30 Lecture 30 Information loss rate

30.1 Observation with finite resolution
When a dynamical system is given, we wish to describe its state at each instant with a constant precision. Suppose the data with information $H$ is required to describe the current state with a prescribed precision. If we could predict the state with the same precision after one unit time with the aid of this initial data, all the states in any future can be predicted with the same precision in terms of the information $H$ contained in the initial data. However, for chaotic systems the cloud of the ensemble is scrambled and the precise locations of the individual points in the phase space are lost. If we wish to locate the state of the system after one time unit as precisely as we can locate it now, we have to increase the precision of the description of the present state. This implies two things:

1. We need extra information $\Delta H$ to describe time one future as accurately as the present. That is, we need more information required at present to maintain the precision of the description.
2. Since the at any one time in the steady state we need the same amount $H$ to have ‘equi-precision’ description, $\Delta H$ means the loss of information due to time evolution.

Let us try to quantify the information loss per time.

30.2 Interval map information loss rate
Consider a discrete measure-theoretical dynamical system $(f, \mu, I)$ on the interval $I$ with a piecewise continuous and differentiable map $f$ with an absolutely continuous invariant measure $\mu$. The average amount $h_\mu(F)$ of the extra information required for the equi-precise description of the states is given by the following formula:

$$h_\mu(F) = \int_I \mu(dx) \log |F'(x)|.$$  (30.1)

Henceforth, we will often write $\mu(dx) = d\mu(x)$. The generalized form that is also

---

331 'piecewise continuous' implies that a map consists of several continuous pieces. ‘Piecewise continuously differentiable’ implies that each piece is differentiable inside, and, at its ends, one-sided derivatives from inside are well-defined.
correct for non-absolutely continuous invariant measures is (30.3).\textsuperscript{333}

### 30.3 Derivation of Rokhlin’s formula

For simplicity, let us assume that $f$ is unimodal (its graph has only one peak and no valley; Fig. 30.1). The invariance condition (28.2) for the measure $\mu$ reads, in this case (see Fig. 30.1),

$$\mu(dx_1) + \mu(dx_2) = \mu(dx). \quad (30.2)$$

Here, $dx$ denotes a small interval and its preimage (the set mapped to $dx$ by $F$) is written as $dx_1 \cup dx_2$ ($F^{-1}(dx) = dx_1 \cup dx_2$). For example, if $F$ is applied once to $dx_1$, it is stretched and mapped onto $dx$, so without any doubt the precision to specify the location is reduced. That is, with the equal probability $\mu(dx_1)/\mu(dx) = \mu(dx_1)/\mu(F(dx_1))$ they are spread over $dx$. Consequently, the information was lost by $-\log(\mu(dx_1)/\mu(dx))$. Therefore, the expectation value of information deficit required to predict one time step later with the same precision as the current state must be obtained by averaging this over $x_1$ with the aid of the (probability) measure $\mu$. We must also do an analogous calculation for $dx_2$. Consequently, the rate of information deficit is written as

$$h_\mu(F) = -\int I \mu(dx) \log \frac{\mu(dx)}{\mu(F(dx))}. \quad (30.3)$$

This loss must be compensated with the extra information ($\Delta H$ in the above) for equi-precise description of the system.

If $\mu$ is absolutely continuous, we can introduce the invariant density $g$ as $g(x)\lambda(dx) = \mu(dx)$, where $\lambda$ is the Lebesgue measure. Hence,

$$\frac{\mu(dx_i)}{\mu(dx)} = \frac{\mu(dx_i) \lambda(dx_i)}{\lambda(dx_i) \mu(dx)} = \frac{g(x_i)}{g(F(x_i))|F'(x_i)|}. \quad (30.4)$$

Noting that $F(x_i) = x$ with the aid of (30.4), we may rewrite (30.3) as

$$h_\mu(F) = \int I \lambda(dx) g(x) \log \frac{g(F(x))|F'(x)|}{g(x)}. \quad (30.5)$$

of measure theoretical dynamical systems had been obtained by the Russian mathematicians long before chaos became popular among the US physicists in the 1980s.

\textsuperscript{333}F. Ledrappier, "Some properties of absolutely continuous invariant measures on an interval," Ergodic Theor. Dynam. Syst. 1, 77 (1981) proved the following theorem for $C^2$-maps (actually, for $C^1$-maps with some conditions):

**Theorem** Let $f$ be a $C^2$-endomorphism on an interval. A necessary and sufficient condition for an ergodic invariant measure to be absolutely continuous is that the Kolmogorov-Sinai entropy is given by Rohlin’s formula (30.4).
Here, the integral with respect to the Lebesgue measure is explicitly written as \( \int \lambda(dx) \), but it is simply \( \int dx \) with the usual notation. With the aid of the elementary property of the \( \delta \)-function,\(^{334}\) regarding \( y \) as the independent variable, we have

\[
\delta(x - F(y)) = \frac{1}{|F'(x_1)|}\delta(y - x_1) + \frac{1}{|F'(x_2)|}\delta(y - x_2),
\]

where \( F(x_1) = F(x_2) = x \). With the help of this formula (30.2) can be rewritten as

\[
g(x) = \int_I dy \, g(y) \delta(x - F(y)),
\]

which is called the Perron-Frobenius equation. Using this, we see

\[
\int_I dy \, g(y) \log F(y) = \int_I dy \int_I dz \, g(y) \delta(z - F(y)) \log g(z) = \int_I dz \, g(z) \log g(z),
\]

so indeed (30.3) is nothing but (30.1).

\(^{334}\)Let \( a \) be an isolated (real) zero point (i.e., \( f(a) = 0 \)), and \( f \) be differentiable around it. In a sufficiently small neighborhood of \( a \) with the aid of the variable change: \( y = f(x) \) we get for sufficiently small positive \( \epsilon \)

\[
\int_{a-\epsilon}^{a+\epsilon} \delta(f(x)) \varphi(x) dx = \frac{1}{|f'(a)|} \varphi(a).
\]

Here, \( \varphi \) is a sufficiently smooth test function. Therefore, if \( f \) has a several isolated real zeros \( a_i \), we can add each contribution to get

\[
\int_{-\infty}^{\infty} \delta(f(x)) \varphi(x) dx = \sum_i \frac{1}{|f'(a_i)|} \varphi(a_i).
\]