Dynamical Properties of Pseudo-Anosov Maps
Student-Directed Colloquium

1 Preliminaries

1.1 Thurston's Classification

Thurston showed that all diffeomorphisms $f : \mathbb{T}^2 \to \mathbb{T}^2$ of the two-dimensional torus ae isotopic to one of the following toral automorphisms $A \in \mathrm{SL}(2, \mathbb{Z})$:

- A has distinct (conjugate) complex eigenvalues ($\lambda_1 = \overline{\lambda}_2$, $|\lambda_i| = 1$), and A is of finite order;
- A has repeated eigenvalues of 1 or -1, respectively resulting (after a change of coordinates) in one of the following *Dehn twists*:

$$A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad A = \begin{pmatrix} -1 & a \\ 0 & -1 \end{pmatrix}$$

- A has distinct irrational eigenvalues whose product is 1, making A an *Anosov diffeomorphism.*

This classification of toral diffeomorphisms can be extended to compact manifolds like so:

Theorem 1 (Nielson-Thurston Classification)**.** *Any diffeomorphism g of a compact surface M is isotopic to a map $f : M \to M$ satisfying one of the following properties:*

- *f is of finite order (that is, $f^n = \mathrm{id}$ for some $n \geq 1$);*
- *f is reducible, leaving invariant a closed curve;*
- *f is pseudo-Anosov.*

As with many classification theorems in geometry and analysis, one can think of these results respectively as being "elliptic" (stable), "parabolic", and "hyperbolic".

1.2 Uniformly Hyperbolic Systems

Definition 1. Let $U \subset M$ be open so that $f : U \to f(U)$ is a diffeomorphism. A compact f-invariant set $\Lambda \subset U$ is a **hyperbolic set** if there is a $\lambda \in (0, 1)$, $C > 0$, and a splitting $T_xM = E^s(x) \oplus E^u(x)$ at each tangent plane for $x \in \Lambda$ so that:

1. $\|Df_x^n v\| \leq C\lambda^n \|v\|$ for every $v \in E^s(x)$, $n \geq 0$;
2. $\|Df_x^{-n} v\| \leq C\lambda^n \|v\|$ for every $v \in E^u(x)$, $n \geq 0$;
3. $Df_x (E^s(x)) = E^s(f(x))$ and $Df_x (E^u(x)) = E^u(f(x))$.

If $U = M$, then $f : M \to M$ is an **Anosov diffeomorphism**.

NOTE: Hyperbolic sets are common. Anosov diffeomorphisms are not.

Definition 2. Let N be a nilpotent simply connected real Lie group, and let $\Gamma \subseteq N$ be a closed subgroup. Then the quotient manifold $M = N/\Gamma$ is a **nilmanifold.**

Let F be a finite set of automorphisms, and define the semidirect product $N \rtimes F$. An orbit space of N by a discrete subgroup of $N \rtimes F$ acting freely on N is a **infranilmanifold**.

All known Anosov systems are (up to conjugation with a homeomorphism) either:

- hyperbolic matrices in $\mathrm{SL}(n, \mathbb{Z})$ acting on $\mathbb{T}^n$; or
- induced systems on nilmanifolds (quotiented over *discrete* subgroups Γ) or infranilmanifolds.

Theorem 2. *The only 2-manifold admitting an Anosov diffeomorphism is* $\mathbb{T}^2$.

Conjecture 1. *All Anosov diffeomorphisms* $f: M \to M$ *are* topologically transitive*: there exists a dense orbit of* f*, or equivalently (on a manifold), there are open* $U, V \subseteq M$ *such that* $f^n(U) \cap V \neq \varnothing$ *for some* $n \in \mathbb{N}$.

Example 1 (Cat map)**.** Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z})$, which has eigenvalues $0 < \lambda^{-1} < 1 < \lambda$. Then A induces a map f on $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$.

Theorem 3. *If* $f : M \to M$ *is an Anosov diffeomorphism, then* M *has two transverse foliations* $\mathcal{F}^s = \{W^s(x)\}$, $\mathcal{F}^u = \{W^u(x)\}$ *so that for* $x \in M$,

- $W^s(x) = \left\{ y \in M : d\left(f^n(x), f^n(y)\right) \xrightarrow{n\to\infty} 0 \right\}$;
- $W^u(x) = \left\{ y \in M : d\left(f^{-n}(x), f^{-n}(y)\right) \xrightarrow{n\to-\infty} 0 \right\}$;
- $T_x W^s(x) = E^s(x)$ *and* $T_x W^u(x) = E^u(x)$;
- $f\left(W^s(x)\right) = W^s(f(x))$ *and* $f\left(W^u(x)\right) = W^u(f(x))$.

For $M = \mathbb{T}^2$, these *stable* and *unstable* manifolds are the affine eigenspaces of the hyperbolic toral automorphism $A \in \mathrm{SL}(2, \mathbb{Z})$.

For $\varepsilon > 0$, the *local stable/unstable submanifolds* at a point $x \in M$ are:

$$W^s_\varepsilon(x) := \{y \in M : d\left(f^n(x), f^n(y)\right) \le \varepsilon \, \forall n \ge 0\}$$

and

$$W^u_\varepsilon(x) := \left\{y \in M : d\left(f^{-n}(x), f^{-n}(y)\right) \le \varepsilon \, \forall n \ge 0\right\}$$

And,

$$W^s(x) = \bigcup_{n\ge 0} W^s_\varepsilon(f^n(x)) \text{ and } W^u(x) = \bigcup_{n \ge 0} W^u_\varepsilon(f^{-n}(x)).$$

Theorem 4 (Local product structure)**.** *There is* $\varepsilon_0, \delta_0 > 0$ *such that if* $x, y \in M$ *have* $d(x, y) < \delta_0$*, then* $[x, y] := W^s(x) \cap W^u(y)$ *consists of a single point.*

Given a finite (or countable) alphabet of symbols $\{R_1, \ldots, R_n\}$, the **symbolic shift space** is the dynamical system (Ω_n, σ), where $\Omega_R = \{R_1, \ldots, R_n\}^{\mathbb{Z}}$ (with the product topology of discrete topologies) and $\sigma(\omega) = \sigma\left(\omega_i\right)_{i\ge 0} = \left(\omega_{i+1 \ge 0}\right)_{i \ge 0}$. A **subshift of finite type** is a σ-invariant closed subset Ω_A of Ω_R.

Definition 3. A **Markov partition** of a smooth dynamical system $f : M \to M$ is a collection of disjoint open sets $\{R_1, \ldots, R_n\}$, called "rectangles", so that:

- $\mu\left(\bigcup_{i=1}^n R_i\right) = \mu(M)$, where μ is the Riemannian volume of M;
- There is a subshift of finite type Ω_A of Ω_R such that, for any $\omega \in \Omega_R$ (with $\omega_i = R_{k_i}$ for each i), the intersection $\bigcap_{i=-\infty}^{\infty} f^{-i}\left(R_{k_i}\right)$ is a single point $x = \pi(\omega)$, and the map $\pi : \Omega_A \to M$ is a semiconjugacy: $\pi \circ \sigma = f \circ \pi$.

In this sense, a Markov partition allows us to analyze our smooth system as a symbolic system almost everywhere.

For Anosov systems or systems with local product structure, the rectangles are formed so the edges are leaves of these stable and unstable foliations.

Theorem 5. *Anosov systems* $f : M \to M$ *admit Markov partitions of arbitrarily small diameter.*

2 Measured Foliations

Definition 4. A *measured foliation with singularities* is a triple $(\mathcal{F}, S, \nu)$, where:

- $S = \{x_1, \ldots, x_m\}$ is a finite set of points in M, called *singularities*;
- $\mathcal{F} = \widetilde{\mathcal{F}} \uplus \mathcal{S}$ is a singular foliation of M, where $\widetilde{\mathcal{F}}$ is a collection of C^∞ curves and $\mathcal{S}$ is a partition of S into points;
- ν is a *transverse measure*; in other words, ν is a measure defined on each curve on M transverse to the leaves of $\widetilde{\mathcal{F}}$;

and the triple satisfies the following properties:

1. There is a finite atlas of C^∞ charts $\varphi_k : U_k \to \mathbb{C}$ for $k = 1, \ldots, \ell$, $\ell \geq m$.
2. For each $k = 1, \ldots, m$, there is a number $p = p(k) \geq 2$ of elements of $\mathcal{F}$ meeting at x_k (these elements are called *prongs* of x_k) such that:
   (a) $\varphi_k(x_k) = 0$ and $\varphi_k(U_k) = D_{a_k} := \{z \in \mathbb{C} : |z| \leq a_k\}$ for some $a_k > 0$;
   (b) if $C \in \widetilde{\mathcal{F}}$, then the components of $C \cap U_k$ are mapped by φ_k to sets of the form
   $$\left\{z \in \mathbb{C} : \operatorname{Im}\left(z^{p/2}\right) = \text{constant}\right\} \cap \varphi_k(U_k);$$
   (c) the measure $\nu | U_k$ is the pullback under φ_k of
   $$\left|\operatorname{Im}\left(dz^{p/2}\right)\right| = \left|\operatorname{Im}\left(z^{(p-2)/2} dz\right)\right|.$$
3. For each $k > m$, we have:
   (a) $\varphi_k(U_k) = (0, b_k) \times (0, c_k) \subset \mathbb{R}^2 \approx \mathbb{C}$ for some $b_k, c_k > 0$;
   (b) If $C \in \widetilde{\mathcal{F}}$, then components of $C \cap U_k$ are mapped by φ_k to lines of the form
   $$\{z \in \mathbb{C} : \operatorname{Im} z = \text{constant}\} \cap \varphi_k(U_k).$$
   (c) The measure $\nu | U_k$ is given by the pullback of $|\operatorname{Im} dz|$ under φ_k.

Remark 1. Properties (2) and (3) in the above definition ensure that ν is holonomy-invariant. In particular, if γ and γ' are isotopic curves in $M \setminus S$, and the initial and terminal points in γ and γ' lie in the same leaf of $\widetilde{\mathcal{F}}$, then $\nu(\gamma) = \nu(\gamma')$.

Definition 5. A surface homeomorphism f of a manifold M is *pseudo-Anosov* if there are measured singular foliations $(\mathcal{F}^s, S, \nu^s)$ and $(\mathcal{F}^u, S, \nu^u)$ (with the same finite set of singularities $S = \{x_1, \ldots, x_m\}$) and an atlas of C^∞ charts $\varphi_k : U_k \to \mathbb{C}$ for $k = 1, \ldots, \ell$, $\ell > m$, satisfying the following properties:

1. f is differentiable, except on S.
2. There are two measured singular foliations $(\mathcal{F}^s, S, \nu^s)$ and $(\mathcal{F}^u, S, \nu^u)$, which share the same singular set S on which f is not differentiable, and for each $x_k \in S$, $\mathcal{F}^s$ and $\mathcal{F}^u$ have the same number $p(k)$ of prongs at x_k.
3. The leaves of $\mathcal{F}^s$ and $\mathcal{F}^u$ intersect transversally at nonsingular points;
4. both singular foliations $\mathcal{F}^s$ and $\mathcal{F}^u$ are f-invariant;
5. There is a constant $\lambda > 1$ such that
   $$f(\mathcal{F}^s, \nu^s) = (\mathcal{F}^s, \nu^s/\lambda) \quad \text{and} \quad f(\mathcal{F}^u, \nu^u) = (\mathcal{F}^u, \lambda\nu^u);$$

6. For each $k = 1, \ldots, m$, we call U_k a *singular neighborhood*, where we have:
   (a) $\varphi_k(x_k) = 0$ and $\varphi_k(U_k) = D_{a_k}$ for some $a_k > 0$;
   (b) if C is a curve leaf in $\mathcal{F}^s$, then the components of $C \cap U_k$ are mapped by φ_k to sets of the form
   $$\left\{z \in \mathbb{C} : \operatorname{Re}\left(z^{p/2}\right) = \text{constant}\right\} \cap D_{a_k};$$
   (c) if C is a curve leaf in $\mathcal{F}^u$, then the components of $C \cap U_k$ are mapped by φ_k to sets of the form
   $$\left\{z \in \mathbb{C} : \operatorname{Im}\left(z^{p/2}\right) = \text{constant}\right\} \cap D_{a_k};$$
   (d) the measures $\nu^s|U_k$ and $\nu^u|U_k$ are given by the pullbacks of
   $$\left|\operatorname{Re}\left(dz^{p/2}\right)\right| = \left|\operatorname{Re}\left(z^{(p-2)/2}dx\right)\right|$$
   and
   $$\left|\operatorname{Im}\left(dz^{p/2}\right)\right| = \left|\operatorname{Im}\left(z^{(p-2)/2}dx\right)\right|$$
   under φ_k, respectively.
7. For each $k > m$, we call U_k a *regular neighborhood*, where we have:
   (a) $\varphi_k(U_k) = (0, b_k) \times (0, c_k) \subset \mathbb{R}^2 \approx \mathbb{C}$ for some $b_k, c_k > 0$;
   (b) If C is a curve leaf in $\mathcal{F}^s$, then components of $C \cap U_k$ are mapped by φ_k to lines of the form
   $$\{z \in \mathbb{C} : \operatorname{Re} z = \text{constant}\} \cap \varphi_k(U_k);$$
   (c) If C is a curve leaf in $\mathcal{F}^u$, then components of $C \cap U_k$ are mapped by φ_k to lines of the form
   $$\{z \in \mathbb{C} : \operatorname{Im} z = \text{constant}\} \cap \varphi_k(U_k);$$
   (d) the measures $\nu^s|U_k$ and $\nu^u|U_k$ are given by the pullbacks of $|\operatorname{Re} dz|$ and $|\operatorname{Im} dz|$ under φ_k, respectively.

Proposition 1. *A pseudo-Anosov homeomorphism $f : M \to M$ is smooth except at its singularities. For $x \in M \setminus S$, $T_xM = T_x\mathcal{F}^s(x) \oplus T_x\mathcal{F}^u(x)$, and in these coordinates, $Df_x(\xi^s, \xi^u) = (\xi^s/\lambda, \lambda\xi^u)$, where ξ^s and ξ^u are nonzero vectors in $T_x\mathcal{F}^s(x)$ and $T_x\mathcal{F}^u(x)$, $\mathcal{F}^s(x)$ and $\mathcal{F}^u(x)$ represent the curve containing x in the respective foliation, and λ is the dilation factor for f.*

Proposition 2. *A pseudo-Anosov surface homeomorphism $f : M \to M$ preserves a smooth invariant probability measure ν defined locally as the product of ν^s on $\mathcal{F}^u$-leaves with ν^u on $\mathcal{F}^s$-leaves. This probability measure ν has a density with respect to Lebesgue measure m, which vanishes at singularities.*

Proposition 3. *Every pseudo-Anosov homeomorphism of a surface M admits a finite Markov partition of arbitrarily small diameter. Conjugated to the symbolic system induced by this Markov partition, with the measure ν as in the preceding proposition, (M, f, ν) is the full Bernoulli shift (i.e. is maximally chaotic).*

I want to emphasize that these are more powerful in a sense than Anosov systems in that they are defined on essentially any manifold in some capacity. However, by first Thurston classification, all hyperbolic maps of the torus (i.e. pseudo-Anosov candidates) are regular Anosov diffeomorphisms.

<u>Construction:</u> One can construct pseudo-Anosov systems by taking a linear map of the torus and lifting it to a map on a covering space. But like Anosov systems, pseudo-Anosov systems are hard to construct explicitly.

3 Smooth Realizations

Let $p = p(x_0)$, and let $\varphi_0 : U_0 \to \mathbb{C}$ be the chart described in part (6) of definition (5). The *stable* and *unstable prongs* at x_0 are the leaves P^s_j and P^u_j, $j = 0, \ldots, p-1$ of $\mathcal{F}^s$ and $\mathcal{F}^u$, respectively, whose endpoints meet at x_0. Locally, they are given by:

$$P^s_j = \varphi_0^{-1}\left\{\rho e^{i\tau} : 0 \le \rho < a_0,\ \tau = \frac{2j+1}{p}\pi\right\},$$
$$\text{and} \quad P^u_j = \varphi_0^{-1}\left\{\rho e^{i\tau} : 0 \le \rho < a_0,\ \tau = \frac{2j}{p}\pi\right\}.$$

For simplicity, assume $f(P^s_j) \subseteq P^s_j$ for all $j = 1, \ldots, p$. Furthermore, we define the *stable* and *unstable sectors* at x_0 to be the regions in U_0 bounded by the stable (resp. unstable) prongs:

$$S^s_j = \varphi_0^{-1}\left\{\rho e^{i\tau} : 0 \le \rho < a_0,\ \frac{2j-1}{p}\pi \le \tau \le \frac{2j+1}{p}\pi\right\},$$
$$\text{and} \quad S^u_j = \varphi_0^{-1}\left\{\rho e^{i\tau} : 0 \le \rho < a_0,\ \frac{2j}{p}\pi \le \tau \le \frac{2j+2}{p}\pi\right\}.$$

The strategy for creating our diffeomorphism g is adapted from section 6.4.2 of [**?**]. In each stable sector, we apply a "slow-down" of the trajectories, followed by a change of coordinates ensuring the resulting diffeomorphism g preserves the measure induced by a convenient Riemannian metric.

Let $F : \mathbb{C} \to \mathbb{C}$ be the map $s_1 + is_2 \mapsto \lambda s_1 + is_2/\lambda$. Note F is the time-1 map of the vector field V given by

$$\dot{s}_1 = (\log\lambda)s_1, \quad \dot{s}_2 = -(\log\lambda)s_2.$$

Let $0 < r_1 < r_0 < a_0$, and define $\widetilde{r}_0$ and $\widetilde{r}_1$ by $\widetilde{r}_j = (2/p)r_j^{p/2}$ for $j = 0, 1$. Define a "slow-down" function Ψ_p for the p-pronged singularity on the interval $[0, \infty)$ so that:

(1) $\Psi_p(u) = (p/2)^{(2p-4)/p}u^{(p-2)/p}$ for $u \le \widetilde{r}_1^2$;

(2) Ψ_p is C^∞ except at 0;

(3) $\Psi_p'(u) \ge 0$ for $u > 0$;

(4) $\Psi_p(u) = 1$ for $u \ge \widetilde{r}_0^2$.

Consider the vector field V_{Ψ_p} defined by

$$\dot{s}_1 = (\log\lambda)s_1\Psi_p\left(s_1^2 + s_2^2\right) \quad \text{and} \quad \dot{s}_2 = -(\log\lambda)s_2\Psi_p\left(s_1^2 + s_2^2\right).$$

Let G_p be the time-1 map of the vector field V_{Ψ_p}. Assume r_1 is chosen to be small enough so that $G_p = F$ on a neighborhood of the boundary of $D_{\widetilde{r}_0}$, and assume r_0 is chosen to be small enough so that D_{r_0} is disjoint from the other open sets in the atlas defined in Definition 5 parts (6) and (7).

Let $\widetilde{a}_0 = (2/p)a_0^{p/2}$, and define the coordinate change $\Phi_j : \varphi_0 S^s_j \to \{z : \mathrm{Re} z \ge 0\} \cap D_{\widetilde{a}_0}$ by

$$\Phi_j(z) = (2/p)z^{p/2} = w = s_1 + is_2.$$

Define $g : M \to M$ by $g(x) = f(x)$ outside D_{r_0}, and meanwhile define g on each sector S^s_j by

$$g(x) = \varphi_0^{-1}\Phi_j^{-1}G_p\Phi_j\varphi_0(x).$$

Theorem 6. *The function g defined above is well-defined on the unstable prongs and singularity. It is in fact a diffeomorphism topologically conjugate to f, and for any $\varepsilon > 0$, r_0 and r_1 can be chosen so that $\|f - g\|_{C^0} < \varepsilon$. In particular, g admits a Markov partition of arbitrarily small diameter.*

Next we define a Riemannian metric $\zeta = \langle \cdot, \cdot \rangle$ on $M \setminus \{x_0\}$ with respect to which the map g is area-preserving. In the stable sector $S_j^s \cap \varphi_0^{-1}(D_{\widetilde{a}_0})$, we consider the coordinates $w = s_1 + is_2$ given by $\Phi_j \circ \varphi_0$ defined above. Outside of this neighborhood, we use the coordinates $z = s_1 + is_2$. In both sets of coordinates, the stable and unstable transversal measures are $\nu^s = |ds_1|$ and $\nu^u = |ds_2|$. On stable sectors in $M \setminus \{x_0\}$, we define the Riemannian metric ζ to be the pullback of $\left(ds_1^2 + ds_2^2\right) / \Psi\left(s_1^2 + s_2^2\right)$ under $\Phi_j \circ \varphi_0$. In regular neighborhoods (U_k, φ_k), we define $\zeta = \varphi_k^* \left(ds_1^2 + ds_2^2\right)$. Since $\widetilde{r}_0$ is chosen so that $\varphi_0^{-1}\left(D_{\widetilde{r}_0}\right)$ is disjoint from regular neighborhoods, and $\Psi(u) \equiv 1$ for $u \geq \widetilde{r}_0^2$, ζ is consistently defined on chart overlaps. One can further show that ζ agrees with the Euclidean metric in $\varphi_0^{-1}\left(D_{\widetilde{r}_0}\right)$. So ζ can be extended to a Riemannian metric on M.

Theorem 7. *The diffeomorphism $g : M \to M$ is area preserving with respect to the Riemannian metric ζ defined above.*

4 Thermodynamics

My main research objective is to effect *thermodynamic formalism*. Given a probability measure μ on a compact manifold M with a μ-preserving map $f : M \to M$, the *entropy* of the map is $h_\mu(f)$.

Without getting into details, $h_\mu(f)$ is the *amount of "randomness"* in the system in the following sense: suppose $\varphi : M \to \mathbb{R}$ is continuous. We can think of φ as a random variable (an *observable*) on the probability space (M, μ). Then $X_n := \varphi \circ f^n$ is a stochastic process, and greater entropy leads to more scattered or "chaotic" values of X_n as $n \to \infty$.

We often have a *potential function* $\varphi : M \to \mathbb{R}$. In principle, this potential function can be any continuous function. But the potential function I'm most interested in is $\varphi_t(x) = -t \log \left|df|_{E^u(x)}\right|$. This is the *geometric potential*: it's the potential function of interest for a dynamical system, since it shows how expansive the dynamics are.

For example, for pseudo-Anosov homeomorphisms, $\left|df|_{E^u(x)}\right| = \lambda > 1$ for every $x \in M \setminus S$, so $\varphi_t(x) \equiv -t \log \lambda$.

But the geometric potential is a little bit more subtle for globally smooth pseudo-Anosov models, in that it vanishes at the singularities but may fail to be Hölder continuous there.

In general, we are interested in measures μ that maximize the quantity

$$h_\mu(f) + \int_M \varphi \, d\mu.$$

In a sense, the integral of the potential represents the "total energy" the system possesses with that potential energy field (strictly speaking times a negative constant, which we usually absorb into φ). So optimizing this quantity is a mathematical formalization of the thermodynamic principle that nature "maximizes entropy and minimizes energy".