8

Chaos (A Short Introduction)

In addition to simple attractors such as equilibrium and fixed points, and closed periodic orbits, a dynamical system may have “strange”, “chaotic” attractors that exhibit a much more complex shape. The system is said to exhibit chaos if, loosely speaking, its solutions are bounded, are asymptotically irregular and aperiodic, with a strong dependence on the initial conditions. In this last chapter, we briefly and superficially survey some of the main features of a chaotic system, the most well known examples of such system being the Lorenz system (in continuous time), the logistic, Bernouilli (doubling) and tent maps (in discrete time).

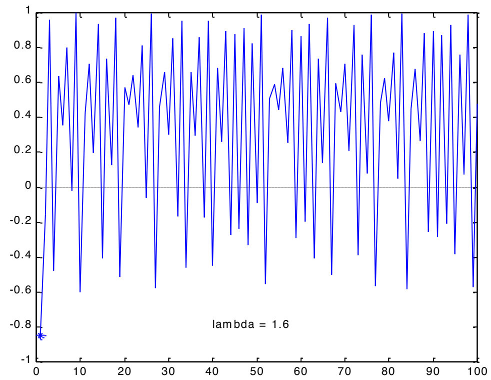

Figure 8.1: 200 iterations of the logistic map (8.1) with $\lambda = 1.6$, starting at $x(0) = 0.86$.

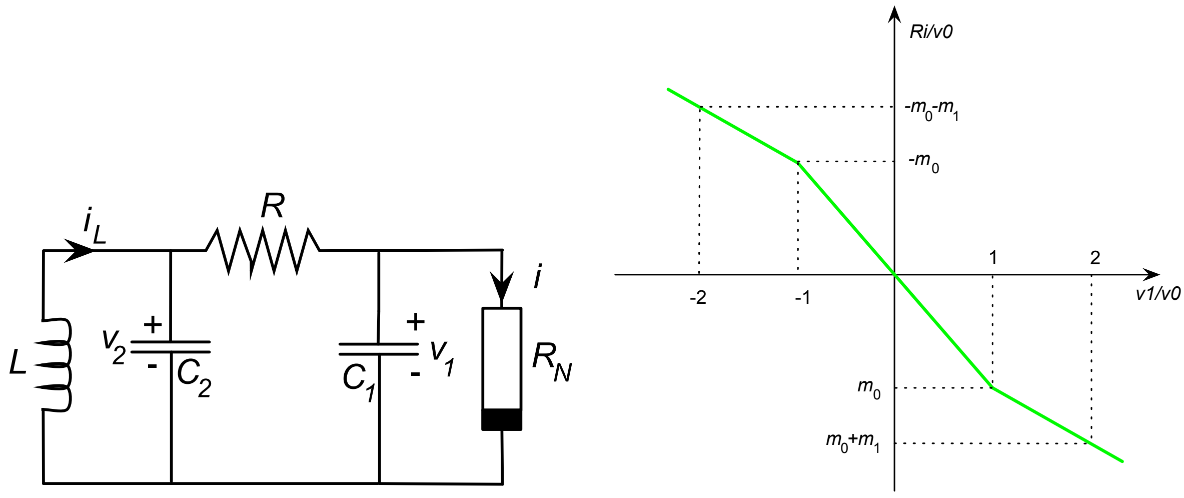

Figure 8.2: Chua's circuit (left) with the characteristic of the nonlinear resistor R_N (right).

8.1 First Property: Irregular and Aperiodic Trajectories

At first sight, this is the most striking property of chaos. A system with very simple state equations can have very complicated trajectories is the logistic map $F(x) = 1 - \lambda x^2$ whose iterations

$$x(t+1) = 1 - \lambda x^2(t) \tag{8.1}$$

for $\lambda = 1.6$ are shown in Figure 8.1.

Autonomous continuous time systems can also have irregular, non-periodic asymptotic behavior. Because of Poincaré-Bendixson's theorem, 2-dimensional continuous-time systems can have as asymptotic behavior only constant (equilibrium points) and periodic solutions, the simplest continuous-time chaotic systems have dimension 3. The most famous example is the Lorenz system (Lorenz, 1963) originating from atmospheric physics, and another famous example is Chua's circuit, formulated in terms of electronics. Let us consider the latter. The circuit diagram is drawn in Figure 8.2. All the components of this circuit are linear, except the nonlinear resistor R_N whose characteristic $i = g(v_1)$ given on the right of Figure 8.2.

The state equations of this circuit are

$$\begin{aligned} C_1 \frac{dv_1}{dt} &= \frac{1}{R}(v_2 - v_1) - g(v_1) \\ C_2 \frac{dv_2}{dt} &= -\frac{1}{R}(v_2 - v_1) + i_L \\ C_2 \frac{di_L}{dt} &= -v_2. \end{aligned}$$

Normalize the currents and voltages as $x_1 = v_1/v_0$, $x_2 = v_2/v_0$, $x_3 = Ri_L/v_0$, where v_0 is the voltage where the characteristic of the nonlinear resistor has a breakpoint, the system of equations becomes

$$\dot{x}_1 = \alpha(-x_1 - f(x_1) + x_2) \tag{8.2}$$

$$\dot{x}_2 = x_1 - x_2 + x_3 \tag{8.3}$$

$$\dot{x}_3 = -\beta x_2. \tag{8.4}$$

where $\alpha = C_2/C_1$, $\beta = R^2 C_2/L$ and the piecewise characteristic $g(\cdot)$ of the nonlinear resistor becomes

$$f(x) = \begin{cases} m_1 x - m_0 + m_1 & \text{if} \quad x < -1 \\ m_0 x & \text{if} \quad -1 \le x \le 1 \\ m_1 x + m_0 - m_1 & \text{if} \quad x > 1, \end{cases}$$

where $m_0 < m_1 < 0$. For the standard choice $\alpha = 9$, $\beta = 100/7$, $m_0 = -8/7$, $m_0 = -5/7$, the trajectories are irregular, non-periodic, as shown in Figure 8.3, although nothing in the state equations indicates the presence of such complex solutions at first sight.

Chaotic systems have irregular, aperiodic solutions, but such solutions are not an unmistakable sign of chaos. The signal

$$y(t) = \sin(2t) + 0.7\cos(2\pi t) + 1.3\sin(\sqrt{2}t) \tag{8.5}$$

is shown in Figure 8.4. It looks irregular, and it is not periodic, since the frequencies in the sin and cos functions do not have rational ratios. Such signals are called quasi-periodic. The signal (8.5) is

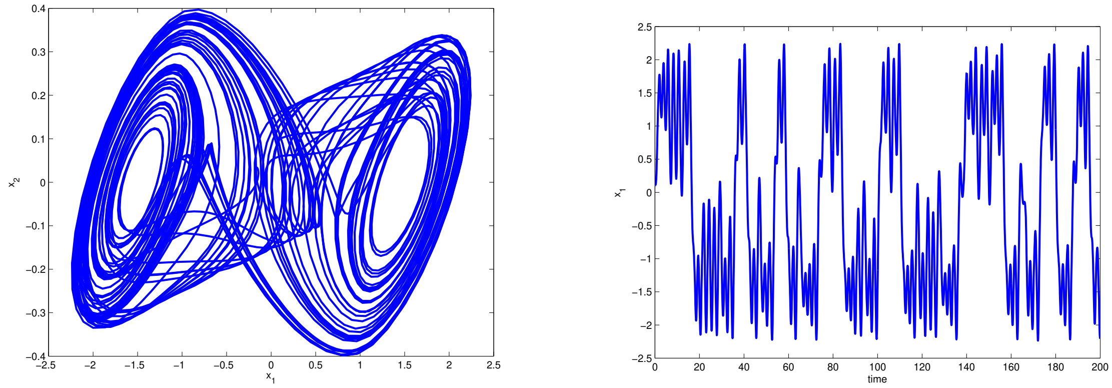

Figure 8.3: Projection of the orbit of Chua's circuit with the parameters $\alpha = 9$, $\beta = 100/7$, $m_0 = -8/7$, $m_0 = -5/7$ onto the (x_1, x_2)-plane. The time interval represented is $[0, 200]$ and the initial state is $(x_1(0), x_2(0), x_3(0)) = (0.1, 0.1, 0.1)$ (left). The trajectory of the state component $x_1(t)$ as a function of time t (right).

the output signal of the linear autonomous continuous-time system

$$\begin{aligned}
\dot{x}_1 &= 2x_2 \\
\dot{x}_2 &= -2x_1 \\
\dot{x}_3 &= 2\pi x_4 \\
\dot{x}_4 &= -2\pi x_3 \\
\dot{x}_5 &= \sqrt{2}x_6 \\
\dot{x}_6 &= -\sqrt{2}x_5 \\
y &= x_1 + x_3 + x_5.
\end{aligned}$$

This system has its eigenfrequencies on the imaginary axis (they are $\pm 2j, \pm 2\pi j, \pm\sqrt{2}j$) and therefore it is stable (without being asymptotically stable). Thus, it clearly would not be reasonable to classify this system as being chaotic.

8.2 Second Property: Sensitivity to initial Conditions

8.2.1 Introduction and Observations

The property that characterizes chaos best is the so-called sensitivity to initial conditions, together with the property that the solutions are bounded. In its weakest form, this simply means that all solutions are unstable. Therefore, loosely speaking, any two solutions that start close together will eventually separate from each other. In a stronger form, it is required that this separation is exponentially fast, at least as long as they are close.

Unstable linear systems have also sensitivity to initial conditions, but almost all their solutions are unbounded. In general, nonlinear systems with bounded solutions may have simultaneously stable and unstable solutions, the unstable ones being the exception. As an example, take a 2-dimensional continuous-time system with two asymptotically stable equilibrium points, which are (topologically equivalent to) stable focus, and one unstable equilibrium point, which is (topologically equivalent to)

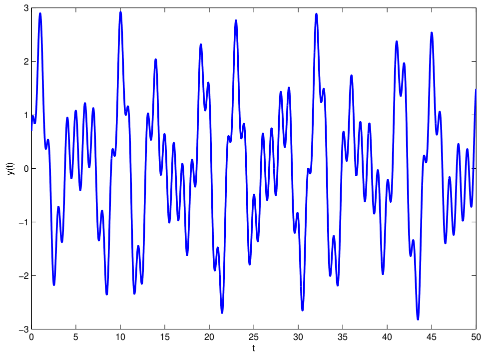

Figure 8.4: Quasi periodic signal given by (8.5) in the time interval $[0, 50]$ starting from the initial state $(0, 1, 0.7, 0, 0, -1.3)$.

a saddle point. The basins of attraction of the stable focuses are separated by the stable manifold of the saddle point. All solutions are asymptotically stable, except those that are starting on the stable manifold of the unstable equilibrium point (a saddle point). They converge to the saddle point, but an arbitrarily small perturbation off the stable manifold will cause such a solution to converge to either of the stable equilibrium points and thus become asymptotically stable.

It is not difficult to imagine that when all solutions are repelling each other and when they nevertheless stay in a bounded region of the state space (or at least those that start in a bounded region) they must have a very disordered behavior (the first property of chaos).

The sensitivity to initial conditions is illustrated in Figure 8.5 for the iterations of the logistic map (8.1) and in Figure 8.6 for Chua's circuit. Figure 8.5 shows for the exponentially fast separation of trajectories that start from close initial conditions. Two trajectories of the logistic map system starting close drift exponentially fast apart. When the initial separation is 10^{-3} and it takes about 10 iterations until the separation becomes visible, whereas if the initial separation is 10^{-6}, and it takes about 30 iterations until the two trajectories can be distinguished in the plot.

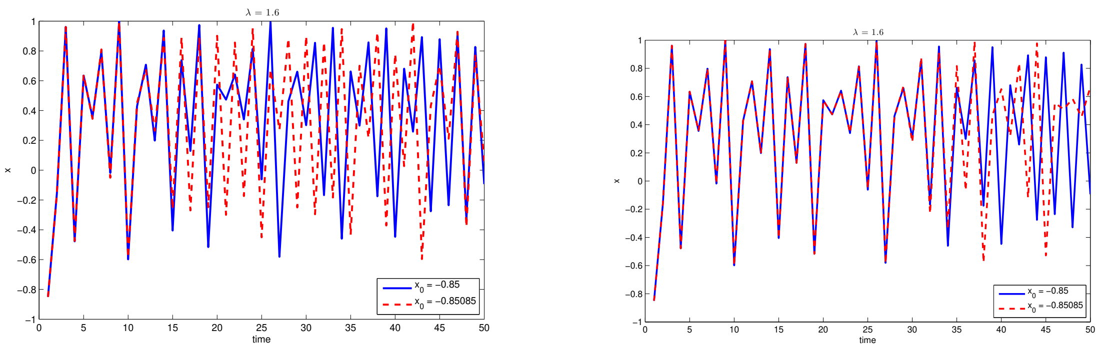

Figure 8.5: Trajectories of the logistic map starting at -0.85 (blue) and -0.85085 (red) at time $t = 0$ (left). Within the resolution of this figure, they are indistinguishable until about time $t = 15$. Trajectories starting at -0.85 (blue) and -0.85000085 (red) at time $t = 0$ (left). Within the resolution of this figure, they are indistinguishable until about time $t = 35$. If the two trajectories were continued in time, they would become less and less correlated.

While the attractor of the logistic map system in the case of chaotic behavior covers part of the interval $[-1, 1]$, in the case of Chua's circuit it is a complex lower dimensional geometric object, whose dimension is strictly less than 3, and that can be seen in a 2-dimensional projection in Figure 8.6 to a certain approximation. This attractor governs the asymptotic behavior of the solutions. There are infinitely many solutions in the attractor. Even if their asymptotic behavior is not unique, the long time behavior of the solutions is nevertheless similar, as seen in Figure 8.6, showing two solutions starting at two initial states whose first component is different by 10^{-3}. They separate rapidly, but their general aspect is similar.

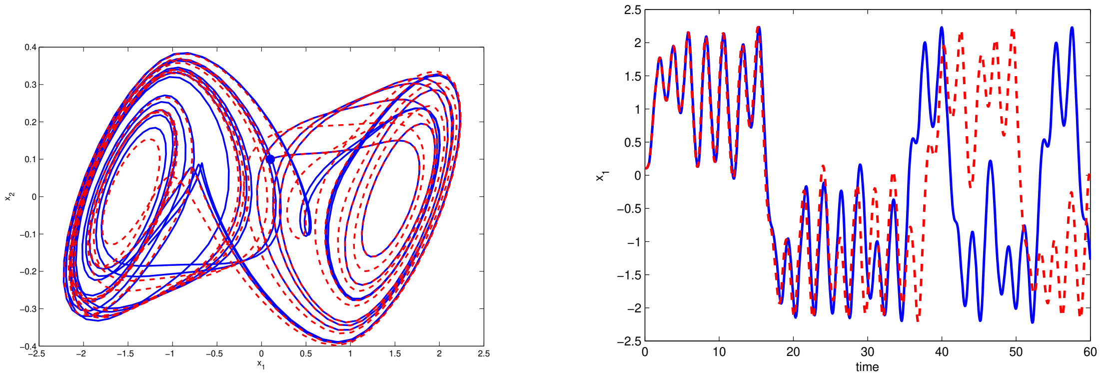

Figure 8.6: Projection of two orbits of Chua's circuit with standard parameters, starting from the initial state $(0.1, 0.1, 0.1)$ (blue) and $(0.1001, 0.1, 0.1)$ (red). For a short time, the two orbits cannot be distinguished in the figure, but then they become different. The general nature of both orbits is however the same (left). The two trajectories of the state component $x_1(t)$ as a function of time t (right).

8.2.2 Lyapunov Exponents for 1-dim. Maps

Since sensitivity to initial conditions is such an important property, is it possible to quantify it? Yes, it is possible to compute the exponential speed of separation of two nearby solutions, as long as they remain close and the linear approximation of the time evolution of the differences is valid. This naturally recalls the use of the variational equations that were successful in deciding the stability or instability of periodic solutions. In fact, exactly the same formalism can be applied to arbitrary solutions, not only periodic ones.

Definition and Computation

We discuss here only 1-dimensional discrete-time systems

$$x(t + 1) = F(x(t)) \tag{8.6}$$

where $F : \mathbb{R} \to \mathbb{R}$ is continuously differentiable.

Let $x(t)$ be an arbitrary solution, and $\tilde{x}(t)$ be another solution such that $|\Delta x(0)| = |\tilde{x}(0) - x(0)|$ is small. Then the time evolution of the increment $\Delta x(t) = \tilde{x}(t) - x(t) = \Phi(t, \tilde{x}(0)) - \Phi(t, x(0))$ is, up to first order approximation, given by

$$\Delta x(t) = M(t)\Delta x(0) \tag{8.7}$$

where $M(t)$ is the Jacobian matrix of Φ with respect to x_0 at the point $x(0)$ and for a given time t, which is here

$$M(t) = \frac{d\Phi}{dx_0}(t, x(0))$$

and is given by the solution of the variational equation

$$M(t+1) = \frac{dF}{dx}(x(t))M(t) \tag{8.8}$$

with $M(0) = 1$, since here the Jacobian matrix $J(x(t))$ of F at $x(t)$ is given by

$$J\left(x(t)\right) = \frac{dF}{dx}(x(t)).$$

Iterating (8.15), we find that

$$M(t) = \frac{dF}{dx}(x(t-1)) \cdot \frac{dF}{dx}(x(t-2)) \cdots \frac{dF}{dx}(x(1)) \cdot \frac{dF}{dx}(x(0)),$$

and therefore the relative growth or shrinking of the increment is given, to a first order approximation, by

$$\frac{|\Delta x(t)|}{|\Delta x(0)|} \approx \left|\frac{dF}{dx}(x(t-1))\right| \cdot \left|\frac{dF}{dx}(x(t-2))\right| \cdots \left|\frac{dF}{dx}(x(1))\right| \cdot \left|\frac{dF}{dx}(x(0))\right|.$$

We are here interested in the exponential speed or convergence of the two solutions and therefore we set

$$\frac{|\Delta x(t)|}{|\Delta x(0)|} \approx \exp(\alpha(t)t)$$

and therefore

$$\alpha(t) \approx \frac{1}{t}\sum_{\tau=0}^{t-1} \ln\left|\frac{dF}{dx}(x(\tau))\right|$$

represents the (time-)average exponential speed of growth or contraction in the time interval $[0, t]$ along the solution $x(t)$. It can be calculated along any solution $x(t)$, and its limit for $t \to \infty$, if it exists, is the Lyapunov exponent of $x(t)$.

Definition 8.1. *The* Lyapunov exponent *of a solution $x(t)$ of the autonomous 1-dim discrete-time system (8.6) is given by*

$$\alpha = \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \ln\left|\frac{dF}{dx}(x(\tau))\right| \tag{8.9}$$

if the limit exists.

To compute the limit (8.9), it is useful to use probabilistic methods and tools from ergodic theory. Instead of attempting to make a prediction of the long-term behavior of a solution from a deterministically chosen initial condition, which is anyway impossible in a chaotic system when the initial data have limited accuracy, we predict the average long-term behavior of a solution from a randomly chosen initial condition. Instead of a single trajectory, all trajectories are considered simultaneously, weighted with a probability measure P. If the system is ergodic, we can then approximate the long term behavior of a typical solution by any single solution with asymptotically good accuracy. The main ingredients of ergodic theory of dynamical system are summarized in Section 8.4, a more detailed treatment goes however beyond the introductory treatment of this chapter, and the interested reader is referred to e.g. the textbook by M. Pollicott and M. Yuri, *Dynamical systems and ergodic theory*, Cambridge University Press, 1998.

The following theorem, which is an application of Birkhoff's ergodic theorem (Theorem 8.5), states that the limit in (8.9) exists and its value if P is an invariant measure under $F(\cdot)$ (which means that $P(F^{-1}(A)) = P(A)$ for any subset $A \subseteq \Sigma$, where Σ is a σ-algebra on Ω, where $F^{-1}(\cdot)$ is the inverse transformation of $F(\cdot)$ and where $\Omega \subseteq \mathbb{R}$ is the state space of the system; see Section 8.4. A subset A of the σ-algebra of subsets of Ω corresponds to an event in probabilistic terms). If in addition $F(\cdot)$ is an ergodic transformation, then the limit does not depend on the solution $x(t)$ used in (8.9).

Theorem 8.1. *If P is an invariant measure under $F(\cdot)$, then for P-almost all solutions $x(t)$, the Lyapunov exponent (8.9) exists. If, in addition, P is ergodic with respect to $F(\cdot)$, then for P-almost all solutions, the Lyapunov exponent (8.9) is the same and its value is given by*

$$\alpha = \int_{-\infty}^{\infty} \ln \left| \frac{dF}{dx}(x) \right| dP(x). \tag{8.10}$$

The proof follows by applying Theorem 8.5 for $\Omega = \mathbb{R}$ and $f = \ln |dF/dx|$.

If the ergodic invariant measure is given by a density $\rho(x)$, i.e. $dP(x) = \rho(x)dx$, then (8.10) becomes

$$\alpha = \int_{-\infty}^{\infty} \ln \left| \frac{dF}{dx}(x) \right| \rho(x)dx. \tag{8.11}$$

This means that if we choose an initial condition at random on the real line (or in an interval Ω, if the system and the invariant measure is restricted to an interval Ω), we will obtain, by calculating the right hand side of (8.11) for sufficiently large t, a good approximation of the Lyapunov exponent. Since determining an invariant measure most of the time cannot be done explicitly, this is in fact the way Lyapunov exponents are computed.

Clearly, a positive Lyapunov exponent of a solution $x(t)$ implies that the solution is unstable, whereas a negative one implies that the solution is asymptotically stable. However, a solution with a 0 Lyapunov exponent can be unstable, stable or even asymptotically stable. In fact, the Lyapunov exponent distinguishes only exponential speeds of expansion or contraction, but not anything slower.

Examples of Lyapunov Exponents

Fixed Point. When the solution $x(t)$ converges to a fixed point $\overline{x}$, then (8.9) implies that its Lyapunov exponent equals

$$\alpha = \ln \left| \frac{dF}{dx}(\overline{x}) \right| . \tag{8.12}$$

This implies that all solutions starting in the basin of attraction of an asymptotically stable fixed point have the same (negative) Lyapunov exponent.

Periodic Solution. Similarly, when the solution $x(t)$ converges to the T-periodic solution $\xi = (\xi_1, \xi_2, \ldots, \xi_T)$ then (8.9) becomes

$$\alpha = \frac{1}{T} \sum_{i=1}^{T} \ln \left| \frac{dF}{dx}(\xi_i) \right| . \tag{8.13}$$

Again, if the periodic solution is asymptotically stable, then all solutions starting in the basin of attraction of this solution, as well as a cyclic permutation of it, have the same (negative) Lyapunov exponent.

These two examples show that the notion of Lyapunov exponent is not really interesting for fixed points and periodic solutions, since it coincides with the notion of eigenvalue of the Jacobian or a product of Jacobians (here just numbers). However, it is a good tool to distinguish between chaotic and non-chaotic behavior, as shown by the following examples.

Bernouilli Map. The Bernouilli map $F(\cdot)$ on $\Omega = [0,1[= [0,1)$ is defined by

$$F(x) = 2x \mod 1 = \begin{cases} 2x & \text{if} \quad 0 \leq x < 1/2 \\ 2x-1 & \text{if} \quad 1/2 \leq x < 1. \end{cases} \tag{8.14}$$

Let P be the measure specified by the intervals, such that for any $0 \leq a \leq b \leq 1$, $P([a,b]) = P([a,b)) = P((a,b]) = P((a,b)) = b-a$ (hence $dP(x) = dx$: this is the Lebesgue measure). It is invariant under $F(\cdot)$ since the total length of the two intervals of the pre-image of $[a,b]$ is $(b-a)$ (see Section 8.4). The function (8.14) has a discontinuity at $x = 1/2$, but (Lebesgue-) almost all trajectories do not pass through this point.

Therefore the Lyapunov exponent of almost all trajectories exists, and since $dF/dx(x) = 2$ for all $x \in \Omega \setminus \{1/2\}$, the Lyapunov exponent $\alpha = \ln 2$ is for almost all trajectories. As this is a positive number, the behavior of the system is chaotic because the state space is compact (all solutions remain bounded for all time $t \in \mathbb{N}$).

Logistic Map.

Figure 8.7 (left) displays the Lyapunov exponent $\alpha(t)$ of a few solutions of the logistic map (8.1) for $\lambda = 0.9$, where there is an asymptotically stable period-2 solution (to be precise, there are two of them). The theoretical (negative) value for the solutions converging to the 2-periodic solution according to (8.13) is indicated by a horizontal black line. The only exception is the solution starting exactly at the unstable fixed point (or at its other pre-image), which is indicated by the horizontal red line. It remains constant at a positive value, as it should be. How is this compatible with Theorem 8.1? The fact is that the two exceptional initial points, i.e. the unstable fixed point and its pre-image, have Lebesgue measure 0 and thus for Lebesgue-almost all initial points in the interval $[-1,1]$, we get the same value for the Lyapunov exponent, namely the value corresponding to the period-2 solution. This is the case for the two other solutions shown in the figure, which nicely converge to this value, with respective initial conditions at $x(0) = 0.37$ (blue) and $x(0) = \overline{x} - 0.0001 \simeq 0.548903$, where $\overline{x} = (-1+\sqrt{1+4\lambda})/(2\lambda)$ is the (unstable) fixed point of the map for $\lambda = 1.6$ (Magenta). Since the initial point is so close to the unstable fixed point, but that both have different Lyapunov exponents, the convergence to the Lyapunov exponent of the 2-periodic solution is longer.

Figure 8.7 (right) displays the Lyapunov exponent $\alpha(t)$ of two solutions of the logistic map (8.1) for $\lambda = 1.6$, one starting at $x(0) = -0.8$ and the other at $x(0) = 0.2$. They appear to converge to the same positive asymptotic value. One could choose many other initial points at random to obtain the same positive asymptotic value. This is an indication of chaos. Note that the two functions are not very smooth. This is also typical of chaos. It comes from the fact that within the chaotic attractor there are infinitely many unstable periodic solutions, as we will see in Section 8.3, each one with a somewhat different Lyapunov exponent. Even though the set of initial points that lead to periodic solutions has Lebesgue measure 0, these solutions are everywhere present and perturb the numerical computation of the Lyapunov exponent of the chaotic solutions.

8.2.3 Lyapunov Exponents for Higher Dimensional Systems

We can generalize the notion of Lyapunov exponent to higher dimensional discrete-time and to continuous-time systems using the same basic ideas and constructions. We again consider two close solutions $\tilde{x}(t)$ and $x(t)$ of the dynamical system, and we define the increments $\Delta x(t) = \tilde{x}(t) - x(t)$. Up to first order approximation, the time evolution of the increments is given by the variational equations

$$M(t+1) = J\left(x(t)\right) M(t) \tag{8.15}$$

for a discrete-time map and by

$$\dot{M}(t) = J\left(x(t)\right) M(t) \tag{8.16}$$

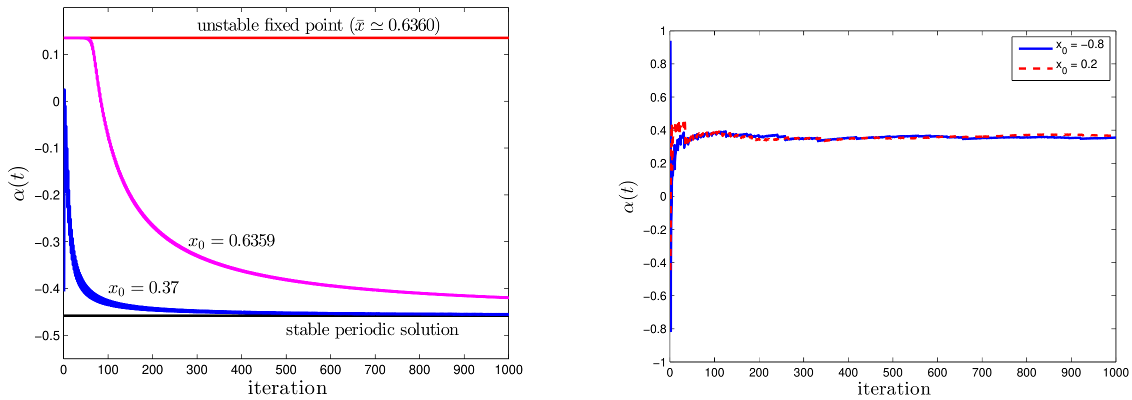

Figure 8.7: Computation of the Lyapunov exponent $\alpha(t)$ when $\lambda = 0.9$ (left) and $\lambda = 1.6$ (right).

for a continuous-time dynamical system, with $J(x(t)) = \partial F/\partial x(x(t))$ and $M(0) = I_n$ in both cases. One has then

$$\Delta x(t) \approx M(t)\Delta x(0). \tag{8.17}$$

Therefore, the expansion or contraction of the initial increment vector is approximated by

$$\frac{\|\Delta x(t)\|^2}{\|\Delta x(0)\|^2} = \frac{\Delta x^T(t)\Delta x(t)}{\Delta x^T(0)\Delta x(0)} = \frac{\Delta x^T(0)M^T(t)M(t)\Delta x(0)}{\Delta x^T(0)\Delta x(0)}.$$

and is thus dictated by the matrix $M^T(t)M(t)$, which describes different exponential expansion/contraction in different directions in state space. Since $M^T(t)M(t)$ is a positive definite matrix, the matrix

$$\Lambda(t) = \frac{1}{2t}\ln\left(M^T(t)M(t)\right)$$

is well defined. Taking its limit for $t \to \infty$, if it exists, gives the Lyapunov exponents of the solution $x(t)$.

Definition 8.2. *The* Lyapunov exponents *of a solution $x(t)$ of an autonomous n-dim system are the eigenvalues of the matrix*

$$\Lambda = \lim_{t\to\infty}\frac{1}{2t}\ln\left(M^T(t)M(t)\right) \tag{8.18}$$

if the limit exists, where $M(t)$ is the solution of the variational equations (8.15) or (8.16).

For a 1-dim system, the matrix (8.18) becomes the scalar (8.9); the factor 2 in (8.18) comes from the squares used in the above expressions. In contrast, the expansion or contraction of the initial increment leading to (8.9) was the absolute value of a scalar (and not its square).

One can similarly generalize Theorem 8.1, using Oseledec's ergodic theorem, which is itself an extension of Birkhoff's ergodic theorem. The only part of Theorem 8.1 that does not generalize is (8.10)

Theorem 8.2. *If P is an invariant measure with respect to a dynamical systems, then for P-almost all solutions $x(t)$, the Lyapunov exponents given by the eigenvalues of (8.18) exist. If, in addition, P is ergodic, then for P-almost all solutions, the Lyapunov exponents are the same.*

In general, invariant measures cannot be determined explicitly in most dynamical systems. Therefore, one chooses a solution at random and one computes the Lyapunov exponents by applying Definition 8.2. The presence of at least one positive Lyapunov exponent indicates the presence of chaos.

In the case of Chua's circuit with the standard parameters, one obtains $\alpha_1 = 0.23$, $\alpha_2 = 0$ and $\alpha_3 = -1.75$.

Here, the positive value of α_1 indicates chaos.

One should note that continuous systems with bounded solutions have always a zero Lyapunov exponent, which corresponds to increments in the direction of the solution. Indeed let $\tilde{x}(t) = x(t+\tau)$ for some small $\tau > 0$. Then

$$\begin{aligned}\Delta x(0) &= x(\tau) - x(0) \approx \frac{dx}{dt}(0) \cdot \tau = F(x(0)) \cdot \tau \\ \Delta x(t) &= x(t+\tau) - x(t) \approx \frac{dx}{dt}(t) \cdot \tau = F(x(t)) \cdot \tau\end{aligned}$$

The solution $x(t)$ being bounded, so is $F(x(t))$ and therefore, this increment neither expands exponentially nor contracts exponentially, yielding a zero Lyapunov exponent, which is here α_2.

Finally, it can be shown that volumes in state space contract/expand with exponential speed $\alpha_1 + \alpha_2 + \ldots + \alpha_n$. In the case of Chua's circuit with standard parameters this sum is $\alpha_1 + \alpha_2 + \alpha_3 = -1.52$, which indicates an exponentially fast contraction of volumes. Therefore, the attractor cannot occupy a positive volume in the state space $\mathbb{R}^3$. In fact, it is a complicated lower dimensional geometric object. The 2-dimensional projection of a typical trajectory, as represented on the left part of Figure 8.3, gives only a poor glimpse of the "thin" and complicated geometry, full of layers and holes of the attractor in $\mathbb{R}^3$. Such attractors are loosely called *strange attractors*.

The computation of Lyapunov exponents is a challenge for numerical analysis. On the one hand, convergence is slow because of the perturbation by unstable periodic solutions, and on the other hand, $M(t)$ in (8.18) either grows or shrinks exponentially fast and becomes singular. There is a considerable literature on this subject.

8.3 Third Property: Presence of a Dense Set of (Unstable) Periodic Solutions

8.3.1 The Period-Doubling Road to Chaos

The asymptotic behavior of the iterations of the logistic map (8.1) as a function of the parameter λ is visualized in the bifurcation diagram of Figure 8.8. For each value of λ between 0.5 and 2, the solution $x(t)$ is computed from suitable initial states $x(0)$. After a number of iterations such that the transient effects have died out, 200 points are represented vertically. Thus, an asymptotically stable T-periodic solution will be represented by the T points of its orbit. One can clearly see the flip bifurcation at $\lambda = 0.75$ where the fixed point becomes unstable and an asymptotically stable 2-periodic solution is born; another flip bifurcation at $\lambda = 1.25$ where the 2-periodic bifurcation becomes unstable and an asymptotically stable 4-periodic solution is born, etc.

This process of flip bifurcation continues indefinitely, and the values of the parameter λ at the bifurcation points are (the values λ_n for $n \geq 3$ must be computed numerically)

$$\lambda_1 = 0.75, \lambda_2 = 1.25, \lambda_3 \approx 1.368099, \lambda_4 \approx 1.394050, \lambda_5 \approx 1.399631, \ldots,$$

Observe that as this sequence $\{\lambda_n\}$ converges to a specific value $\lambda_\infty \approx 1.401155$. Moreover, we see that the sequence

$$\frac{\lambda_n - \lambda_{n-1}}{\lambda_{n+1} - \lambda_n}$$

also converges to a specific value δ, which is known as the Feigenbaum constant

$$\lim_{n\to\infty} \frac{\lambda_n - \lambda_{n-1}}{\lambda_{n+1} - \lambda_n} = \delta \approx 4.669202. \tag{8.19}$$

Indeed this constant δ is universal because it does not depend on the particular map, but it is the same for a large class of maps containing unimodal maps on an interval $\Omega = [a, b]$ (which are maps such that $F(a) = F(b) = 0$ and with a unique critical point between a and b).

Period doubling has been empirically observed in real systems (temperature of a box of liquid heated from below, with a transition between motionless liquid, heated only by conduction, and liquid subject to convection with different variation speed following a period doubling cascade of bifurcations). The number extracted from the sequence of period doubling bifurcations was in good agreement with the Feigenbaum constant (8.19).

After approximately $\lambda = \lambda_\infty \approx 1.401155$, chaos appears where all 200 points are distinct.

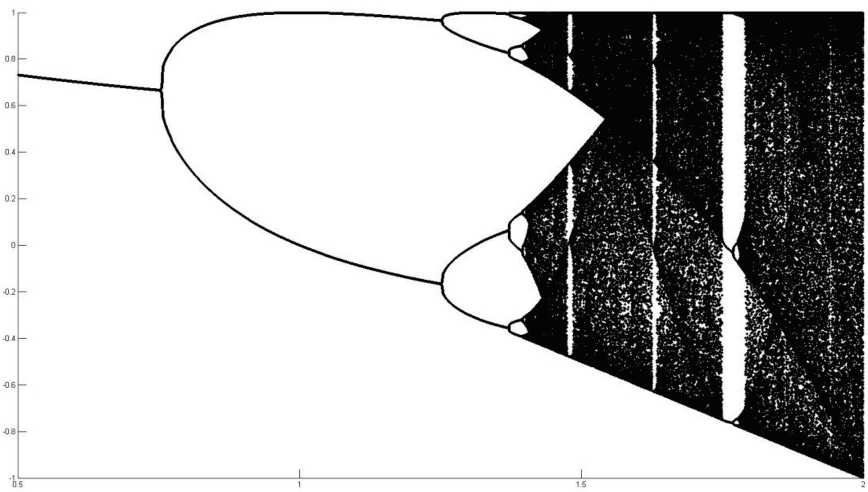

Figure 8.8: Bifurcation diagram of the logistic map system. From the left, it clearly shows the succession of asymptotically stable 2^n-periodic solutions, followed by chaos and other asymptotically stable periodic solutions in the windows in chaos.

8.3.2 Sarkoviskii's Ordering

Sweeping Figure 8.8 from small values of λ to large values of λ, we have have thus a succession of asymptotically stable 2^n-periodic solutions until $\lambda = \lambda_\infty \approx 1.401155$, after which chaos appears. In the same parameter region, however, there are also small subintervals where there is no chaos, but again asymptotically stable periodic solutions. They are called *windows in chaos*. The largest window contains a 3-periodic solution that undergoes again a cascade of flip bifurcations that create $6-$, $12-$, $24-$, $\ldots$-periodic solutions. After the flip bifurcations from a T-periodic to a $2T$-periodic solution, the asymptotically stable T-periodic solution continues to exist, but becomes unstable. Therefore, in the chaotic region, infinitely many unstable periodic solutions are present.

One also observes that the window with the 3-periodic solution is not only the largest, but is also the last one to appear. This is a consequence of a remarkable theorem proven by Sarkoviskii, which shows that for a large class of 1-d maps, which includes the logistic map, periodic solutions appear in a particular order that starts with 3 and the odd integers and ends by powers of 2, in the sense that if an m-periodic solution exists, then all l-periodic solutions exist as well if l appears after m in Sarkoviskii's ordering, which is the following ($m \triangleright l$ means that m appears before l in that ordering.

Definition 8.3. *Sarkoviskii's ordering of the set of natural numbers $\mathbb{N}^*$ is*

$$3 \triangleright 5 \triangleright 7 \triangleright 9 \triangleright \ldots \triangleright 2\cdot 3 \triangleright 2\cdot 5 \triangleright 2\cdot 7 \triangleright \cdots \triangleright 2^2\cdot 3 \triangleright 2^2\cdot 5 \triangleright 2^2\cdot 7 \triangleright \cdots \triangleright 2^m\cdot 3 \triangleright 2^m\cdot 5 \triangleright 2^m\cdot 7 \triangleright \cdots \triangleright 2^3 \triangleright 2^2 \triangleright 2 \triangleright 1.$$

In particular, we see that 3 is the first element in Sarkoviskii's ordering, yielding that the system admits m-periodic solutions of all orders $m \in \mathbb{N}^\star$ if there exists a 3-periodic solution. This lead to a paper by Li and Yorke entitled "Period three implies Chaos", and the following theorem.

Theorem 8.3. *Let Ω be an interval and $F : \Omega \to \Omega$ be a continuous function. If the system $x(t+1) = F(x(t))$ has a 3-periodic solution, then it has an m-periodic solution for all $m \in \mathbb{N}^\star$.*

The proof of this theorem relies on the intermediate value theorem, which states that if $F(\cdot) : [a,b] \to \mathbb{R}$ is a continuous function, then for any y between $F(a)$ and $F(b)$ (i.e., such that either $F(a) \le y \le F(b)$ or $F(b) \le y \le F(a)$), there exists $c \in [a,b]$ such that $F(c) = y$ (Note: we denote by $F([a,b])$ the image of $[a,b]$.

The two following lemma extends the intermediate value theorem from a single point $y \in F([a,b])$ to an interval $I \subset F([a,b])$.

Lemma 8.1. *Let X be an interval and $F : X \to \mathbb{R}$ be a continuous function. For any closed interval $I \subset F(X)$, there is a closed interval $J \subset X$ such that $F(J) = I$.*

Lemma 8.2. *Let X be an interval and $F : X \to X$ be a continuous function. Let $\{I_n; n \in \mathbb{N}\}$ a sequence of closed and bounded intervals such that $I_n \subseteq X$ and $I_{n+1} \subseteq F(I_n)$ for all $n \in \mathbb{N}$. Then there is a sequence of closed and bounded intervals $\{J_n; n \in \mathbb{N}\}$ such that $J_{n+1} \subseteq J_n \subseteq I_0$ and $F^{(n)}(J_n) = I_n$ for all $n \in \mathbb{N}$.*

Proof:

We proof this lemma by induction on $n \in \mathbb{N}$.

For $n = 0$, define $J_0 = I_0$. Then clearly, $F^{(0)}(J_0) = J_0 = I_0$.

Next, let $n \ge 1$ and suppose that $F^{(n-1)}(J_{n-1}) = I_{n-1}$. The assumption in the lemma implying that $I_n \subseteq F(I_{n-1})$, we get that $I_n \subseteq F(I_{n-1}) = F(F^{(n-1)}(J_{n-1})) = F^{(n)}(I_{n-1})$. Since $F^{(n)}(\cdot)$ is a continuous function from I_{n-1} to $\mathbb{R}$, Lemma 8.1 applied to function $F^{(n)}(\cdot)$ and interval $I_n \subseteq F^{(n)}(I_{n-1})$ implies that there is a closed interval $J_n \subset J_{n-1}$ such that $F^{(n)}(J_n) = I_n$.

∎

We can prove Theorem 8.3.

Proof:

Suppose that the system has a 3-cycle $\{\xi, F(\xi), F^{(2)}(\xi)\}$. We can rename these elements as $\{a, b = F(a), c = F(b)\}$ and without loss of generality, assume $a < b < c$. Either $F(a) = b$ or $F(a) = c$. Suppose Let $J = [a,b]$ and $L = [b,c]$, and for any integer $k \in \mathbb{N}^*$, let us define a sequence of closed intervals $\{I_n; n \in \mathbb{N}\}$ such that $I_n = L$ for $n \in \{0, 1, \ldots, k-2\}$, $I_{k-1} = L$. and then $I_{n+k} = I_n$ for all $n \in \mathbb{N}^\star$.

If $k = 1$, then $I_n = L$ for all or all $n \in \mathbb{N}^*$. Since $F(a) = b$ and $F(b) = c$, the intermediate value theorem implies that $J = [a,b] \subseteq F([b,c]) = F(L)$ and similarly, since $F(c) = a$ and $F(b) = c$, that $L = [b,c] \subseteq F([a,b]) = F(J)$.

Therefore for any $k \in \mathbb{N}^*$, we can apply Lemma 8.2 to produce a sequence of closed and bounded intervals $\{J_n; n \in \mathbb{N}\}$ such that $J_k \subseteq J_0 = L$ and $F^{(k)}(J_k) = I_k = L$. Consequently, $L \subseteq F^{(k)}(L)$.

Now the system $x(t+1) = F^{(k)}(x(t))$ admits always a fixed point in L because again of the intermediate value theorem, which apply to the continuous function $F^{(k)} : L \to L$. This fixed point of the system with the map $F^{(k)}$ corresponds to a k-periodic solution of the original system with the map F.

■

8.3.3 Dense Set of Periodic Solutions and Definition of Chaos

The orbits of the unstable periodic solutions are actually dense in the chaotic attractors. This is illustrated in the following figures. The chaotic trajectory in Figure 8.9 is first approximated by 16-periodic trajectory (Figure 8.10), then by an 8-periodic trajectory (Figure 8.11) and then then by a trajectory of period 11 (Figure 8.12). This can be continued. Thus, the chaotic trajectory navigates between the periodic trajectories, getting close to one, then to another, etc.

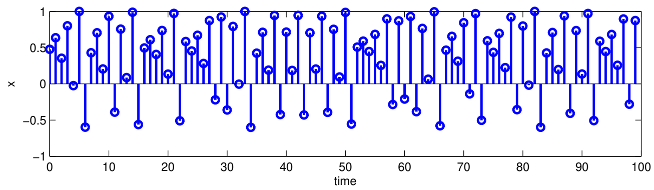

Figure 8.9: Trajectory (chaotic) of the logistic map system with $\lambda = 1.6$ starting at $x(1) = 0.4772$.

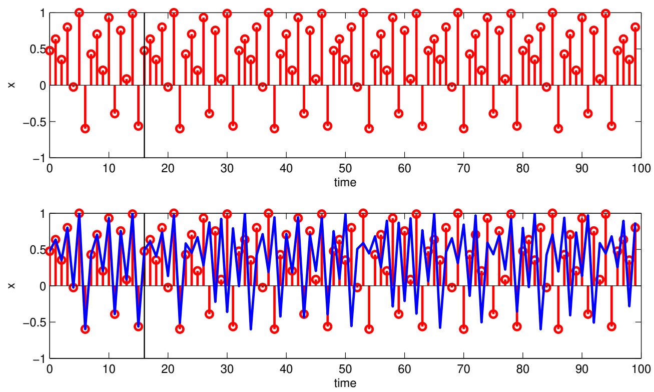

Figure 8.10: Approximation of a portion of a chaotic trajectory by a periodic trajectory. Top: 16-periodic trajectory of the same system. Bottom: Superposition of the chaotic and of the 16-periodic trajectories. In the resolution of this figure, both trajectories are almost indistinguishable from time $t = 1$ to $t = 18$. After that they separate visibly.

There are many possible definitions of chaos. As we saw in the previous subsection, the sensitivity to initial conditions measured by a positive Lyapunov exponent for almost all solutions is a good indicator that a system is chaotic, provided some other conditions are met, such as a compact state space (so that all solutions will remain bounded). However it might problematic to use in a definition, because it is a necessary but not always sufficient condition for a system to be chaotic.

One of the most widely accepted definitions of chaotic system for a 1-dimensional discrete-time system $x(t+1) = F(x(t))$ over a finite interval $I = [a, b]$, due to R. Devaney (*An Introduction To Chaotic*

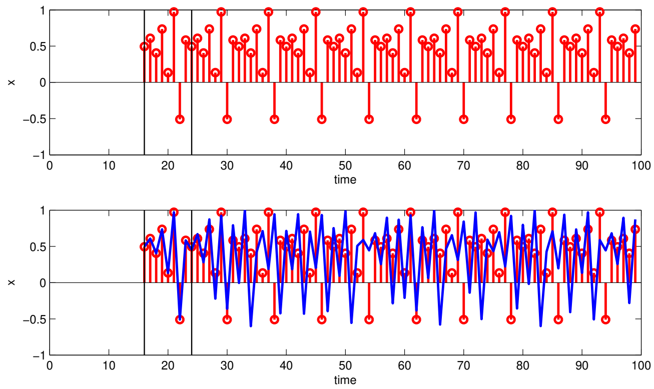

Figure 8.11: Approximation of a portion of a chaotic trajectory by a periodic trajectory (continued). Top: 8-periodic trajectory. Bottom: Superposition of the chaotic and of the 8-periodic trajectories. They are indistinguishable in this figure from $t = 17$ to $t = 25$.

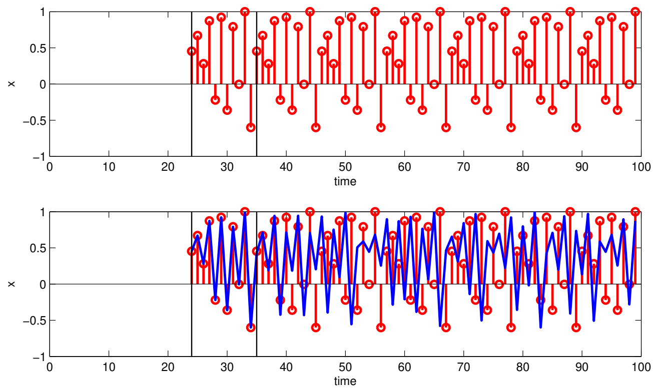

Figure 8.12: Approximation of a portion of a chaotic trajectory by a periodic trajectory (continued). Top: 11-periodic trajectory. Bottom: superposition of the chaotic trajectory and the 11-periodic trajectory. In this figure they are indistinguishable from $t = 25$ until $t = 37$.

Dynamical Systems, CRC Press, 1989) requires the existence of a dense set of periodic solutions and of the transitivity of the map F.

We call ξ a m-periodic point of F if and only if $F^{(m)}(\xi) = \xi$, and its prime or minimal period is the smallest integer $m \in \mathbb{N}^*$ that verifies this equality.

Remember that a set $A \subset B$ is *dense* in B if any point of B is either a point of A, or is arbitrary close to a point of B. More formally, $A \subset B$ is dense in B if for each point $x \in B$ and each $\varepsilon > 0$, there exists $y \in A$ such that $\|x - y\| < \varepsilon$. If B is an interval, A is dense in B if for each point $x \in B$ and each $\varepsilon > 0$ the open interval $(x - \varepsilon, x + \varepsilon)$ contains a point $y \in A$

A *transitive* map has points that will eventually move under iterations of the system, from one arbitrarily small sub-interval of I to any other small sub-interval of I. More formally, F is (topologically) transitive, if for any two open subsets U_1 and U_2 of I, there is some point $x_0 \in U_1$ and some $m \in \mathbb{N}^*$ such that $F^{(m)}(x_0) \in U_2$.

Definition 8.4. *Let Ω be the interval $\Omega = [a, b]$ and $F : \Omega \to \Omega$. The system $x(t + 1) = F(x(t))$ is chaotic if and only if*

1. *Periodic points of F are dense in the interval $\Omega = [a, b]$;*
2. *F is (topologically) transitive;*
3. *F has sensitive dependence on Ω, i.e., there is some sensitivity constant $\beta > 0$ such that for any point $x_0 \in \Omega$ and any open sub-interval U containing x_0, there is some $y_0 \in U$ and $T \in \mathbb{N}^*$ such that $\left|F^{(T)}(x_0) - F^{(T)}(y_0)\right| > \beta$.*

Quite surprisingly, the third condition is implied by the two first conditions (J. Banks et al, "On Devaney's definition of chaos", *The American mathematical monthly*, vol. 99(4), pp. 332–334, 1992), and hence is redundant in the definition. Let us apply this definition to the Bernouilli map given by (8.14)

Bernouilli Map (continued)

Since the Bernouilli map $F(\cdot)$ on $\Omega = [0, 1)$ is defined by

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

we find that

$$F^{(m)}(x) = 2^m x \mod 1 = \begin{cases} 2^m x & \text{if} \quad 0 \le x < 1/2^m \\ 2^m x - 1 & \text{if} \quad 1/2^m \le x < 2/2^m \\ & \dots \\ 2^m x - 2^{m-1} & \text{if} \quad (2^m - 1)/2^m \le x < 1 \end{cases}$$

Since $F^{(m)}(\cdot)$ maps each of the sub-intervals $I_k = [k/2^m, (k+1)/2^m)$ for $k \in \{0, 1, 2, \dots, 2^m - 1\}$ onto the interval $\Omega = [0, 1)$, its graph crosses the diagonal line $F(x) = x$ inside $\Omega \times \Omega$ at some point in this interval, and therefore there is an m-periodic point of F in each of the the sub-intervals I_k. Since the length of each of these intervals is $1/2^m$, it follows that any point $x \in \Omega$ can be made arbitrarily close to an m-periodic point of F if we take m large enough. Therefore the set of the periodic points of F is dense in the interval $\Omega = [0, 1)$.

Let $U_1 \subseteq \Omega$ be any open interval in $\Omega = [0, 1)$. By taking again m large enough, we can always find some $k \in \{0, 1, 2, \dots, 2^m - 1\}$ such that the interval $I_k = [k/2^m, (k + 1)/2^m) \subset U_1$. Therefore the image of U_1 through $F^{(m)}$ is the whole interval $\Omega = [0, 1)$: $F^{(m)}(U_1) = \Omega$. Hence, for any open interval $U_2 \subseteq \Omega$, we can find a point $x_0 \in U_1$ such that $F^{(m)}(x_0) \in U_2$, which shows that F is (topologically) transitive.

Therefore the Bernouilli Map is chaotic.

8.3.4 Topological Conjugacy

Topological conjugacy allows to map one chaotic system to another, and therefore to conclude that the system $x(t+1) = G(x(t))$ is chaotic if the system $x(t+1) = F(x(t))$ is chaotic and the maps F and G are topologically conjugate.

Definition 8.5. *Let I and J be two intervals, and consider the two maps $F : I \to I$ and $G : J \to J$. The maps F and G are conjugate if there is a homeomorphism $H : I \to J$ such that*

$$H \circ F = G \circ H.$$

In other words, for all $x \in I$, $H(F(x)) = G(H(x))$. This implies by induction that for all $t \in \mathbb{N}^*$,

$$H(F^{(t)}(x)) = (H \circ F)((F^{(t-1)})(x)) = (G \circ H)((F^{(t-1)})(x)) = G(H(F^{(t-1)})(x) = \ldots = G^{(t)}(H(x)),$$

so that the conjugacy takes orbits of the system $x(t+1) = F(x(t))$ onto orbits of the system $x(t+1) = G(x(t))$. Similarly, H^{-1} takes orbits of the system $x(t+1) = G(x(t))$ onto orbits of the system $x(t+1) = F(x(t))$.

Theorem 8.4. *Let I and J be two closed intervals of finite lengths. If $F : I \to I$ and $G : J \to J$ are conjugate via H, then the system $x(t+1) = G(x(t))$ is chaotic if the system $x(t+1) = F(x(t))$ is chaotic.*

Proof:

Since the system $x(t+1) = F(x(t))$ is chaotic, periodic points of F are dense in I and F is transitive.

(i) Let U be an open sub-interval of J, and let us consider the open set $H^{-1}(U) \subseteq I$. Periodic points of F are dense in I, hence there is at least one periodic point ξ of F which is in $H^{-1}(U)$. Let m be its period. Then conjugacy implies that

$$G^{(m)}(H(\xi)) = H(F^{(m)}(\xi)) = H(\xi),$$

which yields that $H(\xi)$ is an m-periodic point of G. As $H(\xi) \in U$, it shows that any open sub-interval $U \subset J$ contains an m-periodic point of G, hence that the periodic points of G are dense in J.

(ii) Let U_1 and U_2 be any two open sub-intervals of J. Then $H^{-1}(U_1)$ and $H^{-1}(U_2)$ are two open subsets of I. Since F is transitive, there is some point $x_1 \in H^{-1}(U_1)$ and some $m \in \mathbb{N}^*$ such that $F^{(m)}(x_1) \in H^{-1}(U_2)$. But then $H(x_1) \in U_1$ and by conjugacy $G^{(m)}(H(x_1)) = H(F^{(m)}(x_1)) \in U_2$, which shows that G is also transitive.

Combining (i) and (ii) yields that $x(t+1) = G(x(t))$ is chaotic.

■

The requirement that the conjugacy H is one-to-one (bijective) can be relaxed and replaced by the requirement that the conjugacy H is a continuous m-to-one function (for some finite m) Definition 8.5. We say then that F and G are *semi-conjugate*. Semi-conjugacy still maps cycle to cycle, but without preserving their minimal periods, and preserves chaotic behavior on intervals of finite length.

A Variant of the Logistic Map with $\lambda = 4$.

Let us consider a variant of the logistic map $F_\lambda(x) = \lambda x(1-x)$ (which is itself conjugate with the variant of logistic map given in (8.1)). When $\lambda = 4$, its iterations are given by

$$x(t+1) = 4x(t)(1-x(t)) \tag{8.20}$$

and the state space is $\Omega = [0, 1]$. Then one can check that the logistic map $F_4(x) = 4x(1 - x)$ and the Bernouilli map $F(\cdot)$ given by (8.14) are semi-conjugate, with conjugacy $H(x) = (1 - \cos(2\pi x))/2$, which maps the interval $\Omega = [0, 1]$ in a two-to-one fashion over $\Omega = [0, 1]$, except at 1/2, which is the only point mapped to $H(1/2) = 1$. One computes indeed that $H(F(x)) = G(H(x))$. Therefore the system (8.20) is chaotic.

8.3.5 Symbolic Analysis

We will new consider the shift map $S : \Omega \to \Omega$ that acts on binary sequences of 0's and 1's, and which is therefore defined on a state space Ω that is the set of all binary sequences. A "point" $\omega \in \Omega$ is therefore an infinite binary sequence of the form $\omega = (\omega_0, \omega_1, \omega_2, \omega_3 \ldots)$. The distance between two points $a, b \in \Omega$ is given by

$$d(a, b) = \sum_{i=0}^{\infty} |a_i - b_i| \cdot 2^{-i}. \tag{8.21}$$

It is easy to check that (8.21) verifies the three requirements to be a distance function, and that it always converge to a value between 0 and 2. Moreover, if $a_i = b_i$ for $0 \leq i \leq m$, then (8.21) becomes

$$d(a, b) = \sum_{i=m+1}^{\infty} |a_i - b_i| \cdot 2^{-i} \leq 2^{-(m+1)} \sum_{j=0}^{\infty} 2^{-j} = 2^{-m}$$

and conversely, if for some $m \in \mathbb{N}^*$, $d(a, b) \leq 2^{-m}$, then we must have $a_i = b_i$ for $0 \leq i \leq m$ as otherwise, if $a_i \neq b_i$ for some $0 \leq i \leq m$, $d(a, b) \geq |a_i - b_i| \cdot 2^{-i} = 2^{-i} \geq 2^{-m}$.

The (left) *shift map* $S : \Omega \to \Omega$ is defined by

$$S(\omega_0, \omega_1, \omega_2, \omega_3 \ldots) = (\omega_1, \omega_2, \omega_3, \omega_4 \ldots). \tag{8.22}$$

The shift map S has two fixed points, which we denote by $\overline{0} = (0, 0, 0, \ldots)$ and $\overline{1} = (1, 1, 1, \ldots)$. Its k-periodic solutions (also called k-periodic points) are obtained by repeating blocks of length k, and are denoted by

$$\overline{s_0 s_1 \ldots s_k} = (s_0, s_1, s_2, \ldots, s_k, s_0, s_1, s_2, \ldots, s_k, \ldots).$$

The shift map defines a mixing transformation (see Example 2 in Section 8.4), and is chaotic, which can be assessed using the different methods introduced in the previous sections.

(i) One can follow a direct approach to show that the map is transitive and that its periodic points are dense in Ω.

To show transitivity, we can construct a sequence $s^\star$ that will come arbitrarily close to another other sequence $a \in \Omega$ by successively listing in $s^\star$ all possible blocks of 0's and 1's of all possible lengths, starting from length 1 (0 and 1), then 2 $(00, 01, 10, 11)$, 3 and so forth:

$$s^\star = (0,\ 1,\quad 00,\ 01,\ 10,\ 11,\quad 000,\ 001,\ \ldots). \tag{8.23}$$

Let $a = (a_0, a_1, a_2, \ldots) \in \Omega$. Because of the construction of $s^\star$, the first m bits $a_0, a_1, a_2, \ldots, a_m$ of a will be found in a block located in the sequence (8.23) – more precisely, between the $(2^m - 1)$th and $2(2^m - 1)$th blocks of the sequence (8.23). Therefore there is some k, with $(2^m - 1) \leq k \leq 2(2^m - 1)$, such that the kth iterate of the shift map will output a sequence starting with $a_0, a_1, a_2, \ldots, a_m$, i.e.,

$$S^{(k)}(s^\star) = (a_0, a_1, a_2, \ldots, a_m, s^\star_{m+1}, s^\star_{m+2}, \ldots),$$

whose distance from a is therefore

$$d(a, S^{(k)}(s^\star)) = \sum_{i=m+1}^{\infty} |a_i - s^\star_i| \cdot 2^{-i} \leq 2^{-m}.$$

This shows that the orbit of $s^\star$ will come arbitrarily close to any point $a \in \Omega$, since we can pick m as large as we want. In particular, it will pass through any two open subsets $U_1, U_2 \subseteq \Omega$, which establishes transitivity.

Any open subset of Ω contains a sequence $a = (a_0, a_1, a_2, \ldots)$ whose distance with the m-periodic point

$$\overline{a_0 a_1 \ldots a_m} = (a_0, a_1, a_2, \ldots, a_m, a_0, a_1, a_2, \ldots, a_m, \ldots)$$

is

$$d(a, S^{(m)}(\overline{a_0 a_1 \ldots a_m})) = \sum_{i=m+1}^{\infty} |a_i - a_{i-m}| \cdot 2^{-i} \leq 2^{-m}.$$

This shows that any point $a = (a_0, a_1, a_2, \ldots) \in \Omega$ is arbitrarily close to an m-periodic point $\overline{a_0 a_1 \ldots a_m}$ of the shift map by taking m large enough, hence that the set of periodic points of S is dense in Ω.

(ii) We can also follow an indirect approach that couples, by conjugacy, the (left) shift map S on the set of binary sequences Ω with the Bernoulli map F on the interval $[0, 1)$ given by (8.14). The homeomorphism $H : [0, 1) \to \Omega$ is then simply the binary expansion of a real $x \in [0, 1)$ in the sequence $(0, b_1, b_2, \ldots)$, and which is defined by

$$x = \sum_{i=1}^{\infty} b_i 2^{-i}, \tag{8.24}$$

so that $H(x) = (0, b_1, b_2, \ldots)$ (note that $b_0 = 0$ because $0 \leq x < 1$), and $H^{-1}(0, b_1, b_2, \ldots) = \sum_{i=1}^{\infty} b_i 2^{-i}$. Clearly H is one-to-one, continuous and with continuous inverse.

We show that $H \circ F = H \circ S$. Indeed, on the one hand

$$F(x) = 2x \mod 1 = \sum_{i=1}^{\infty} b_i 2^{-i+1} \mod 1 = \left(b_1 + \sum_{i=2}^{\infty} b_i 2^{-i+1} \right) \mod 1 = \left(\sum_{i=2}^{\infty} b_i 2^{-i+1} \right) \mod 1,$$

whence

$$F(x) = \sum_{j=1}^{\infty} b_{j+1} 2^{-j} \mod 1, \tag{8.25}$$

which implies that $H(F(x)) = (b_1, b_2, b_3, \ldots)$. On the other hand, if the binary expansion of x is $H(x) = (0, b_1, b_2, \ldots)$, then the left shift is $S(H(x)) = (b_1, b_2, \ldots)$. This also shows that S is chaotic since F is chaotic.

A Variant of the Logistic Map with $\lambda > 4$

The conjugacy of other maps with the shift map allows to use the simplicity of the trajectories in the latter map, which is called symbolic analysis. This is the case of the variant of the logistic map $F_\lambda = \lambda x(1 - x)$ when the parameter λ is no longer 4 like in (8.20), but larger than 4

$$x(t + 1) = \lambda x(t)(1 - x(t)) \tag{8.26}$$

Unlike the case of $0 < \lambda \leq 4$, the interval $I = [0, 1)$ is no longer forward invariant when $\lambda > 4$. Indeed, if $x(t) \in A_0$ where

$$A_0 = \left(\frac{1}{2} - \frac{\sqrt{\lambda^2 - 4\lambda}}{2\lambda}, \frac{1}{2} + \frac{\sqrt{\lambda^2 - 4\lambda}}{2\lambda} \right),$$

then $F_\lambda(x) = \lambda x(1 - x) > 1$ and $F_\lambda^t(x) < 0$ for all $t \geq 2$, so that all trajectories diverge to $-\infty$. The same applies of course to the pre-image $A_1 = F_\lambda^{-1}(A_0)$ of A_0, as all trajectories in A_1 will also eventually diverge to $-\infty$ after transiting in A_0. Now, A_1 consists of two open sub-intervals, one on each side of A_0. We proceed next with the pre-image of A_1 and we see that the pre-image of each of the two sub-intervals of A_1 is again a pair of disjoint open sub-intervals, so that $A_2 = F_\lambda^{-1}(A_1) = F_\lambda^{-2}(A_0)$

consists of four disjoint open sub-intervals. Let $A_m = F_\lambda^{-m}(A_0)$ be mth pre-image of A_0: A_m consists of 2^m disjoint open sub-intervals of $I = [0, 1)$. The set of points whose orbits remain in $I = [0, 1)$ is therefore given by

$$\Lambda = I \setminus \bigcup_{m=0}^{\infty} A_m.$$

We limit ourselves to the set of points that are in the invariant set Λ, whose dynamics can be captured by symbolic analysis. Let I_0 and I_1 denote the two closed intervals

$$I_0 = \left[0, \frac{1}{2} - \frac{\sqrt{\lambda^2 - 4\lambda}}{2\lambda}\right] \quad \text{and} \quad I_1 = \left[\frac{1}{2} + \frac{\sqrt{\lambda^2 - 4\lambda}}{2\lambda}, 1\right]$$

so that I_0 is located to the left of A_0, I_1 is located to the right of A_0, and $I_0 \cup I_1 = I \setminus A_0$. Let $x_0 \in \Lambda$. Then $x(t) = F_\lambda^{(t)}(x_0) \in \Lambda \subset I_0 \cup I_1$ for all $t \in \mathbb{N}$ and we can associate a binary sequence $b = (b_0, b_1, b_2, \ldots)$ to this trajectory $x(t), t \in \mathbb{N}$ by setting $b_t = 0$ if and only if $x(t) \in \Lambda \cap I_0$ and $b_t = 1$ if and only if $x(t) \in \Lambda \cap I_1$. We denote by B this map from Λ to the set of all binary sequences Ω.

It takes some work to show that B is continuous and one-to-one if $\lambda > 2 + 2\sqrt{5}$, and considerable work to show it when $4 < \lambda \leq 2 + 2\sqrt{5}$, and we skip this part.

We only show that $B \circ F_\lambda = S \circ B$ where S is the left shift map given by (8.22). Observe that if the binary sequence corresponding to x_0 is $B(x_0) = (b_0, b_1, b_2, \ldots)$, then it means that $x(0) = x_0 \in I_{b_0}$, $x(1) = F_\lambda(x_0) \in I_{b_1}$, $x(2) = F_\lambda^{(2)}(x_0) \in I_{b_2}$, and so on. Therefore, since $F_\lambda(x_0) \in I_{b_1}, F_\lambda^{(2)}(x_0) \in I_{b_2}$, etc, it yields that $B(F_\lambda(x_0)) = (b_1, b_2, \ldots)$, and thus that $B(F_\lambda(x_0)) = S(B(x_0))$ for all $x_0 \in \Lambda$. This shows that the logistic map F_λ given by (8.26) and the shift map S given by (8.22) are conjugate, and therefore that (8.26) is a chaotic system if $\lambda > 4$.

8.4 Appendix: Some Elements from the Theory of Ergodic Dynamical Systems

This appendix is a summary of useful notions in Ergodic Theory, from the chapter authored by Ali Ajdari Rad and Martin Hasler.

8.4.1 Probability Space

Let us first recall the notion of a probability space and make a few definitions.

Definition 8.6 (Probability Space)**.** *A probability space* (Ω, Σ, P) *is composed of:*

- *the sample space* Ω*. The elements of* Ω *are the (outcomes of the) elementary (atomic) events;*
- *a* σ*-algebra* Σ *on* Ω*, which is a collection of subsets of* Ω*, which contains* $\varnothing$ *and* Ω*, and which is closed under the countable set operations of union, intersection and set complement. The elements of* Σ *are the events;*
- *a probability measure* $P : \Sigma \rightarrow [0,1]$*, which assigns to each event a real between 0 and 1, and which has the properties that* $P(\varnothing) = 0$ *and that if* $\Sigma_i \in \Sigma$*, with* $i \in \mathbb{N}$*, is a countable collection of disjoint sets in* Σ*, then*
$$P\left(\bigcup_{i=0}^{\infty} \Sigma_i\right) = \sum_{i=0}^{\infty} P\left(\Sigma_i\right). \tag{8.27}$$

Let us make first some remarks about σ-algebras.

- Because of de Morgan's laws in set theory, stating that the complement of the countable union of sets is the intersection of their complements, and vice-versa:
$$\begin{aligned}\overline{\bigcup_i \Sigma_i} &= \bigcap_i \overline{\Sigma}_i \\ \overline{\bigcap_i \Sigma_i} &= \bigcup_i \overline{\Sigma}_i,\end{aligned}$$
it is sufficient for a collection of subsets of Ω to contain $\varnothing$ and to be closed under complementation and countable unions, to be a σ-algebra.
- If Ω is a countable set, we can take as Σ the power set of all its subsets: it is the largest σ-algebra on Ω. For a countable set, we will take the power set of its subsets as the default σ-algebra on the set. There are however other σ-algebras that can be defined on the set. For example, we can also take $\Sigma = \{\varnothing, \Omega\}$: it is the smallest σ-algebra on Ω.
- If A is any (collection of) subset(s) of Ω, the σ-algebra generated by A is the smallest σ-algebra on Ω that contains A. For example, if $\Omega = \{a, b, c\}$ then the σ-algebra on Ω generated by subset $\{a\}$ is $\Sigma = \{\varnothing, \{a\}, \{b, c\}, \{a, b, c\}$. Note that it contains $\varnothing$ and Ω, and that the union, intersection and complement of any subset in Σ is another element of Σ, implying that it is indeed a σ-algebra on Ω.
- If Ω is a non countable set, we need to specify the σ-algebra. For example, if $\Omega = \mathbb{R}$ or a finite open interval in $\mathbb{R}$, a natural choice is the σ-algebra generated by the collection of all open intervals in $\mathbb{R}$.

Let us make next some remarks about P measures on the σ-algebra Σ.

- The measure P is in general a function $P : \Sigma \to [0, \infty]$, which is such that $P(\varnothing) = 0$ and which satisfies the countable additivity condition (8.27) for any countable collection of disjoint sets in Σ. For example, if $\Omega = \mathbb{R}$ or a finite open interval in $\mathbb{R}$, the Lebesgue measure on the σ-algebra generated by the collection of all open intervals is their length, i.e. $P((a, b)) = b - a$ for any open interval $(a, b) \in \Sigma$. Since the complement and countable union of open intervals are also in Σ, it yields that $P([a, b]) = P([a, b)) = P((a, b]) = b - a$ as well. The space (Ω, Σ, P) is a metric space. When $P(\Omega) = 1$, then P is called a probability measure and the metric space (Ω, Σ, P) is a probability space - this is the default setting with which we work in this chapter and in Definition 8.6.
- If (Ω, Σ, P) is a probability space, and if $A \in \Sigma$, then $P(A)$ is the probability that A occurs.

Example 1: Single Die Throwing

The probability space (Ω, Σ, P) describes the experiment of throwing a die once. An elementary event is the outcome of the experiment. Hence

- $\Omega = \{1, 2, 3, 4, 5, 6\}$;
- Σ is the power set of Ω (i.e., the set of all its subsets);
- P is the uniform distribution on Ω: $P(\{1\}) = P(\{2\}) = \ldots = P(\{6\}) = 1/6$.

Example 2: Repeated Die Throwing

The probability space (Ω, Σ, P) describes the experiment of throwing a die repeatedly. An elementary event is the sequence of outcomes when the die is thrown over and over and over again. Hence

- $\Omega = \{\omega \mid \omega = (\omega_1, \omega_2, \ldots), \omega_i \in \{1, 2, 3, 4, 5, 6\}\}$;
- Σ is the σ-algebra generated by the sets obtained by fixing a finite set of coordinates. Such elementary events, called *cylinder sets*, have the form

$$S_{(i_1,j_1)(i_2,j_2)\ldots(i_n,j_n)} = \{\omega \mid \omega_{i_1} = j_1, \omega_{i_2} = j_2, ..., \omega_{i_n} = j_n\}, \tag{8.28}$$

  for some finite $n \in \mathbb{N}^*$, where $i_1, i_2, \ldots, i_n$ are n natural numbers satisfying $i_1 < i_2 < \ldots < i_n$ and where $j_1, j_2, \ldots, j_n \in \{1, 2, 3, 4, 5, 6\}$;
- P is defined by its specification for the cylinder sets:

$$P\left(S_{(i_1,j_1)(i_2,j_2)\ldots(i_n,j_n)}\right) = \left(\frac{1}{6}\right)^n. \tag{8.29}$$

Example 3: Uniform Distribution on the Unit Interval

We consider the interval $[0, 1) = [0, 1[$ with the Lebesgue measure. The probability space (Ω, Σ, P) describes the experiment of picking a point uniformly at random in the interval $[0, 1)$. Hence

- $\Omega = [0, 1)$;
- Σ is the smallest σ-algebra containing the intervals $[a, b], [a, b), (a, b], (a, b)$ for any $0 \le a \le b \le 1$ (which is known as the Borel σ-algebra of the interval $[0, 1)$);
- P is the Lebesgue measure defined by $P([a, b]) = P([a, b)) = P((a, b]) = P((a, b)) = b - a$ (for any $0 \le a \le b \le 1$). It is the measure given by the length of intervals.

8.4.2 Measurable and Measure-Preserving Transformations

One can define functions between probability spaces (Ω, Σ, P) and (Ω', Σ', P'). A measurable function is a mapping $f : \Omega \to \Omega'$ such that $f^{-1}(A') = A$ for all $A' \in \Sigma'$. For example, a measurable function $f : \Omega \to \mathbb{R}$ is called a *random variable*.

A function F from Ω onto Ω (i.e. $F : \Omega \to \Omega$) is called a *transformation* on Ω. Functions arising in the right hand side of the state equations of dynamical systems are such transformations, they can enjoy a number of important properties, which we now review.

Definition 8.7 (Measurable, Invariant Transformations, Invariant Sets)**.** *Let (Ω, Σ, P) be a probability space and $F : \Omega \to \Omega$ be a given transformation.*

- *F is* measurable *if $F^{-1}(A) \in \Sigma$ for all $A \in \Sigma$.*
- *F is* measure preserving *if F is measurable and if $P(F^{-1}(A)) = P(A)$ for all $A \in \Sigma$. One also says that P is an* invariant measure *under F.*
- *A set $A \in \Sigma$ is an* invariant set *under F if $F^{-1}(A) = A$.*

Some remarks about invariant measures and sets:

- Note that we use the inverse of F rather than F itself in Definition 8.7. When F is bijective, this is equivalent to using directly F in the definition. However, if F is not bijective, the definition could lead to undesirable constraints. Indeed, if F is not bijective, one can possibly find two different sets A_1 and A_2 with the same image , hence $F(A_1) = F(A_2) = F(A_1 \cup A_2)$. If F is measure preserving, with a definition that would have declared the transformation to be measure preserving if $P(F(A)) = P(A)$ for all $A \in \Sigma$, then $P(A_1) = P(F(A_1)) = P(F(A_1 \cup A_2)) = P(A_1 \cup A_2)$, even though $A_1 \subsetneq A_1 \cup A_2$, which is not a desirable constraint, and explains why F^{-1} is used in the definition rather than F.
- The notion of *invariant measure* is equivalent to that of stationarity in terms of stochastic processes.
- Note the difference between an invariant *measure* under F, and an invariant *set* under F. Clearly, if a set is invariant under F, its measure is also invariant, but the converse is not true.
- The empty set $\varnothing$ and the whole space Ω are always invariant, they are called the trivial invariant sets.

Example 1 (continued)

Let F_1 and F_2 be the two permutations given in Table 8.1. From Definition 8.7, one can see that F_1 and F_2 are both measurable and measure-preserving. F_2 has no other invariant set than the trivial invariant sets $\varnothing$ and $\Omega = \{1, 2, 3, 4, 5, 6\}$, contrary to F_1 which has also $\{1, 2, 3, 4\}$ and $\{5, 6\}$ as non-trivial invariant sets.

ω	1	2	3	4	5	6
$F_1(\omega)$	3	4	1	2	6	5

ω	1	2	3	4	5	6
$F_2(\omega)$	2	3	4	5	6	1

Table 8.1: Permutations F_1 (left) and F_2 (right).

Example 2 (continued)

Let F be the left shift transformation:

$$F((\omega_1, \omega_2, \ldots)) = (\omega_2, \omega_3, \ldots).$$

Clearly

$$F^{-1}\left(S_{(i_1,j_1)(i_2,j_2)\ldots(i_n,j_n)}\right) = S_{(i_1+1,j_1)(i_2+1,j_2)\ldots(i_n+1,j_n)}$$

By extension from the cylinder sets to the whole σ-algebra Σ that is generated by them, this implies that F is measurable and because of (8.29), is also measure preserving.

The cylinder sets (8.28) are never left-shift invariant. An example of a left-shift invariant set is $\{\omega \mid \text{ there exists } n \geq 1 \text{ such that } \omega_i = 1 \text{ for all } i \geq n\}$.

Example 3 (continued)

In general, invariant measures cannot be determined explicitly in most dynamical systems. An exception is the Bernouilli map $F(\cdot)$ on $\Omega = [0, 1)$, defined by (8.14), which we recall here

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

The inverse image or pre-image of the interval $[a, b]$ is

$$F^{-1}([a, b]) = \left[\frac{a}{2}, \frac{b}{2}\right] \cup \left[\frac{a+1}{2}, \frac{b+1}{2}\right], \tag{8.30}$$

which yields that F is measurable. Since the total length of the two intervals of the pre-image of $[a, b]$ is $(b - a)$, F is also measure preserving. Finally, (8.30) also implies that any interval other than the empty interval, or the whole of $[0, 1)$, is never invariant under the Bernoulli map. An example of an invariant set is the set of all rational numbers in $[0, 1)$, i.e. $\mathbb{Q} \cap [0, 1)$.

8.4.3 Ergodic Transformations

Roughly speaking, a measure-preserving transformation F is ergodic if computing a statistics (e.g., the mean) along the multiple iterations of one trajectory of $x(t+1) = F(x(t))$, starting from a particular point, or computing it over the entire set of initial conditions but for only one iteration, gives the same result. Intuitively, consider the following real-life example [1]. Suppose you want to find out what the most visited parks in a city are. One idea is to take a snapshot at a particular time, to see how many people are in this moment in park A, how many are in park B and so on. Another idea is to look at one individual and to follow him/her for a certain period of time, e.g. a year, and to observe how often the individual is going to park A, how often (s)he is going to park B and so on. The former method provides a statistical analysis over the entire ensemble of people at a certain moment in time, and the latter gives the statistical analysis for one person over a long period of time. The first one may not be representative for a longer period of time, whereas the second one may not be representative for all the people. The ensemble is ergodic if the two methods for computing the statistics give the same result. We now state the formal definition of an ergodic transformation.

Definition 8.8 (Ergodic Transformation). *Let (Ω, Σ, P) be a probability space and $F : \Omega \to \Omega$ be a measure preserving transformation. F is an* ergodic *transformation if for every set $A \in \Sigma$ that is invariant under F, either $P(A) = 0$ or $P(A) = 1$. One also says that P is an* ergodic *measure for F.*

[1] taken from Vlad Tarko, “What is Ergodicity?”, `http://archive.news.softpedia.com/news/What-is-ergodicity-15686.shtml`

Some remarks about ergodic transformations:

- One can show that an invariant measure is ergodic if and only if it cannot be decomposed as the convex combination of two orthogonal measures. More formally, F is not ergodic if and only if there exist a non trivial invariant set $\Omega_1 \subset \Omega$, two probability measures $P_1 \neq P_2$ that are invariant under F such that $P_1(\Omega_1) = 1$ and $P_2(\Omega_1) = 0$, and a real $0 < \lambda < 1$ such that $P = \lambda P_1 + (1 - \lambda)P_2$.
- A transformation whose only invariant sets are the trivial invariant sets $\varnothing$ and Ω is ergodic with respect to any invariant measure.

A stronger property than ergodicity is mixing.

Definition 8.9 (Mixing Transformation)**.** *Let (Ω, Σ, P) be a probability space and $F : \Omega \to \Omega$ be a measure preserving transformation. F is a* mixing *transformation for P if for every sets $A, B \in \Sigma$*

$$\lim_{N\to\infty} P\left(A \cap F^{-N}(B)\right) = P(A)P(B). \tag{8.31}$$

- The meaning of Definition 8.9 is the following. If we consider all trajectories that start in $A \in \Sigma$ and check where they are after N iterations, it turns out that they are distributed over the whole space Ω approximately according to the probability measure P when the number of iterations N is large, and this property is independent of A.
- The mixing property is indeed stronger than ergodicity. Suppose a probability space (Ω, Σ, P) and a measure preserving transformation $F : \Omega \to \Omega$ are given. If F is mixing, then F is ergodic. To prove this property, let A be an invariant set under F. Then $F^{-N}(A) = A$ and thus

  $$\lim_{N\to\infty} P(A \cap F^{-N}(A)) = \lim_{N\to\infty} P(A) = P(A).$$

  Hence, it follows from (8.31) with $B = A$ that $P(A) = P^2(A)$, which is only possible if $P(A) = 0$ or $P(A) = 1$. This proves that F is ergodic.

Equipped with these definitions, one can the state the main theorem of ergodic theory.

Theorem 8.5 (Birkhoff's Ergodic Theorem)**.** *Let (Ω, Σ, P) be a probability space and $F : \Omega \to \Omega$ be a measure preserving transformation. Let $f : \Omega \to \mathbb{R}$ be a P-integrable random variable (i.e., such that the expectation of $|g|$ is finite). Then for P-almost $\omega \in \Omega$ the limit*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f\left(F^n(\omega)\right)$$

exists. If, in addition, F is ergodic, then for P-almost $\omega \in \Omega$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f\left(F^n(\omega)\right) = \int_\Omega f(\omega)dP(\omega). \tag{8.32}$$

Some remarks about Birkhoff's Ergodic Theorem.

- In the theorem, taking P-almost $\omega \in \Omega$ means all $\omega \in \Omega$ except possibly a set M of measure zero (i.e. such that $P(M) = 0$).

- The integral on the right hand side of (8.32) is the Lebesgue integral, which extends the Riemann integral that is seen in first year calculus courses, to a more general class of functions, such as discontinuous functions (a typical example is the indicator function of rational numbers (i.e. $f(\omega) = 1$ if $\omega \in \mathbb{Q}$ and $f(\omega) = 0$ otherwise), whose Lebesgue integral is zero but which is not Riemann integrable). In probabilistic terms,

$$\int_{\Omega} f(\omega) dP(\omega) = \mathbb{E}[f]$$

  where $\mathbb{E}$ is the expectation operator, which is here applied to the random variable f.

- The meaning of the theorem is that for almost all initial conditions of the discrete-time dynamical system generated by the iterations of F, the average of the function f along the trajectory issued from one initial condition (the left hand side of (8.32)) is equal to its expected value with respect to the invariant probability measure P (the right hand side of (8.32)).

- Let $A \in \Sigma$ and let f be the indicator function of A, i.e. $f(\omega) = 1$ if $\omega \in A$ and $f(\omega) = 0$ if $\omega \notin A$); f is called a Bernouilli random variable. Then for P-almost $\omega \in \Omega$, (8.32) becomes

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f\left(F^n(\omega)\right) = \int_{\Omega} f(\omega) dP(\omega) = P(A). \tag{8.33}$$

  The left hand side of (8.33) is nothing else than the proportion of points of the trajectory $\omega, F(\omega), F(F(\omega)), \ldots$ that are lying in A. Then (8.33) shows that this proportion converges to the probability of A for P-almost $\omega \in \Omega$.

Example 1 (continued)

F_1 is not ergodic. Indeed, $\Omega_1 = \{1, 2, 3, 4\}$ is a non-trivial invariant set of F, whose measure is $P(\Omega_1) = 2/3$. Now let us define the two probability measures P_1 and P_2 by

$$P_1(\omega) = \begin{cases} 1/4 & \text{if} \quad \omega \in \{\{1\}, \{2\}, \{3\}, \{4\}\} \\ 0 & \text{if} \quad \omega \in \{\{5\}, \{6\}\} \end{cases}$$

and

$$P_2(\omega) = \begin{cases} 0 & \text{if} \quad \omega \in \{\{1\}, \{2\}, \{3\}, \{4\}\} \\ 1/2 & \text{if} \quad \omega \in \{\{5\}, \{6\}\}. \end{cases}$$

One easily verifies that both measures P_1 and P_2 are invariant under F, and that $P_1(\Omega_1) = 1$ and $P_2(\Omega_1) = 0$. As one can write $P = \lambda P_1 + (1 - \lambda) P_2$ with $\lambda = 2/3$, F_1 is not ergodic.

In contrast, as F_2 has no non-trivial invariant subsets, it is ergodic with respect to P. However it is not mixing: set $A = B = \{1\}$. Then $F^{-N}(B) = \{i\}$ with $i \neq 1$ if $N \mod 6 \neq 0$. Hence

$$A \cap F^{-N}(B) = \begin{cases} \varnothing & \text{if} \quad N \mod 6 \neq 0 \\ \{1\} & \text{if} \quad N \mod 6 = 0, \end{cases}$$

which implies that $\lim_{N\to\infty} P(A \cap F^{-N}(B))$ does not exist andthat F is only ergodic, but not mixing.

Example 2 (continued)

We prove the mixing properties for cylinder sets. Let

$$\begin{aligned} A &= S_{(i_1,j_1)(i_2,j_2)\ldots(i_n,j_n)} \\ B &= S_{(k_1,l_1)(k_2,l_2)\ldots(k_m,l_m)} \end{aligned}$$

where $n, m \in \mathbb{N}^*$, $i_1, i_2, \ldots, i_n$ are n natural numbers satisfying $i_1 < i_2 < \ldots < i_n$, $k_1, k_2, \ldots, k_m$ are m natural numbers satisfying $k_1 < k_2 < \ldots < k_m$, $j_1, j_2, \ldots, j_n, l_1, l_2, \ldots, l_m \in \{1, 2, 3, 4, 5, 6\}$. Then

$$F^{-N}(B) = F^{-N}\left(S_{(k_1,l_1)(k_2,l_2)\ldots(k_m,l_m)}\right) = S_{(k_1+N,l_1)(k_2+N,l_2)\ldots(k_m+N,l_m)}.$$

and for $k_1 + N < i_n$

$$A \cap F^{-N}(B) = S_{(i_1,j_1)(i_2,j_2)\ldots(i_n,j_n)(k_1+N,l_1)(k_2+N,l_2)\ldots(k_m+N,l_m)}$$

Therefore, for sufficiently large N,

$$P(A \cap F^{-N}(B)) = 6^{-n+m} = P(A)P(B),$$

which proves the mixing property for cylinder sets. It is possible to show that this actually implies that the system is mixing for all sets in Σ.

Example 3 (continued)

Let the binary expansion of $\omega = x \in [0, 1)$ be defined by (8.24), which we recall here:

$$x = \sum_{i=1}^{\infty} b_i 2^{-i}.$$

and which yields (8.25), i.e.

$$F(x) = 2x \mod 1 = \sum_{j=1}^{\infty} b_{j+1} 2^{-j} \mod 1,$$

which shows that if we associate with x the binary sequence $b = (b_1, b_2, \ldots)$, then the transformation F amounts to a left shift of binary sequence. Having established that the Bernoulli map on the unit interval is equivalent to the left shift on the binary sequences, we can then apply the same reasoning as in the previous example to prove that F is mixing.