22 Lecture 22. Interval maps

22.1 Interval maps = interval endomorphism

Let I be an interval (often closed). A map $f : I \to I$ is called an interval map, or an interval endomorphism. Starting from $x \in I$, we can define (at least a one-sided) sequence $\{f^n(x)\}_{n\in\mathbb{N}}$.footnote$f^n(x) = (f \circ f \circ \cdots \circ f)(x) = f(f(f \cdots (f(x)) \cdots))$. n f's show up in each expression. We have already encountered with a nontrivial examples in Figs. 18.12, 19.4 and 20.3. In these cases the relations of these maps to the original higher dimensional (often time continuous) systems are 'natural' (not very artificially contrived), so from the maps we can learn a lot about the original systems as Lorenz demonstrated.

The word 'chaos' was introduced into mathematics (and subsequently into physics) by Li and Yorke (see **22.2**). The simplest nontrivial example may be $x_{n+1} = 2x_n$ (mod 1) defined on $[0, 1]$ (see **22.4**).

Utida[239] used such a discrete systems to describe the population dynamics of insects (although empirically 'cyclic fluctuations' were observed, no simulation results were reported):

$$P_{n+1} = P_n \left(\frac{1}{b + cP_n} - \sigma \right), \tag{22.1}$$

where P_n is the population of the generation n and b, c, σ are non-negative parameters. The existence of complicate behavior in nonlinear difference equations was reviewed by May[240], who discussed

$$P_{n+1} = \lambda P_n (1 + P_n)^{-\beta} \tag{22.2}$$

and compared with some experimental and observational results (Fig. 22.1).

[239] S Utida, Population fluctuation, an experimental and theoretical approach, Cold Spring Harbor Symposia on Quantitative Biology, 22 139 (1957).

[240] R M May Simple mathematical models with very complicated dynamics, Nature 261 459 (1976)

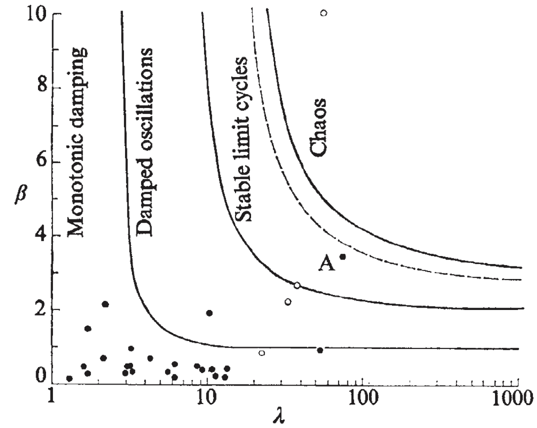

Figure 22.1: The solid lines demarcate the stability domains for the density dependence parameter β and the population growth rate λ in (22.2); the dashed line shows where 2-point cycles give way to higher cycles of period 2^n. The solid circles come from analyses of life table data on field populations, and the open circles from laboratory populations [Fig. 6 of May N 261 459 (1976)]

22.2 Period three implies 'chaos'[241]

Stimulated by Lorenz's map 19.3, Li and Yorke proved the following famous theorem.
Theorem. Let J be an interval and let $F : J \to J$ be continuous. Assume there is a point $a \in J$ for which the points $b = F(a)$, $c = F^2(a)$ and $d = F^3(a)$, satisfy

$$d \le a < b < c \text{ (or } d \ge a > b > c).$$

Then

T1: for every $k = 1, 2, \cdots$ there is a periodic point in J having period k.

Furthermore,

T2: there is an uncountable set $S \subset J$ (containing no periodic points), which satisfies the following conditions:
(A) For every $p, q \in S$ with $p \neq q$,

$$\limsup_{n\to\infty} |F^n(p) - F^n(q)| > 0 \tag{22.3}$$

and

$$\liminf_{n\to\infty} |F^n(p) - F^n(q)| = 0. \tag{22.4}$$

[241] T-Y Li and J A Yorke, Period three implies chaos. Amer Math Month 82 985 (1975).

(B) For every $p \in S$ and periodic point $q \in J$,

$$\limsup_{n\to\infty} |F^n(p) - F^n(q)| > 0. \tag{22.5}$$

The authors added:
REMARKS. Notice that if there is a periodic point with period 3, then they will be satisfied.

The uncountable set S is called a 'scrambled set.' Their definition of chaos is the existence of a scrambled set:
A system exhibits the Li-Yorke chaos, if the system has a scrambled set.

We know S is generally non-measurable, and if measurable, it is measure zero (i.e., the inner measure of S is always zero).[242]
Thus, observable chaos (numerically detectable chaos) cannot be characterized as LI-Yorke chaos. Therefore, in these lecture notes, we do not demonstrate the theorem and will adopt a more natural definition of chaos.

22.3 Observability

Definition [Observability] We say a set B is observable with respect to a given dynamical system (f, M), if the totality of the points on the trajectories that can reach B has a positive Lebesgue measure. In other words, B is observable, if B has a basin with a positive Lebesgue measure or the set $\{x : \exists n \geq 0, f^n(x) \in B, x \in \Gamma\}$ has a positive Lebesgue measure.
In short, if you throw at the collection of what you wish to observe, and if you can hit what you wish to see with a finite probability, it is observable.

---

[242] Y Baba, I Kubo and Y Takahashi, Li-Yorke's scrambled sets have measure zero, Nonlinear Anal 25 1611 (1996). Smital [A chaotic function with some extremal properties, PAMS 87 54 (1983)] constructed a scrambled set S of outer measure 1 for the tent map and mentioned also that measurable scrambled sets have measure zero for this case.

If a map is only continuous (not differentiable), then there are examples of scrambled sets that are measurable and with positive measure: see, for example, I. Kan, "A chaotic function possessing a scrambled set with positive Lebesgue measure," Proc. Amer. Math. Soc. **92**, 45 (1984) or J. Smital, "A chaotic function with a scrambled set of positive Lebesgue measure," Proc. Amer. Math. Soc. **92**, 50 (1984).
V. J. Lopez, "Paradoxical functions on the interval," Proc. Amer. Math. Soc., 120, 465 (1994) proves the following: If a map f from an interval I to I is expansive, then the dynamical system cannot have a measure positive scrambled set. However, if in $I \times I$, x and y are both in the scrambled set, then the totality $\mathrm{Ch}(f)$ of $\{x, y\}$ is always measurable. Furthermore, if f is expansive and its derivative is piecewisely Lipshitz, then $\mathrm{Ch}(f)$ has a positive measure.

22.4 The simplest genuine chaotic dynamical system[243]

The purpose of this unit is to give a preview and outline of our logic with the aid of perhaps the simplest example of chaos.

A map T from $[0, 1]$ into itself is defined as

$$Tx = 2x \mod 1. \tag{22.6}$$

If we use the symbol $\{r\}$ that extracts the fractional part of a real number r, we may write $Tx = \{2x\}$ (i.e., multiply 2 and then remove the integer part). Important points are summarized in Fig. 22.2.

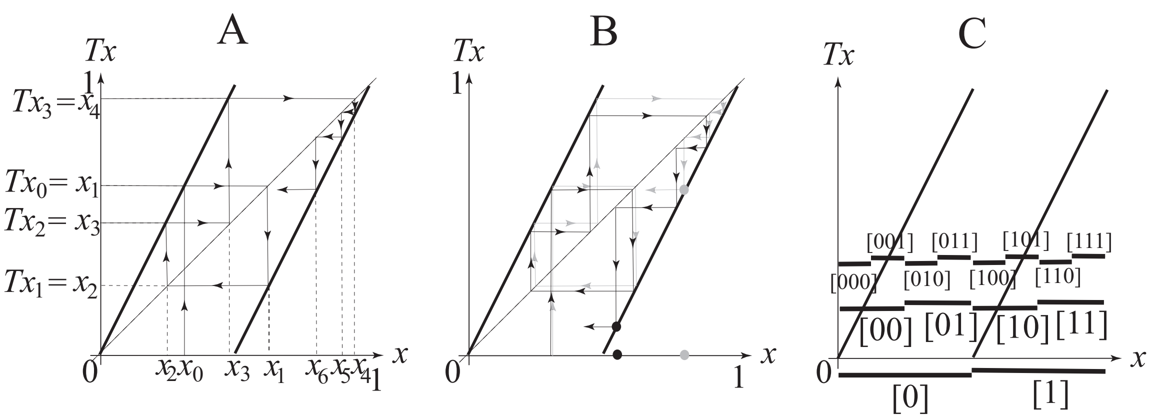

Figure 22.2: A simple map $Tx = \{2x\}$ that produces chaos.

Fig. 22.2 A simple map $Tx = \{2x\}$ that produces chaos

A: ⟪**How to chase history graphically**⟫ A method to chase the trajectory which is determined by the initial condition x_0 on the graph is illustrated. The broken line diagonal denotes $y = x$ where the values on the horizontal and vertical axes coincide. Oblique thick parallel lines denote the graph of $y = Tx$. If an initial condition x_0 is given on the horizontal axis, look vertically upward to find the point on the graph of T. Its vertical coordinate is x_1. To find x_2, we must find x_1 on the horizontal axis, and then $Tx_1 = x_2$ may be obtained just as before. To this end, with the aid of the diagonal, we can fold the vertical axis onto the horizontal axis and locate x_1 on the latter. Therefore, if we chase the vertical or horizontal lines with an arrow we can successively find $x_2, x_3, \cdots$.

B: ⟪**Exponential separation of nearby trajectories**⟫ Doubling of the gap is shown due to each

[243] YO TNW Appendix 2.1A

application of T between the gray and black trajectories that are initially very close. The positions after 6 applications of T are denoted by the gray and black small disks on the horizontal axis.
C: ⟪**Coding of trajectory**⟫ How to convert a trajectory into a 01 symbol sequence is illustrated. The interval $[0, 1]$ is divided into two and each is named [0] or [1]?[0] is further subdivided into [00] and [01]. T maps [00] onto [0] and [01] onto [1]. If the subdivision of [10] is named [100] or [101] as in the figure, T maps [100] onto [00] and [101] onto [01]. If we recursively use this prescription, each point in $[0, 1]$ becomes correspondent to a particular 01 infinite sequence. This is nothing but the binary expansion of a number in [0,1] (however, $[111\cdots]$ is not identified with [0]).

⟪**History is determined by initial condition**⟫ If an initial condition is given, an indefinitely long sequence $\{x_n\}$ may be constructed as $x_1 = Tx_0$, $x_2 = Tx_1 = T^2x_0, \cdots$; the future is perfectly determined by x_0. How to chase this trajectory graphically is illustrated in Fig. 22.2A.

⟪**Chaos directly connects the unknown world to our world**⟫ Fig. 22.2B illustrates how chaos expands the world of microscopic scales and connect it to the world we can directly observe. In this example, the microscopic world is doubled every time T is applied. Therefore, although the system we consider at present is a deterministic system, we lose our predictive power. Roughly speaking, for large n x_n becomes indistinguishable from a random number.

⟪**Coding trajectories or correspondence to number sequence**⟫ To see the random nature of chaotic trajectories more explicitly, their discrete coding $s_0, s_1, \cdots$ is introduced in terms of symbols s_n that take 0 or 1. It is a rather obvious transformation in this case; the interval $[0, 1]$ is divided into two intervals, $[0, 1/2]$ and $(1/2, 1]$, and we call them, [0] and [1], respectively. Generally, $[s_0s_1, \cdots s_n]$ is defined as a set $\{x : T^kx \in [s_k]$ for $k = 0, \cdots n\}$, where s_k are 0 or 1. In Fig. 22.2C,[244] we see that if we apply T once to the interval, say, [011], it is mapped onto [11].
One more application of T to it coincide [1] (it is easy to see this, if we chase the end points as explained in Fig. mod2mapA). That is, [011] is a bundle of trajectories for which x_0 is in [0], x_1 is in [1] and x_2 is in [1] (Such a bundle of trajectories is called a cylinder set.).

⟪**How prediction becomes difficult**⟫ If we follow the way to construct small intervals in Fig. 22.2C, we see that, for example, the bundle of trajectories named as [0011010011] (there are 10 digits) is mapped by T successively as [011010011] $\to$ [11010011] $\to$ [1010011] $\to$ [010011] $\to$ [10011] $\to \cdots \to$ [011] $\to$ [11] $\to$ [1](digits are lost one by one from the left end). One more application of T makes the trajectories 'all over' $[0, 1]$. The lesson we have learned is that even if the initial condition is in

[244]The reader might worry about to which interval the boundary points belong, but in this example, no careful assignment of symbols is important.

the interval [0011010011] of width $1/2^{10} \sim 10^{-3}$, after 10 consecutive T map applications we will totally lose the location information. Perhaps, the reader might guess that we will do much better if we specify the initial point (the cylinder set) more accurately. If we wish to keep some knowledge of the initial position for 20 seconds, we need the accuracy of unrealistic $1/2^{20} \sim 10^{-6}$. In short, sooner or later we will fail to predict the behavior of the system.
Even if we fail to make any prediction, it does not imply the end of the world. Beyond the predictable time range what determines x_n? The system is deterministic, so it is determined by the far right portion of the 01 sequence obtained by coding of the initial condition, which we can never know beforehand. Isn't it virtually the same as an arbitrary 01 sequence? Then, after a while, chaotic behavior would be indistinguishable from the head-tail sequence obtained by tossing a coin. This is the intuitive meaning of the definition of chaos given later.

⟨⟨Chaos is a random deterministic behavior⟩⟩ After all, it is a natural idea that the essence of chaos is randomness with a tint of initial condition effects due to determinism. However, without carefully reflecting on the concept of randomness, we cannot express this intuition precisely. After a rather 'heavy' preparation, eventually we will arrive at the conclusion that chaos is a phenomenon that deterministic trajectories exhibit randomness.

⟨⟨Quantitative correspondence of chaos and randomness⟩⟩ The reader may think the above argument provides only a qualitative characterization of chaos. However, there is a way to quantify randomness, which allows us to make a quantitative correspondence between chaos and randomness. This quantification is realized through quantifying the needed information to predict the future with a predetermined fixed accuracy. This is not hard in terms of the model being considered. How the information we know at the initial time becomes insufficient to describe the system behavior can be seen almost explicitly from Fig. 22.2C and loss of digits from the 'cylinder sets' due to the application of T. The information is lost by 1 bit every time T is applied. Suppose we wish to predict the position of the point at time t in the future with the same accuracy we describe the system now ($t = 0$). The information we must prepare now increases by 1 bit, if we push the future time t further to $t + 1$. The increasing rate of the needed information (1 bit per unit time in the present example) is called the Kolmogorov-Sinai entropy (already discussed informally). On the other hand, to describe an arbitrary 01 sequence we of course need 1 bit per digit. That is, the needed information to describe a trajectory is 1 bit per unit time. This equality of the amounts of information needed to predict the future and to describe the trajectory is a general assertion of Brudno's theorem (**24.4**). Thus, we may conclude that our definition of chaos is right on the mark.

22.5 Chaos in the logistic map

The logistic map: $[0,1] \to [0,1]$ is given by

$$F(x) = 4x(1-x). \tag{22.7}$$

We can define a dynamical system on $[0,1]$ as $x_{n+1} = F(x_n)$. Run the system for some intial conditions x_0. You will find quite erratic sequences $\{x_k\}_{k=0}^{\infty}$.

There is a very important warning: whether we can have an explicit analytical expression for x_n as a function of x_0 and n and whether the system exhibits chaotic behaviors are generally not logically related. Indeed, for the logistic map we have an explicit formula

$$x_n = \sin^2[2^n \mathrm{Arcsin}\sqrt{x_0}]. \tag{22.8}$$

You might think that these sequences can be predicted for any far future time thanks to such analytic expressions, but try to evaluate this for three digits with $n = 100$.

There are many other examples.[245]

22.6 Coin-tossing, deterministic or not

Throwing a coin many times, we can construct a sequence of heads and tails. If we denote 'head' by 1 and 'tail' by 0, we have a 01 sequence. In this case, the phase space is $\Gamma = \{0,1\}$ and the time set is $\boldsymbol{T} = \mathbb{N}^+ \equiv \{1, 2, \cdots\}$ (i.e., discrete). As its path space Ω we may choose the totality of one-sided 01 infinite sequences $\Omega = \{0,1\}^{\mathbb{N}^+}$, because heads and tails can appear in any order. The resultant dynamical system (σ, Ω) is called the coin-tossing process. Let $\omega \in \Omega$. If $\omega(n)$ happens to be 0, this implies that the n-th outcome is a tail. Even if we know a history of this system up to time n: $\omega(1), \cdots, \omega(n)$, no one can predict anything beyond. Therefore, "the system is not deterministic."

22.4 already mentioned that the sequences we can get from 'the simplest' chaotic map $Tx = \{2x\}$ correspond (one-to-one) to all the possible outcomes of a 'coin tossing process.' You might say one is deterministic, and the other stochastic. However, there is a basic question: Can you distinguish deterministic and non-deterministic systems from their trajectories alone?

It is possible to regard the difference between the deterministic and non-deterministic dynamical systems is due to the difference of view points. Take a discrete dynamical system (one-sided), whose history starting from $\omega(0)$ $\omega \in \mathcal{D}$ may be written as

[245] K. Umeno, "Method of constructing exactly solvable chaos," Phys. Rev. E **55**, 5280 (1997).

$\omega(0)\omega(1)\omega(2)\cdots$. $\omega(t)$ is the state at time t. In our simplest example. The whole sequence $\omega(0)\omega(1)\omega(2)\cdots$ is just the binary expansion of the initial position $x \in [0,1]$, perfectly deterministic, although, of course, you cannot tell even $\omega(1)$ from $\omega(0)$ alone.

Now, let us define the shift operator (or simply, *shift*) $\sigma : \mathcal{D} \to \mathcal{D}$ as

$$(\sigma\omega)(t) = \omega(t+1). \tag{22.9}$$

The shift is a vehicle to experience the history in the chronological order. Suppose, for example, the phase space is $\Gamma = \{0,1\}$ and its one possible history ω starting from $\omega(0) = 0$: $001010011101001101\cdots$, where the leftmost number is interpreted as the state we observe at present. Its time evolution is

$$\omega = 001010011101001101\cdots \tag{22.10}$$

$$\sigma\omega = 010100111010011010\cdots \tag{22.11}$$

$$\sigma^2\omega = 101001111100110101\cdots. \tag{22.12}$$

The currently observed state evolves step by step. The state $\omega(t+1)$ may not be determined by the states up to time t: $\cdots, \omega(t-1), \omega(t)$, so for the observer who is observing the current and the (recorded) past states only the system behavior does not look deterministic.

Instead of interpreting the shift as a vehicle to experience a history chronologically, however, if we interpret ω as a whole history (chronicle), the shift maps one chronicle to another: $\sigma\omega = \omega_1$, where ω_1 gives the chronicle one time unit ahead of ω; indeed, for all time $t \in \boldsymbol{T}$ $\omega_1(t) = \omega(t+1)$. σ is understood as a map from $\mathcal{D}$ into itself; it defines a discrete dynamical system on the path space: $(\sigma, \mathcal{D})$. (22.10)-(22.12) illustrated its time evolution. Here, we observe the history not around present, but observe histories from God's point of view. Notice that ω completely determines ω_1. The dynamical system $(\sigma, \mathcal{D})$ is obviously deterministic.

22.7 Actual coin tossing

J. STRZALKO, J. GRABSKI, A. STEFANSKI, P. PERLIKOWSKI AND T. KAPITANIAK, Understanding coin-tossing, Math Intel 32(4), 54 (2010) illustrates the basin of attraction of the actual coin.

Fig. 22.3 Basins of attraction indicating the face of the coin which is up after the n-th collision: (a) n = 0, (b) n = 3, (c) n = 10, heads and tails are indicated in black and white, respectively. The initial conditions are $x_0 = y_0 = 0$; $\dot{x}_0 = \dot{y}_0 = \dot{z}_0 = 0$; $\varphi_0 = \psi_0 = 0$; $\theta_0 = 7$ deg, $\omega_{\zeta_0} = 0$, $\omega_{\eta_0} = 40.15$ rad/s.
Right: 3-dimensional model of the coin and its orientation in space.

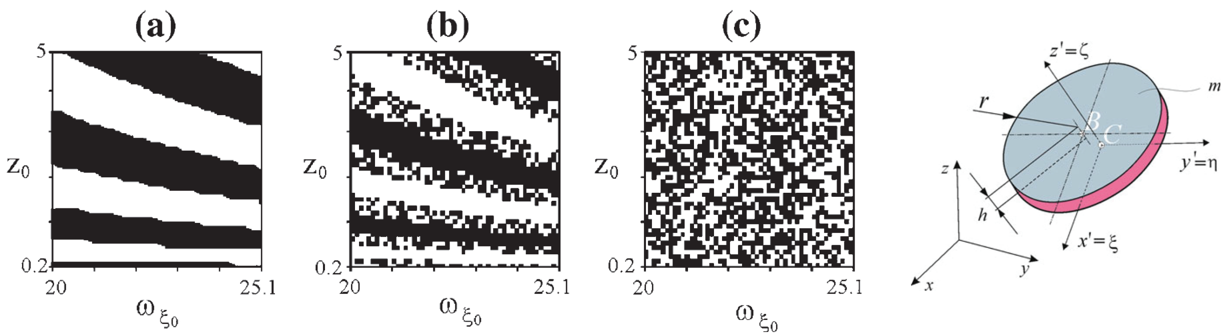

Figure 22.3: [Fig. 1,3 of STRZALKO et al., Math Intel 32(4), 54 (2010)]

22.8 No true randomness without quantum mechanics?

As you have realized, the randomness in a deterministic systems is totally in the initial conditions. Therefore, whether we can sample the intial condition ‘totally randomly’ or not is a crucial question as to the (un)predictability of chaos. Some people[246] assert that without quantum mechanics (so, essentially, the Born rule) there cannot be any true randomness.

Is this assertion meaningful?

22.9 Use of ‘symbolic dynamics’

As we will discuss later we study a dynamical system with the aid of a certain coding system for the system *(called a symbolic dynamics). $\{\sigma, \{0,1\}\}$ is the symbolic dynamical expression (in this case very faithful: isomorphic) of $\{T, [0,1]\}$, where $Tx = \{2x\}$. Thus, we study symbol sequences, so the relevance of information coded in the history is obvious. In our simple example, all the symbols are equally probable, so are all the words (finite sequences of symbols). However, this is not generally the case, so we must consider ‘natural’ probabilities associated with words. Thus, we are injecting a new ingredient in our study of nonlinear systems, measures.

22.10 Definition of chaos[247]

Let f be a map from the phase space M into itself (i.e., an endomorphism of Γ).[248]

[246] For example, T. Sagawa (of (quantum) information thermodynamics). When I heard this assertion from him for the first time, I thought it is a deep idea. Is it really so?

[247] The following several units are taken from TNW Chapter 2.

[248] In this definition, maps are not understood measure-theoretically; they are pointwise transfor-

Choose $n \in \mathbb{N}$ and construct f^n that applies f n-times,[249] and restrict it to its invariant set $A \subset M$ ($f^{-n}(A) = A$). If it is isomorphic to the shift dynamical system $(\sigma, \{0,1\}^{\mathbb{N}})$, then we say the dynamical system (f, M) exhibits chaos (or the system is chaotic). That is, if φ is a one to one map and the following diagram becomes commutative,[250] we say the dynamical system (f, M) exhibits chaos.

$$\begin{array}{ccc} A & \xrightarrow{f^n} & A \\ \downarrow \varphi & & \downarrow \varphi \\ \{0,1\}^{\mathbb{N}} & \xrightarrow{\sigma} & \{0,1\}^{\mathbb{N}} \end{array}$$

In words, if an appropriate one-to-one coding scheme φ of points (i.e., states) in A (we can decode uniquely the sequences) allows to transform the original dynamical system (restricted to A) into the full shift on two symbols, we say the dynamical system is chaotic.

Roughly speaking, if the behavior of a dynamical system (restricted on an invariant set) is coded with symbols 0 and 1, and the result cannot be distinguished from the totality of the outcomes of the coin-tossing process, we wish to say the system is chaotic or is a chaotic dynamical system.

A continuous time dynamical system is chaotic, if a discrete dynamical system constructed from the original system in a 'natural' fashion is chaotic in the above sense.

The chaos defined here is sometimes called formal chaos, because there is no guarantee of its observability **22.3**. For chaos of one-dimensional map systems, the observability of chaos and the existence of an absolutely continuous invariant measure seem to be equivalent.

22.11 How good is the definition of chaos given above?

It is not easy to find criteria for a definition to be good, but consistency with intuition, close connection to fundamental concepts, equivalence to definitions based on very different points of view, etc., may be counted among them. Our definition relies on our intuition: as long as we assume that "(apparent) random behavior is fundamentally important characteristic of chaos," consistency with intuition is built into the definition. Randomness must be a fundamental concept, so the definition

---

mations.

[249] $f^n(x) = (f \circ f \circ \cdots \circ f)(x) = f(f(f \cdots (f(x)) \cdots))$. n f's show up in each expression.

[250] This means in the present context that the results obtained by following various combinations of arrows never yield different results.

has a close connection with a fundamental concept. This can be seen further from the theorem we demonstrate later. It is also important to check the consistency and relations with other definitions of chaos to confirm the naturalness of the definition.

22.12 Sensitive dependence on initial conditions

As can be seen from (22.8) and as emphasized by Ruelle, Guckenheimer[251] and others, magnification of small effects was regarded as a fundamental significance of chaos (the "butterfly effect"). Intuitively, this is easy to understand in terms of (one-sided) shifts, because, in contrast to the digits on the far right that describes minute differences in states, the digits near the left end correspond to global differences. Thus, movement of the digits to the left by shift corresponds to magnification of small structures. Two initial states whose codes are different only in digits on the far right are very close in the phase space, and time evolution magnifies the difference.

Sensitive dependence on initial conditions, however, does not necessarily imply that the dynamical system is chaotic. This must be obvious to those who know the roulette; the ball jumps around awhile, but eventually it settles down to a fixed point (the system is a multiply stable system). If several attractors coexist, the ultimate fate of the system is determined by which basin its initial condition lies. Even if the long-time (eventual) behavior of the system is not chaotic, if the boundaries between basins are extremely complicated, then sensitive dependence on initial conditions can exist.[252] Obviously, roulettes, dice, and coins must be (at least approximately) such systems. Sensitive dependence on initial conditions is not enough to characterize chaotic dynamical behavior.[253]

22.13 Apparently different definitions of chaos

S.-H. Li proposed a revised version of scrambled set called ω-scrambled set to make the Li-Yorke chaos equivalent to our chaos.[254]

The most popular definition of chaos at present may be due to Devaney.[255] If we take into account S.-H. Li's result,[256] this definition may be stated as

[251] J. Guckenheimer, "Sensitive dependence on initial conditions for one-dimensional maps," Commun. Math. Phys., **70**, 133 (1979).

[252] See, for example, H. E. Nusse and J. A. Yorke, "Basin of attraction", Science **271**, 1376 (1996).

[253] However, if we require sensitivity to perturbation at most instants may characterize chaotic systems.

[254] S.-H. Li, "ω-chaos and topological entropy," Trans. Amer. Math. Soc. **339**, 243 (1993).

[255] R. Devaney, *An introduction to chaotic dynamical systems* (Benjamin/Cummings, 1986).

[256] S.-H. Li, "Dynamical properties of the shift maps on the inverse limit space," Ergodic Theor. Dynam. Syst. **12**, 95 (1992). See also J. Banks, J. Brooks, G. Cairns, G. Davis and P. Stacey, "On

follows. Let (f, X) be a discrete time dynamical system, and D $(\subset X)$ be a closed invariant set (i.e., $f^{-1}(D) \supset D$). If the following two conditions hold, the dynamical system exhibits chaos.
(D1) $f|_D$ (the restriction of f to D) is topologically transitive on D (i.e, $f|_D$ is surjective on D and has an orbit dense in D),
(D2) The totality of the periodic orbit of f is dense in D.
D is called a chaos set. The set A in **22.10** is a chaos set. S.-H. Li showed the equivalence of this definition and our definition, if the phase space is a compact metric space.

Already in 1968[257] Alekseev defined quasirandom dynamical systems in terms of a Markov chain with a positive Kolmogorov-Sinai entropy. It is an example that Russian dynamical systems study was far ahead of the Western counterpart (actually, the use of entropy to classify dynamical systems was the Russian starting point of the 'modern' dynamical systems study). The term 'chaos' may have been good for popularization of the concept, but the term 'quasirandom' summarizes the essence.

22.14 Period $\neq 2^n$ implies chaos

Theorem[258] Let I be a finite interval and $F : I \to I$ be a continuous endomorphism. F exhibits chaos if and only if F has a periodic orbit whose period is not equal to the power of 2.

Remark We have already mentioned the Li-Yorke chaos **22.2**, but the chaos in the above theorem is distinct from the Li-Yorke chaos. The most important distinction is, as note below, with slight weakening of the 'one-to-one' correspondence with A chaotic behaviors are observable; we can choose the set A in **22.10** observable (e.g., measure positive). That is, in contradistinction to Li-Yorke chaos whose 'core feature' = the scrambled set is never observable, our chaos is closely tied to observability, when chaos is observable. We can show that if the system is chaotic in our

---

Devaney's definition of chaos," Am. Math. Month. **99**, 332 (1992). In the latter X need not be bounded.

[257][1968: Prague Spring, The Tet Offensive, Down to the Countryside Movement, Nuclear non-proliferation treaty, Assassinations of Rev. M. L. King and Senator R. F. Kennedy.]

[258]Y. Oono, Period $\neq 2^n$ implies chaos, Prot. Theor. Phys., 59, 1029 (1978). "In the present Letter, we give another definition of chaos which is not directly related to the nonperiodicity of the solution, and sketch the proof of the theorem asserting that period other than 2" implies chaos (Theorem 1). The assertion also holds even if the definition of chaos by Li, Yorke and Nathanson [J. Combinatorial Theor. (A) 22 61 (1977)] is adopted (Theorem 2)." However, the proof in this letter is not very elegant.

sense, the system is also Li-Yorke chaotic, BUT the converse is not generally true. Furthermore, even when chaos is observable, it is not on the scrambled set.

We will show the following theorem:

22.15 General theorem for chaos of C^0-endomorphisms of intervals

Theorem[259] Let I be a finite interval and $F : I \to I$ be a continuous endomorphism. Then, the following (1)-(4) are equivalent.[260]
(1) F exhibits chaos.
(2) F has a periodic orbit whose period is not equal to the power of 2.
(3) There is a positive integer m such that F^m has a mixing invariant measure.
(4) F has an invariant measure whose Kolmogorov-Sinai entropy is positive.

An intuitive explanation of the theorem is given here.

"Invariant measure" is a steady distribution.

"Mixing" implies that the system relaxes toward some steady state.

"Kolmogorov-Sinai entropy" is the required extra information to predict the next time step state as accurately as the current state. Its positivity implies that (since more information is needed to determine the future state) the system behavior in the future becomes increasingly difficult to predict as the future is further away.

Practically, the following Proposition equivalent to (1)-(4) in the theorem is useful:
(5) There are two closed intervals J_1 and J_2 in I that share at most one point such that $f^p(J_1) \cap f^q(J_2) \supset J_1 \cup J_2$ holds for some positive integers p and q.

That Ito's earthquake model in Section 2.1 exhibits chaos is immediately seen from the 'folded paper' model (Fig. 18.12). We may draw a periodic orbit whose period is not a power of 2, but to check (5) may be the easiest.

Is A observable? The set A constructed in the proof is not observable, since it is measure zero and not attractive. However, if we ignore the 1 to 1 correspondence on a measure zero set, we can make A to be an interval (very often, especially when chaos is observable).

22.16 (2) implies (5)

[259] M. Osikawa and Y. Oono, "Chaos in C^0-diffeomorphism of interval," Publ. RIMS **17**, 165 (1981); Y. Oono and M. Osikawa, "Chaos in nonlinear difference equations. I," Prog. Theor. Phys. **64**, 54 (1980) (There is no part II, since the author was 'expelled' from Japan in 1979). Also see: L. Block, "Mappings of the interval with finitely many periodic points have zero entropy," Proc. Amer. Math. Soc. **67**, 357 (1978); "Homoclinic points of mappings of the interval," *ibid.* **72**, 576 (1978).

[260] (1) $\Rightarrow$ (2), (3), (4) is trivial. (4) $\Rightarrow$ (1) is also almost trivial.

If F has a periodic orbit whose period is not equal to the power of 2, then there are two closed intervals I_0 and I_1 in I that share at most one point such that $F^p(I_0) \cap F^q(I_1) \supset I_0 \cup I_1$ holds for some positive integers p and q.

[Demo][261]

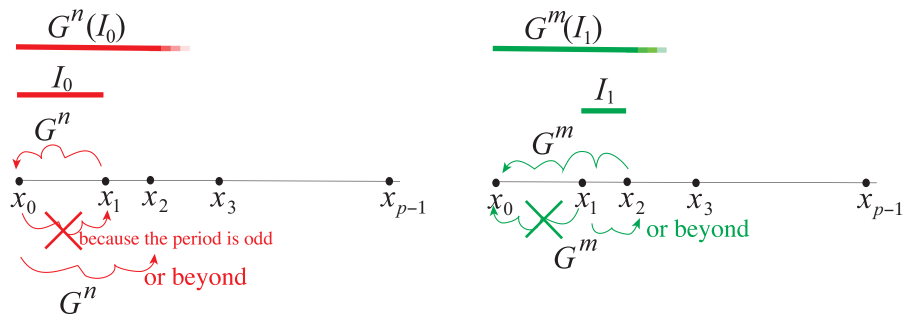

Figure 22.4: Illustrated proof of (2) ⇒ (5). (Due to possible folding, the open of kernels of the images of the intervals may well contain $I_0 \cup I_1$.)

$G \equiv F^{2^n}$ has a periodic orbit with odd periodicity p: $O_p = \{x_0, x_1, \cdots, x_{p-1}\}$, where $x_0 < x_1 < \cdots < x_{p-1}$ (Needless to say, the orbit is not necessarily chronologically in this order).

Let $I_0 = [x_0, x_1]$ and $I_1 = [x_1, x_2]$. There must be an integer n $(0 < n < p)$ such that $G^n(x_1) = x_0$. For this n $G^n(x_0) \neq x_1$; otherwise, $G^{2n}(x_1) = x_1$, that is, $p = 2n$.[262] Therefore, $G^n(x_0) \geq x_2$. That is, $G^n(I_0) \supset I_0 \cup I_1$.

There must be an integer m $(0 < m < p)$ such that $G^m(x_2) = x_0$. For this m obviously $G^m(x_1) \neq x_0$ nor $G^m(x_1) \neq x_1$, so $G^m(x_1) \geq x_2$. That is, $G^m(I_1) \supset I_0 \cup I_1$.

22.17 (5) implies (1)

If there are two closed intervals I_0 and I_1 in I that share at most one point such that $F^p(I_0) \cap F^q(I_1) \supset I_0 \cup I_1$ holds for some positive integers p and q, then F exhibits chaos.

[Demo]

The key to this assertion consists of two parts:

(i) If there are two disjoint closed intervals I_0 and I_1 in I such that $F^n(I_0) \cap F^n(I_1) \supset I_0 \cup I_1$ for some positive integer n, then F exhibits chaos.

(ii) (5) implies the existence of I_0 and I_1 required by (i).

Let us prove (i) first. Set $G = F^n$.

---

[261]Proof due to Hamachi.

[262]Note that $2n < 2p$, the smallest even number that is a multiple of p.

First we note the following elementary fact: (ia) Let I_0 be a closed interval, and $G(I) \supset I_0$ for a continuous function G. There there is a closed interval $Q \subset I$ such that $G(Q) = I_0$. It is essentially the intermediate value theorem.

Next, we note that (ib) if $G(I) \supset I_0, I_1$, which are disjoint closed intervals, then we can find $G(Q_i) = I_i$ ($i = 1$ or 2) such that $Q_i \subset I$ and $Q_1 \cap Q_2 = \emptyset$. The existence of Q_1 and Q_2 in I follows from the preceding elementary fact. If $Q_1 \cap Q_2$ were not empty and had x in it, then $f(x) \in I_0 \cap I_1$, contradicting the disjointness of I_0 and I_1

We can recursively construct a closed interval sequence $\{I_{a_1 a_2 \cdots a_n}\}_{n=1}^{\infty}$ where $a_i \in \{0, 1\}$ as

$$G(I_{a_1 a_2 \cdots a_n}) = I_{a_2 \cdots a_n}, \tag{22.13}$$

where

$$I_{a_1 a_2 \cdots a_{n-1} a_n} \subset I_{a_1 a_2 \cdots a_{n-1}} \tag{22.14}$$

with the aid of (ia).

If s and s' ($\neq s$) are the length n 01 sequences. Then (ib) tells us $I_s \cap I_{s'} = \emptyset$.

Let $b \in [0, 1)$ and its binary expansion $b = 0.a_1 a_2 \cdots a_n \cdots$. We can define I_b as

$$I_b = \cap_{n=1}^{\infty} I_{a_1 a_2 \cdots a_n}, \tag{22.15}$$

which is closed nonempty set thanks to (22.14) and if $b \neq b'$, $I_b \cap I_{b'} = \emptyset$. Thanks to (22.13)

$$G(I_b) = I_{\{2b\}}, \tag{22.16}$$

where $\{x\}$ is the fractional portion of a real x (i.e., $2b = a_1.a_2 \cdots a_n \cdots$, so $\{2b\} = 0.a_2 a_2 \cdots a_n \cdots$)

Define $A = \cup_{b \in [0,1)} I_b$. Introduce an equivalence relation $\sim$ on A as $x \sim x'$ iff $x, x' \in I_b$ for some $b \in [0, 1)$. Then, there is a standard surjection $\tau : A \to A/\sim$, and if we define $G^* = \tau \circ G \circ \tau^{-1} : A/\sim \to A/\sim$, it is isomorphic to the (one-sided) Bernoulli shift on $\{0, 1\}$ of the coin-tossing process.

We must prove (ii). Let us take an inelegant but the simplest path. If we take $m = pq$, obviously

$$F^n(I_0) \cap F^n(I_1) \supset I_0 \cup I_1. \tag{22.17}$$

Let $G = F^m$ and construct $I_{a_1 a_2}$ as in the proof of (i). They must be on I, so we can always choose a pair $I_{a_1 0} = J_0$ and $I_{a_1' 1} = J_1$, where $a_1, a_1' \in \{0, 1\}$m, such that they are not adjacent. Then, if we choose $G^2 = F^{2m}$ (i.e, $n = 2pq$)

$$G^2(J_0) \cap G^2(J_1) \supset J_0 \cup J_1. \tag{22.18}$$

Remark Actually, we can prove the odd periodicity lemma:
If G has a periodic orbit of odd periodicity, then there exists two disjoint closed intervals I_0 and I_1 such that $G^2(I_0) \cap G^2(J_1) \supset I_0 \cup J_1$.

Therefore, if F has a periodic orbit of period $2^n p$, $F^{2^{n+1}}$ has such intervals.

22.18 (1) (2) and (5) are equivalent

This is shown via (1) $\Rightarrow$ (2). Then, **22.16** closes the demonstration with **22.17**.

22.19 (1) and (3) are equivalent

There is an F^m invariant mixing measure μ for some positive integer k:

$$\lim_{n \to infty} \mu(A \cap F^{-nm} B) = \mu(A)\mu(B), \tag{22.19}$$

if and only if F is chaotic.[263]

22.20 (1) and (4) are equivalent

(4) $\Rightarrow$ (1) is shown via (2). The converse is generally hard to prove (proved with H Takahashi's results with his help).

22.21 Time correlation function of chaos

Let x_n be the state at time step n. The, the time correlation function is defined as (here we assume the time average of x_n vanishes or $x_n - \langle x_n \rangle$ is considered)

$$c(n) = \frac{1}{N} \sum_{k=0}^{N-1} x_{n+k} x_k. \tag{22.20}$$

We are interested in this quantity when the system allows a steady state.
If the map (or dynamics) allows an ergodic invariant measure μ,[264],

$$c(n) = \int d\mu(x_0) x_n x_0 \quad \left[\text{or} = \int d\mu(\omega) x_n(\omega) x_0(\omega) \text{ is better}\right]. \tag{22.21}$$

$c(n) \le c(0)$, because the average of $(x_n - x_0)^2$ must be non-negative (use stationality). Therefore, we often normalize this with $c(0)$:

$$C(n) = \frac{\langle x_n x_0 \rangle}{\langle x_0^2 \rangle}. \tag{22.22}$$

[263] This is a special assertion generally true for C^0-endomorphism of an interval.

[264] We assume it is observable; thus, it is almost surely an absolutely continuous invariant measure.

If $C(n) \to 0$ in the large n limit, we say the system is mixing,[265] where $\langle \ \rangle$ implies time or invariant measure average.

We usually think that the faster the time correlation function decays, the more irregular/disordered/chaotic is the dynamical system. However, the situation is not that simple, because isomorphisms of dynamical systems can change the decay rate (or even its algebraic nature), but the KS entropy stays the same.[266]

For the standard tent map its time correlation decays to zero at time 1 (!), but for the logistic map, its decay is not that quick (though exponential). Both are mixing and the KS entropy is log 2, the same.

22.22 Power spectrum

Experimentally (and practically), the time correlation function is computed via its Fourier transform called the power spectrum:

$$\sigma(\nu) = \sum_n C(n) e^{-2\pi i n \nu}. \tag{22.23}$$

This strategy is advantageous, because the Fourier transform $\hat{x}(\nu)$ of a signal can be efficiently computed with the aid of the Fast Fourier Transform algorithm. The square (average) of the $|\hat{x}(\nu)|^2$ is the power spectrum, and its inverse Fourier transform is the time correlation function thanks to the Wiener-Khinchin theorem **22.23**.

22.23 Wiener-Khinchine theorem[267]

Let $x_n(\omega)$ be a sampled trajectory with summable stationary time correlation function $C(n) = \langle x_n x_0 \rangle$ (we normalize the signal). Let us compute its power spectrum $\sigma(\nu)$.

We compute the Fourier transform of the signal as

$$\hat{x}(\nu) = \sum_{-\infty}^{\infty} x_n e^{-2\pi i n \nu}. \tag{22.24}$$

Let us compute

$$\langle \hat{x}(\nu)\hat{x}(\nu') \rangle = \sum_n \sum_m C(n-m) e^{-2\pi i n \nu - 2\pi i m \nu'}, \tag{22.25}$$

[265] A more official definition of mixing will be given later, when we discuss ergodicity.

[266] The mixing property is isomorphism invariant, so the correlation is guaranteed to decay eventually to zero.

[267] The needed normalization is not carefully traced.

$$= \sum_m \sum_p C(p) e^{-2\pi i(\nu+\nu')m - 2\pi i\nu p}, \tag{22.26}$$

$$= \delta_{\nu,\nu'} \sum_p C(p) e^{-2\pi i\nu p} = \delta_{\nu,\nu'}\sigma(\nu). \tag{22.27}$$

That is,

$$C(n) = \sum \sigma(\nu) e^{2\pi i n\nu}. \tag{22.28}$$

This is called *Wiener-Khinchine's theorem*.

This is a practically very important theorem, because the equality (22.28) is the most convenient method to compute the correlation function: the power spectrum is easy to compute numerically from the original data thanks to the fast Fourier transformation (FFT). If you have never heard of the FFT, study the Cooly-Tukey algorithm.[268] 32B.12 of

https://www.dropbox.com/home/ApplMath?preview=AMII-32+FourierTransformation.pdf

explains the principle of FFT.

Note that the time correlation function is positive definite in the sense appearing in Bochner's theorem **22.25**). This implies that if the correlation function is continuous at the origin and if normalized as $C(0) = 1$, it is a characteristic function of a certain probability distribution.

We know that this 'certain function' is the corresponding power spectrum. Therefore, the power spectrum may be interpreted as a probability distribution function (of a certain quantity). Then, the information minimization technique may be used to infer or model the time correlation function.

22.24 Long correlation means sharp spectrum

As we have learned, the power spectrum is easy to observe,[269] so it is advantageous to have some 'feeling' about the relation between correlation functions and power spectra.

A summary statement is:

"A signal that has a short time correlation has a broad band."

This is intuitively obvious, because short time correlation means rapid changes that must have high frequency components.

There may be several mathematical statements that are related to the above assertion. Perhaps the most elementary is to look at

$$C(t) = e^{-\alpha|t|} \leftrightarrow \sigma(\omega) = \frac{2\alpha}{\omega^2 + \alpha^2}. \tag{22.29}$$

[268] As usual, Gauss used this method long before Cooly and Tukey.

[269] Simply turn on a spectrum analyzer.

That is, the half-width of the power spectrum is the reciprocal of the 'life-time' of the signal. (The power spectrum with the Cauchy distribution is often called the *Lorentzian spectrum*)

The result can be understood intuitively with the aid of dimensional analysis $\tau\omega \sim 1$.

Exercise 1. The *Riemann-Lebesgue lemma* tells us that $\lim_{|\omega|\to\infty} \sigma(\omega) = 0$ for time-continuous signals. Learn about this lemma. A related statement was used in the demo of the KAM theorem.□

Exercise 2. The Riemann-Lebesgue lemma also tells us that the decay rate of the power spectrum for large $|\omega|$ gives information about the smoothness of the signal.

Demonstrate that if the signal is m-times continuously differentiable with respect to time, $\sigma(\omega) = o(\omega^{-2m})$ for large ω.

22.25 Bochner's theorem.

A positive definite function φ on $\boldsymbol{R}^n$ that is continuous at the origin and $\varphi(0) = 1$ is a characteristic function of some probability measure on $\boldsymbol{R}^n$.□

Here, 'positive definite' means the following: for any $n \in \mathbb{N}$, for any $a_i \in \mathbb{C}$ $(i = 1, \cdots, n)$ and for any 'time points' $t_i \in \mathbb{R}$

$$\sum_{i,j \leq n} a_i \overline{a_j} \varphi(t_i - t_j) \geq 0. \tag{22.30}$$

22.26 Sarkovskii's theorem[270]

Sharkovsky ordering of the set of natural numbers is given by the following ordering

$$\begin{aligned} &3 \prec 5 \prec 7 \prec 9 \prec \cdots \prec 3 \cdot 2 \prec 5 \cdot 2 \prec 7 \cdot 2 \prec 9 \cdot 2 \prec \\ &\quad \cdots \prec 3 \cdot 2^2 \prec 5 \cdot 2^2 \prec 7 \cdot 2^2 \prec 9 \cdot 2^2 \prec \cdots \prec 2^3 \prec 2^2 \prec 2 \prec 1. \end{aligned} \tag{22.31}$$

Let I be either the real line or an interval and $f : I \to I$ be a continuous map. Let us say $a \prec b$, if a precedes b in the Sharkovskii ordering. The three parts of the full Sharkovsky Theorem are:

[270] A. N. Sarkovskii, "Coexistence of cycles of a continuous map of a line into itself," Ukr. Mat. Z. 16, 61-71 (1964) [See P. Stefan, "A theorem of Sarkovski on the existence of periodic orbits of continuous endomorphisms of the real line," Commun. Math. Phys. 54, 237 (1977) as well]. M. Misiurevicz, "Remarks on Sharkovsky's theorem," Am. Math. Month. 104 864 (1997) is a good summary. Here, the summary follows the third paper.

Theorem 1. Let $f : I \to I$ be a continuous map. If f has a cycle of period n and if $n \prec k$, then f has a cycle of period k.
Theorem 2. For every k there exists a continuous map $f : I \to I$ that has a cycle of period k, but has no cycles of period n for any $n \prec k$.
Theorem 3. There exists a continuous map $f : I \to I$ that has a cycle of period 2^n for every n and has no cycles of any other periods (the existence of the critical map). [This part is actually included in Theorem 2.]

Sharkovskii's theorems are beautiful theorems, and tell us clearly the existence of the critical map. Also they tell us that having a periodic orbit of odd period is the essence of the complicated trajectories as Hamachi's proof of the key lemma **22.16** exploits. However, the theorems tell us virtually nothing about chaos, so I will not give a proof.