Introduction to Ergodic theory

Lecture by Amie Wilkinson
Notes by Clark Butler

October 9, 2014


1. **The Ehrenfest Urn model**
   
   Imagine that we have two urns $A$ and $B$ and a collection of $N$ balls, which we label $1, \ldots, N$. Initially, all $N$ balls are contained in urn $A$. Think about the following procedure: We roll an $N$-sided fair die whose faces are labeled $1, \ldots, N$ (so the probability of rolling an $i$ is $\frac{1}{N}$). If we roll the number $j$, then we move the ball labeled $j$ from the urn it is currently in to the other urn. So if we roll a 6 and the ball labeled 6 is in urn $B$, we move it to urn $A$. We may repeat this procedure indefinitely.

   Question: If we begin with $N$ balls in urn $A$ and begin applying this procedure, will it ever be the case in the future that all of the balls are in urn $A$ again?

   Answer: Obviously not always. For instance, if we roll a 2 initially, and then roll only the number 1 from then on, the ball labeled 2 will never return to $A$.

2. **The Doubling map on the circle**
   
   We think of $S^1$ as the interval $[0,1]$ with the endpoints identified. Consider the map $f : S^1 \to S^1$ defined by

   \[ f(x) = 2x \pmod{1} \]

   By ‘mod 1’ we mean that we take the fractional part of the number $2x$. Iterate this map.

   Question: Suppose we start with a point $x$ such that $|x| < .0001$. Will there be infinitely many $n$ such that $|f^n(x)| < .0001$?

   Answer: Not always. Take $x = \frac{1}{2m+3}$ for a sufficiently large $m > 0$. Then $f^m(x) = \frac{1}{3}$, and $\frac{1}{3}$ is periodic under $f$ with period 2, so for $n \geq m$ we will always have $f^n(x) = \frac{1}{3}$ or $\frac{2}{3}$.

3. **Continued fractions**
   
   For $x \in (0,1)$, write

   \[ x = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}} \]

   for natural numbers $a_i \in \mathbb{N}$, with the convention that if $a_k = 0$, then $a_{k+\ell} = 0$ for every $\ell \geq 0$. This is the continued fraction expansion of $x$. It terminates if and only if $x$ is a rational number. It is eventually periodic if and only if $x$ is the root of a quadratic equation with integer coefficients. For irrational numbers, this expansion is unique, unlike the base $b$ expansions.

   Question: Does the digit $a_1$ ever occur again in the continued fraction expansion of $x$?
Hence by additivity

\[ \mu \text{ properties of a system when we iterate it many times.} \]

For \( T \) \( \text{Ergodic theory is the study of measure-preserving systems.} \)

We want to study the long term statistical

| Definition | posed in terms of a measure-preserving dynamical system. All of the ‘yes’ answers are a consequence of the Poincaré recurrence theorem. All of these problems can be \( \text{other two examples.} \)

The answer to all of the questions posed above is actually \textit{almost always}. The counterexamples we chose somehow represent a negligible proportion of the possibilities.

In the Ehrenfest Urn model, with probability 1 we will come back to the initial configuration of balls, regardless of what that initial configuration was. For the doubling map, if we pick a point uniformly at random in \((0, .0001)\), then with probability 1 we will return to that interval. The answer is similarly yes in the other two examples.

All of the ‘yes’ answers are a consequence of the Poincaré recurrence theorem. All of these problems can be posed in terms of a measure-preserving dynamical system.

**Definition:** A probability space \((X, B, \mu)\) is a set \(X\) together with a \(\sigma\)-algebra \(B\) of subsets of \(X\) and a measure \(\mu\) on \(B\) such that \(\mu(X) = 1\).

**Definition:** A \textit{measure-preserving dynamical system} is a probability space \((X, B, \mu)\) with a measurable map \(T : X \to X\) such that for every \(A \in B\), \(\mu(T^{-1}(A)) = \mu(A)\).

We may write the measure preservation condition for \(T\) as \(T_\ast \mu = \mu\).

Some examples: Take \(X = \{1, \ldots, k\}\) a finite set and \(B\) the set of all subsets of \(X\). Define a probability measure \(\mu\) by \(\mu(\{i\}) = k^{-1}\) for each \(i \in X\). Let \(T : X \to X\) be any permutation. This is a measure-preserving dynamical system.

Another example: Let \(X = S^1 = [0, 1)/(0 \sim 1)\), the unit interval with endpoints identified. Take \(B\) to be the standard Borel \(\sigma\)-algebra on \([0, 1]\) generated by open sets and \(\mu\) to be Lebesgue measure. Take \(T\) to be the doubling map \(T(x) = 2x \mod 1\). Then \(T\) preserves Lebesgue measure, even though it doubles the length of intervals. This can be seen by observing that for an interval \(I \subset (0, 1)\), the preimage \(T^{-1}(I)\) is a pair of disjoint intervals, one located in \((0, \frac{1}{2})\) and the other inside of \((\frac{1}{2}, 1)\), each of which is half the length of \(I\). Hence by additivity \(\mu(T^{-1}(I)) = \mu(I)\).

Ergodic theory is the study of measure-preserving systems. We want to study the long term statistical properties of a system when we iterate it many times. For \(T : X \to X\), we will write \(T^n\) for the \(n\)-fold composition of \(T\) with itself if \(n > 0\), and set \(T^0 = Id_X\). A fundamental type of problem is the following: Take a measurable function \(\varphi : X \to \mathbb{R}\). We think of \(\varphi\) as a random variable (this is actually the definition of a random variable in probability theory). We obtain a sequence of random variables \(\varphi_0, \varphi_1, \varphi_2, \ldots\) by setting \(\varphi_j = \varphi \circ T^j\). The measure preservation condition \(T_\ast \mu = \mu\) implies these random variables are identically distributed with distribution given by the measure \(\varphi_\ast \mu\) on \(\mathbb{R}\). However, these random variables will generally be very far from being independent of one another.
We can nevertheless ask questions about how closely the sequence \((\varphi_j)_{j \in \mathbb{N}}\) resembles a sequence of independent identically distributed random variables. Do they obey a law of large numbers, a central limit theorem, are there large deviation estimates, etc.

**The Poincaré Recurrence theorem:** Let \(T : (X, \mathcal{B}, \mu) \circlearrowright\) be a measure preserving system. Then for any measurable set \(A\) such that \(\mu(A) > 0\),

1. For every \(N > 0\), there is some \(n \geq N\) such that \(\mu(T^{-n}A \cap A) > 0\).
2. For almost every \(x \in A\), there is some \(n > 0\) such that \(T^n(x) \in A\).

**Proof.** We prove (1) and leave (2) as an exercise. The proof consists only of applying our intuition from the pigeonhole principle to measure spaces. For instance, in the example of the permutation given above, the proof literally is the pigeonhole principle; for every \(n\), \(|T^{-n}A| = |A|\), and a finite space only has a finite number of subsets of a given cardinality, so the sets \(T^{-n}A\) cannot all be disjoint.

In general, suppose to a contradiction that there is some \(N\) such that \(\mu(T^{-n}A \cap A) = 0\) for every \(n \geq N\). Fix some \(n \geq N\). Since \(T\) is measure preserving, \(T^n\) is as well, so that for every \(k > \ell \geq 0\),

\[
\mu(T^{-kn}A \cap T^{-\ell n}A) = \mu(T^{-(k-\ell)n}A \cap A) = 0
\]

since \((k-\ell)n \geq n \geq N\). Hence, up to null sets, the sets \(T^{-kn}(A)\) are pairwise disjoint. Thus

\[
\mu \left( \bigcup_{n=0}^{\infty} T^{-kn}A \right) = \sum_{n=0}^{\infty} \mu(T^{-kn}(A)) = \sum_{n=0}^{\infty} \mu(A) = \infty
\]

But we assumed that \(\mu\) was a probability measure, so that \(\mu \left( \bigcup_{n=0}^{\infty} T^{-kn}A \right) \leq 1\), a contradiction. \(\square\)

It is very easy to find counterexamples to the Poincaré Recurrence theorem for dynamical systems with infinite measures. Take for instance \(X = \mathbb{R}\), \(\mu\) to be Lebesgue measure, and \(T(x) = x + 1\).

The proof of Poincaré Recurrence is almost trivial yet yields remarkable conclusions. Much of the work in ergodic theory consists of finding the appropriate nontrivial invariant measure for a system in order to apply results like this.

We can apply the theorem to our examples. First consider example 2, the doubling map on the circle \(T(x) = 2x \pmod{1}\). We showed that \(T_*\mu = \mu\), where \(\mu\) is Lebesgue measure. In example 2, we were interested in recurrence on the set \(A = (0, .0001)\). \(\mu(A) = .0001 > 0\), so the Poincaré recurrence theorem applies and we conclude that for Lebesgue-a.e. \(x \in (0, .0001)\) there are infinitely many \(n\) such that \(T^n(x) \in A\).

Now consider example 3, the continued fraction example. Here \(X = (0, 1]\), and the map we consider is the **Gauss map** \(T : X \to X\),

\[
T(x) = \begin{cases} 
\frac{1}{x} & x \in \mathbb{R} \\
\frac{1}{x} \mod 1 & x \in (0, 1]
\end{cases}
\]

Index the interval \(I_k := \left( \frac{1}{k+1}, \frac{1}{k} \right]\) by the positive integer \(k\). As an exercise, show that if

\[
x = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}
\]

then \(x \in I_{a_1}\), \(T(x) \in I_{a_2}\), and in general \(T^n(x) \in I_{a_{n+1}}\). Thus we can think of \(T\) as generating the continued fraction expansion of \(x\). We would like to apply Poincaré recurrence to conclude that if \(x \in I_{a_1}\), then there are infinitely many \(n\) such that \(T^n(x) \in I_{a_1}\), i.e., there are infinitely many \(n\) such that \(a_{n+1} = a_1\). But we
could only deduce this if we knew that \( T \) preserved a measure \( \mu \) which assigned positive measure to \( I_{a_1} \). There is such a measure, but it is not easy to guess. It is the Gauss measure defined by

\[
\mu(A) = \frac{1}{\log 2} \int_A \frac{dx}{1 + x}
\]

for any Lebesgue-measurable set \( A \).

**Exercise:** Prove that \( \mu \) is invariant with respect to \( T \). It suffices to prove that \( \mu(I) = \mu(T^{-1}(I)) \) for any open interval \( I \subset X \).

Since \( \mu \) is defined by integrating the positive density \( \frac{1}{\log 2} \frac{1}{1 + x} \) with respect to Lebesgue measure, \( \mu \) has the same null sets as Lebesgue measure. Since \( I_{a_1} \) has positive length, it has positive \( \mu \)-measure, and we conclude that for Lebesgue-a.e. point \( x \in I_{a_1} \), the digit \( a_1 \) will recur in the continued fraction expansion of \( x \).

Lastly, we analyze example 1, the Ehrenfest urn model. Here we need to construct an appropriate probability space to model the process. Let \( X = \{1, \ldots, N\}^\mathbb{N} \times (\mathbb{Z}/2\mathbb{Z})^\mathbb{N} \). We will write points of \( X \) as \((\omega, v)\) where \( \omega = (\omega_n)_{n \in \mathbb{N}} \) is a sequence of numbers with \( \omega_n \in \{1, \ldots, N\} \), and \( v \in (\mathbb{Z}/2\mathbb{Z})^\mathbb{N} \) is a vector \( v = (v_1, \ldots, v_n) \) with \( v_i \in \{0, 1\} \). Take 0 to correspond to urn \( A \) and 1 to correspond to urn \( B \). \( v \) represents a configuration of the balls divided between the two urns; ball \( j \) is in urn \( A \) if and only if \( v_j = 0 \), and likewise \( j \) is in \( B \) if and only if \( v_j = 1 \). \( \omega \) represents our sequence of die rolls: at step \( n \), we roll the number \( k \) if \( \omega_n = k \), and then transfer ball \( k \) from its urn to the other one.

We think of a point \((\omega, v) \in X\) as specifying a possible infinite sequence of outcomes for our procedure. \( v \) represents the initial configuration of the balls, and \( \omega \) tells us the sequence of die rolls we will obtain, and hence how we should modify \( v \) at each step. Set \( v_0 = v \), and inductively define \( v_n = v_{n-1} + e_{\omega_j} \), where \( e_k \in (\mathbb{Z}/2\mathbb{Z})^\mathbb{N} \) is the vector which is zero everywhere except in the \( k \)th place, and addition is taken mod 2. \( v_n \) is the configuration of the balls after \( n \) steps. We can thus rephrase our recurrence question as inquiring whether there is some future \( n \) such that \( v_n = v_0 \).

Now we introduce a dynamical system. Let \( \sigma : \{1, \ldots, N\}^\mathbb{N} \to \{1, \ldots, N\}^\mathbb{N} \) be the left shift, which takes a sequence and drops the first term. More formally, \( \sigma(\omega) \) is defined by \( [\sigma(\omega)]_n = \omega_{n+1} \). Then define \( T : X \to X \) by

\[
T(\omega, v) = (\sigma(\omega), v + e_{\omega_1})
\]

Check that \( T^n(\omega, v) = (\sigma^n(\omega), v_n) \).

**Exercise:** Find an interesting invariant measure for \( T \) on \( X \), and prove that with probability 1 with respect to this measure, the initial configuration \( v_0 \) recurs.