Monstrous Menagerie with Vandehey 7/2

Jason Schuchardt

July 3, 2019

1 Why is invariance actually useful? - Poincaré Recurrence

We resume from where we left off last time.

Theorem 1.1 (Poincaré Recurrence)**.** *Let (X, T, μ) be a dynamical system, with μ a T-invariant probability measure.*

Then for any set A with $\mu(A) > 0$, almost all points in A return to A infinitely often (as we iterate T).

Note that this doesn't address how frequently things come back.

Proof. Let F be the set of points in A such that they never return to A.

Consider $T^{-1}F$. This is the set of points that are in A after one application of T, but never return to A again.

So F and $T^{-1}F$ have to be disjoint sets. Likewise, we can take all the sets $T^{-j}F$. Then $T^{-j}F \cap T^{-k}F = \varnothing$ if $j \neq k$.

Now suppose F had positive measure. Then there exists a positive integer $n \in \mathbb{N}$ such that $\mu(F) > \frac{1}{n}$. Then consider the following

$$\mu(F \cup T^{-1}F \cup \cdots \cup T^{-n}F) = \mu(F) + \mu(T^{-1}F) + \cdots + \mu(T^{-n}F) = (n+1)\mu(F) > \frac{n+1}{n} > 1 = \mu(X).$$

This is a contradiction. Thus F has measure 0.

Consider

$$N = \bigcup_{i=0}^{\infty} T^{-i}F.$$

This is the set of all points, which enter A for the last time at some finite iteration n, and never return to A again.

This contains all points of X (not just A) which visit A at least once, but not infinitely often.

Note that $\mu(N) = 0$, since $\mu(F) = 0$, and μ is T-invariant. Then the set of all points of A which return to A infinitely often is

$$A \setminus N,$$

and since $\mu(N) = 0$, $\mu(A \setminus N) = \mu(A)$. ∎

2 Symbolic Shifts

Let $\mathscr{D}$ be a set of digits. Let $\mathscr{D}^{\mathbb{N}}$ be the set of "words in $\mathscr{D}$," i.e., the set of sequences with entries in $\mathscr{D}$.

Let σ denote the left shift on $\mathscr{D}^{\mathbb{N}}$, $(\sigma\omega)_i = \omega_{i+1}$.

Let $\Omega \subseteq \mathscr{D}^{\mathbb{N}}$ be any shift invariant subset, $\sigma\Omega = \Omega$ (note that this is now forwards shifting!).

Let $\mathscr{D} = \{0, 1\}$, and let Ω be the subset where 11 never occurs. This is shift invariant, since if you take a word that doesn't contain 11, its left shift still won't contains 11.

3 Isomorphisms

Let (X, T) correspond to base-b expansion. Consider $(\Omega = \{0, 1, \ldots, b-1\}^{\mathbb{N}}, \sigma)$. There is a natural (almost) bijection between these two spaces,

$$\varphi : X \to \Omega$$
$$\varphi(x) = a_1(x)a_2(x)a_3(x)\cdots .$$

While this isn't a bijection because there are some points with two expansions, this is almost a bijection in the sense that we can ignore a measure zero subset and it becomes a bijection.

φ respects the transformation in the sense that

$$\varphi \circ T = \sigma \circ \varphi,$$

since both T and σ correspond to chopping off the first digit.

Suppose we have an invariant measure on one of these spaces. Then we can get one on the other via φ.

For example, suppose μ is T-invariant on (X, T), and let ν be a measure on (Ω, σ) given by $\nu(A) = \mu(\varphi^{-1}(A))$.
Exercise 1. Prove that ν is a measure.

We'll prove that ν is invariant:

$$\nu(\sigma^{-1}A) = \mu(\varphi^{-1}\sigma^{-1}A) = \mu((\sigma\varphi)^{-1}A) = \mu((\varphi T)^{-1}A) = \mu(T^{-1}(\varphi^{-1}(A))) = \mu(\varphi^{-1}(A)) = \nu(A).$$

Note that if we wanted to, we could produce μ given ν by taking $\mu(A) = \nu(\varphi(A))$.

If we take measures μ, ν satisfying $\nu = \mu \circ (\varphi^{-1})$, we could say that these systems are isomorphic.

4 Bernoulli product measure

Let m be any probability measure on $\mathscr{D}$. Then define ν on $(\Omega = \mathscr{D}^{\mathbb{N}}, \sigma)$ by

$$\nu(C_{[a_1,\ldots,a_k]}) := m(a_1)m(a_2)\cdots m(a_k) = \nu(C_{a_1})\nu(C_{a_2})\cdots\nu(C_{a_k})$$

This produces a nice invariant measure. For example,

$$\nu(\sigma^{-1}(C_s)) = \nu\left(\bigcup_{d\in\mathscr{D}} C_{ds}\right) = \sum_{d\in\mathscr{D}} \nu(C_{ds}) = \sum_{d\in\mathscr{D}} \nu(C_d)\nu(C_s) = \nu(C_s)\sum_{d\in\mathscr{D}} m(d) = \nu(C_s)\cdot 1 = \nu(C_s)$$

Definition 4.1. This measure ν is the *Bernoulli product measure* corresponding to m.

Consider the $\left(\frac{1}{2}, 0, \frac{1}{2}\right)$- Bernoulli product measure on base-3 expansions.

Here we give no weight to the digit 1, and equal weight to the digits 0 and 2. Those familiar with the middle thirds Cantor set will notice that this measure gives the Cantor set full measure.

We can also consider the $\left(\frac{1}{3}, \frac{2}{3}\right)$-measure on base 2.

Picture: crazy bar charts. Split the interval in the middle, and then split those intervals again, and keep drawing the bar with height the measure of the intervals.

5 Ergodicity

History: Ergodic theory is a bit of a strange word. Where does it come from. "Ergos" means work. So ergodic theory is about work being done. This subject started off in statistical mechanics. We have a box full of particles, and want to understand what those particles are doing.

We could freeze time for an instant and look at what all the particles are doing and average them out. Or we could look at a single point, and see where it goes over time, and average it out over a long period of time. These two averages should be equal.

The slogan is: Time average should equal the space average.

Definition 5.1. Let (X, μ, T) be a dynamical system. (Not necessarily measure preserving, though we'll usually want that). Then T is said to be *ergodic* if for every set A for which $T^{-1}A = A$, we have $\mu(A) = 0$, or $\mu(X \setminus A) = 0$.

In other words, if A is an invariant subset it is either basically nothing or basically everything.

This doesn't sound like statistical mechanics. Instead, ergodicity is an indecomposability criterion. Suppose it fails.

If A is T-invariant, then $X \setminus A$ is also T-invariant, so we can break our dynamical system into two pieces. So ergodicity is saying that we can't break our system into two 'big' independent pieces.

It is really, really hard to prove that systems are ergodic. For example, the complex continued fraction system is ergodic, but the proof is so delicate, that shifting the square we are working with at all breaks the proof completely.

Proposition 5.1. *Let (X, μ, T) be a measure preserving probability space. Then the following are equivalent:*

1. *T is ergodic.*
2. *If A satisfies $\mu(T^{-1}A \Delta A) = 0$, then $\mu(A) = 0$, or $\mu(X \setminus A) = 0$. (We can weaken the ergodicity assumption to $T^{-1}A$ is almost the same as A.)*
3. *If $\mu(A) > 0$, then*
$$\mu\left(\bigcup_{n=1}^{\infty} T^{-n}A\right) = 1$$
(Basically everything ends up inside A eventually, so basically everything ends up basically everywhere.)
4. *If $\mu(A)$, $\mu(B) > 0$, then there exists n such that*
$$\mu(T^{-n}A \cap B) > 0.$$
(There is some n, such that the set of points that are in A after n steps has positive measure overlap with B.)

5. *Suppose $\mathscr{A}$ is a semi-algebra generating $\mathcal{A}$. Then for every $A, B \in \mathscr{A}$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(T^{-n}A \cap B) = \mu(A)\mu(B)$$

   (The probability that a point in B ends up in A after a given number of steps is $\mu(A)$ on average.) Alternatively: (On average, being in B vs being in A are independent.)

The first four are fairly easy to prove from each other. The fifth is much harder, and relies on the Ergodic Theorem.

5.1 The Miracle of Ergodicity

We already know

$$\text{measure preserving} \implies \text{recurrence},$$

but now with the additional assumption of indecomposability, we have

$$\text{indecomposability} \implies \text{recurrence, } \textit{and} \text{ how often we recur.}$$

6 Hierarchy of Mixing

Let (X, μ) be a probability space.

1. T is *weakly mixing* if for all $A, B \in \mathcal{A}$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{\infty} \big(\mu(T^{-n}A \cap B) - \mu(A)\mu(B)\big) = 0.$$

   Weak mixing implies ergodicity by equivalent condition 5.

2. T is *strongly mixing* if for all $A, B \in \mathcal{A}$,

$$\lim_{N\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B).$$

   Strong mixing implies weak mixing.

3. T is *Bernoulli* if (X, T, μ) is isomorphic to a full symbolic shift with Bernoulli product measure. (The noun form is Bernoullicity.)

   For cylinder sets, this means that

$$\mu(T^{-n}C_s \cap C_t) = \mu(C_s)\mu(C_t),$$

   for large enough n.

4. T is *exact* if

$$\bigcap_{i=0}^{\infty} T^{-i}\mathcal{A} = \{\varnothing, X\}.$$

   This is implies Bernoullicity.