28 Lecture 28. Measure-theoretical dynamical systems

28.1 Why measure-theoretical dynamical systems?
We have realized that even for a single dynamical system (that is defined by a definite law or rule) its behavior depends on initial conditions. Thus, even if we are told that the system exhibit periodic movements or apparently random (that is, chaotic) behaviors, we may not be lucky enough to observe them. How lucky can we be? That is a question of probability, and probability is a ‘measure’ of our confidence.

For example, for a continuous time dynamical system \( \mathcal{X}^r(M) \) defined on \( M \), a trajectory starting from \( x_0 \in M \) at \( t = 0 \) \( \varphi_t(x_0) \) depends on \( x_0 \). Suppose the system exhibits an unstable periodic orbit, you can observe it only when \( x_0 \) is on the orbit. What is your chance to observe it? If we throw a dart ‘randomly’ on \( M \), you would expect that you would never hit the orbit. However, if there is some device channeling the darts onto the orbit, with probability one, you will observe it. If \( x_0 \) is in the basin of attraction of a stable periodic orbit and the basin is of positive Riemann volume defined on \( M \), then we may have a finite (not infinitesimal) chance to observe the periodic motion. Even in this case there is a fixed point to whose basin of attraction all the initial conditions are channelled, there will be no chance to observe the stable periodic orbit. These trivial considerations clearly tells us that it is very important (especially for physicists) to specify how we can sample the initial condition. This is the question of the sampling measure.

Once the system settles down to a steady state, that is, the system is in a certain attractor, then we may be interested in its average behavior. For example, in the case of a chaotic attractor, we could ask how nearby orbits leave each other on the average (that is, we ask ‘how chaotic’ the system is). Inevitably we need a measure to average the behavior.

Thus, if we ask about observability and about the average behavior, we need (probability) measures.

28.2 Measure
What is a measure? See XXX.

28.3 Remark on family of measurable sets
If you read a respectable book on probability you will see a setup of probability space $(M, \mathcal{F}, \mu)$, where $\mu$ is a (probability) measure defined on $M$ and $\mathcal{F}$ is a set of all the measurable sets on which $\mu$ is defined. Not all the subsets of $M$ can be assumed to have ‘probability.’ This restriction is a must under the usual axioms of mathematics.

Therefore, when we consider a measure-theoretical dynamical system on $M$ with an invariant measure the ‘dynamical rule’ (say, the map defining a discrete time dynamical system) must be compatible with (preserved) $\mathcal{F}$. When we say two dynamical systems are isomorphic (= equivalent), then there must be a correspondence (modulo $\mu$-measure zero sets) between the families of measurable sets for these dynamical systems.

However, in this exposition we will not mention about $\mathcal{F}$.

### 28.4 Invariant measure: introduction
Prepare numerous clones of the dynamical system under consideration, and plot their states in the same phase space. Then, we would see a cloud of points that describe individual clones in the phase space (Fig. 28.1).

![Figure 28.1: Distribution or ensemble. The space is the phase space $\Gamma$. Each point represents an instantaneous state of each system. An invariant measure corresponds to a steady state for which the cloud as a whole becomes time-independent, although individual points may keep wandering around.](image-url)

The distribution described by the cloud, if the total mass is normalized to be unity, may be understood as a probability measure on the phase space. Following the time evolution of the system, the cloud may change its shape. After a sufficiently long time, often the cloud representing an ensemble ceases to change its shape. Then, we say that the ensemble has reached its steady state. Individual points corresponding to the members of the ensemble may still keep wandering around in the cloud, but the cloud as a whole is balanced and the distribution on the phase space becomes invariant. The (probability) distribution corresponding to this invariant cloud is called an *invariant measure* of the dynamical system.

### 28.5 Invariant measure of dynamical system: discrete time
Discrete time case: Measure $\mu$ on the phase space $M$ is an invariant measure of a dynamical system $f \in C^r(M)$, if

$$\mu = \mu \circ f^{-1} \quad (28.1)$$

holds. That is, for an arbitrary ($\mu$-measurable) subset $A \subset M$

$$\mu(A) = \mu(f^{-1}A) \quad (28.2)$$

holds, where $f^{-1}A$ is the totality of points that comes to $A$ after a unit time step (Fig. 28.2).

![Figure 28.2: The preimage $f^{-1}A$ of $A$ by the time evolution operator $f$ need not be connected (in this figure it consists of two connected components). Since $f(f^{-1}A) = A$, $f^{-1}A$ is the totality of the points that come to $A$ after one time step.

For an arbitrary measurable set $A$, if we know $f^{-1}A$ at present, we know everything we can discuss probabilistically that will occur after one time step. (28.2) expresses the condition that the evaluation of the weights of $A$ and $T^{-1}A$ by $\mu$ is consistent with the conservation law of the members of the ensemble.

### 28.6 Invariant measure of dynamical system: continuous time

Continuous time case: Measure $\mu$ on the phase space $M$ is an invariant measure of a dynamical system $\varphi_t$ defined by a vector field $\mathcal{X}(M)$, if

$$\mu = \mu \circ \varphi_t^{-1} \quad (28.3)$$

holds for all $t$. That is, for an arbitrary ($\mu$-measurable) subset $A \subset M$

$$\mu(A) = \mu(\varphi_t^{-1}A) \quad (28.4)$$

holds (quite analogous to the discrete time case).
28.7 Invariant measures are not unique for a given dynamical system
Generally speaking, an invariant measure for a given dynamical systems is not unique (as already suggested in 28.1). For example, $T x = \{2x\}$ discussed previously has actually uncountably many distinct invariant measures.

For a Hamiltonian system we know the phase volume is invariant, but there are many other invariant measures. Furthermore, not all the invariant measures can give ‘natural averages’ that agree with averages obtained by continuous even observations.\(^{318}\)

28.8 Measure-theoretical dynamical system
Let $\mu$ be an $f \in C^r(M)$-invariant measure on $M$. The triplet $(f, \mu, M)$ is called a measure theoretical dynamical system. This may be interpreted as a mathematical expression of a steady state allowed to the dynamical system $f \in C^r(M)$; given a dynamical system $f \in C^r(M)$, for each invariant measure $\mu$, a distinct measure theoretical dynamical system is constructed.

For a continuous time dynamical system defined by $X \in \mathcal{X}$, $(X, \mu, M)$ may be understood as a measure-theoretical dynamical system.

28.9 Absolutely continuous invariant measure
If a measure $\mu$ satisfies $\mu(A) = 0$ for any measure zero set $A$, $\mu$ is called an absolutely continuous measure. For a dynamical system with an absolutely continuous invariant measure, chaos may often be observed by numerical experiments.

If a measure is absolutely continuous, its probability density $g$ may be defined as $d\mu = gd\lambda$, where $\lambda$ is the Lebesgue measure.\(^{319}\)

Absolutely continuous invariant measures are observable (see 22.3) (e.g., computationally).

28.10 Ergodic measure
For any invariant set $A \subset M$ (that is, any set satisfying $A = f^{-1}A$), if $\mu(A) = 0$ or $1$, the measure theoretical dynamical system is said to be ergodic. Note that this is neither the property of the measure nor the property of the dynamical

\(^{318}\)As we will see later, invariance does not mean ergodicity.

\(^{319}\)That is, $g$ is the Radon-Nikodym derivative $d\mu/d\lambda$. See, e.g., Kolmogorov and Fomin cited already.

\(^{320}\)This condition must be, precisely speaking, the equality ignoring the $\mu$-measure zero difference.
Roughly speaking, if a dynamical system is ergodic, the trajectory starting from any initial point\footnote{Precisely speaking, starting from \( \mu \)-almost all initial points. ‘\( \mu \)-almost all’ implies that all except for \( \mu \)-measure zero sets.} can eventually go into any \( \mu \)-measure positive set. In fact, if there is a positive measure set \( D \) to which any trajectory starting from a certain positive measure set \( C \) cannot reach, then there must be an invariant set \( B \) with \( 0 < \mu(B) < 1 \) such that \( B \supset C \) and \( B^c \supset D \). Conversely, if there is an invariant set \( B \) such that \( 0 < \mu(B) < 1 \). Then, we can choose a measure positive set \( C \) with \( B \cap C = \emptyset \) such that any trajectory starting from \( C \) never visits \( B \) (if possible, \( \mu(\bigcup_{k=1}^{\infty} T^{-k} B \cap C) > 0 \), but \( T^{-1} B = B \) and \( B \cap C = \emptyset \), a contradiction).

Let \( \Gamma \) be a unit circle, and \( f \) be a rigid rotation around the center by an angle that is irrational multiple of \( 2\pi \). Since the uniform distribution \( \mu_U \) on the circle is rotation-invariant, we can make a measure-theoretical dynamical system \((T, \mu_U, \Gamma)\), which is ergodic. This not so interesting example illustrates why ergodicity is totally insufficient for modeling irreversible phenomena such as relaxation exhibited by usual many-body systems.

### 28.11 Mixing measure

For any sets\footnote{Precisely speaking, any \( \mu \)-measurable sets. From now on, such statements will not be added.} \( A \) and \( B \) (both \( \subset M \)), if

\[
\lim_{n \to \infty} \mu(f^{-n} A \cap B) = \mu(A) \mu(B)
\]

(28.5)

holds, the measure-theoretical dynamical system is said to be \textit{mixing} index mixing, where \( f^{-n} \) implies to apply \( f^{-1} \) \( n \)-times. \( f^{-n} A \) is the set that agrees with \( A \) after \( n \) time steps. \( f^{-n} A \cap B \) is the totality of the points that is at present in \( B \) and will be in \( A \) after \( n \) time steps. Intuitively, the cloud of points starting from \( B \) spread over the phase space evenly if \( n \) is sufficiently large, so the probability for the cloud to overlap with \( A \) is proportional to the ‘statistical weight’ of \( A \). It is obvious that initial conditions cannot be used to predict the future. Physically, irreversible processes such as relaxation phenomena occur. In particular, the time-correlation function eventually decays to zero.