24 Lecture 24. Characterization of chaos

24.1 Basic motivation

Chaos is characterized by 'complicate' or 'apparently random' trajectories. We have refined the concept of randomness in the preceding section. 'Algorithmic randomness' is the refined concept, but, as noted there, there is no general way (no algorithm) to judge whether a given sample sequence or an object (coded appropriately as a number sequence) is random or not. However, collectively we can determine whether a set consists of mostly random numbers or not. We know almost all the binary expansions of the numbers in $[0,1]$ are algorithmically random (because there are only countably many non-random numbers). Therefore, the basic idea to connect chaos and randomness is to check whether there is a bunch of trajectories that are 'almost surely' random or not. Thus, we wish to claim that a dynamical system which has an invariant set on which most trajectories are algorithmically random (after appropriate discretization = coding).

24.2 Can we be quantitative?

The claim "a dynamical system which has an invariant set on which most trajectories are algorithmically random" is actually already built in into my definition of chaos **22.10** through the coin-tossing process, because most numbers in $[0,1]$ is algorithmically random.

We must recall that the definition of algorithmic random numbers is in terms of complexity K. We wish to define a corresponding number within the theory of dynamical systems and compare it with K (**23.23**). Since K measures how much information we need to specify a symbol in the coding sequence, it is a very natural idea that it must be compared with the Kolmogorov-Sinai entropy, which was intuitively introduced as the information loss rate (or required information to predict the future; **17.18**). The KS entropy h will be discussed in detail later mathematically (-**31.9**), but here let us use its intuitive meaning, and actually show $K = h$, completing our characterization of chaos even quantitatively.

About coding: We have extensively used symbolic dynamics (or shift dynamics) isomorphic or homomorphic to the original dynamical system to analyze it. Some people bitterly criticize this strategy, saying that this approach does not respect how 'random sequences' are actually produced. We could produce the same '01 sequence' from a black box containing a person with a coin. Therefore, if one observes only the coded results, one can never infer the content of the black box (even if it is driven by a tent map). Hence, characterizing the dynamical system in terms of the coded result is impossible. *A fortiori* characterizing chaos with the randomness of the trajectories is flawed. How do you respond?[305]

24.3 Randomness of a trajectory

Let $\mathcal{B} = \{B_1, \cdots, B_k\}$ be a generator (= the best coding scheme; see **31.10**) for a map T defined on M with an invariant measure (= stationary distribution) μ (denoted as (T, μ, M) in standard books). We can code a trajectory starting from x at time 0 as ω in terms of k symbols with the rule $\omega_n = a$ if the trajectory goes through B_a at time n ($T^n x \in B_a$). Define the randomness of the trajectory starting from x as

$$K(x, T) \equiv \limsup_{n \to \infty} \frac{1}{n} \ell(\omega[n]), \tag{24.1}$$

where $\omega[n]$ is, as before, the first n symbols of ω, and $\ell(z)$ is the code length of the minimal program for z in terms of k symbols as defined in the preceding lecture (there, k was 2).

24.4 Brudno's theorem[306]

Brudno's theorem adapted to the current situation: coding of (T, μ, M) with k symbols reads:

> For μ-almost all $x \in M$
>
> $$K(x, T) = h_\mu(T)/\log k.$$
>
> Here, h_μ is the Kolmogorov-Sinai entropy of (T, μ, M) defined (as usual) in terms of the natural logarithm, but it is divided by $\log k$, so the right-hand side gives the entropy defined in terms of the base k logarithm.

---

[305] Since this is a good discussion topic, no comment should be added, but note that we do not simply treat dynamical systems as black boxes. The correspondence between a shift and a dynamical system must be at least homomorphic. There cannot be any deterministic dynamical system homomorphic to the package of a person + a coin.

[306] A. A. Brudno, THE COMPLEXITY OF THE TRAJECTORIES OF A DYNAMICAL SYSTEM, Russ. Math. Surveys 33 197 (1978); Entropy and the complexity of the trajectories of a dynamical system. Trans. Moscow Math. Soc., 2 127 (1983). See also H. S. White, Algorithmic complexity of points in dynamical systems, Ergod. Th. & Dynam. Sys., 13 807 (1993).

The following exposition is taken from TNW Chapter 2, but probably more readable (reader-friendly).

What is claimed is, qualitatively, the equality between the amount of the extra information required for prediction of the state time t in the future and the amount of information to describe the trajectory for the time span t. It is quite a natural assertion. To predict a chaotic trajectory for a long time, we need a tremendous amount of information initially. Since such an amount of information is needed to single out a trajectory, the coding sequence needed to describe the trajectory cannot be simple and information compression is out of question. Thus, Brudno's theorem confirms the goodness of our definition of chaos.

24.5 Brudno's theorem for symbolic dynamics

The information needed to describe a trajectory may be considered in terms of the symbol sequence after coding. Therefore, the core of Theorem in **24.4** is the following fact about the shift dynamical system (isomorphic to the dynamical system under consideration):

Theorem. Let (σ, Ω) be a certain shift dynamical system with k symbols, and μ its ergodic invariant measure. Then, for μ-almost all $\omega \in \Omega$

$$K(\omega) = h_\mu(\sigma)/\log k,$$

where $h_\mu(\sigma)$ is the Kolmogorov-Sinai entropy of the measure theoretical dynamical system (σ, μ, Ω), and $K(\omega)$ is the randomness defined in (23.4) (as in **24.4**, it is defined for k-symbol sequences instead of binary sequences and h uses natural log; that is why $\log k$ appears).

The above theorem is proved by showing the following two statements:
(i) ω satisfying $K(\omega) < h_\mu(\sigma)/\log k$ is μ-measure zero.
(ii) μ-almost surely (= for μ-almost all ω) $K(\omega) \le h_\mu(\sigma)/\log k$.

24.6 Demonstration of (i)

The number of $\omega[n]$ satisfying

$$K(\omega) \sim \ell(\omega[n])/n \le s \tag{24.2}$$

is no more than k^{ns} (in our context ω is the k-symbol sequence). On the other hand, according to the Shannon-McMillan-Breiman theorem **31.13**, the measure of the cylinder set specified by $\omega[n]$ is estimated as $e^{-nh_\mu(\sigma)}$. Therefore, the measure of all ω satisfying (24.3)

$$K(\omega) < h_\mu(\sigma)/\log k \tag{24.3}$$

is bounded by $e^{n(s \log k - h_\mu(\sigma))}$. This exponent is negative ($s \log k < h_\mu(\sigma)$; do not forget that h_μ is defined with the natural logarithm), so the upper bound converges to zero in the large n limit. Therefore the possibility of (24.3) is almost surely ignored.

24.7 Demonstration of (ii)

Next, we wish to show that μ-almost surely (= for μ-almost all ω)

$$K(\omega) \le h_\mu(\sigma)/\log k. \tag{24.4}$$

If this is demonstrated, then, since we just showed that the cases with $K(\omega) < h_\mu(\sigma)/\log k$ may be ignored almost surely, only the equality remains.

Since we have only to estimate the upper limit of $K(\omega)$, let us estimate the upper limit of $\ell(\omega[n])$. $\omega[n]$ is decomposed as follows in terms of q length-m-symbol-sequences ω_i^m ($i = 1, \cdots, M$, where M is the total number of distinct length-m-symbol-sequences; $n = mq + r$, i.e., $q = [n/m]$ and r is the residue):

$$\omega[n] = \omega_0^r \omega_{i_1}^m \omega_{i_2}^m \cdots \omega_{i_q}^m. \tag{24.5}$$

Here, ω_i^m is the ith kind of length-m-symbol-sequence, which is assumed to appear s_i times. With this representation, $\omega[n]$ can be uniquely specified by r, m, $s_1, \cdots, s_M$, ω_0^r and the arrangement of q length-m-symbol sequences $\omega_{i_1}^m \omega_{i_2}^m \cdots \omega_{i_q}^m$. Therefore, the needed information to specify $\omega[n]$ is given by (or, the length of the shortest required program with k symbols, or the information measured with the base k logarithm is given by)

$$\ell(\omega[n]) \le \ell(r) + R + \ell(m) + \ell(q) + \sum_{j=1}^{M} \ell(s_j) + H(\omega_{i_1}^m \omega_{i_2}^m \cdots \omega_{i_q}^m), \tag{24.6}$$

where $H(\omega_{i_1}^m \omega_{i_2}^m \cdots \omega_{i_q}^m)$ is the information needed to specify the arrangement (ordering) of q length-m-symbol sequences $\omega_{i_1}^m \omega_{i_2}^m \cdots \omega_{i_q}^m$ and R is the information required to specify ω_0^r, which is bounded by a constant independent of n. That is, except for the last term, all the terms are $o[n]$ and unrelated to the randomness. Hence,

$$K(\omega) \le \limsup_{n \to \infty} H(\omega_{i_1}^m \omega_{i_2}^m \cdots \omega_{i_q}^m)/n. \tag{24.7}$$

$H(\omega_{i_1}^m \omega_{i_2}^m \cdots \omega_{i_q}^m)$ is bounded by the information (in terms of base k logarithm) carried by the possible sequences under the assumption that all such sequences appear with equal probability. Therefore, it cannot exceed the logarithm (base k) of the number of sequences that can appear as $\omega[n]$. That is, $K(\omega) \le \lim_{n\to\infty}[\log_k N(n)]/n$. [Beyond this point you need rudiments of KS entropy: Lecture 31] $N(n)$ is equal to the number of non-empty elements $\vee_{k=0}^{n}\sigma^{-k}\mathcal{B}$ for a generator $\mathcal{B}$. According to the Shannon-McMillan-Breiman theorem, for the cylinder sets contributing to entropy $\mu(\omega[n])/e^{-nh_\mu(\sigma)}$ must not vanish in the $n \to \infty$ limit. Therefore, the number of cylinder sets we must count must be, since the sampling probability of each cylinder set is $e^{-nh_\mu(\sigma)}$, the order of its inverse: $N(n) \sim e^{nh_\mu(\sigma)}$. Thus, we can understand (24.4).

As can be seen from the explanation here, Brudno's theorem is based on very crude estimates, so it is a natural theorem. Such a theorem should have been discovered by theoretical physicists without any help of mathematicians. Most physicists in the US did not know this theorem well into the 1990s.

24.8 What is chaos, after all?

Out conclusion is that chaos is a deterministic dynamical system whose trajectories are algorithmically random. Actually, equivalently, we can say that a dynamical system with a positive KS entropy invariant measure is a chaotic dynamical system.

Is this a satisfactory outcome? That a trajectory is random is, with the quantification in terms of the Kolmogorov-Sinai entropy, invariant under the isomorphism of the dynamical systems. Isomorphism is a crude correspondence ignoring even the topology of the phase space, so the characterization we have pursued has nothing at all to do with how the correlation function decays or what the shape of the attractor is. Even the observability of chaos by computer experiments is not invariant under isomorphism.

Perhaps we should conclude that chaos is very common random phenomenon exhibited by deterministic dynamical systems and is far more basic than the exponential decay of the correlation function, or the invariant sets being fractal.

24.9 Does chaos exist in Nature?

Since we started this chapter with a simple realizable example, the question whether there is actually chaos may sound strange. However, it is difficult to tell whether the actual apparently chaotic phenomenon is really chaos or not due to the existing noise.[307] For example, it is easy to make an example that apparently exhibits observable chaos, even though there is no observable chaos without noise. It is difficult

[307] Here, 'noise' need not mean the effect of the unknown scale, but any unwanted external disturbance as usual.

to answer affirmatively the question whether there is really chaos without external noise in a system for which the existence of chaos is experimentally confirmed (this is in principle impossible[308]). Therefore, whether the concept of chaos is meaningful in natural science or not depends on whether it is useful as an ideal concept to understand the real world just as points and lines in elementary geometry. The relevance of chaos to the instabilities in some engineering systems or instabilities in numerical computation shows that it is a useful ideal concept.

For actual systems, what is important is its response to small perturbations. For a deterministic chaos, its practically important aspect is almost exhausted by the exponential separation of nearby trajectories. Whether the system is deterministic or not is unimportant. What is practically important is that the phase space is bounded and the trajectory itinerates irregularly various 'key' points in the phase space that are crucial to the system behavior. However, in order to model a system that easily exhibits such trajectories with small external perturbations, use of a chaotic system is at least metaphorically effective.[309]

[308]For example, for a one dimensional map, indefinitely small modification of the map can change it to have a stable fixed point. Such a system behaves just as before the modification, if a small noise is added. M. Cencini, M. Falcioni, E. Olbrich, H. Kantz, and A. Vulpiani, "Chaos or noise: difficulties of a distinction," Phys. Rev. E **62**, 427 (2000) recommend a more practical attitude toward chaos.

[309]See for chaotic itinerancy view of brain by I. Tsuda, "Hypotheses on the functional roles of chaotic transitory dynamics," Chaos **19**, 015113 (2009).