Lecture 6: Invariant Measures for Dynamical Systems

Def^{n}. Suppose T is a continuous map on a compact metric space (X,d). A Borel probability measure μ is called T-invariant if $\mu \circ T^{-1} = \mu$, i.e.

$$\mu(T^{-1}E) = \mu(E) \quad \text{for each } E \in \mathcal{B}.$$

Remark: We saw last time that for any prob. measure μ and a bounded measurable f,

$$\int f \, d\mu\circ T^{-1} = \int f\circ T \, d\mu$$

Thus μ is T-inv iff $\int f\circ T\, d\mu = \int f \, d\mu$ $\forall f$ bounded & Borel.

Exercise. Use the uniqueness part of the Riesz representation theorem to show that μ is T-inv iff $\int f\circ T \, d\mu = \int f \, d\mu$ $\forall f$ continuous.

Example 1. Let m denote Lebesgue's measure on S^1, i.e. the unique measure on S^1 which represents the functional

$$m(f) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta}) \, d\theta \, , \quad f \in C(S^1).$$

(a) m is invariant for the rotation $R(e^{i\theta}) = e^{i(\theta+\alpha)}$

(b) ——————— " ——————— angle doubling $R(e^{i\theta}) = e^{2i\theta}$.

Proof of (a) By the exercise, it's sufficient to check that for each continuous $f : S^1 \to \mathbb{R}$, $\int_0^{2\pi} f(e^{i(\theta+\alpha)})\,d\theta = \int_0^{2\pi} f(e^{i\theta})\,d\theta$. Indeed,

$$\int_0^{2\pi} f(e^{i(\theta+\alpha)})\,d\theta = \int_0^{2\pi-\alpha} f(e^{i(\theta+\alpha)})\,d\theta + \int_{2\pi-\alpha}^{2\pi} f(e^{i(\theta+\alpha)})\,d\theta$$

$$= \int_\alpha^{2\pi} f(e^{i\theta})\,d\theta + \int_0^{\alpha} f(e^{i\theta})\,d\theta = \int_0^{2\pi} f(e^{i\theta})\,d\theta .$$

Proof of (b) Caution! $R(e^{i\theta}) = e^{2i\theta}$ is not invertible on S^1. But it is invertible on

$$A = \{ e^{i\theta} : 0 \leq \theta < \pi \}$$

$$B = \{ e^{i\theta} : \pi \leq \theta < 2\pi \}$$

$$\int_0^{2\pi} f(e^{2i\theta})\,d\theta = \int_0^{\pi} f(e^{2i\theta})\,d\theta + \int_{\pi}^{2\pi} f(e^{2i\theta})\,d\theta$$

$$= \int_0^{2\pi} f(e^{2i(\eta/2)})\,d(\eta/2) + \int_0^{2\pi} f(e^{2i(\pi+\eta/2)})\,d(\pi+\eta/2)$$

$$= \frac{1}{2}\int_0^{2\pi} f(e^{i\eta})\,d\eta + \frac{1}{2}\int_0^{2\pi} f(e^{2i\eta})\,d\eta = \int_0^{2\pi} f(e^{2i\eta})\,d\eta .$$

Remark. The angle doubling map has many other inv. measures (e.g. δ_0, $\frac{1}{2}(\delta_0 + \delta_{1/2})$). The irrational rotation doesn't.

<u>Example 2</u> Let m denote Lebesgue's measure on $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, the unique measure s.t.

$$m(f) = \int_0^1\int_0^1 f(x,y)\,dx\,dy \quad \forall f \in C(\mathbb{T}^2).$$

This is an invariant measure for the Cat Map

$$T_A\left[\begin{pmatrix}x\\y\end{pmatrix} + \mathbb{Z}^2\right] = A\begin{pmatrix}x\\y\end{pmatrix} + \mathbb{Z}^2, \quad A = \begin{pmatrix}2 & 1\\ 1 & 1\end{pmatrix}.$$

<u>Proof</u>. Again, it's enough to check continuous test functions

$$\int_{\mathbb{T}^2} (f\circ T_A)\,dm = \int_0^1\int_0^1 f\left(A\begin{pmatrix}x\\y\end{pmatrix} + \mathbb{Z}^2\right)dx\,dy$$

$$= \int_0^1\int_0^1 f(2x+y \bmod 1,\ x+y \bmod 1)\,dx\,dy$$

We'd like to change variables

$$\begin{pmatrix}u\\v\end{pmatrix} = \begin{pmatrix}2x+y \bmod 1\\ x+y \bmod 1\end{pmatrix}$$

- Since $T_A : \mathbb{T}^2 \to \mathbb{T}^2$ is a bijection, $\begin{pmatrix}u\\v\end{pmatrix}$ ranges over $[0,1)\times[0,1)$
- The Jacobian $\left|\dfrac{\partial(u,v)}{\partial(x,y)}\right| = \det\begin{pmatrix}2 & 1\\ 1 & 1\end{pmatrix} = 1.$

So $\displaystyle\int_0^1\int_0^1 f(2x+y \bmod 1,\ x+y \bmod 1)\,dx\,dy =$

$$= \int_0^1\int_0^1 f(u,v)\left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv = \int_0^1\int_0^1 f(u,v)\,du\,dv = \int_{\mathbb{T}^2} f\,dm.$$

Example 3 (Liouville's Thm). Suppose $H(q,p)$ is continuously diff twice on $\mathbb{R}^2$. Let $\varphi^t : \mathbb{R}^2 \to \mathbb{R}^2$ be the flow associated to the ODE

$$(*) \quad \begin{cases} \dot{q} = \dfrac{\partial H}{\partial p}(q,p) \\[2ex] \dot{p} = -\dfrac{\partial H}{\partial q}(q,p) \end{cases}$$

Then φ^t preserves Lebesgue's measure on $\mathbb{R}^2$.
(This is an infinite measure. Often, the level sets $\{(q,p) : H(q,p) < M\}$ are invariant open sets with finite measure.)

Proof. Fix t small and let

$$\begin{pmatrix} Q_t(q,p) \\ P_t(q,p) \end{pmatrix} = \varphi^t \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} \text{sol}^{\underline{n}} \text{ to } (*) \text{ at} \\ \text{time } t \text{ with initial} \\ \text{condition } (q,p) \end{pmatrix}.$$

The map $(q,p) \mapsto (Q,P)$ is a bijection, because φ^t is a flow, so $\varphi^t \circ \varphi^{-t} = id$.

As before, we just need to check that

$$\left| \frac{\partial(Q_t, P_t)}{\partial(q,p)} \right| = 1 \quad \text{for all } t.$$

Taylor's expansion tells us that

$$Q_t(q,p) = Q_0(q,p) + t\cdot \frac{d}{dt}\Big|_{t=0} Q_t(q,p) + \frac{1}{2}t^2 \frac{d^2}{dt^2}\Big|_{t=\xi_t(q,p)} Q_t(q,p)$$

$$= q + t\,\dot q(q,p) + \frac{1}{2}t^2 \times (\text{bounded smooth function of } q,p,t).$$

$$\overset{(*)}{=} q + t\,\frac{\partial H}{\partial p}(q,p) + \frac{1}{2}t^2 \times (\text{———''———})$$

$$P_t(q,p) = \cdots\cdots$$

$$= p - t\frac{\partial H}{\partial q}(q,p) + \frac{1}{2}t^2 \times (\text{bounded smooth function of } q,p,t$$

$$\Longrightarrow \quad \frac{\partial(Q_t,P_t)}{\partial(q,p)} = \begin{pmatrix} 1 + t\frac{\partial^2 H}{\partial q\partial p} + O(t^2) & t\frac{\partial^2 H}{\partial p^2} + O(t^2) \\ -\frac{\partial^2 H}{\partial q^2} + O(t^2) & 1 - t\frac{\partial^2 H}{\partial p\partial q} + O(t^2) \end{pmatrix}$$

$$= I + t\underbrace{\begin{pmatrix} \frac{\partial^2 H}{\partial q\partial p} & * \\ * & -\frac{\partial^2 H}{\partial p\partial q} \end{pmatrix}}_{\text{trace} = 0\,!} + O(t^2)$$

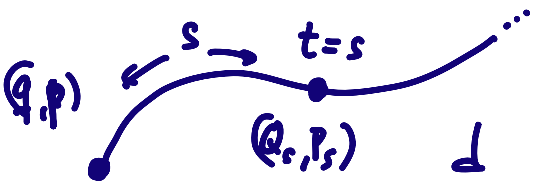
$$\underset{\text{check!}}{\Longrightarrow} \quad \det\left[\frac{\partial(Q_t,P_t)}{\partial(q,p)}\right] = 1 + O(t^2)$$

$$\Longrightarrow \quad \frac{d}{dt}\Big|_{t=0}\left[\det\left(\frac{\partial(Q_t,P_t)}{\partial(q,p)}(q,p)\right)\right] = 0$$

Because of the flow property,

$$\frac{d}{dt}\Big|_{t=s} \left[\det \frac{\partial(Q_t, P_t)}{\partial(q,p)}(q,p) \right] =$$

$$= \frac{d}{dt}\Big|_{t=0} \left[\det \frac{\partial(Q_t, P_t)}{\partial(q,p)}\big(Q_s(q,p), P_s(q,p)\big) \right] = 0$$



So we prove that

$$\frac{d}{dt}\left(\det\left[\frac{\partial(Q_t, P_t)}{\partial(q,p)}(q,p) \right] \right) = 0 \quad \forall t.$$

Thus the Jacobian of φ^t doesn't depend on t. Since $\varphi^t = \mathrm{Id}$ at $t=0$, the Jacobian $\equiv 1$, and the flow preserves Lebesgue's measure. $\square$

<u>Exercise</u>: Show that any flow associated to an ODE

$$\begin{cases} \dot{x} = \varphi_1(x,y) \\ \dot{y} = \varphi_2(x,y) \end{cases}$$

where $\vec{\varphi} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}$ satisfies $\mathrm{div}(\vec{\varphi}) = 0$ preserves Lebesgue's measure on $\mathbb{R}^2$.

Krylov – Bogolyubov Theorem: Any continuous map T on <u>compact</u> metric space (X,d) admits at least one invariant Borel probability measure.

<u>Proof</u>. Fix some $x_0 \in X$ and define the functionals

$$\varphi_n(f) := \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x_0)$$

Since $|\varphi_n(f)| \leq \frac{1}{n} \sum_{k=0}^{n-1} \|f\|_\infty = \|f\|_\infty$, $\|\varphi\| \leq 1$.

By the Banach-Alaoglu theorem, $\exists n_k$ s.t.

$$\varphi_{n_k} \xrightarrow[k\to\infty]{} \varphi \quad \text{weak star.}$$

Clearly $f \geq 0 \Rightarrow \varphi_{n_k}(f) \geq 0 \Rightarrow \varphi(f) \geq 0$,
and $\varphi(1) = 1$.

By the Riesz representation theorem, there is a unique Borel probability measure μ s.t.

$$\varphi(f) = \int_X f \, d\mu \qquad (f \in C(X)).$$

Observe that for every $f \in C(X)$,

$$\int_X f \circ T \, d\mu = \varphi(f \circ T) = \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k-1} (f\circ T)(T^j x_0)$$

$$= \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k-1} f(T^{j+1} x_0)$$

$$= \lim_{k\to\infty} \frac{1}{n_k} \left[\sum_{j=0}^{n_k-1} f(T^j x_0) + f(T^{n_k} x_0) - f(x_0) \right]$$

$$= \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k-1} f(T^j x_0)$$

$$= \lim_{k\to\infty} \varphi_{n_k}(f) = \varphi(f) = \int_X f \, d\mu .$$

So $\int_X f \circ T \, d\mu = \int_X f \, d\mu \quad \forall f \in C(X)$. $\quad (*)$

Consider now $\varphi, \psi \in C(X)^*$ given by

$$\varphi(f) = \int_X f \circ T \, d\mu = \int_X f \, d\mu \circ T^{-1}$$

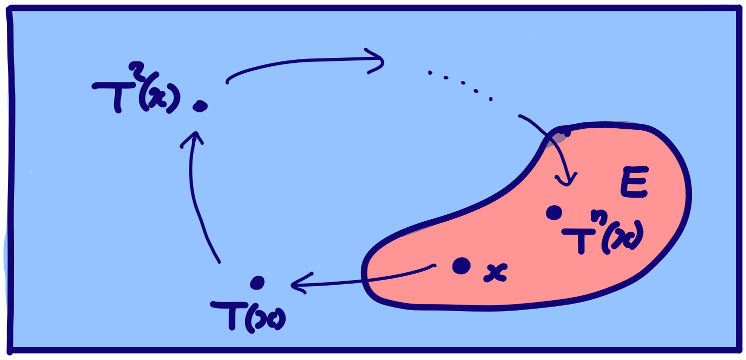
$$\psi(f) = \int_X f \, d\mu$$

By $(*)$, $\varphi = \psi$. Since φ is represented by $\mu \circ T^{-1}$, ψ is represented by μ, and the representing measure in Riesz's theorem is unique, $\mu \circ T^{-1} = \mu$. $\square$

The Dynamical Significance of Invariant Measures

Almost Everywhere: Suppose $(X, \mathcal{F}, \mu)$ is a probability space. We say that a property P of $x \in X$ holds <u>almost everywhere in E</u> if $\Omega := \{x \in E : P \text{ holds at } x\} \in \mathcal{F}$, and $\mu(E \setminus \Omega) = 0$.

Poincaré's Recurrence Theorem Suppose μ is a T-invariant <u>probability</u> measure, and E is a measurable set with positive measure. Then for μ-almost every $x \in E$ $\exists n \geq 1$ s.t. $T^n(x) \in E$.



<u>Proof</u>. Let $W := \{x \in E : \forall n \geq 1 \; T^n(x) \notin E\}$.
For every $n \geq 1$, $T^{-n} W \cap W = \emptyset$, because

$$x \in T^{-n} W \Rightarrow T^n(x) \in W \subseteq E \Rightarrow T^n(x) \in E$$
$$x \in W \quad \Rightarrow \quad T^n(x) \notin E.$$

It follows that $W, T^{-1}W, T^{-2}W, \ldots$ are pairwise disjoint:

$$T^{-n} W \cap T^{-(n+k)} W = T^{-n}(W \cap T^{-k} W) = T^{-n}(\emptyset) = \emptyset.$$

It follows that

$$1 = \mu(X) \geq \mu\left(\bigcup_{n=0}^{\infty} T^{-n} W\right) = \sum_{n=0}^{\infty} \mu(T^{-n} W) = \sum_{n=0}^{\infty} \mu(W) = \infty \cdot \mu(W).$$

Necessarily, $\mu(W) = 0$. $\square$

Corollary. Let T be a continuous map on a compact metric space, and let μ be an invariant probability measure. Then μ-a.e. x is recurrent.

Proof Fix $\varepsilon_n \geq 0$ s.t. $\varepsilon_n \to 0$.

Claim: For all n, $\mu\{x : \exists k_n \geq 1 \text{ s.t. } d(T^{k_n}x, x) < \varepsilon_n\} = 1$.

Proof: By compactness, X can be covered by finitely many balls of radius $\varepsilon_n/3$, $B_1^{(n)}, \dots, B_{N_n}^{(n)}$.

By Poincaré's recurrence theorem, for μ-almost every $x \in B_i^{(n)}$, $\exists k_n \geq 1$ s.t. $T^{k_n}(x) \in B_i^{(n)}$, and then the sets

$$\Omega_i^{(n)} := \{x \in B_i^{(n)} : \exists k_n \geq 1 \;\; d(T^{k_n}(x), x) < \operatorname{diam} B_i^{(n)} < \varepsilon_n\}$$

satisfy $\mu(B_i^{(n)} \setminus \Omega_i^{(n)}) = 0$. It follows that

$$\Omega := \bigcup_{i=1}^{N_n} \Omega_i^{(n)} \quad \text{satisfies} \quad \mu(X \setminus \Omega) = 0$$

[Because: $\mu(X \setminus \Omega) = \mu(\bigcup B_i^{(n)} \setminus \bigcup_i \Omega_i^{(n)}) \leq \sum_{i=1}^{N_n} \mu(B_i^{(n)} \setminus \bigcup_j \Omega_j^{(n)})$

$\leq \sum_{i=1}^{N_n} \mu(B_i^{(n)} \setminus \Omega_i^{(n)}) = \sum_{i=1}^{N_n} 0 = 0$.]

So for μ-a.e. $x \in X$ (including all $x \in \Omega$) $\exists k_n \geq 1$ s.t. $d(T^{k_n}x, x) < \varepsilon_n$.

Let $A_n := \{x : \exists k \geq 1 \ d(T^k(x), x) < \varepsilon_n\}$, then A_n is measurable (even open!), and by the claim $\mu(A_n) = 1$.

Let $A := \bigcap_{n=1}^{\infty} A_n$; then

(a) $\mu(A) = 1$, because $\mu(A^c) = \mu\left(\bigcup_{n=1}^{\infty} A_n^c\right) \leq \sum_{n=1}^{\infty} \mu(A_n^c) = 0$.

(b) $\forall x \in A \ \forall n \ \exists k_n \geq 1$ s.t. $d(T^{k_n}(x), x) < \varepsilon_n$.

- If $k_n \to \infty$, x is recurrent.

- If $k_n \not\to \infty$ then $\exists n_i \to \infty$ s.t. $k_{n_i} = k$, and then
$$d(T^k x, x) = d(T^{k_{n_i}} x, x) < \varepsilon_{n_i} \to 0.$$
This implies that $T^k(x) = x$, and this (of course) implies that x is recurrent ($T^{kn} x \to x$).

In summary, every $x \in A$ is recurrent, and $\mu(A) = 1$. $\square$

Ergodicity

Now we'd like more quantitative info on the n_k s.t. $T^{n_k}(x) \in E$.

Def^n. A T-invariant prob measure is called <u>ergodic</u>, if every measurable set E s.t. $T^{-1}E = E$ satisfies $\mu(E) = 0$ or $\mu(E) = 1$.

<u>The Ergodic Thm</u>: Suppose μ is an ergodic T-invariant prob. measure. Then:

(1) for every measurable function $f : X \to \mathbb{R}$ s.t. $\int |f| \, d\mu < \infty$, for μ-a.e. x,

$$\frac{1}{N} \sum_{n=1}^{N} f(T^n(x)) \longrightarrow \int f \, d\mu .$$

(2) for every measurable set E, for μ-a.e. x,

$$\frac{1}{N} \# \{ 1 \le n \le N : T^n(x) \in E \} \xrightarrow[N \to \infty]{} \mu(E) .$$

The proof is given, e.g., in the lecture notes on ergodic theory available on my home page.

Unique Ergodicity

Def$^{\underline{n}}$. A continuous map on a compact metric space is called <u>uniquely ergodic</u>, if it has exactly <u>one</u> invariant probability measure.

Thm. A continuous map on a compact metric space is uniquely ergodic with invariant prob. measure μ iff for every $f: X \to \mathbb{R}$ continuous,

$$(*) \quad \frac{1}{N} \sum_{n=1}^{N} f(T^n(x)) \xrightarrow[N \to \infty]{} \int f \, d\mu \quad \text{uniformly.}$$

Proof. Suppose μ is the unique inv. prob. measure of T and assume by way of contradiction that $(*)$ is false for some f. Then $\exists\, \epsilon > 0$ s.t. for all n there exists $N_n > n$ and $x_n \in X$ s.t.

$$\left| \frac{1}{N_n} \sum_{k=1}^{N_n} f(T^k(x_n)) - \int f \, d\mu \right| > \epsilon .$$

Let $\varphi_n(\cdot)$ be the functional

$$\varphi_n(g) = \frac{1}{N_n} \sum_{k=1}^{N_n} g(T^k(x_n)) .$$

As in the proof of Krylov–Bogolyubov's Thm, $\exists\, n_j \to \infty$ s.t.

$$\varphi_{n_j} \longrightarrow \varphi \quad \text{weak-star}$$

and $\varphi(g) = \int g\, d\nu$, where ν is a T-inv prob measure. By unique ergodicity, $\nu = \mu$.

But by construction,

$$|\nu(f) - \mu(f)| = \lim_{i\to\infty} \underbrace{\left| \frac{1}{N_{n_i}} \sum_{k=1}^{N_{n_i}} f(T^k(x_{n_i})) - \mu(f) \right|}_{> \epsilon} \geq \epsilon.$$

So $\nu \neq \mu$, and we obtained a contradiction. We proved $(\Longrightarrow)$.

The proof of $(\Longleftarrow)$ is simple: Had ν been some other invariant prob. measure, then $\forall f \in C(X)$

$$\nu(f) = \frac{1}{N} \sum_{k=1}^{N} \underbrace{\nu(f\circ T^k)}_{=\nu(f)} = \nu\Big(\frac{1}{N}\sum_{k=1}^{N} f\circ T^k\Big)$$

$$= \nu\Big(\int f\, d\mu\Big) + O\Big(\nu\Big(\underbrace{\Big\|\frac{1}{N}\sum_{k=1}^{N} f\circ T^k - \int f\, d\mu\Big\|_\infty}_{\longrightarrow 0}\Big)\Big)$$

$$= \int f\, d\mu + o(1).$$

Necessarily, $\nu(f) = \mu(f)$ for all $f \in C(X)$.
By the uniqueness of the representing function in Riesz's theorem, $\nu = \mu$. So $\exists !$ inv. prob. meas. $\square$