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 is no 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. Therefore, the basic idea to connect chaos and randomness is whether there is a bunch of trajectories that are ‘almost surely’ random or not.

24.2 Kolmogorov-Sinai entropy and randomness: a guess
Let us take a measure-theoretical dynamical system \((T, \mu, \Gamma)\), where \(T\) is an endomorphism of a compact phase space \(\Gamma\) with an invariant measure \(\mu\) (Section 2.6). We assume \(\mu\) is ergodic (Section 2.6). We have seen in Section 2.7 that its Kolmogorov-Sinai entropy \(h_\mu(T)\) is a good quantifier of chaos, or of the extent of difficulty of predicting its future, because \(h_\mu(T)\) measures the needed extra information for equi-pre-cise description of the states. The extent of the prediction difficulty must be related to the complicate nature of the dynamics (or trajectories). Then, it is a natural guess that the algorithmic randomness \(K(\omega)\) of the symbol sequence \(\omega\) obtained by encoding the trajectories of a dynamical system using a generator and its Kolmogorov-Sinai entropy \(h_\mu(T)\) must have an intimate relationship. In fact, Brudno’s theorem asserts that they are equal. This is the ultimate justification of our characterization of chaos given in Section 2.3.

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?\(^{278}\)

\(^{278}\)Since this is a 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 ho-
24.3 Randomness of a trajectory
Let \( \mathcal{B} = \{B_1, \ldots, B_k\} \) be a generator (see Section 2.7) for \((T, \mu, \Gamma)\). We can code a trajectory starting from \( x \) at time 0 as \( \omega \) 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

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

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

24.4 Brudno’s theorem
Brudno’s theorem adapted to the current situation: coding of \((T, \mu, \Gamma)\) with \( k \) symbols reads:

For \( \mu \)-almost all \( x \in \Gamma \)

\[
K(x, T) = h_\mu(T)/\log k.
\]

Here, \( h_\mu \) is the Kolmogorov-Sinai entropy of \((T, \mu, \Gamma)\) 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.

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 momorphic to the package of a person + a coin.
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 $(\sigma, \Omega)$ be a certain shift dynamical system with $k$ symbols, and $\mu$ its ergodic invariant measure. Then, for $\mu$-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 $(\sigma, \mu, \Omega)$, 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; that is why $\log k$ appears).

The above theorem is proved by showing the following two statements:

(i) $\omega$ satisfying $K(\omega) < h_\mu(\sigma) / \log k$ is $\mu$-measure zero.

(ii) $\mu$-almost surely (= for $\mu$-almost all $\omega$) $K(\omega) \leq h_\mu(\sigma) / \log k$.

### 24.6 Demonstration of (i)

The number of $\omega[n]$ satisfying

$$K(\omega) \sim \ell(\omega[n])/n \leq s \quad (24.2)$$

is no more than $k^{ns}$ (in our context $\omega$ is the $k$-symbol sequence). On the other hand, according to the Shannon-McMillan-Breiman theorem (Theorem 2.7A.3), the measure of the cylinder set specified by $\omega[n]$ is estimated as $e^{-n h_\mu(\sigma)}$. Therefore, the measure of all $\omega$ satisfying (24.3)

$$K(\omega) < h_\mu(\sigma) / \log k \quad (24.3)$$

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 $h_\mu$ is defined with the natural logarithm), so the upper bound converges to zero in the large $n$ limit. The possibility of (24.3) is almost surely ignored.

### 24.7 Demonstration of (ii)

Next, we wish to show that $\mu$-almost surely (= for $\mu$-almost all $\omega$)

$$K(\omega) \leq h_\mu(\sigma) / \log k. \quad (24.4)$$

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 $q$ $m$-symbol-sequences $\omega^m_i$ ($i = 1, \ldots, M$, where $M$ is the total number of distinct $m$-symbol-sequences; $n = mq + r$, i.e., $q = [n/m]$ and $r$ is the residue):

$$\omega[n] = \omega^m_r \omega^m_1 \omega^m_2 \cdots \omega^m_q. \quad (24.5)$$

Here, $\omega^m_i$ is the $i$th kind of $m$-symbol-sequence, which is assumed to appear $s_i$ times. With this representation, $\omega[n]$ can be uniquely specified by $r, m, s_1, \ldots, s_M, \omega^r_0$ and the arrangement of $q$ $m$-symbol sequences $\omega^m_1 \omega^m_2 \cdots \omega^m_q$. 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)

$$\ell(\omega[n]) \leq \ell(r) + R + \ell(m) + \sum_{j=1}^M \ell(s_j) + h(\omega^m_1 \omega^m_2 \cdots \omega^m_q), \quad (24.6)$$

where $h(\omega^m_1 \omega^m_2 \cdots \omega^m_q)$ is the information needed to specify the arrangement of $q$ $m$-symbol sequences $\omega^m_1 \omega^m_2 \cdots \omega^m_q$ and $R$ is the information required to specify $\omega^r_0$, 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,

$$K(\omega) \leq \limsup_{n \to \infty} h(\omega^m_1 \omega^m_2 \cdots \omega^m_q)/n. \quad (24.7)$$

$h(\omega^m_1 \omega^m_2 \cdots \omega^m_q)$ 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, $K(\omega) \leq \lim_{n \to \infty} [\log_k N(n)]/n$. $N(n)$ is equal to the number of non-empty elements $\bigvee_{k=0}^n \sigma^{-k}\mathcal{B}$ 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.

24.8 What is chaos, after all?
The final result is that chaos is a deterministic dynamical system whose trajectories
are algorithmically random.

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 (Note 2.7A.4) 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. Chaos as random behavior of deterministic dynamical systems is a common phenomenon and 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. 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 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). 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

---

279 Here, ‘noise’ need not mean the effect of the unknown scale, but any unwanted external disturbance as usual.

280 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.
is at least metaphorically effective.\textsuperscript{281}