Decay of Correlations
Probabilistic Methods in Dynamics

M. Smith
Department of Mathematics
University of Utah

22 March 2016 / GSAC
Outline

1. Introduction
   - Definitions and Examples
   - Tracking Orbits

2. Probability in Dynamics
   - Comparisons
   - Central Limit Theorem
   - Spectral Gap Method

3. Going Further
Outline

1. Introduction
   - Definitions and Examples
   - Tracking Orbits

2. Probability in Dynamics
   - Comparisons
   - Central Limit Theorem
   - Spectral Gap Method

3. Going Further
What is Dynamics?

A **dynamical system** is a deterministic rule for "time evolution".

\[ t_0 \rightarrow t_1 \]
Suppose we have a map $T : X \rightarrow X$.

- Applying $T$ to $X$ moves time forward by one step.
- Notation:
  \[ T^n = T \circ T \circ \cdots \circ T \ (n \text{ times}). \]
- For $x \in X$, $T^n(x)$ is the position of $x$ at time $n$. 
Other Types of Time Evolution

- If $T$ is invertible, $T^{-n} = (T^{-1})^n$, $n \geq 0$.
- $\mathbb{Z}$ acts on $X$ by $n \mapsto T^n$.
- A group $G$ acting on $X$ gives "time evolution".
- $\mathbb{R}$-actions arise from differential equations.
- $\mathbb{Z}$-actions are similar to $\mathbb{R}$-actions.
Examples

Discrete time:
- $X = \text{your favorite space}, \ T = \text{id}_X$.
- $X = \text{a circle}, \ T = \text{rotation by some angle}$.
- $X = \text{a compact manifold}, \ T = \text{a diffeomorphism}$.

Continuous time:
- Flows on manifolds.
- Frictionless motion with constraints.
Main Example: Doubling

- $X = S^1 = \mathbb{R}/\mathbb{Z} = [0, 1)$.
- Define the **doubling map**:
  \[ D(x) = 2x \pmod{1} \]
- Isomorphic to $z \mapsto z^2$ for the unit circle in $\mathbb{C}$.
- $D^n(x) = 2^n x \pmod{1}$.
- Exponential "complexity".
- This example is an extreme case!
Doubling Graph

\[ D(x) \]

\[ \begin{align*}
0 & \quad 0.5 & \quad 1 \\
1 & & \\
\end{align*} \]
Repeated Doubling

\[ D^2(x) \]

- The graph shows the plot of \( D^2(x) \) against \( x \) with values ranging from 0 to 1 on both axes.
- The graph illustrates the concept of repeated doubling in probability or dynamics.
Repeated Doubling

\[ D^3(x) \]

\[ 0 \quad 1 \]

\[ 0 \quad 1 \]

M. Smith

Decay of Correlations
Outline

1. Introduction
   - Definitions and Examples
   - Tracking Orbits

2. Probability in Dynamics
   - Comparisons
   - Central Limit Theorem
   - Spectral Gap Method

3. Going Further
Suppose $x \in X$, and consider its **forward orbit**:

$$\mathcal{O}^+(x) = \{ T^n(x) \}_{n=0}^\infty$$

- **Question**: What does the orbit of $x$ look like?
- What about orbits of subsets of $X$?
- This is very difficult in general.
Orbits Under Doubling

- Any $\frac{j}{2^k}$ is sent to 0 by $D^k$.
- Similarly $\frac{j}{3^k}$ has a periodic orbit.
- Both kinds are dense, but only countable.
- **Refined Question:** What is the orbit of a "generic point" $x$?
- Generic means we should have a notion of "size".
Measurable Dynamics

Let $\mu$ be a probability measure on $X$ so that $T$ is $\mu$-measurable.

- $T$ is $\mu$-measure preserving if:

$$\mu(T^{-1}A) = \mu(A)$$

for all $\mu$-measurable sets $A$.

- Invariant measures are the easiest to understand.
Doubling Invariant Measures

\[ D \text{ preserves } \lambda, \text{ Lebesgue measure on } S^1. \]

Warning: \( D \) has many preserved measures!
Let \((X, \mu, T)\) be a probability measure preserving system.

- If \(T^{-1}A = A\) implies \(\mu(A) = 0\) or \(\mu(A) = 1\), \(T\) is **ergodic** with respect to \(\mu\).
- Ergodicity means indecomposibility.
- \(D\) is \(\lambda\)-ergodic (by general representation theory).
Birkhoff Ergodic Theorem

Let \((X, \mu, T)\) be an ergodic probability measure preserving system.

- If \(F \in L^1(\mu)\), then for \(\mu\)-almost every \(x \in X\):
  \[
  \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} F(T^n x) = \int_X F \, d\mu
  \]

- Interpretation: time average = space average.
Equidistribution

- Take \( F = \chi_A \) for \( A \subseteq X \).
- \( \chi_A(T^n x) = 1 \) if \( T^n x \) is in \( A \), otherwise 0.
- The ergodic theorem tells us, for \( \mu \)-almost every \( x \):

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \chi_A(T^n x) = \mu(A).
\]

- The \( \mu \)-typical point "spends \( \mu(A) \) of its time in \( A \)" (after a long time).
The doubling map is "chaotic".

- Almost every point equidistributes (for Lebesgue measure).
- Periodic points are dense.
- Points which are eventually fixed at 0 are dense.
- Very sensitive to initial conditions!
Outline

1. Introduction
   - Definitions and Examples
   - Tracking Orbits

2. Probability in Dynamics
   - Comparisons
   - Central Limit Theorem
   - Spectral Gap Method

3. Going Further
Let \((X, \mu, T)\) be a probability measure preserving system.

- For \(F \in L^1(\mu)\) consider the sequence \(F_n = F \circ T^n\).
- \(T\) is measure preserving, so \(F_n\) are identically distributed.
- Unlikely to be independent, \(T\) is deterministic.
- **Question**: Can we approximate \(F_n\) by IID?
Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space.

Let \(\{X_n\}_{n=0}^{\infty}\) be a sequence of IID real random variables such that \(\mathbb{E}(|X_0|)\) exists and is finite.

Then almost surely:

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} X_n = \mathbb{E}(X_0).
\]

This is a special case of the Birkhoff ergodic theorem!
Ergodic systems have SLLN.

**Question:** can we get any other theorems from probability?

More theorems implies a system is "essentially random" (over a long time).

Dynamical systems are deterministic, so this is hard!
Let \((X, \mu, T)\) be a probability measure preserving system.

- \(T\) is \(\mu\)-mixing if for any measurable \(A, B\) we have:
  \[
  \lim_{n \to \infty} \mu(A \cap T^{-n}(B)) = \mu(A)\mu(B).
  \]

- After a long time, \(B\) is essentially independent of \(A\).
- Diffusion-like behavior.
- Mixing implies ergodicity.
- "Most" systems are not mixing.
Doubling is Mixing

A

B

Decay of Correlations
$T$ is $\mu$-mixing if and only if for any $F, G \in L^2(\mu)$ we have:

$$\lim_{n \to \infty} \left[ \int_X F \cdot G \circ T^n d\mu - \int_X F d\mu \int_X G d\mu \right] = 0.$$ 

Consider the correlation between $F$ and $G \circ T^n$ (up to scaling).

Mixing is equivalent to observables being uncorrelated (after a long time).
Outline

1. Introduction
   - Definitions and Examples
   - Tracking Orbits

2. Probability in Dynamics
   - Comparisons
   - Central Limit Theorem
   - Spectral Gap Method

3. Going Further
Let \( \{X_m\}_{m=0}^{\infty} \) be a sequence of IID real random variables.

- Suppose \( \mathbb{E}[X_0] = 0 \) and \( 0 < \text{Var}[X_0] = \sigma^2 < \infty \).
- Then the sequence:

\[
Y_M = \frac{1}{\sqrt{M\sigma^2}} \sum_{m=0}^{M-1} X_m
\]

converges in distribution to \( N(0, 1) \).
Let \((X, \mu, T)\) be a probability measure preserving system.

- A function \(F \in L^2_0(\mu)\) has the CLT under \(T\) if:

\[
\lim_{M \to \infty} \mu \left( x : \frac{1}{\sqrt{M \sigma^2}} \sum_{m=0}^{M-1} F(T^m x) \leq z \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt
\]

for some \(0 < \sigma^2 < \infty\).

- Here \(L^2_0(\mu)\) means integral zero.

- Tells us certain orbits occur with positive measure.
Proving CLT

- $F \circ T^m$ is not an IID sequence in general.
- Probability: Stationary processes with fast mixing of sets still have CLT.
- $F \circ T^m$ is a stationary process if $T$ preserves $\mu$.
- If $D$ has fast mixing of sets, then $D$ will have CLT!
Suppose $\gamma_n \to 0$ is a sequence of positive numbers. Then there exist $A, B \subseteq [0, 1)$ so that:

$$|\lambda(A \cap D^{-n}(B)) - \lambda(A)\lambda(B)| \geq \gamma_n$$

That is, $D$ has sets with arbitrarily slow mixing.
Outline

1. Introduction
   - Definitions and Examples
   - Tracking Orbits

2. Probability in Dynamics
   - Comparisons
   - Central Limit Theorem
   - Spectral Gap Method

3. Going Further
Changing the Problem

- We could not prove CLT for $D$ and $L^2_0(S^1)$.
- $L^2$ functions can be arbitrarily complicated.
- **Idea:** "nicer" functions could still have CLT.
- More difficult to prove.
Define the Ruelle transfer operator $P_T$ by:

$$\int_X F \cdot G \circ T \, d\mu = \int_X P_T F \cdot G \, d\mu$$

- $P_T$ is well defined on various $L^p$ spaces.
- For $D$, we have:

$$P_D F(x) = \frac{1}{2} \left[ F \left( \frac{x}{2} \right) + F \left( \frac{x}{2} + \frac{1}{2} \right) \right]$$
Step 2: Banach Spaces

$P_T$ acts nicely on $L^2(\mu)$.

- If $F$ is constant, then $P_T F = F$.
- If $F$ has integral 0, then $P_T F$ also has integral zero.
- $L^2 = \mathbb{R} \oplus L^2_0$, each piece is $P_T$-invariant.

**Question:** what does $P_T$ do to $L^2_0$?
Step 3: Spectral Gap

Suppose that $B \subseteq L^\infty_0$ is a Banach space such that:

- $\|P_T F\|_B \leq r \|F\|_B$ with $0 < r < 1$.
- $\|F\|_1 \leq C \|F\|_B$ for some $C > 0$.

Then we have:

$$\left| \int_X F \cdot G \circ T^n \, d\mu \right| \leq r^n C \|F\|_B \|G\|_\infty$$

so functions in $B$ have exponential decay of correlations.
Step 3: Spectral Gap

- Contraction by $r$ bounds the spectrum of $P_T$ away from 1.
- A "gap" in the spectrum of $P_T$ makes it a contraction on $B$. 

\[ r < 1 \]
Step 4: CLT, at last

If $P_T$ has a spectral gap, then we are done.

- Suppose every $F \in B$ has exponential decay of correlations
- Each $F \in B$ then has some appropriate $0 < \sigma^2 < \infty$.
- The sequence of functions $F \circ T^m$ satisfy CLT with variance $\sigma^2$.
- The proof is "similar" to the proof of the usual CLT.
For $D$, take $B = \text{Lip}_0 = \text{Lipschitz functions of integral zero.}$

- $\|F\|_B$ is the Lipschitz constant for $F$.
- $P_D$ is a contraction by $\frac{1}{2}$ on $B$:

\[
\frac{1}{2} \left| F \left( \frac{x}{2} + c \right) - F \left( \frac{y}{2} + c \right) \right| \leq \frac{\|F\|_B}{4} |x - y|
\]

\[
|P_D F(x) - P_D F(y)| \leq 2 \frac{\|F\|_B}{4} |x - y| = \frac{\|F\|_B}{2} |x - y|
\]

- So $D$ has CLT for Lipschitz functions!
Outline

1. Introduction
   - Definitions and Examples
   - Tracking Orbits

2. Probability in Dynamics
   - Comparisons
   - Central Limit Theorem
   - Spectral Gap Method

3. Going Further
The doubling map belongs to this class of maps on $S^1$. 

![Diagram](image)
There are three ways we can go further.

- Piecewise expanding maps have a "nice" invariant measure. If this measure is mixing, we get CLT for functions in BV (bounded variation).
- Stronger results like the Almost Sure Invariance Principle are possible.
- We can try to get CLT in applied systems.
Theorems from probability are still true for deterministic systems, especially chaotic ones.

Functional analysis gives a lot of information in dynamics.

For continuous or smooth systems, invariant measures interact with topological or smooth structures in nice ways.
References and Further Reading

General ergodic theory


CLT for the doubling map

- <https://vaughnclimenhaga.wordpress.com/2013/01/30/spectral-methods-in-dynamics/>
- Split into separate posts, there are links at the bottom.

Slowly mixing sets

- <https://vaughnclimenhaga.wordpress.com/2014/04/22/slowly-mixing-sets/>
CLT for piecewise expanding maps