Furstenberg's Proof of Szemerédi's Theorem

Dylan Wallace

April 25, 2025

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Colorings and Van der Waerden's theorem

$$\mathbb{N} = 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 11,\ 12,\ 13,\ 14,\ \ldots$$

Colorings and Van der Waerden's theorem

Coloring

A *coloring* of $\mathbb{N}$ is a pattern of assigning a color to each number $n \in \mathbb{N}$.
If only a finite number of colors is used in a coloring, it is called a *finite coloring*.

$$\mathbb{N} = 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 11,\ 12,\ 13,\ 14,\ \ldots$$

Colorings and Van der Waerden's theorem

Coloring

A *coloring* of $\mathbb{N}$ is a pattern of assigning a color to each number $n \in \mathbb{N}$.
If only a finite number of colors is used in a coloring, it is called a *finite coloring*.

$$\mathbb{N} = 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 11,\ 12,\ 13,\ 14,\ \ldots$$

Van der Waerden's Theorem (1927): For any finite coloring of the natural numbers and any integer k, there exists a k-term arithmetic progression in one color

Colorings and Van der Waerden's theorem

Coloring

A *coloring* of $\mathbb{N}$ is a pattern of assigning a color to each number $n \in \mathbb{N}$.
If only a finite number of colors is used in a coloring, it is called a *finite coloring*.

$$\mathbb{N} = 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 11,\ 12,\ 13,\ 14,\ \ldots$$

Van der Waerden's Theorem (1927): For any finite coloring of the natural numbers and any integer k, there exists a k-term arithmetic progression in one color

1-term arithmetic progression : 1

2-term a.p. : $2, 6$

3-term a.p. : $1, 4, 7$

4-term a.p. : $5, 8, 11, 14$

$\ldots$

and so on

Erdos and Turán (1936)

- For a subset $S \subseteq \{1, \ldots, n\}$, we can consider S as having a *density* of $d_n(S) = \frac{\#(S)}{n}$ = fraction of elements in $\{1, \ldots, n\}$ that are in S.

Erdos and Turán (1936)

- For a subset $S \subseteq \{1, \ldots, n\}$, we can consider S as having a *density* of $d_n(S) = \frac{\#(S)}{n}$ = fraction of elements in $\{1, \ldots, n\}$ that are in S.
- ...but how do we assign a density to an infinite set $S \subseteq \mathbb{N}$?

Erdos and Turán (1936)

- For a subset $S \subseteq \{1, \ldots, n\}$, we can consider S as having a *density* of $d_n(S) = \frac{\#(S)}{n}$ = fraction of elements in $\{1, \ldots, n\}$ that are in S.
- ...but how do we assign a density to an infinite set $S \subseteq \mathbb{N}$?

Upper density

For a set $S \subseteq \mathbb{N}$, define the *upper density* of S as
$\overline{d}(S) = \limsup_{n \to \infty} d_n(S \cap \{1, \ldots, n\})$

- $P = \{1, 4, 9, 16, \ldots\}$ = {Set of all perfect squares} has $\overline{d}(P) = 0$, because the chance of seeing a perfect square $\to 0$ as we look at bigger numbers
- $E = \{2, 4, 6, 8, \ldots\}$ = {Set of all even numbers} has $\overline{d}(E) = \frac{1}{2}$, because every other number is an even number, even for large numbers

Erdos and Turán (1936)

FACT: If we have a finite coloring of $\mathbb{N}$, then at least one color has an upper density > 0.

Erdos and Turán (1936)

FACT: If we have a finite coloring of $\mathbb{N}$, then at least one color has an upper density > 0.

$\implies$ Some people (Erdos and Turán) thought that Van der Waerden's theorem was true because **some color with positive upper density has all the k-term arithmetic progressions**

Erdos and Turán (1936)

FACT: If we have a finite coloring of $\mathbb{N}$, then at least one color has an upper density > 0.

$\implies$ Some people (Erdos and Turán) thought that Van der Waerden's theorem was true because **some color with positive upper density has all the k-term arithmetic progressions**

Erdos-Turán Conjecture (1936)

Any subset $S \subseteq \mathbb{N}$ with $\overline{d}(S) > 0$ has a k-term arithmetic progression for all k.

Erdos and Turán (1936)

FACT: If we have a finite coloring of $\mathbb{N}$, then at least one color has an upper density > 0.

$\implies$ Some people (Erdos and Turán) thought that Van der Waerden's theorem was true because **some color with positive upper density has all the k-term arithmetic progressions**

~~Erdos-Turán Conjecture (1936)~~ Szemerédi's Theorem (1975)!

Any subset $S \subseteq \mathbb{N}$ with $\overline{d}(S) > 0$ has a k-term arithmetic progression for all k.

History of Szemerédi's Theorem

1. 1936: Conjectured by Erdos and Turán

History of Szemerédi's Theorem

1. 1936: Conjectured by Erdos and Turán
2. 1975: Proven by Szemerédi (combinatorics)

History of Szemerédi's Theorem

1. 1936: Conjectured by Erdos and Turán
2. 1975: Proven by Szemerédi (combinatorics)
3. 1977: Proven by Furstenberg (ergodic theory)

History of Szemerédi's Theorem

1. 1936: Conjectured by Erdos and Turán
2. 1975: Proven by Szemerédi (combinatorics)
3. 1977: Proven by Furstenberg (ergodic theory)
4. 2001: Proven by Gowers (fourier analysis)

History of Szemerédi's Theorem

1. 1936: Conjectured by Erdos and Turán
2. 1975: Proven by Szemerédi (combinatorics)
3. 1977: Proven by Furstenberg (ergodic theory)
4. 2001: Proven by Gowers (fourier analysis)
5. 2006: Proven by Nagle et al. (hypergraph removal)

History of Szemerédi's Theorem

1. 1936: Conjectured by Erdos and Turán
2. 1975: Proven by Szemerédi (combinatorics)
3. 1977: **Proven by Furstenberg (ergodic theory)** ← Today's Talk
4. 2001: Proven by Gowers (fourier analysis)
5. 2006: Proven by Nagle et al. (hypergraph removal)

What is Ergodic Theory?

What is Ergodic Theory?

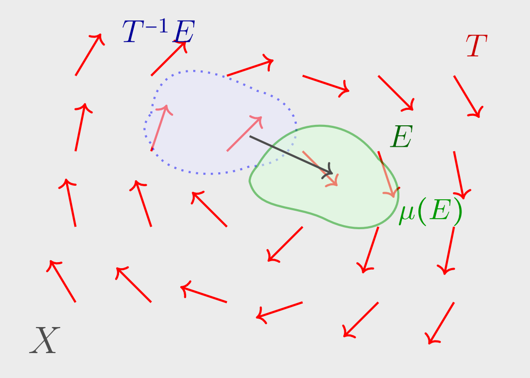
Q: When do systems exhibit mixing?

Figure: Coffee Mixing. Credit: *Shtetl-Optimized (Scott Aaronson)*

What is Ergodic Theory?

Measure-Preserving System

Given a space X, a transformation $T : X \to X$ and a measure μ on X, a triple (X, T, μ) is called a *measure-preserving system* (MP-system) if $\mu(E) = \mu(T^{-1}E)$ for any* subset $E \subseteq X$.



We can think of a measure-preserving system as a dynamical system where the "size" of sets are preserved

Furstenberg's Proof

1. Find an appropriate measure-preserving system to model our problem
2. Prove that every measure-preserving system has a particular property (called "SZ")
3. Use the fact that our problem's measure-preserving system being SZ to prove Szemerédi's Theorem

- $\{$Sets of subsets of $\mathbb{N}\} \cong \{$Set of binary sequences$\} = 2^{\mathbb{N}}$

$$S = \{1, 3, 4, 6, \dots\}$$
$$\Leftrightarrow$$
$$\{s_i\} = 1, 0, 1, 1, 0, 1, \dots$$

Constructing our Measure-Preserving System

We want a condition for a binary sequence $\{s_i\}$ having an arithmetic progression

Sets	Sequences
$S \subseteq \mathbb{N}$ has a k-term a.p. if there are $n, m \in \mathbb{N}$ with $n, n+m, n+2m, \ldots, n+(k-1)m \in S$	$\{s_i\} \in 2^{\mathbb{N}}$ has a k-term a.p. if there are $n, m \in \mathbb{N}$ with $s_n = s_{n+m} = \ldots s_{n+(k-1)m} = 1$

Constructing our Measure-Preserving System

We want a condition for a binary sequence $\{s_i\}$ having an arithmetic progression

Sets	Sequences
$S \subseteq \mathbb{N}$ has a k-term a.p. if there are $n, m \in \mathbb{N}$ with $n, n+m, n+2m, \dots, n+(k-1)m \in S$	$\{s_i\} \in 2^{\mathbb{N}}$ has a k-term a.p. if there are $n, m \in \mathbb{N}$ with $s_n = s_{n+m} = \dots s_{n+(k-1)m} = 1$

WANT: Some $T : 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ and a way to express RHS in terms of T and set inclusion

Constructing our Measure-Preserving System

Shift Operator

Define the *shift operator* $T : 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ by $T(\{s_1, s_2, s_3, \dots\}) = \{s_2, s_3, \dots\}$.

- i.e. the shift operator "shifts" the sequence to the left by one
- $T^n(\{s_1, s_2, \dots,\}) = \{s_{n+1}, s_{n+2}, \dots\}$

Constructing our Measure-Preserving System

Shift Operator

Define the *shift operator* $T : 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ by $T(\{s_1, s_2, s_3, \dots\}) = \{s_2, s_3, \dots\}$.

- i.e. the shift operator "shifts" the sequence to the left by one
- $T^n(\{s_1, s_2, \dots, \}) = \{s_{n+1}, s_{n+2}, \dots\}$
- Define $[1] = \{$Binary sequences $\{s_i\}$ with $s_1 = 1\} \Leftrightarrow S \subseteq \mathbb{N}$ with $1 \in S$.
- Then, $T^{-n}[1] = \{$Binary sequences with $s_{n+1} = 1\}$

Constructing our Measure-Preserving System

- Begin for a moment that $\{s_i\} \in [1]$ and our k-term arithmetic progression starts with 1.
- Then $\{s_i\}$ has a k-term arithmetic progression if there exists $m \in \mathbb{N}$ with $\{s_i\} \in [1], \{s_i\} \in T^{-m}[1], \ldots, \{s_i\} \in T^{-(k-1)m}[1]$

Constructing our Measure-Preserving System

- Begin for a moment that $\{s_i\} \in [1]$ and our k-term arithmetic progression starts with 1.
- Then $\{s_i\}$ has a k-term arithmetic progression if there exists $m \in \mathbb{N}$ with $\{s_i\} \in [1], \{s_i\} \in T^{-m}[1], \ldots, \{s_i\} \in T^{-(k-1)m}[1]$
  - $\implies \{s_i\} \in [1] \cap T^{-m}[1] \cap T^{-2m}[1] \cap \cdots \cap T^{-(k-1)m}[1]$

Constructing our Measure-Preserving System

- Begin for a moment that $\{s_i\} \in [1]$ and our k-term arithmetic progression starts with 1.
- Then $\{s_i\}$ has a k-term arithmetic progression if there exists $m \in \mathbb{N}$ with $\{s_i\} \in [1], \{s_i\} \in T^{-m}[1], \dots, \{s_i\} \in T^{-(k-1)m}[1]$
  - $\implies \{s_i\} \in [1] \cap T^{-m}[1] \cap T^{-2m}[1] \cap \cdots \cap T^{-(k-1)m}[1]$
  - $\implies \{s_i\} \in \bigcap_{i=0}^{k-1} T^{-im}[1] = I_m$

Constructing our Measure-Preserving System

- Begin for a moment that $\{s_i\} \in [1]$ and our k-term arithmetic progression starts with 1.
- Then $\{s_i\}$ has a k-term arithmetic progression if there exists $m \in \mathbb{N}$ with $\{s_i\} \in [1], \{s_i\} \in T^{-m}[1], \dots, \{s_i\} \in T^{-(k-1)m}[1]$
  - $\implies \{s_i\} \in [1] \cap T^{-m}[1] \cap T^{-2m}[1] \cap \cdots \cap T^{-(k-1)m}[1]$
  - $\implies \{s_i\} \in \bigcap_{i=0}^{k-1} T^{-im}[1] = I_m$
- Then if our arithmetic progression doesn't start with 1, then we can just "shift" it to start at 1
  - $\implies$ exists $n \geq 0$ (our "shift" amount) and $m \in \mathbb{N}$ with $T^n\{s_i\} \in I_m$

Our Desired Property

The SZ property

A measure-preserving system (X, T, μ) is *SZ* if for any $E \subseteq X$ with $\mu(E) > 0$, the following holds true:

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu\left(\bigcap_{i=0}^{k-1} T^{-in} E \right) > 0$$

Our Desired Property

The SZ property

A measure-preserving system (X, T, μ) is *SZ* if for any $E \subseteq X$ with $\mu(E) > 0$, the following holds true:

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu\left(\bigcap_{i=0}^{k-1} T^{-in} E\right) > 0$$

- SZ $\implies$ for some $m \in \mathbb{N}$, we have $\mu\left(\bigcap_{i=0}^{k-1} T^{-im} E\right) > 0$
- so if we choose a measure so that $\mu([1]) > 0$, then we have $\mu\left(\bigcap_{i=0}^{k-1} T^{-im}[1]\right) = \mu(I_m) > 0$

Our Desired Property

The SZ property

A measure-preserving system (X, T, μ) is *SZ* if for any $E \subseteq X$ with $\mu(E) > 0$, the following holds true:

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu\left(\bigcap_{i=0}^{k-1} T^{-in} E\right) > 0$$

- SZ $\implies$ for some $m \in \mathbb{N}$, we have $\mu\left(\bigcap_{i=0}^{k-1} T^{-im} E\right) > 0$
- so if we choose a measure so that $\mu([1]) > 0$, then we have $\mu\left(\bigcap_{i=0}^{k-1} T^{-im}[1]\right) = \mu(I_m) > 0$
- **It turns out all measure-preserving systems are SZ!** (Furstenberg's proof is 50 pages)

Recap

We have:

- An *almost*-MP system $(2^{\mathbb{N}}, T)$, missing a measure
- A condition for a sequence having a k-term arithmetic progression, namely $\exists n \geq 0, m \in \mathbb{N}$ with $T^n\{s_i\} \in I_m$
- A guarantee that if we can find a measure μ on $2^{\mathbb{N}}$, then $(2^{\mathbb{N}}, T, \mu)$ is SZ

Recap

We have:

- An *almost*-MP system $(2^{\mathbb{N}}, T)$, missing a measure
- A condition for a sequence having a k-term arithmetic progression, namely $\exists n \geq 0, m \in \mathbb{N}$ with $T^n\{s_i\} \in I_m$
- A guarantee that if we can find a measure μ on $2^{\mathbb{N}}$, then $(2^{\mathbb{N}}, T, \mu)$ is SZ

Next step: Find a specific measure so that $\mu([1]) > 0$

Recap

We have:

- An *almost*-MP system $(2^{\mathbb{N}}, T)$, missing a measure
- A condition for a sequence having a k-term arithmetic progression, namely $\exists n \geq 0, m \in \mathbb{N}$ with $T^n\{s_i\} \in I_m$
- A guarantee that if we can find a measure μ on $2^{\mathbb{N}}$, then $(2^{\mathbb{N}}, T, \mu)$ is SZ

Next step: Find a specific measure so that $\mu([1]) > 0$

- Many measures that satisfy this

Recap

We have:

- An *almost*-MP system $(2^{\mathbb{N}}, T)$, missing a measure
- A condition for a sequence having a k-term arithmetic progression, namely $\exists n \geq 0, m \in \mathbb{N}$ with $T^n\{s_i\} \in I_m$
- A guarantee that if we can find a measure μ on $2^{\mathbb{N}}$, then $(2^{\mathbb{N}}, T, \mu)$ is SZ

Next step: Find a specific measure so that $\mu([1]) > 0$

- Many measures that satisfy this
- Specifically, we want to find a measure so that we can infer $\mu(I_m) > 0 \implies T^n\{s_i\} \in I_m$

Construction of our measure

For the sake of clarity, we'll write $s = \{s_i\}$ from here.

- Define $\mu_N(E) = \frac{1}{N} \cdot \#(E \cap \{s, Ts, T^2 s, \dots, T^{N-1}s\})$:= the fraction of $\{s, Ts, \dots, T^{N-1}s\}$ that E contains.

Construction of our measure

For the sake of clarity, we'll write $s = \{s_i\}$ from here.

- Define $\mu_N(E) = \frac{1}{N} \cdot \#(E \cap \{s, Ts, T^2 s, \ldots, T^{N-1} s\})$:= the fraction of $\{s, Ts, \ldots, T^{N-1} s\}$ that E contains.
- Then let $\mu = \lim_{N \to \infty} \mu_N$ (the "weak limit")

Construction of our measure

For the sake of clarity, we'll write $s = \{s_i\}$ from here.

- Define $\mu_N(E) = \frac{1}{N} \cdot \#(E \cap \{s, Ts, T^2 s, \ldots, T^{N-1}s\})$:= the fraction of $\{s, Ts, \ldots, T^{N-1}s\}$ that E contains.
- Then let $\mu = \lim_{N \to \infty} \mu_N$ (the "weak limit")

Then $\mu(E) > 0$ only if E has a high density in $\mathrm{Orb}(s) = \{T^n s \mid n \in \mathbb{N}\}$:= "shifted" elements of s, so μ is shift invariant and $\mu([1]) > 0$.

Construction of our measure

For the sake of clarity, we'll write $s = \{s_i\}$ from here.

- Define $\mu_N(E) = \frac{1}{N} \cdot \#(E \cap \{s, Ts, T^2 s, \ldots, T^{N-1} s\})$:= the fraction of $\{s, Ts, \ldots, T^{N-1} s\}$ that E contains.
- Then let $\mu = \lim_{N \to \infty} \mu_N$ (the "weak limit")

Then $\mu(E) > 0$ only if E has a high density in $\mathrm{Orb}(s) = \{T^n s \,|\, n \in \mathbb{N}\}$:= "shifted" elements of s, so μ is shift invariant and $\mu([1]) > 0$.

$\Longrightarrow$ So $(2^{\mathbb{N}}, T, \mu)$ is an MP-system, so $(2^{\mathbb{N}}, T, \mu)$ is SZ, so $\mu\,(I_m) > 0$ for some $m \in \mathbb{N}$.

Almost at the end...

- $\mu(E) > 0$, so (by construction of measure) there exists some $s' = \lim_{j \to \infty} T^{n_j} s \in I_m$.

Almost at the end...

- $\mu(E) > 0$, so (by construction of measure) there exists some $s' = \lim_{j \to \infty} T^{n_j} s \in I_m$.
- But I_m is open in $2^{\mathbb{N}}$, so there's some $J \in \mathbb{N}$ such that $T^{n_j} s \in I_m$ for $j > J$ (definition of limit).

Almost at the end...

- $\mu(E) > 0$, so (by construction of measure) there exists some $s' = \lim_{j\to\infty} T^{n_j} s \in I_m$.
- But I_m is open in $2^{\mathbb{N}}$, so there's some $J \in \mathbb{N}$ such that $T^{n_j} s \in I_m$ for $j > J$ (definition of limit).
- Hence if we let $n = n_j$ for $j > J$, then $T^n s \in I_m = \bigcap_{i=0}^{k-1} T^{-im}[1]$! ☺

What have we learned?

What have we learned?

- We can model some combinatorics problems as dynamical systems, then use ergodic theory to solve them

What have we learned?

- We can model some combinatorics problems as dynamical systems, then use ergodic theory to solve them
- More generally, we can apply fields of math to different fields

What have we learned?

- We can model some combinatorics problems as dynamical systems, then use ergodic theory to solve them
- More generally, we can apply fields of math to different fields
- It can be useful to re-prove results using different approaches