Introduction to Ergodic theory

Lecture by Amie Wilkinson
Notes by Clark Butler

November 4, 2014

Hyperbolic dynamics studies the iteration of maps on sets with some type of Lipschitz structure used to measure distance. In a hyperbolic system, some directions are uniformly contracted and others are uniformly expanded. Examples include expanding maps on manifolds, Anosov diffeomorphisms, and the shift map discussed in previous lectures. Hyperbolic dynamical behavior often gives rise to iterated function systems of contracting maps which create complicated fractal limit sets for orbits. An important example from Riemannian geometry is given by the geodesic flow on the unit tangent bundle of a closed negatively curved manifold. Many of these systems have the common feature that they can be *coded* as subshifts of finite type.

Recall: Given a finite alphabet $\Sigma = \{1, \dots, n\}$, we defined the shift space $\Sigma^{\mathbb{Z}}$ to be the set of all functions $\omega : \mathbb{Z} \to \Sigma$, which we thought of as sequences with a decimal point to mark the zero place. We defined the shift on this space by

$$\sigma : \Sigma^{\mathbb{Z}} \circlearrowleft, \ (\sigma(\omega))[n] = \omega(n+1)$$

We introduce new notation. $\Sigma_n := \{1, \dots, n\}^{\mathbb{Z}}$ is the set of two-sided sequences, and $\Sigma_n^+ = \{1, \dots, \}^{\mathbb{N}}$ the set of one-sided infinite sequences. σ acts on Σ_n^+ by the same formula as above, but note now that σ is no longer invertible if $n \geq 2$; in fact σ is n-to-1.

A *subshift* $X \subset \Sigma_n$ is a compact σ-invariant subset. We make the same definition for $X \subset \Sigma_n^+$. As an example of a subshift, take the closure of the orbit of any point of Σ_n under σ.

For a matrix $A \in \mathrm{Mat}_{n \times n}(\{0,1\})$, we define

$$\Sigma_A = \left\{ \omega \in \Sigma_n | \ A_{\omega(i)\omega(i+1)} = 1 \text{ for each } i \right\}$$

An analogous definition of Σ_A^+ is given by taking $\omega \in \Sigma_n^+$ instead. Subshifts Σ_A arising from a matrix A in this fashion are called *subshifts of finite type.*

Comment: Technically, we would also like to use the term 'subshift of finite type' for $X \subset \Sigma_n$ if it is topologically conjugate to a subshift $\Sigma_A \subset \Sigma_N$ corresponding to a $0-1$ matrix A in a (possibly larger) alphabet Σ_N. This extended definition includes subshifts $\Sigma_{\mathcal{A}} \subset \sigma_n$ corresponding to an m-dimensional $0-1$ array $\mathcal{A}$, where

$$\Sigma_{\mathcal{A}} = \left\{ \omega \in \Sigma_n | \ \mathcal{A}_{\omega(i)\omega(i+1)\dots\omega_{i+m-1}} = 1 \text{ for each } i \right\}$$

This allows for the possibility of specifying which finite strings we allow, instead of just which strings of length 2 we allow. However, by extending the alphabet, $\Sigma_{\mathcal{A}}$ can be made into a subshift of finite type corresponding to a $0-1$ matrix A, and the theory is essentially the same.

One way to think of Σ_A is to imagine a directed graph with n vertices labeled $1, \dots, n$, such that there is a directed arrow from vertex i to vertex j if and only if $A_{ij} = 1$. Then Σ_A corresponds to all possible walks on this graph with a marked starting point corresponding to $0 \in \mathbb{Z}$, and such that the walk extends infinitely both forward and backward in time. When we use subshifts of finite type to code hyperbolic maps, we will see that the choice of marked starting point corresponds to a choice of a point x in the space, and the walk on the graph corresponds to the orbit $\{f^n(x) : n \in \mathbb{Z}\}$ of x under the transformation f.

The words generated by a subshift of finite type form a regular language, but not all regular languages arise from subshifts of finite type.

We say that Σ_A is *transitive* if for every i, j, there is some n such that $A_{ij}^n > 0$. We say that Σ_A is *mixing* if there is some n such that $A_{ij}^n > 0$ for every i, j. Note the order of the quantifiers; mixing implies transitive, but transitive does not imply mixing. A typical example of a subshift of finite type that is transitive but not mixing is given by the matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Exercise: Suppose $f : X \to X$ is a continuous self map of a compact metric space X. Show that the following are equivalent,

1. There is an $x \in X$ such that $\overline{\{f^n(x) : n \in \mathbb{N}\}} = X$.
2. For every pair of nonempty open subsets U and V of X, there is some $n > 0$ such that $f^{-n}(U) \cap V \neq \emptyset$.

Such an f is said to be (topologically) transitive.

Exercise:Suppose as before that $f : X \to X$ is a continuous self map of a compact metric space X. Show that the following are equivalent,

1. For every increasing subsequence $(n_i)_i \subset \mathbb{N}$, there is an $x \in X$ such that $\overline{\{f^{n_i}(x) : n \in \mathbb{N}\}} = X$.
2. For every pair of nonempty open subsets U and V of X, there is some $N > 0$ such that for every $n \geq N$, $f^{-n}(U) \cap V \neq \emptyset$.

Such an f is said to be topologically mixing.

Now let M be a compact manifold. A C^1 map $f : M \to M$ is *expanding* if there is a Riemannian metric on M whose distance d satisfies, for some $\theta > 1$ and $\delta > 0$,

$$d(x, y) \leq \delta \Rightarrow d(f(x), f(y)) \geq \theta d(x, y)$$

It is straightforward to check that this is equivalent to the infinitesmal condition

$$\inf_{x \in M} \inf_{\|v\|=1} \|Df_x(v)\| \geq \theta > 1$$

The main example we will work with is $M = \mathbb{R}/\mathbb{Z}$, $f(x) = m \cdot x \pmod 1$, where $m \in \mathbb{Z}$, $m \neq 0, \pm 1$.

Some history of expanding maps: Franks proved that on the torus $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$, any expanding map $F : \mathbb{T}^n \to \mathbb{T}^n$ is topologically conjugate to its linear action on the homology torus

$$F_* : H_1(\mathbb{T}^n; \mathbb{R})/H_1(\mathbb{T}^n; \mathbb{Z}) \to H_1(\mathbb{T}^n; \mathbb{R})/H_1(\mathbb{T}^n; \mathbb{Z})$$

Recall that $H_1(\mathbb{T}^n; \mathbb{R}) \cong \mathbb{R}^n$, $H_1(\mathbb{T}^n; \mathbb{Z}) \cong \mathbb{Z}^n$, so F_* is in fact a self map of a torus, and Franks showed that up to a continuous change of coordinates, F_* and F are the same. This conclusion was later extended to compact nilmanifolds, which can either be thought of as iterated circle bundles over a torus, or as quotients of nilpotent Lie groups by a lattice, and then later to infranilmanifolds as well, which are nilpotent Lie groups quotiented by a combination of the left action of the group and the automorphism group of the Lie group. When the Lie group is $\mathbb{R}^n$, these additional possibilities are commonly known as the crystallographic groups.

Franks and Shub proved that if you have a compact manifold that admits an expanding map, then its fundamental group has polynomial growth, i.e., if one takes a finite list of generators for the fundamental group then the number of group elements which can be written as a word of length n in these generators grows polynomially with n. Gromov subsequently proved his famous theorem on groups of polynomial growth, which states that any finitely generated group of polynomial growth has a finite index subgroup which is nilpotent. By combining Gromov's theorem with the results of Franks and Shub, a classification of expanding maps was reached: all expanding maps of compact manifolds are topologically conjugate to an algebraic expanding map on an infranilmanifold.

We return to $M = \mathbb{R}/\mathbb{Z}$, f an expanding map, but note that the techniques below generalize in a straightforward manner to higher dimensional tori. We fix the flat Riemannian metric on $\mathbb{R}/\mathbb{Z}$, and note that f is expanding if and only if $|f'| > 1$ everywhere. Furthermore, since we can write

$$\deg f = \int_{\mathbb{R}/\mathbb{Z}} f'(x)\, dx$$

we see that if $|f'| > 1$ everywhere, then $|\deg(f)| > 1$, so that expanding maps are nontrivial covering maps of M. For each $m \in \mathbb{Z}$, let $E_m(x) = m \cdot x \pmod 1$ be the standard map of degree m on $\mathbb{R}/\mathbb{Z}$.

Theorem 0.1. *Let $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ be an expanding map. Then there is a homeomorphism $h : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ such that $f(h(x)) = h(E_m(x))$, where $m = \deg f$.*

Proof. We give an outline of the proof. The proof is easily adapted from $\deg f > 1$ to $\deg f < -1$, so we will assume throughout that $m = \deg f > 1$. Let $\pi : \mathbb{R} \to \mathbb{R}/\mathbb{Z}$ be the projection. We begin by lifting f to a map $F : \mathbb{R} \to \mathbb{R}$ such that $F(x+1) = F(x) + m$, $(\pi \circ F)(x) = (f \circ \pi)(x)$, and $F(0) \in [0, 1)$. Our goal is to find an injective continuous map $H : \mathbb{R} \to \mathbb{R}$ such that $H(F(x)) = mH(x)$, $H(0) \in [0, 1)$, and $H(x+1) = H(x)+1$. If we are able to find such an H, then H descends to a homeomorphism $h : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ satisfying $f(h(x)) = h(E_m(x))$, and we are done.

We set this up as a fixed point problem in a certain function space. We are looking for an H which satisfies $H = \frac{1}{m} H \circ F$ with some additional properties. We interpret the right hand side as an operator on continuous funcitons; H is then a fixed point of this operator. More precisely, define

$$\mathcal{G} = \{G : \mathbb{R} \to \mathbb{R} |\ G \text{ is continuous, nondecreasing, } G(0) \in [0, 1], \text{ and } G(x+1) = G(x) + 1\}$$

The four conditions that functions in $\mathcal{G}$ must satisfy are clearly closed under uniform limits. The last periodicity condition ensures that $\mathcal{G}$ is actually a complete metric space with respect to the uniform metric on continuous functions (since each function G is determined by its values on $[0, 1]$). For $G \in \mathcal{G}$, let $\Gamma(G) = \frac{1}{m} G \circ F$.

Exercise: $\Gamma(\mathcal{G}) \subset \mathcal{G}$.

Exercise: Γ is a contracting map in the C^0 norm (i.e., the uniform norm) on $\mathcal{G}$. In fact, for every $G_1, G_2 \in \mathcal{G}$,

$$\|\Gamma(G_1) - \Gamma(G_2)\|_0 \leq \frac{1}{m} \|G_1 - G_2\|_0$$

The contraction mapping theorem then implies that Γ has a unique fixed point in $\mathcal{G}$. Thus we have a continuous nondecreasing function $H : \mathbb{R} \to \mathbb{R}$ which satisfies $\Gamma(H) = H$, $H(x + 1) = H(x) + 1$, and $H(0) \in [0, 1]$. If $H(0) = 1$, we can translate by -1 so that $H(0) \in [0, 1)$.

It only remains to show that H is a homeomorphism. So far we have only actually used the facts that F was increasing and that $F(x + 1) = F(x) + m$, i.e., F came from a degree m map of $\mathbb{R}/\mathbb{Z}$. We can define another operator

$$\Lambda(G) = F^{-1} \circ G \circ \widetilde{E}_m$$

on $\mathcal{G}$, where we use $\widetilde{E}_m$ to denote the lift of E_m with $\widetilde{E}_m(0) = 0$. Since F expands distances uniformly, F^{-1} contracts distances uniformly, and so Λ is a contracting map with a unique fixed point K as well. Then

$$H \circ K \circ \tilde{E}_m = H \circ F \circ K = \frac{1}{m} \circ H \circ K$$

So that $H \circ K$ is the unique fixed point of the operator Γ in the special case $F = \tilde{E}_m$. But we know that the identity is also a fixed point of Γ in this case, so we get that $H \circ K = Id$. Similar reasoning shows that $K \circ H$ is the identity as well, so $K = H^{-1}$ and the proof is complete. $\square$

It's easy to see from this proof that the homeomorphism h satisfying $mx = h(f(h^{-1}(x)))$ constructed there is unique. Our next natural question is then, how nice is h? In general we cannot expect h to be differentiable. But we can always establish that h is Hölder continuous, with Hölder exponent depending on how far $|f'|$ deviates from being constant. First note that there exist $1 < \theta \leq m \leq \Theta$ such that

$$\theta^n \leq \frac{d(F^n(x), F^n(y))}{d(x,y)} \leq \Theta^n,$$

these distances being measured in the universal cover, with F the lift of f as described in the proof of the preceding theorem. Given points x and y, choose n such that $d(F^n(x), F^n(y)) \asymp 1$, i.e., $d(F^n(x), F^n(y))$ differs from 1 by a multiplicative factor close to 1 that does not depend on x, y, or n. Let H be the lift of h. Then

$$\widetilde{E}_m^n = H \circ F^n \circ H^{-1}$$

and therefore $H = \frac{1}{m^n} H \circ F$. H increasing and satisfying $H(x+1) = H(x)+1$ implies that $d(H(F^n(x)), H(F^n(y))) \asymp 1$ as well. So we conclude that

$$d\left(\frac{1}{m^n} H(F^n(x)), \frac{1}{m^n} H(F^n(y))\right) \asymp \frac{1}{m^n}$$

Therefore $d(H(x), H(y)) = O(m^{-n})$. On the other hand, n was chosen so that $d(F^n(x), F^n(y)) \asymp 1$, and therefore $d(x,y) \geq \Theta^{-n}$. Collecting this all together, we get a bound

$$d(H(x), H(y)) \leq O(d(x,y)^{\beta})$$

where $\beta = \frac{\log m}{\log \Theta}$.

An alternative method of proof of Hölder continuity is to show that the contraction Γ actually preserves Hölder continuous maps with a certain exponent.

One of the main corollaries of our theorem that we will be interested in is the following: For any expanding map f of degree m, we can code f by a one-sided shift on an alphabet with m letters. More precisely, there is a surjective continuous map $g : \Sigma_m^+ \to \mathbb{R}/\mathbb{Z}$ which is injective off of a countable subset of Σ_m^+, such that $f \circ g = g \circ \sigma$.

Exercise: Derive this corollary by combining the topological conjugacy of f to the standard expanding map E_m of degree m with the coding of the map E_m given by interpreting points of Σ_m^+ as base m expansions of real numbers in $[0, 1]$.

If we fix a metric on $\mathbb{R}/\mathbb{Z}$, we get a function $|f'| : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$ which we can pull back via g to a *potential* $\varphi : \Sigma_m^+ \to \mathbb{R}_{>0}$ defined by

$$\varphi(\omega) = \log |f'(g(\omega))|$$

We will see that the properties of this potential are intimately related to the statistical behavior of f.