THE STRUCTURE OF THE SPACE OF INVARIANT MEASURES

VAUGHN CLIMENHAGA

Broadly, a dynamical system is a set X with a map $f\colon X\circlearrowleft$. This is discrete time. Continuous time considers a flow $\varphi_t\colon X\circlearrowleft$. We will mostly consider discrete time.

Often X has some extra structure that the map f respects.

- X a smooth manifold, f a diffeomorphism
- X a metric space, f continuous
- (X, μ) a measure space, f measure-preserving

f **is measure-preserving /** μ **is** f**-invariant:** $\mu(f^{-1}E) = \mu(E)$ for all measurable $E \subset X$. Equivalently, $\int \varphi \circ f\, d\mu = \int \varphi\, d\mu$ for all $\varphi \in L^1$.

Classical source of examples: X is a smooth manifold, φ_t is the flow of a conservative vector field. Then each φ_t both respects smooth structure and preserves volume.

Smooth manifolds have many measures, not just volume. But having an **invariant** measure opens up the rich toolbox of ergodic theory. For example, "time average = space average" (Birkhoff ergodic theorem).

Aside: What about the dissipative case? What measure should we use instead of volume, when volume is not invariant? Big question, skip for now.

Connections between topological and measure-theoretic structure are illustrated by two "toy" examples on $X = S^1 \subset \mathbb{C}$.

(1) $R_\alpha\colon z \mapsto ze^{2\pi i\alpha}$ for α an irrational parameter.
(2) $T_2\colon z \mapsto z^2$.

These represent two extremes of dynamical behaviour: R_α is **elliptic**, T_2 is **hyperbolic**.

Date: February 25, 2013.

First consider these topologically. Both are **topologically transitive** – any two open sets can be connected by an orbit. This is an irreducibility criterion.

Aside: Transitivity equivalent to existence of a dense orbit. Weaker than **minimality** – *every* orbit is dense. R_α is minimal, T_2 is not.

What about invariant measures? For both, Lebesgue measure is invariant and **ergodic**: every f-invariant set E has $\mu(E) = 0$ or 1.

This implies, via **Birkhoff ergodic theorem**: if $\varphi \in L^1$, then for Leb-a.e. x,

$$\frac{1}{n} S_n \varphi(x) = \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^n x) \to \int \varphi \, dx.$$

This is the **law of large numbers** for the "random variables" $\varphi, \varphi \circ f, \varphi \circ f^2, \dots$" What about other statistical properties, and the nature of this convergence?

- Is this convergence uniform in x?
- How quickly does convergence happen? Look at $E_n := \{x \mid \frac{1}{n} S_n \varphi(x) > \epsilon\}$. How quickly does the measure of E_n go to 0?

Fact: *Although ergodicity of Lebesgue measure determines the asymptotic behaviour of Lebesgue-a.e. trajectory for both R_α and T_2, the nature of the convergence to this asymptotic behaviour is strongly contingent on the presence of other invariant measures.*

R_α: Lebesgue is the **only** invariant measure.

T_2: There are many, many others. Any periodic orbit supports an invariant (ergodic) measure, and there are 2^n fixed points of T_2^n.

Given $f \colon X \circlearrowleft$, let $\mathcal{M}_f$ be the collection of f-invariant Borel probability measures on X, and $\mathcal{M}_f^e$ the set of ergodic measures.

Geometrical interpretation: $\mathcal{M}_f^e$ is the set of extreme points of $\mathcal{M}_f$, and $\mathcal{M}_f$ is a **simplex** – elements of $\mathcal{M}_f$ are in 1-1 correspondence with probability measures on $\mathcal{M}_f^e$ (**ergodic decomposition**).

Consider R_α on k concentric circles. Each circle has exactly one ergodic measure. $\mathcal{M}_f$ is a $(k-1)$-simplex.

Question: When do two systems have the same $\mathcal{M}_f$ and $\mathcal{M}_f^e$? (Up to affine homeomorphism.)

First invariant: Number of extreme points (ergodic measures). Finite-dimensional simplices are affinely homeomorphic iff same number of extreme points. Also distinguishes countable/uncountable.

Consider R_α on countably many concentric circles, and R_α on unit disc. First has countable $\mathcal{M}_f^e$, second has uncountable.

Second invariant: Topology of $\mathcal{M}_f^e$. Becomes important when $\mathcal{M}_f^e$ infinite. All examples of R_α have $\mathcal{M}_f^e$ closed, while T_2 has $\mathcal{M}_f^e$ dense in $\mathcal{M}_f$.

Last property is important. Simplex with dense extreme points constructed in 1961 by E Poulsen. Abstract construction, no dynamics.

Universality of Poulsen simplex: In 1978, J Lindenstrauss, G Olsen, Y Sternfeld showed that if two simplices both have dense extreme points then they are affinely homeomorphic.

The extreme set of Poulsen's simplex is path-connected. So two conclusions from fact that (countable) set of periodic measures is dense in $\mathcal{M}_f$ for T_2:

- existence of uncountably many other ergodic measures;
- path-connectedness of $\mathcal{M}_f^e$.

Questions: How to describe other ergodic measures concretely? For which other systems is $\mathcal{M}_f$ the Poulsen simplex? What is connection between this fact and statistical properties?

Aside: Natural to ask for example of system where $\mathcal{M}_f^e$ is path-connected but not dense. R_α on disc does it but in a silly way - disjoint union of closed subsystems, and $\mathcal{M}_f^e$ is only one-dimensional.

A more sophisticated example is the **Dyck shift**. $X \subset \{0, 1, 2, 3\}^{\mathbb{Z}}$ defined by syntax rules on brackets, identifying $0, 1, 2, 3$ with $(,), [,]$. Map f is the left shift. Can show $\mathcal{M}_f^e$ connected but not dense.

Return to questions. Useful to think of other **symbolic systems** where $X \subset \Sigma_2^+ := \{0, 1\}^{\mathbb{N}}$ and $f = \sigma$. Connect to maps such as T_2 by fixing a partition of S^1 into two subsets and labelling each subset with 0 or 1.

For T_2, get $X = \Sigma_2^+$. Measure μ defined by $\mu[w]$, where $w \in \{0, 1\}^*$ and $[w]$ is set of sequences starting with w. Two important classes:

- $p_1 + p_2 = 1 \Rightarrow$ **Bernoulli** measure $\mu[w] = p_{w_1} \cdots p_{w_n}$.

- stochastic 2×2 matrix $\Rightarrow$ **Markov** $\mu[w] = p_{w_1} P_{w_1 w_2} \cdots P_{w_{n-1} w_n}$, where p a left eigenvector for P.

For T_2, no restrictions on what symbol sequences can appear. Corresponds to configurations on lattice: each site can be on or off, + or -, $\uparrow$ or $\downarrow$. Suggests language of **statistical mechanics**.

Can code R_α by $X \subset \Sigma_2$. Many restrictions, some very long-range.

Interactions of uniformly bounded range: **subshift of finite type**. More generally, **specification** property.

- Transitivity for shift space X means any set of words can be concatenated by putting some "buffers" in between.
- Specification means the buffers are uniformly short.

In 1970, K Sigmund showed that specification implies $\mathcal{M}_f^e$ is dense, hence $\mathcal{M}_f$ is the Poulsen simplex.

The space of invariant measures is often very large – how do we select a distinguished measure?

Topological entropy: exponential growth rate of number of words of length n. Call it $h(X)$.

Measure-theoretic entropy: growth rate of number of words of length n *needed to get to mass* $\frac{1}{2}$. Call it $h(\mu)$.

Variational principle: $h(X) = \sup\{h(\mu) \mid \mu \in \mathcal{M}_f^e\}$.

Pressure: Give words weights according to a potential function $\varphi \in C(X)$. Still get variational principle. Measure achieving supremum is an **equilibrium state**.

Aside: For smooth systems, another notion of distinguished measure is **SRB measure**. I have active research on these.

Various properties of $\mathcal{M}_f$ and $\mathcal{M}_f^e$:

- (C) $\mathcal{M}_f^e$ is path-connected.
- (D) $\mathcal{M}_f^e$ is dense in $\mathcal{M}_f$.
- (H) $\mathcal{M}_f^e$ is **entropy-dense** in $\mathcal{M}_f$ – can approximate in weak* and in entropy.
- (E) There exists a dense subspace $V \subset C(X)$ such that each $\varphi \in V$ has a unique equilibrium state.

SFT $\Rightarrow$ specification $\Rightarrow$ (E), (H), (D)

Conjecture: (E) implies (H). *(The idea is that (E) gives a way to map a very large vector space homeomorphically into $\mathcal{M}_f^e$. The image should be "large enough".)*

(E) implies various multifractal results. (VC, Nonlinearity)

(H) and (E) are important for large deviations properties: recall sets $E_n = \{x \mid \frac{1}{n} S_n\varphi(x) > \epsilon\}$, where $\int \varphi \, dx = 0$.

$$\lim_{n\to\infty} \frac{1}{n} \mathrm{Leb}(E_n) = \sup\{h(\mu) - \log 2 \mid \int \varphi \, d\mu > \epsilon\}.$$

Can get similar results anytime (E) holds (H Comman, J Rivera–Letelier 2010).

Problem: Specification is a very uniform phenomenon, and hence somehow rare. What non-uniform versions still give (E), LDP, etc?

Example: Fix $\beta > 1$, let $T_\beta \colon x \mapsto \beta x \pmod 1$. Code this into $X_\beta \subset \Sigma_b^+$, where $b = \lceil \beta \rceil$. Typically specification fails. But X_β has (E). (VC, DJ Thompson, 2013) Can use this to get LDP. (VC, DJ Thompson, K Yamamoto, in progress)