Forcing

The statements of the Li-Yorke theorem and of the Sarkovski theorem seem to imply that period is the fundamental property in implication chains among orbits. However, forcing relations are more fine-grained. Admittedly, the existence of a period-5 orbit forces the existence of at least a period-7 orbit. However, an orbit of the period-5 pair 5_1 of the logistic map does not force every period-7 orbit of this map. Actually, it forces the two 7_1 orbits *but is forced* by the two 7_2 orbits. As we see below, it is in fact the permutation associated with a given orbit that determines which orbits it forces.

In Section 2.6.1, we noted that the permutation associated with a periodic orbit \mathcal{O} can be represented graphically by a piecewise-linear map such as shown in Fig. 2.18. This map is conjugate to the Markov chain that describes the action of the map on the periodic points of the orbit (see Section 2.6.2). If the Markov chain has positive topological entropy, it is chaotic and so is the piecewise-linear map which has thus an infinite number of unstable periodic orbits. These orbits comprise the minimal set of orbits that must exist in a map having the original orbit \mathcal{O} as one of its periodic orbits. The existence of these orbits is forced by that of \mathcal{O} .

As an example, we show that the period-5 orbit 01011 (born in bifurcation 7_1) forces exactly one pair of period-7 orbits, and determine the symbolic code of this orbit. To this end, we apply the techniques outlined in Section 2.5.6 to the transition matrix 2.94. Using the non-commutative symbols I_1, I_2, I_3 and I_4 , we first rewrite it as

$$\mathcal{A} = \begin{bmatrix} 0 & 0 & 0 & I_1 \\ 0 & 0 & I_2 & 0 \\ I_3 & 0 & I_3 & 0 \\ I_4 & I_4 & 0 & 0 \end{bmatrix} \quad (2.100)$$

Then we compute

$$\text{tr } \mathcal{A}^7 = [I_1 I_3 I_3 I_2 I_4 I_1 I_4] + [I_1 I_3 I_3 I_3 I_2 I_4] + (I_3)^7 \quad (2.101)$$

where $[W]$ represents the n cyclic permutations of a length- n word W . The minimal model compatible with the period-5 orbit 01011 has thus 2 period-7 orbits, whose itineraries in the Markov partition can be read from 2.101. Since the interval I_1 is located left of the critical point (Sec. 2.6.2) and the intervals I_2, \dots, I_4 right of it, the canonical symbolic names of these orbits can be obtained through the recoding $I_1 \rightarrow 0, \{I_2, I_3, I_4\} \rightarrow 1$. The two period-7 orbits forced by the period-5 orbit under study are thus 0111101 and 0111101.

The reader can check in Table 2.1 that these are the two orbits born in the saddle-node bifurcation 7_1 and that this bifurcation indeed occurs before 5_1 in the universal sequence. By repeating the above calculation for all low-period orbits of the logistic map, the whole universal sequence of Table 2.1 can be reproduced.

It is not a coincidence that the two period-7 orbits forced in the example are saddle-node partners and thus are associated with the same permutation. As was suggested by the simple example above, orbit forcing in one-dimensional maps is naturally expressed as an order on permutations rather than on periods. This is illustrated in Section 9.3.1. In unimodal maps, each bifurcation corresponds to a different permutation (Section 2.6.1) so that this order on permutations induces a total order on bifurcations, which corresponds to the universal sequence discussed in Section 2.4.5.

As is discussed in Chapter 9, there is no longer a total ordering of bifurcations in two-dimensional maps (and three-dimensional flows). The topological structure of a periodic orbit then is not specified by a permutation but by a braid type. There is a forcing relation on braid types. However, several saddle-node bifurcations can be associated with the same braid type, so that the induced order on bifurcations is only a partial order.