Szemerédi's Theorem via Ergodic Theory

Rotation Project Final Presentation

Kristóf Huszár (PhD student)
Project supervised by: László Erdős

Institute of Science and Technology Austria

July 24, 2014

Historical Overview of Results

Theorem (B. L. van der Waerden, 1927)

If we color $\mathbb{Z}$ with finitely many colors, then for any $k \in \mathbb{N}_+$ there exists a monochromatic k-term arithmetic progression.

Historical Overview of Results

Theorem (B. L. van der Waerden, 1927)

If we color $\mathbb{Z}$ with finitely many colors, then for any $k \in \mathbb{N}_+$ there exists a monochromatic k-term arithmetic progression.

Definition. $(a_1, \ldots, a_k) \subset \mathbb{Z}$ ordered k-tuple is called a ***k*-term arithmetic progression**, if $a_{i+1} - a_i = d$ for all $i = 1, \ldots, k-1$.

Historical Overview of Results

Theorem (B. L. van der Waerden, 1927)

If we color $\mathbb{Z}$ with finitely many colors, then for any $k \in \mathbb{N}_+$ there exists a monochromatic k-term arithmetic progression.

Definition. $(a_1, \ldots, a_k) \subset \mathbb{Z}$ ordered k-tuple is called a ***k*-term arithmetic progression**, if $a_{i+1} - a_i = d$ for all $i = 1, \ldots, k-1$.

Example

$\ldots, \mathbf{5}, \mathbf{6}, \mathbf{7}, \mathbf{8}, \mathbf{9}, \mathbf{10}, \mathbf{11}, \mathbf{12}, \mathbf{13}, \mathbf{14}, \mathbf{15}, \mathbf{16}, \mathbf{17}, \mathbf{18}, \mathbf{19}, \mathbf{20}, \mathbf{21}, \ldots$

Historical Overview of Results

Theorem (B. L. van der Waerden, 1927)

If we color $\mathbb{Z}$ with finitely many colors, then for any $k \in \mathbb{N}_+$ there exists a monochromatic k-term arithmetic progression.

Definition. $(a_1, \ldots, a_k) \subset \mathbb{Z}$ ordered k-tuple is called a ***k*-term arithmetic progression**, if $a_{i+1} - a_i = d$ for all $i = 1, \ldots, k-1$.

Example

$\ldots, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, \ldots$

Historical Overview of Results

Theorem (B. L. van der Waerden, 1927)

If we color $\mathbb{Z}$ with finitely many colors, then for any $k \in \mathbb{N}_+$ there exists a monochromatic k-term arithmetic progression.

Definition. $(a_1, \dots, a_k) \subset \mathbb{Z}$ ordered k-tuple is called a ***k*-term arithmetic progression**, if $a_{i+1} - a_i = d$ for all $i = 1, \dots, k-1$.

Example

$\dots, \mathbf{5}, \mathbf{6}, \mathbf{7}, \mathbf{8}, \mathbf{9}, \mathbf{10}, \mathbf{11}, \mathbf{12}, \mathbf{13}, \mathbf{14}, \mathbf{15}, \mathbf{16}, \mathbf{17}, \mathbf{18}, \mathbf{19}, \mathbf{20}, \mathbf{21}, \dots$

$k = 1$: trivial;

Historical Overview of Results

Theorem (B. L. van der Waerden, 1927)

If we color $\mathbb{Z}$ with finitely many colors, then for any $k \in \mathbb{N}_+$ there exists a monochromatic k-term arithmetic progression.

Definition. $(a_1, \ldots, a_k) \subset \mathbb{Z}$ ordered k-tuple is called a ***k*-term arithmetic progression**, if $a_{i+1} - a_i = d$ for all $i = 1, \ldots, k-1$.

Example

$\ldots, \mathbf{5}, \mathbf{6}, \mathbf{7}, \mathbf{8}, \mathbf{9}, \mathbf{10}, \mathbf{11}, \mathbf{12}, \mathbf{13}, \mathbf{14}, \mathbf{15}, \mathbf{16}, \mathbf{17}, \mathbf{18}, \mathbf{19}, \mathbf{20}, \mathbf{21}, \ldots$

$k = 1$: trivial; $k = 2$: trivial, e. g. $(\mathbf{6}, \mathbf{7}), (\mathbf{8}, \mathbf{9}), (\mathbf{10}, \mathbf{18}), \ldots$

Historical Overview of Results

Theorem (B. L. van der Waerden, 1927)

If we color $\mathbb{Z}$ with finitely many colors, then for any $k \in \mathbb{N}_+$ there exists a monochromatic k-term arithmetic progression.

Definition. $(a_1, \ldots, a_k) \subset \mathbb{Z}$ ordered k-tuple is called a ***k*-term arithmetic progression**, if $a_{i+1} - a_i = d$ for all $i = 1, \ldots, k-1$.

Example

$\ldots, \mathbf{5}, \mathbf{6}, \mathbf{7}, \mathbf{8}, \mathbf{9}, \mathbf{10}, \mathbf{11}, \mathbf{12}, \mathbf{13}, \mathbf{14}, \mathbf{15}, \mathbf{16}, \mathbf{17}, \mathbf{18}, \mathbf{19}, \mathbf{20}, \mathbf{21}, \ldots$

$k = 1$: trivial; $k = 2$: trivial, e. g. $(\mathbf{6}, \mathbf{7}), (\mathbf{8}, \mathbf{9}), (\mathbf{10}, \mathbf{18}), \ldots$
$k = 3$: non-trivial, e. g. $(\mathbf{10}, \mathbf{15}, \mathbf{20})$;

Historical Overview of Results

Theorem (B. L. van der Waerden, 1927)

If we color $\mathbb{Z}$ with finitely many colors, then for any $k \in \mathbb{N}_+$ there exists a monochromatic k-term arithmetic progression.

Definition. $(a_1, \ldots, a_k) \subset \mathbb{Z}$ ordered k-tuple is called a ***k*-term arithmetic progression**, if $a_{i+1} - a_i = d$ for all $i = 1, \ldots, k-1$.

Example

$\ldots, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, \ldots$

$k = 1$: trivial; $k = 2$: trivial, e. g. $(6, 7), (8, 9), (10, 18), \ldots$
$k = 3$: non-trivial, e. g. $(10, 15, 20)$; $k = 5$: $(5, 9, 13, 17, 21)$.

Historical Overview of Results

Theorem (B. L. van der Waerden, 1927)

If we color $\mathbb{Z}$ with finitely many colors, then for any $k \in \mathbb{N}_+$ there exists a monochromatic k-term arithmetic progression.

Definition. $(a_1, \ldots, a_k) \subset \mathbb{Z}$ ordered k-tuple is called a ***k*-term arithmetic progression**, if $a_{i+1} - a_i = d$ for all $i = 1, \ldots, k-1$.

Example

$\ldots, \mathbf{5}, \mathbf{6}, \mathbf{7}, \mathbf{8}, \mathbf{9}, \mathbf{10}, \mathbf{11}, \mathbf{12}, \mathbf{13}, \mathbf{14}, \mathbf{15}, \mathbf{16}, \mathbf{17}, \mathbf{18}, \mathbf{19}, \mathbf{20}, \mathbf{21}, \ldots$

$k = 1$: trivial; $k = 2$: trivial, e. g. $(\mathbf{6}, \mathbf{7}), (\mathbf{8}, \mathbf{9}), (\mathbf{10}, \mathbf{18}), \ldots$
$k = 3$: non-trivial, e. g. $(\mathbf{10}, \mathbf{15}, \mathbf{20})$; $k = 5$: $(\mathbf{5}, \mathbf{9}, \mathbf{13}, \mathbf{17}, \mathbf{21})$.

Conjecture of P. Erdős & P. Turán (1936)**.** True reason: at least one of the color classes has positive **upper density**.

Historical Overview of Results

Definition. The **upper density** $\overline{\delta}(A)$ of a set $A \subseteq \mathbb{Z}$ is defined as

$$\overline{\delta}(A) := \limsup_{n\to\infty} \frac{1}{2n+1} |A \cap \{-n, -n+1, \ldots, n-1, n\}|.$$

Historical Overview of Results

Definition. The **upper density** $\overline{\delta}(A)$ of a set $A \subseteq \mathbb{Z}$ is defined as

$$\overline{\delta}(A) := \limsup_{n\to\infty} \frac{1}{2n+1} |A \cap \{-n, -n+1, \ldots, n-1, n\}|.$$

Conjecture of P. Erdős & P. Turán (1936)**.** If $A \subseteq \mathbb{Z}$, such that $\overline{\delta}(A) > 0 \Rightarrow A$ contains a k-term arithmetic progression for each k.

Historical Overview of Results

Definition. The **upper density** $\overline{\delta}(A)$ of a set $A \subseteq \mathbb{Z}$ is defined as

$$\overline{\delta}(A) := \limsup_{n\to\infty} \frac{1}{2n+1} |A \cap \{-n, -n+1, \ldots, n-1, n\}|.$$

Conjecture of P. Erdős & P. Turán (1936). If $A \subseteq \mathbb{Z}$, such that $\overline{\delta}(A) > 0 \Rightarrow A$ contains a k-term arithmetic progression for each k.

Theorem (K. F. Roth, 1953; E. Szemerédi, 1969 and 1975)

- K. F. Roth: $k = 3$ (1953)
- E. Szemerédi: $k = 4$ (1969), $\forall k \in \mathbb{N}_+$ (1975).

Roth: Fields Medal (1958)
Szemerédi: Abel Prize (2012)

Historical Overview of Results

Definition. The **upper density** $\overline{\delta}(A)$ of a set $A \subseteq \mathbb{Z}$ is defined as

$$\overline{\delta}(A) := \limsup_{n\to\infty} \frac{1}{2n+1} |A \cap \{-n, -n+1, \ldots, n-1, n\}|.$$

Conjecture of P. Erdős & P. Turán (1936)**.** If $A \subseteq \mathbb{Z}$, such that $\overline{\delta}(A) > 0 \Rightarrow A$ contains a k-term arithmetic progression for each k.

Theorem (K. F. Roth, 1953; E. Szemerédi, 1969 and 1975)

- K. F. Roth: $k = 3$ (1953)
- E. Szemerédi: $k = 4$ (1969), $\forall k \in \mathbb{N}_+$ (1975).

$\Downarrow$

Roth: Fields Medal (1958)
Szemerédi: Abel Prize (2012)

Szemerédi's Theorem

Historical Overview of Results

Definition. The **upper density** $\overline{\delta}(A)$ of a set $A \subseteq \mathbb{Z}$ is defined as

$$\overline{\delta}(A) := \limsup_{n\to\infty} \frac{1}{2n+1} |A \cap \{-n, -n+1, \dots, n-1, n\}|.$$

Conjecture of P. Erdős & P. Turán (1936). If $A \subseteq \mathbb{Z}$, such that $\overline{\delta}(A) > 0 \Rightarrow A$ contains a k-term arithmetic progression for each k.

Theorem (K. F. Roth, 1953; E. Szemerédi, 1969 and 1975)

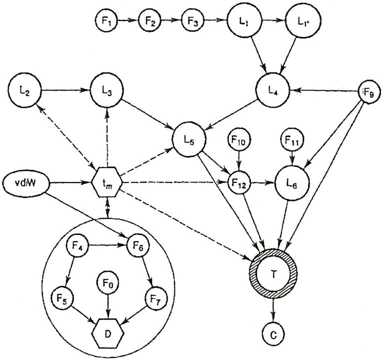
- K. F. Roth: $k = 3$ (1953) $\Longleftarrow$ **Goal of the talk.**
- E. Szemerédi: $k = 4$ (1969), $\forall k \in \mathbb{N}_+$ (1975).

Roth: Fields Medal (1958)
Szemerédi: Abel Prize (2012)

$\Downarrow$

Szemerédi's Theorem

Roadmap of Szemerédi's Original Proof



The diagram represents an approximate flow chart for the accompanying proof of Szemerédi's theorem. The various symbols have the following meanings: F_k ≡ Fact k, L_k ≡ Lemma k, T ≡ Theorem, C ≡ Corollary, D ≡ Definitions of B, S, P, α, β, etc., t_m ≡ Definition of t_m, vdW ≡ van der Waerden's theorem, F_0 ≡ "If $f: R^+ \to R^+$ is subadditive then $\lim_{n\to\infty} \frac{f(n)}{n}$ exists".

A Surprising Connection

In 1977 **H. Furstenberg** gave a new proof via ergodic theory.

$$\text{Additive Combinatorics} \xleftrightarrow{\textit{connection}} \text{Dynamical Systems}$$

A Surprising Connection

In 1977 **H. Furstenberg** gave a new proof via ergodic theory.

$$\text{Additive Combinatorics} \overset{\textit{connection}}{\Longleftrightarrow} \text{Dynamical Systems}$$

Outline of the Talk

1. **Ergodic Theory.** Basic concepts, notation, ergodic theorems, ergodic decomposition.
2. **Correspondence Principle.** Converting problems in additive combinatorics into problems about dynamical systems.

---

3. **Weak Mixing & Compact Systems.** Two extreme cases ("pseudorandom" & "structured" systems).
4. **Roth's Theorem** ($k = 3$)**.** Putting the pieces together.

A Surprising Connection

In 1977 **H. Furstenberg** gave a new proof via ergodic theory.

$$\text{Additive Combinatorics} \overset{\text{connection}}{\Longleftrightarrow} \text{Dynamical Systems}$$

Outline of the Talk

1. **Ergodic Theory.** Basic concepts, notation, ergodic theorems, ergodic decomposition.
2. **Correspondence Principle.** Converting problems in additive combinatorics into problems about dynamical systems.

---

3. **Weak Mixing & Compact Systems.** Two extreme cases ("pseudorandom" & "structured" systems).
4. **Roth's Theorem** ($k = 3$). Putting the pieces together.

Remark. We mainly follow the exposition of the essay Szemerédi's Theorem via Ergodic Theory by **Yufei Zhao** (2011).

Ergodic Theory I. – Dynamical Systems

Ergodic Theory I. – Dynamical Systems

Definition. X some set, $T\colon X \to X$ (shift map). The pair (X, T) is called a **dynamical system**. [We assume T is invertible.]

Ergodic Theory I. – Dynamical Systems

Definition. X some set, $T\colon X \to X$ (shift map). The pair (X, T) is called a **dynamical system**. [We assume T is invertible.]

Question. What is the **evolution** of the system, as we iterate T?

Ergodic Theory I. – Dynamical Systems

Definition. X some set, $T\colon X \to X$ (shift map). The pair (X, T) is called a **dynamical system**. [We assume T is invertible.]

Question. What is the **evolution** of the system, as we iterate T?

Notation. $T^n := T \circ \cdots \circ T$, $T^{-n} := T^{-1} \circ \cdots \circ T^{-1}$ (n times).
$T^n E := \{T^n(x) \mid x \in E\}$; $T^n f(x) := f(T^{-n}(x))$, if $f\colon X \to X$.

Ergodic Theory I. – Dynamical Systems

Definition. X some set, $T\colon X \to X$ (shift map). The pair (X, T) is called a **dynamical system**. [We assume T is invertible.]

Question. What is the **evolution** of the system, as we iterate T?

Notation. $T^n := T \circ \cdots \circ T$, $T^{-n} := T^{-1} \circ \cdots \circ T^{-1}$ (n times). $T^n E := \{T^n(x) \mid x \in E\}$; $T^n f(x) := f(T^{-n}(x))$, if $f\colon X \to X$.

Topological Dynamical Systems

X is a compact metric space,
$T\colon X \to X$ is homeomorphism.
$\Rightarrow$ **Topological Dynamics**

Ergodic Theory I. – Dynamical Systems

Definition. X some set, $T\colon X \to X$ (shift map). The pair (X, T) is called a **dynamical system**. [We assume T is invertible.]

Question. What is the **evolution** of the system, as we iterate T?

Notation. $T^n := T \circ \cdots \circ T$, $T^{-n} := T^{-1} \circ \cdots \circ T^{-1}$ (n times).
$T^n E := \{T^n(x) \mid x \in E\}$; $T^n f(x) := f(T^{-n}(x))$, if $f\colon X \to X$.

Topological Dynamical Systems

X is a compact metric space,
$T\colon X \to X$ is homeomorphism.
$\Rightarrow$ **Topological Dynamics**

Measure Preserving Systems

$(X, \mathcal{X}, \mu, T)$, where

- X compact metric space,
- $\mathcal{X}$ σ-algebra of measurable sets,
- μ prob. measure: $\mu(X) = 1$,
- T measure preserving: $\forall E \in \mathcal{X}$, $\forall n \in \mathbb{Z}$: $\mu(T^n E) = \mu(E)$.

$\Rightarrow$ **Ergodic Theory**

Ergodic Theory I. – Dynamical Systems

Definition. X some set, $T\colon X \to X$ (shift map). The pair (X, T) is called a **dynamical system**. [We assume T is invertible.]

Question. What is the **evolution** of the system, as we iterate T?

Notation. $T^n := T \circ \cdots \circ T, T^{-n} := T^{-1} \circ \cdots \circ T^{-1}$ (n times).
$T^n E := \{T^n(x) \mid x \in E\}$; $T^n f(x) := f(T^{-n}(x))$, if $f\colon X \to X$.

Topological Dynamical Systems

X is a compact metric space,
$T\colon X \to X$ is homeomorphism.
$\Rightarrow$ **Topological Dynamics**

In this talk: **system** :=
measure preserving system.
All functions are measurable.

Measure Preserving Systems

$(X, \mathcal{X}, \mu, T)$, where

- X compact metric space,
- $\mathcal{X}$ σ-algebra of measurable sets,
- μ prob. measure: $\mu(X) = 1$,
- T measure preserving: $\forall E \in \mathcal{X}$, $\forall n \in \mathbb{Z}$: $\mu(T^n E) = \mu(E)$.

$\Rightarrow$ **Ergodic Theory**

Ergodic Theory I. – Dynamical Systems

Definition. X some set, $T: X \to X$ (shift map). The pair (X, T) is called a **dynamical system**. [We assume T is invertible.]

Question. What is the **evolution** of the system, as we iterate T?

Notation. $T^n := T \circ \cdots \circ T$, $T^{-n} := T^{-1} \circ \cdots \circ T^{-1}$ (n times).
$T^n E := \{T^n(x) \mid x \in E\}$; $T^n f(x) := f(T^{-n}(x))$, if $f: X \to X$.

Topological Dynamical Systems

X is a compact metric space,
$T: X \to X$ is homeomorphism.
$\Rightarrow$ **Topological Dynamics**

In this talk: **system** :=
measure preserving system.
All functions are measurable.

Measure Preserving Systems

$(X, \mathcal{X}, \mu, T)$, where

- X compact metric space,
- $\mathcal{X}$ σ-algebra of measurable sets,
- μ prob. measure: $\mu(X) = 1$,
- T measure preserving: $\forall E \in \mathcal{X}$, $\forall n \in \mathbb{Z}$: $\mu(T^n E) = \mu(E)$.

$\Rightarrow$ **Ergodic Theory**

Ergodic Theory I. – Dynamical Systems

Definition. X some set, $T\colon X \to X$ (shift map). The pair (X, T) is called a **dynamical system**. [We assume T is invertible.]

Question. What is the **evolution** of the system, as we iterate T?

Notation. $T^n := T \circ \cdots \circ T$, $T^{-n} := T^{-1} \circ \cdots \circ T^{-1}$ (n times). $T^n E := \{T^n(x) \mid x \in E\}$; $T^n f(x) := f(T^{-n}(x))$, if $f\colon X \to X$.

Topological Dynamical Systems

X is a compact metric space,
$T\colon X \to X$ is homeomorphism.
$\Rightarrow$ **Topological Dynamics**

Measure Preserving Systems

$(X, \mathcal{X}, \mu, T)$, where

- X compact metric space,
- $\mathcal{X}$ σ-algebra of measurable sets,
- μ prob. measure: $\mu(X) = 1$,
- T measure preserving: $\forall E \in \mathcal{X}$, $\forall n \in \mathbb{Z}\colon \mu(T^n E) = \mu(E)$.

$\Rightarrow$ **Ergodic Theory**

In this talk: **system** := measure preserving system. All functions are measurable.

Ergodic Theory II. – Basic Concepts

Notation. $\mathcal{X}^T := \{E \in \mathcal{X} \mid TE = E\}$ (sub-σ-algebra of $\mathcal{X}$).

Ergodic Theory II. – Basic Concepts

Notation. $\mathcal{X}^T := \{E \in \mathcal{X} \mid TE = E\}$ (sub-σ-algebra of $\mathcal{X}$).

Definition I. A system $X = (X, \mathcal{X}, \mu, T)$ is called **ergodic** if for all $E \in \mathcal{X}^T$ we have $\mu(E) = 0$ or $\mu(E) = 1$.

Ergodic Theory II. – Basic Concepts

Notation. $\mathcal{X}^T := \{E \in \mathcal{X} \mid TE = E\}$ (sub-σ-algebra of $\mathcal{X}$).

Definition I. A system $X = (X, \mathcal{X}, \mu, T)$ is called **ergodic** if for all $E \in \mathcal{X}^T$ we have $\mu(E) = 0$ or $\mu(E) = 1$. $\iff$

Definition II. [...] **ergodic** if every f s. t. $Tf = f$ is constant a. e.

Ergodic Theory II. – Basic Concepts

Notation. $\mathcal{X}^T := \{E \in \mathcal{X} \mid TE = E\}$ (sub-σ-algebra of $\mathcal{X}$).

Definition I. A system $X = (X, \mathcal{X}, \mu, T)$ is called **ergodic** if for all $E \in \mathcal{X}^T$ we have $\mu(E) = 0$ or $\mu(E) = 1$. $\iff$

Definition II. [...] **ergodic** if every f s. t. $Tf = f$ is constant a. e.

$L^2 := L^2(X, \mathcal{X}, \mu) := \left\{ f \colon X \to \mathbb{R} \mid \int_X |f|^2 \, \mathrm{d}\mu < \infty \right\} / \sim$, where $f \sim g \iff f = g$ a. e. L^2 is a **Hilbert space**; $\langle f, g \rangle := \int_X fg \, \mathrm{d}\mu$.

Ergodic Theory II. – Basic Concepts

Notation. $\mathcal{X}^T := \{E \in \mathcal{X} \mid TE = E\}$ (sub-σ-algebra of $\mathcal{X}$).

Definition I. A system $X = (X, \mathcal{X}, \mu, T)$ is called **ergodic** if for all $E \in \mathcal{X}^T$ we have $\mu(E) = 0$ or $\mu(E) = 1$. $\iff$

Definition II. [...] **ergodic** if every f s. t. $Tf = f$ is constant a. e.

$L^2 := L^2(X, \mathcal{X}, \mu) := \left\{f \colon X \to \mathbb{R} \mid \int_X |f|^2 \, \mathrm{d}\mu < \infty\right\} / \sim$, where $f \sim g \iff f = g$ a. e. L^2 is a **Hilbert space**; $\langle f, g\rangle := \int_X fg \, \mathrm{d}\mu$.

$\mathcal{Y} \subseteq \mathcal{X}$ is a σ-alg. $\Longrightarrow L^2(X, \mathcal{Y}, \mu) \leq L^2(X, \mathcal{X}, \mu)$ closed subspace.

Ergodic Theory II. – Basic Concepts

Notation. $\mathcal{X}^T := \{E \in \mathcal{X} \mid TE = E\}$ (sub-σ-algebra of $\mathcal{X}$).

Definition I. A system $X = (X, \mathcal{X}, \mu, T)$ is called **ergodic** if for all $E \in \mathcal{X}^T$ we have $\mu(E) = 0$ or $\mu(E) = 1$. $\iff$

Definition II. [...] **ergodic** if every f s. t. $Tf = f$ is constant a. e.

$L^2 := L^2(X, \mathcal{X}, \mu) := \left\{f\colon X \to \mathbb{R} \mid \int_X |f|^2 \,\mathrm{d}\mu < \infty\right\} / \sim$, where $f \sim g \iff f = g$ a. e. L^2 is a **Hilbert space**; $\langle f, g\rangle := \int_X fg \,\mathrm{d}\mu$.

$\mathcal{Y} \subseteq \mathcal{X}$ is a σ-alg. $\implies L^2(X, \mathcal{Y}, \mu) \leq L^2(X, \mathcal{X}, \mu)$ closed subspace. The **orthogonal projection** $\mathbb{E}(\cdot|\mathcal{Y})\colon L^2(X, \mathcal{X}, \mu) \to L^2(X, \mathcal{Y}, \mu)$ is called **conditional expectation**.

Ergodic Theory II. – Basic Concepts

Notation. $\mathcal{X}^T := \{E \in \mathcal{X} \mid TE = E\}$ (sub-σ-algebra of $\mathcal{X}$).

Definition I. A system $X = (X, \mathcal{X}, \mu, T)$ is called **ergodic** if for all $E \in \mathcal{X}^T$ we have $\mu(E) = 0$ or $\mu(E) = 1$. $\iff$

Definition II. [...] **ergodic** if every f s. t. $Tf = f$ is constant a. e.

$L^2 := L^2(X, \mathcal{X}, \mu) := \left\{f\colon X \to \mathbb{R} \mid \int_X |f|^2 \,\mathrm{d}\mu < \infty\right\} / \sim$, where $f \sim g \iff f = g$ a. e. L^2 is a **Hilbert space**; $\langle f, g\rangle := \int_X fg \,\mathrm{d}\mu$.

$\mathcal{Y} \subseteq \mathcal{X}$ is a σ-alg. $\implies L^2(X, \mathcal{Y}, \mu) \leq L^2(X, \mathcal{X}, \mu)$ closed subspace.

The **orthogonal projection** $\mathbb{E}(\cdot|\mathcal{Y})\colon L^2(X, \mathcal{X}, \mu) \to L^2(X, \mathcal{Y}, \mu)$ is called **conditional expectation**.

Proposition. If a system X is ergodic, then $\mathbb{E}(f|\mathcal{X}^T) = \mathbb{E}(f)$, where $\mathbb{E}(f) = \int_X f \,\mathrm{d}\mu$ is the usual expectation.

Ergodic Theory III. – Ergodic Theorems

"General Form" of Ergodic Theorems.

$$\mathrm{Av}_N(T^n f) := \frac{1}{N}\sum_{n=0}^{N-1} T^n f \quad \xrightarrow[\text{some sense}]{(N\to\infty)} \quad \mathbb{E}(f|\mathcal{X}^T)$$

Definition. $\mathrm{Av}_N(T^n f)$: **time average**; $\mathbb{E}(f)$: **space average**.

Ergodic Theory III. – Ergodic Theorems

"General Form" of Ergodic Theorems.

$$\mathrm{Av}_N(T^n f) := \frac{1}{N}\sum_{n=0}^{N-1} T^n f \quad \xrightarrow[\text{some sense}]{(N\to\infty)} \quad \mathbb{E}(f|\mathcal{X}^T)$$

Definition. $\mathrm{Av}_N(T^n f)$: **time average**; $\mathbb{E}(f)$: **space average**.

Proposition. If a system X is ergodic, then $\mathbb{E}(f|\mathcal{X}^T) = \mathbb{E}(f)$.

Ergodic Theory III. – Ergodic Theorems

"General Form" of Ergodic Theorems.

$$\mathrm{Av}_N(T^n f) := \frac{1}{N}\sum_{n=0}^{N-1} T^n f \quad \xrightarrow[\text{some sense}]{(N\to\infty)} \quad \mathbb{E}(f|\mathcal{X}^T)$$

Definition. $\mathrm{Av}_N(T^n f)$: **time average**; $\mathbb{E}(f)$: **space average**.

Proposition. If a system X is ergodic, then $\mathbb{E}(f|\mathcal{X}^T) = \mathbb{E}(f)$.

Theorem (von Neumann mean ergodic theorem)

Let $X = (X, \mathcal{X}, \mu, T)$ be a system, and $f \in L^2(X, \mathcal{X}, \mu)$. Then

$$\mathrm{Av}_N(T^n f) \longrightarrow \mathbb{E}(f|\mathcal{X}^T) \quad \text{in } L^2$$

as $N \to \infty$. If X is ergodic, then the limit equals $\mathbb{E}(f)$.

Ergodic Theory III. – Ergodic Theorems

"General Form" of Ergodic Theorems.

$$\mathrm{Av}_N(T^n f) := \frac{1}{N}\sum_{n=0}^{N-1} T^n f \quad \xrightarrow[\text{some sense}]{(N\to\infty)} \quad \mathbb{E}(f|\mathcal{X}^T)$$

Definition. $\mathrm{Av}_N(T^n f)$: **time average**; $\mathbb{E}(f)$: **space average**.

Proposition. If a system X is ergodic, then $\mathbb{E}(f|\mathcal{X}^T) = \mathbb{E}(f)$.

Theorem (von Neumann mean ergodic theorem)

Let $X = (X, \mathcal{X}, \mu, T)$ be a system, and $f \in L^2(X, \mathcal{X}, \mu)$. Then

$$\mathrm{Av}_N(T^n f) \longrightarrow \mathbb{E}(f|\mathcal{X}^T) \quad \text{in } L^2 \quad (\Rightarrow \text{also } \textbf{weakly})$$

as $N \to \infty$. If X is ergodic, then the limit equals $\mathbb{E}(f)$.

Ergodic Theory IV. – Ergodic Decomposition

Goal. Decompose an arbitrary system into **ergodic components**.

Ergodic Theory IV. – Ergodic Decomposition

Goal. Decompose an arbitrary system into **ergodic components**.

Why? Ergodic theorems have **simple** forms for ergodic systems.

Ergodic Theory IV. – Ergodic Decomposition

Goal. Decompose an arbitrary system into **ergodic components**.

Why? Ergodic theorems have **simple** forms for ergodic systems.

Theorem (Ergodic Decomposition)

Let $(X, \mathcal{X}, \mu, T)$ be a system. Let $\mathcal{E}(X)$ denote the set of ergodic measures on X. There exists a probability measure ρ_μ on $\mathcal{E}(X)$ such that

$$\mu = \int_{\mathcal{E}(X)} \nu \rho_\mu(\mathrm{d}\nu).$$

Ergodic Theory IV. – Ergodic Decomposition

Goal. Decompose an arbitrary system into **ergodic components**.

Why? Ergodic theorems have **simple** forms for ergodic systems.

Theorem (Ergodic Decomposition)

Let $(X, \mathcal{X}, \mu, T)$ be a system. Let $\mathcal{E}(X)$ denote the set of ergodic measures on X. There exists a probability measure ρ_μ on $\mathcal{E}(X)$ such that

$$\mu = \int_{\mathcal{E}(X)} \nu \rho_\mu(\mathrm{d}\nu).$$

Finite decomposition. $\mu = \sum_{i=1}^n \alpha_i \mu_i$, with $\sum_{i=1}^n \alpha_i = 1$, $\alpha_i \geq 0$, where the system $(X, \mathcal{X}, \mu_i, T)$ is ergodic for $i = 1, \dots, n$.

Ergodic Theory IV. – Ergodic Decomposition

Goal. Decompose an arbitrary system into **ergodic components**.

Why? Ergodic theorems have **simple** forms for ergodic systems.

Theorem (Ergodic Decomposition)

Let $(X, \mathcal{X}, \mu, T)$ be a system. Let $\mathcal{E}(X)$ denote the set of ergodic measures on X. There exists a probability measure ρ_μ on $\mathcal{E}(X)$ such that

$$\mu = \int_{\mathcal{E}(X)} \nu \rho_\mu(\mathrm{d}\nu).$$

Finite decomposition. $\mu = \sum_{i=1}^{n} \alpha_i \mu_i$, with $\sum_{i=1}^{n} \alpha_i = 1$, $\alpha_i \geq 0$, where the system $(X, \mathcal{X}, \mu_i, T)$ is ergodic for $i = 1, \ldots, n$.

$\Longrightarrow$ **We can assume ergodicity of systems in certain types of proofs** (including the proof of Szemerédi's Theorem).

Correspondence Principle I. – Bernoulli Systems

Goal. Transforming Szemerédi's Theorem into a problem in ergodic theory.

Correspondence Principle I. – Bernoulli Systems

Goal. Transforming Szemerédi's Theorem into a problem in ergodic theory. $\Longrightarrow$ We define an equivalent problem for **systems**!

Correspondence Principle I. – Bernoulli Systems

Goal. Transforming Szemerédi's Theorem into a problem in ergodic theory. $\Longrightarrow$ We define an equivalent problem for **systems**!

Bernoulli systems. $X = \mathcal{P}(\mathbb{Z})$ (the power set of $\mathbb{Z}$), and $T \colon X \to X$ is defined as $T(B) = B + 1 = \{b + 1 \mid b \in B \subseteq \mathbb{Z}\}$.

Correspondence Principle I. – Bernoulli Systems

Goal. Transforming Szemerédi's Theorem into a problem in ergodic theory. $\Longrightarrow$ We define an equivalent problem for **systems**!

Bernoulli systems. $X = \mathcal{P}(\mathbb{Z})$ (the power set of $\mathbb{Z}$), and $T \colon X \to X$ is defined as $T(B) = B + 1 = \{b + 1 \mid b \in B \subseteq \mathbb{Z}\}$.

- $X \cong \{0,1\}^{\mathbb{Z}}$, which we equip with the product topology (each $\{0,1\}$ is a discrete space).

Correspondence Principle I. – Bernoulli Systems

Goal. Transforming Szemerédi's Theorem into a problem in ergodic theory. $\Longrightarrow$ We define an equivalent problem for **systems**!

Bernoulli systems. $X = \mathcal{P}(\mathbb{Z})$ (the power set of $\mathbb{Z}$), and $T\colon X \to X$ is defined as $T(B) = B + 1 = \{b + 1 \mid b \in B \subseteq \mathbb{Z}\}$.

- $X \cong \{0, 1\}^{\mathbb{Z}}$, which we equip with the product topology (each $\{0, 1\}$ is a discrete space).
- By **Tychonoff's Theorem** X is compact.
  [Main reason for choosing $\{0, 1\}^{\mathbb{Z}}$ instead of $\mathbb{Z}$.]

Correspondence Principle I. – Bernoulli Systems

Goal. Transforming Szemerédi's Theorem into a problem in ergodic theory. $\Longrightarrow$ We define an equivalent problem for **systems**!

Bernoulli systems. $X = \mathcal{P}(\mathbb{Z})$ (the power set of $\mathbb{Z}$), and $T\colon X \to X$ is defined as $T(B) = B + 1 = \{b + 1 \mid b \in B \subseteq \mathbb{Z}\}$.

- $X \cong \{0,1\}^{\mathbb{Z}}$, which we equip with the product topology (each $\{0,1\}$ is a discrete space).
- By **Tychonoff's Theorem** X is compact.
  [Main reason for choosing $\{0,1\}^{\mathbb{Z}}$ instead of $\mathbb{Z}$.]
- X is also metrizable.

Correspondence Principle I. – Bernoulli Systems

Goal. Transforming Szemerédi's Theorem into a problem in ergodic theory. $\Longrightarrow$ We define an equivalent problem for **systems**!

Bernoulli systems. $X = \mathcal{P}(\mathbb{Z})$ (the power set of $\mathbb{Z}$), and $T \colon X \to X$ is defined as $T(B) = B + 1 = \{b + 1 \mid b \in B \subseteq \mathbb{Z}\}$.

- $X \cong \{0,1\}^{\mathbb{Z}}$, which we equip with the product topology (each $\{0,1\}$ is a discrete space).
- By **Tychonoff's Theorem** X is compact. [Main reason for choosing $\{0,1\}^{\mathbb{Z}}$ instead of $\mathbb{Z}$.]
- X is also metrizable.

$\Longrightarrow$ We have a **topological dynamical system**.

Correspondence Principle I. – Bernoulli Systems

Goal. Transforming Szemerédi's Theorem into a problem in ergodic theory. $\Longrightarrow$ We define an equivalent problem for **systems**!

Bernoulli systems. $X = \mathcal{P}(\mathbb{Z})$ (the power set of $\mathbb{Z}$), and $T\colon X \to X$ is defined as $T(B) = B + 1 = \{b + 1 \mid b \in B \subseteq \mathbb{Z}\}$.

- $X \cong \{0,1\}^{\mathbb{Z}}$, which we equip with the product topology (each $\{0,1\}$ is a discrete space).
- By **Tychonoff's Theorem** X is compact.
  [Main reason for choosing $\{0,1\}^{\mathbb{Z}}$ instead of $\mathbb{Z}$.]
- X is also metrizable.

$\Longrightarrow$ We have a **topological dynamical system**.

Idea. Working in an appropriate subspace of X, we shall turn it into a **measure space** via a T-invariant measure μ. $\rightsquigarrow$ **Goal.**

Correspondence Principle II. – Arithmetic Progressions

$\exists$ arithmetic progressions in $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$.	$\Longleftrightarrow$	$\exists$ **arithmetic progressions** in $E \subseteq X$ with $\mu(E) > 0$.

Correspondence Principle II. – Arithmetic Progressions

$\exists$ arithmetic progressions in $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$.	$\Longleftrightarrow$	$\exists$ **arithmetic progressions** in $E \subseteq X$ with $\mu(E) > 0$.

***k*-term arithmetic progression:** $x, T^n x, T^{2n} x, \ldots, T^{(k-1)n} x \in E$.

Correspondence Principle II. – Arithmetic Progressions

∃ arithmetic progressions in $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$.	$\Longleftrightarrow$	∃ **arithmetic progressions** in $E \subseteq X$ with $\mu(E) > 0$.

***k*-term arithmetic progression:** $x, T^n x, T^{2n} x, \ldots, T^{(k-1)n} x \in E$.

Question. Given a system, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$, can we show $E \cap T^n E \cap \cdots \cap T^{(k-1)n} \neq \emptyset$ always for some $n > 0$?

Correspondence Principle II. – Arithmetic Progressions

$\exists$ arithmetic progressions in $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$.	$\iff$	$\exists$ **arithmetic progressions** in $E \subseteq X$ with $\mu(E) > 0$.

***k*-term arithmetic progression:** $x, T^n x, T^{2n} x, \ldots, T^{(k-1)n} x \in E$.

Question. Given a system, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$, can we show $E \cap T^n E \cap \cdots \cap T^{(k-1)n} \neq \emptyset$ always for some $n > 0$?

We shall prove more: $\mu(E \cap T^n E \cap \cdots \cap T^{(k-1)n}) > 0$. $\Longrightarrow$ This would give an **affirmative answer** for the above question.

Correspondence Principle II. – Arithmetic Progressions

$\exists$ arithmetic progressions in $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$.	$\Longleftrightarrow$	$\exists$ **arithmetic progressions** in $E \subseteq X$ with $\mu(E) > 0$.

***k*-term arithmetic progression:** $x, T^n x, T^{2n} x, \ldots, T^{(k-1)n} x \in E$.

Question. Given a system, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$, can we show $E \cap T^n E \cap \cdots \cap T^{(k-1)n} \neq \emptyset$ always for some $n > 0$?

We shall prove more: $\mu(E \cap T^n E \cap \cdots \cap T^{(k-1)n}) > 0$. $\Longrightarrow$ This would give an **affirmative answer** for the above question.

Note. For $k = 2$, the claim is a trivial also in this setting. (Also compare with: **Poincaré Recurrence Theorem**.)

Correspondence Principle III. – Multiple Recurrence

Theorem (E. Szemerédi, 1975)

If $A \subseteq \mathbb{Z}$, such that $\overline{\delta}(A) > 0 \Longrightarrow A$ contains a k-term arithmetic progression for each $k \in \mathbb{N}_+$.

Correspondence $\Updownarrow$ **Principle**

Multiple Recurrence Theorem (H. Furstenberg, 1977)

Let $(X, \mathcal{X}, \mu, T)$ be a system, and $k \in \mathbb{N}_+$. Then for any $E \in \mathcal{X}$ with $\mu(E) > 0$ there exists some $n \in \mathbb{N}_+$ such that

$$\boxed{\mu\left(E \cap T^n E \cap \cdots \cap T^{(k-1)n} E\right) > 0.}$$

Furstenberg $\Longrightarrow$ Szemerédi

Fix $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$.

Furstenberg $\Longrightarrow$ Szemerédi

Fix $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$. **Represent** $A \leftrightsquigarrow a \in \{0,1\}^{\mathbb{Z}}$ in the Bernoulli system $(\{0,1\}^{\mathbb{Z}}, T)$, where $T \leftrightsquigarrow (B \mapsto B+1)$.

Furstenberg $\Longrightarrow$ Szemerédi

Fix $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$. **Represent** $A \leftrightsquigarrow a \in \{0,1\}^{\mathbb{Z}}$ in the Bernoulli system $(\{0,1\}^{\mathbb{Z}}, T)$, where $T \leftrightsquigarrow (B \mapsto B+1)$.

Let $X := \overline{\{T^n a \mid n \in \mathbb{Z}\}}$, and $E = \{b \in X \mid b_0 = 1\}$.

Furstenberg $\Longrightarrow$ Szemerédi

Fix $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$. **Represent** $A \leftrightsquigarrow a \in \{0,1\}^{\mathbb{Z}}$ in the Bernoulli system $(\{0,1\}^{\mathbb{Z}}, T)$, where $T \leftrightsquigarrow (B \mapsto B+1)$.

Let $X := \overline{\{T^n a \mid n \in \mathbb{Z}\}}$, and $E = \{b \in X \mid b_0 = 1\}$.

If there was a μ T-invariant measure on X, s. t. $\mu(E) > 0$, then by **Furstenberg** we would get $\mu(E \cap T^n E \cap \dots \cap T^{(k-1)n} E) > 0$ for some $n \in \mathbb{N}_+$.

Furstenberg $\Longrightarrow$ Szemerédi

Fix $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$. **Represent** $A \leftrightsquigarrow a \in \{0,1\}^{\mathbb{Z}}$ in the Bernoulli system $(\{0,1\}^{\mathbb{Z}}, T)$, where $T \leftrightsquigarrow (B \mapsto B+1)$.

Let $X := \overline{\{T^n a \mid n \in \mathbb{Z}\}}$, and $E = \{b \in X \mid b_0 = 1\}$.

If there was a μ T-invariant measure on X, s. t. $\mu(E) > 0$, then by **Furstenberg** we would get $\mu(E \cap T^n E \cap \cdots \cap T^{(k-1)n} E) > 0$ for some $n \in \mathbb{N}_+$. $\Longrightarrow \emptyset \neq E \cap T^n E \cap \cdots \cap T^{(k-1)n} E \ni T^{-m} a$ for some $m \in \mathbb{Z}$.

Furstenberg $\Longrightarrow$ Szemerédi

Fix $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$. **Represent** $A \leftrightsquigarrow a \in \{0,1\}^{\mathbb{Z}}$ in the Bernoulli system $(\{0,1\}^{\mathbb{Z}}, T)$, where $T \leftrightsquigarrow (B \mapsto B+1)$.

Let $X := \overline{\{T^n a \mid n \in \mathbb{Z}\}}$, and $E = \{b \in X \mid b_0 = 1\}$.

If there was a μ T-invariant measure on X, s. t. $\mu(E) > 0$, then by **Furstenberg** we would get $\mu(E \cap T^n E \cap \cdots \cap T^{(k-1)n} E) > 0$ for some $n \in \mathbb{N}_+$. $\Longrightarrow \emptyset \neq E \cap T^n E \cap \cdots \cap T^{(k-1)n} E \ni T^{-m} a$ for some $m \in \mathbb{Z}$.

Then $(T^{-m}a)_0 = (T^{-n-m}a)_0 = \cdots = (T^{-(k-1)n-m}a)_0 = 1$.

Furstenberg $\Longrightarrow$ Szemerédi

Fix $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$. **Represent** $A \leftrightsquigarrow a \in \{0,1\}^{\mathbb{Z}}$ in the Bernoulli system $(\{0,1\}^{\mathbb{Z}}, T)$, where $T \leftrightsquigarrow (B \mapsto B+1)$.

Let $X := \overline{\{T^n a \mid n \in \mathbb{Z}\}}$, and $E = \{b \in X \mid b_0 = 1\}$.

If there was a μ T-invariant measure on X, s. t. $\mu(E) > 0$, then by **Furstenberg** we would get $\mu(E \cap T^n E \cap \cdots \cap T^{(k-1)n} E) > 0$ for some $n \in \mathbb{N}_+$. $\Longrightarrow \emptyset \neq E \cap T^n E \cap \cdots \cap T^{(k-1)n} E \ni T^{-m} a$ for some $m \in \mathbb{Z}$.

Then $\underbrace{(T^{-m}a)_0}_{m \in A} = \underbrace{(T^{-n-m}a)_0}_{n+m \in A} = \cdots = \underbrace{(T^{-(k-1)n-m}a)_0}_{(k-1)n+m \in A} = 1.$ $\square$

Furstenberg $\implies$ Szemerédi

Fix $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$. **Represent** $A \leftrightsquigarrow a \in \{0,1\}^{\mathbb{Z}}$ in the Bernoulli system $(\{0,1\}^{\mathbb{Z}}, T)$, where $T \leftrightsquigarrow (B \mapsto B+1)$.

Let $X := \overline{\{T^n a \mid n \in \mathbb{Z}\}}$, and $E = \{b \in X \mid b_0 = 1\}$.

If there was a μ T-invariant measure on X, s. t. $\mu(E) > 0$, then by **Furstenberg** we would get $\mu(E \cap T^n E \cap \cdots \cap T^{(k-1)n} E) > 0$ for some $n \in \mathbb{N}_+$. $\implies \emptyset \neq E \cap T^n E \cap \cdots \cap T^{(k-1)n} E \ni T^{-m} a$ for some $m \in \mathbb{Z}$.

Then $\underbrace{(T^{-m}a)_0}_{m \in A} = \underbrace{(T^{-n-m}a)_0}_{n+m \in A} = \cdots = \underbrace{(T^{-(k-1)n-m}a)_0}_{(k-1)n+m \in A} = 1.$ □

Existence. $\mu_N := \frac{1}{2N+1} \sum_{n=-N}^{N} \delta_{T^n a}.$

Furstenberg $\Longrightarrow$ Szemerédi

Fix $A \subseteq \mathbb{Z}$ with $\overline{\delta}(A) > 0$. **Represent** $A \leftrightsquigarrow a \in \{0,1\}^{\mathbb{Z}}$ in the Bernoulli system $(\{0,1\}^{\mathbb{Z}}, T)$, where $T \leftrightsquigarrow (B \mapsto B+1)$.

Let $X := \overline{\{T^n a \mid n \in \mathbb{Z}\}}$, and $E = \{b \in X \mid b_0 = 1\}$.

If there was a μ T-invariant measure on X, s. t. $\mu(E) > 0$, then by **Furstenberg** we would get $\mu(E \cap T^n E \cap \cdots \cap T^{(k-1)n} E) > 0$ for some $n \in \mathbb{N}_+$. $\Longrightarrow \emptyset \neq E \cap T^n E \cap \cdots \cap T^{(k-1)n} E \ni T^{-m} a$ for some $m \in \mathbb{Z}$.

Then $$\underbrace{(T^{-m}a)_0}_{m \in A} = \underbrace{(T^{-n-m}a)_0}_{n+m \in A} = \cdots = \underbrace{(T^{-(k-1)n-m}a)_0}_{(k-1)n+m \in A} = 1.$$ $\square$

Existence. $\mu_N := \frac{1}{2N+1} \sum_{n=-N}^{N} \delta_{T^n a}$. **Homework:** The sequence $(\mu_N)_{N \in \mathbb{N}}$ has some T-invariant weak limit μ, for which $\mu(E) > 0$.

[Use the assumption $\overline{\delta}(A) > 0$ & the **Banach–Alaoglu Theorem**.]

Szemerédi Systems (SZ Systems)

Multiple Recurrence Theorem (alternative form)

$(X, \mathcal{X}, \mu, T)$ is a system, and $k \in \mathbb{N}_+$. $\forall E \in \mathcal{X}$ with $\mu(E) > 0$,

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

Szemerédi Systems (SZ Systems)

Multiple Recurrence Theorem (alternative form)

$(X, \mathcal{X}, \mu, T)$ is a system, and $k \in \mathbb{N}_+$. $\forall E \in \mathcal{X}$ with $\mu(E) > 0$,

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

Definition. A system $X = (X, \mathcal{X}, \mu, T)$ is **SZ of level *k*** if

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \int_X f \cdot T^n f \cdot T^{2n} f \cdots T^{(k-1)n} f \, \mathrm{d}\mu > 0,$$

whenever $f \in L^\infty(X)$, $f \geq 0$, and $\mathbb{E}(f) > 0$.

Szemerédi Systems (SZ Systems)

Multiple Recurrence Theorem (alternative form)

$(X, \mathcal{X}, \mu, T)$ is a system, and $k \in \mathbb{N}_+$. $\forall E \in \mathcal{X}$ with $\mu(E) > 0$,

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

Definition. A system $X = (X, \mathcal{X}, \mu, T)$ is **SZ of level *k*** if

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \int_X f \cdot T^n f \cdot T^{2n} f \cdots T^{(k-1)n} f \, \mathrm{d}\mu > 0,$$

whenever $f \in L^\infty(X)$, $f \geq 0$, and $\mathbb{E}(f) > 0$. A system X is **SZ** if it is SZ of every level.

Szemerédi Systems (SZ Systems)

Multiple Recurrence Theorem (alternative form)

$(X, \mathcal{X}, \mu, T)$ is a system, and $k \in \mathbb{N}_+$. $\forall E \in \mathcal{X}$ with $\mu(E) > 0$,

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

Definition. A system $X = (X, \mathcal{X}, \mu, T)$ is **SZ of level k** if

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \int_X f \cdot T^n f \cdot T^{2n} f \cdots T^{(k-1)n} f \, d\mu > 0,$$

whenever $f \in L^\infty(X)$, $f \geq 0$, and $\mathbb{E}(f) > 0$. A system X is **SZ** if it is SZ of every level. **Ultimate Goal:** Every system is SZ.

Ideas of the Proof – Types of Systems

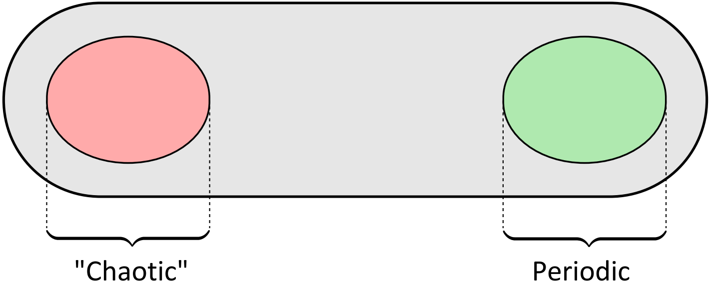
Measure Preserving Systems

(according to the behavior of T)

Ideas of the Proof – Types of Systems

Measure Preserving Systems

(according to the behavior of T)



Ideas of the Proof – Two Extreme Cases

T is "chaotic". If $E \in \mathcal{X}$ is any event in the probability space $(X, \mathcal{X}, \mu)$, then $E, TE, T^2E, \ldots$ are all independent.

Ideas of the Proof – Two Extreme Cases

T is "chaotic". If $E \in \mathcal{X}$ is any event in the probability space $(X, \mathcal{X}, \mu)$, then $E, TE, T^2E, \ldots$ are all independent. $\Longrightarrow$
$\mu\left(E \cap T^n E \cap \cdots \cap T^{(k-1)n} E\right) = \prod_{i=0}^{k-1} \mu\left(T^{i \cdot n} E\right) = \mu(E)^k > 0$, whenever $\mu(E) > 0$. ✓

Ideas of the Proof – Two Extreme Cases

T is "chaotic". If $E \in \mathcal{X}$ is any event in the probability space $(X, \mathcal{X}, \mu)$, then $E, TE, T^2E, \ldots$ are all independent. $\Longrightarrow$
$\mu\left(E \cap T^nE \cap \cdots \cap T^{(k-1)n}E\right) = \prod_{i=0}^{k-1} \mu\left(T^{i \cdot n}E\right) = \mu(E)^k > 0$, whenever $\mu(E) > 0$. ✓

T is periodic. For every $E \in \mathcal{X}$ there is an $r \in \mathbb{N}_+$ (may depend on E), such that $T^rE = E$.

Ideas of the Proof – Two Extreme Cases

T is "chaotic". If $E \in \mathcal{X}$ is any event in the probability space $(X, \mathcal{X}, \mu)$, then E, TE, $T^2E, \dots$ are all independent. $\Longrightarrow$
$\mu\left(E \cap T^nE \cap \dots \cap T^{(k-1)n}E\right) = \prod_{i=0}^{k-1} \mu\left(T^{i \cdot n}E\right) = \mu(E)^k > 0$, whenever $\mu(E) > 0$. ✓

T is periodic. For every $E \in \mathcal{X}$ there is an $r \in \mathbb{N}_+$ (may depend on E), such that $T^rE = E$. $\Longrightarrow$
$\mu\left(E \cap T^nE \cap \dots \cap T^{(k-1)n}E\right) = \mu(E) > 0$, whenever $r \mid n$. ✓

Ideas of the Proof – Two Extreme Cases

T is "chaotic". If $E \in \mathcal{X}$ is any event in the probability space $(X, \mathcal{X}, \mu)$, then $E, TE, T^2E, \ldots$ are all independent. $\Longrightarrow$
$\mu\left(E \cap T^nE \cap \cdots \cap T^{(k-1)n}E\right) = \prod_{i=0}^{k-1} \mu\left(T^{i \cdot n}E\right) = \mu(E)^k > 0$, whenever $\mu(E) > 0$. ✓

T is periodic. For every $E \in \mathcal{X}$ there is an $r \in \mathbb{N}_+$ (may depend on E), such that $T^rE = E$. $\Longrightarrow$
$\mu\left(E \cap T^nE \cap \cdots \cap T^{(k-1)n}E\right) = \mu(E) > 0$, whenever $r \mid n$. ✓

Problem. The above assumptions are very restrictive, and give solution only for special cases. We need to weaken them!

Ideas of the Proof – Two Extreme Cases

***T* is "chaotic".** If $E \in \mathcal{X}$ is any event in the probability space $(X, \mathcal{X}, \mu)$, then $E, TE, T^2E, \ldots$ are all independent. $\Longrightarrow$

$\mu\left(E \cap T^n E \cap \cdots \cap T^{(k-1)n} E\right) = \prod_{i=0}^{k-1} \mu\left(T^{i \cdot n} E\right) = \mu(E)^k > 0,$ whenever $\mu(E) > 0$. ✓

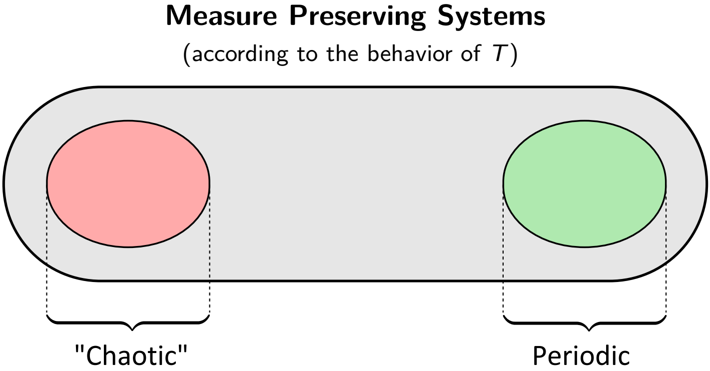
***T* is periodic.** For every $E \in \mathcal{X}$ there is an $r \in \mathbb{N}_+$ (may depend on E), such that $T^r E = E$. $\Longrightarrow$

$\mu\left(E \cap T^n E \cap \cdots \cap T^{(k-1)n} E\right) = \mu(E) > 0$, whenever $r \mid n$. ✓

Problem. The above assumptions are very restrictive, and give solution only for special cases. We need to weaken them! $\rightsquigarrow$

Weak mixing and **Almost Periodic/Compact** systems.

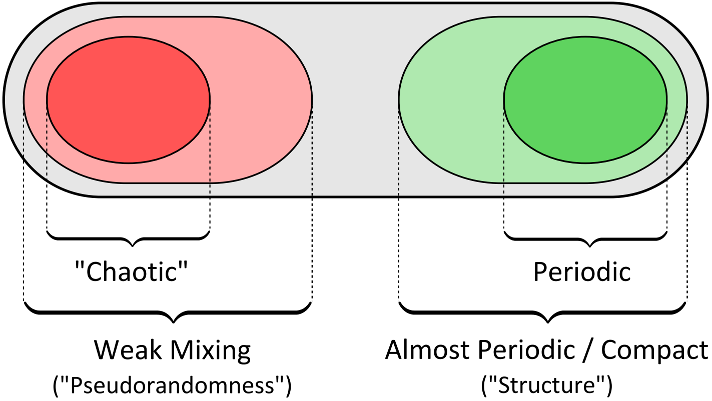
Ideas of the Proof – Types of Systems



Ideas of the Proof – Types of Systems

Measure Preserving Systems

(according to the behavior of T)



Weak Mixing Systems

Intuition. The events $E, TE, T^2E, \ldots$ are not independent, but E and T^nE become nearly uncorrelated in **some sense** as $n \to \infty$.

Weak Mixing Systems

Intuition. The events $E, TE, T^2E, \ldots$ are not independent, but E and T^nE become nearly uncorrelated in **some sense** as $n \to \infty$.

Definition. $v \in V$, $(v_n)_{n\in\mathbb{N}} \subset V$ normed. **$\text{D-lim}_{n\to\infty} v_n = v$**, if for any $\varepsilon > 0$ we have $\overline{\delta}\left(\{n \in \mathbb{N} \mid \|v_n - v\| > \varepsilon\}\right) = 0$.

Weak Mixing Systems

Intuition. The events $E, TE, T^2E, \ldots$ are not independent, but E and T^nE become nearly uncorrelated in **some sense** as $n \to \infty$.

Definition. $v \in V$, $(v_n)_{n\in\mathbb{N}} \subset V$ normed. **D-lim**$_{n\to\infty}$ $v_n = v$, if for any $\varepsilon > 0$ we have $\overline{\delta}\left(\{n \in \mathbb{N} \mid \|v_n - v\| > \varepsilon\}\right) = 0$.

Definition. $(X, \mathcal{X}, \mu, T)$ is **weak mixing** if for any $A, B \in \mathcal{X}$

$$\operatorname*{D-lim}_{n\to\infty} \mu(T^nA \cap B) = \mu(A)\mu(B). \qquad \Big| \qquad \operatorname*{D-lim}_{n\to\infty} \langle T^nf, g\rangle = \mathbb{E}(f)\mathbb{E}(g).$$

Weak Mixing Systems

Intuition. The events $E, TE, T^2E, \ldots$ are not independent, but E and T^nE become nearly uncorrelated in **some sense** as $n \to \infty$.

Definition. $v \in V$, $(v_n)_{n\in\mathbb{N}} \subset V$ normed. $\textbf{D-lim}_{n\to\infty} \boldsymbol{v_n = v}$, if for any $\varepsilon > 0$ we have $\overline{\delta}\left(\{n \in \mathbb{N} \mid \|v_n - v\| > \varepsilon\}\right) = 0$.

Definition. $(X, \mathcal{X}, \mu, T)$ is **weak mixing** if for any $A, B \in \mathcal{X}$

$$\operatorname*{D-lim}_{n\to\infty} \mu(T^n A \cap B) = \mu(A)\mu(B). \qquad \Big| \qquad \operatorname*{D-lim}_{n\to\infty} \langle T^n f, g\rangle = \mathbb{E}(f)\mathbb{E}(g).$$

Comparing weak mixing and ergodic systems:

Proposition. Weak mixing $\Longrightarrow$ ergodicity (but not vica versa).

Proposition. X w. m. $\Longleftrightarrow$ $X \times X$ w. m. $\Longleftrightarrow$ $X \times X$ ergodic.

Remark. X ergodic $\not\Longrightarrow$ $X \times X$ ergodic [irrational rotation of S^1].

Weak Mixing Functions

Definition. In a system $(X, \mathcal{X}, \mu, T)$ a function $f \in L^2(X)$ is called **weak mixing** if $\text{D-lim}_{n\to\infty} \langle T^n f, f\rangle = 0$.

Weak Mixing Functions

Definition. In a system $(X, \mathcal{X}, \mu, T)$ a function $f \in L^2(X)$ is called **weak mixing** if $\text{D-lim}_{n\to\infty}\langle T^n f, f\rangle = 0$.

Intuition. f is w. m. if the "shifts" $T^n f$ eventually become orthogonal to f (for which T displays "mixing" behavior).

Weak Mixing Functions

Definition. In a system $(X, \mathcal{X}, \mu, T)$ a function $f \in L^2(X)$ is called **weak mixing** if $\text{D-lim}_{n\to\infty}\langle T^n f, f\rangle = 0$.

Intuition. f is w. m. if the "shifts" $T^n f$ eventually become orthogonal to f (for which T displays "mixing" behavior).

Characterization of w. m. systems by w. m. functions:

A system $(X, \mathcal{X}, \mu, T)$ is weak mixing $\iff$ every $f \in L^2(X)$ with $\mathbb{E}(f) = 0$ is weak mixing.

Weak Mixing Functions

Definition. In a system $(X, \mathcal{X}, \mu, T)$ a function $f \in L^2(X)$ is called **weak mixing** if $\text{D-lim}_{n\to\infty} \langle T^n f, f \rangle = 0$.

Intuition. f is w. m. if the "shifts" $T^n f$ eventually become orthogonal to f (for which T displays "mixing" behavior).

Characterization of w. m. systems by w. m. functions:

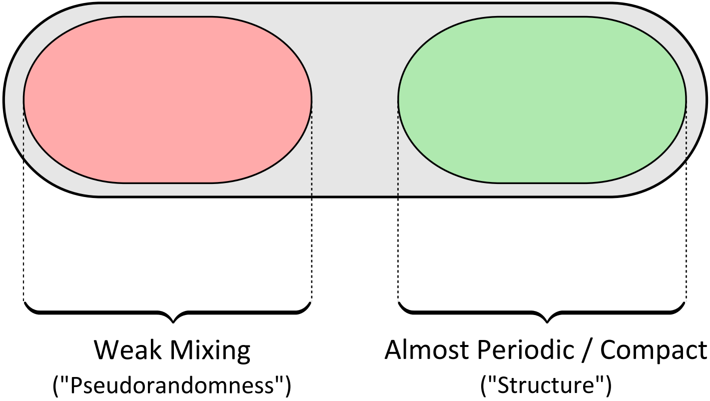
A system $(X, \mathcal{X}, \mu, T)$ is weak mixing $\iff$ every $f \in L^2(X)$ with $\mathbb{E}(f) = 0$ is weak mixing.

Theorem. Every weak mixing system is SZ.

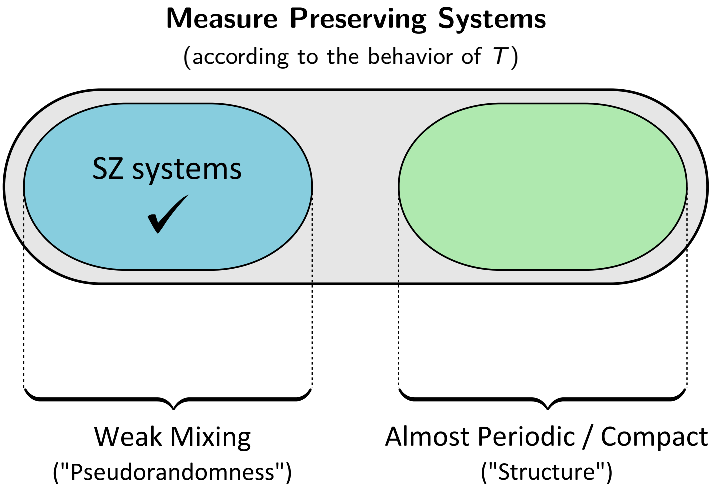
Ideas of the Proof – Types of Systems

Measure Preserving Systems

(according to the behavior of T)



Ideas of the Proof – Types of Systems



Compact Systems & Almost Periodic Functions

Definition. A function $f \in L^2(X)$ is **almost periodic** if for every $\varepsilon > 0$, the set $S_\varepsilon = \{n \in \mathbb{Z} \mid \|f - T^n f\|_2 < \varepsilon\}$ has bounded gaps, which means $\exists N > 0 \colon S_\varepsilon \cap [m, m+N] \neq \emptyset$ for all $m \in \mathbb{Z}$.

Compact Systems & Almost Periodic Functions

Definition. A function $f \in L^2(X)$ is **almost periodic** if for every $\varepsilon > 0$, the set $S_\varepsilon = \{n \in \mathbb{Z} \mid \|f - T^n f\|_2 < \varepsilon\}$ has bounded gaps, which means $\exists N > 0 \colon S_\varepsilon \cap [m, m+N] \neq \emptyset$ for all $m \in \mathbb{Z}$.

Definition. A system $(X, \mathcal{X}, \mu, T)$ is called **compact** if every $f \in L^2(X)$ is almost periodic.

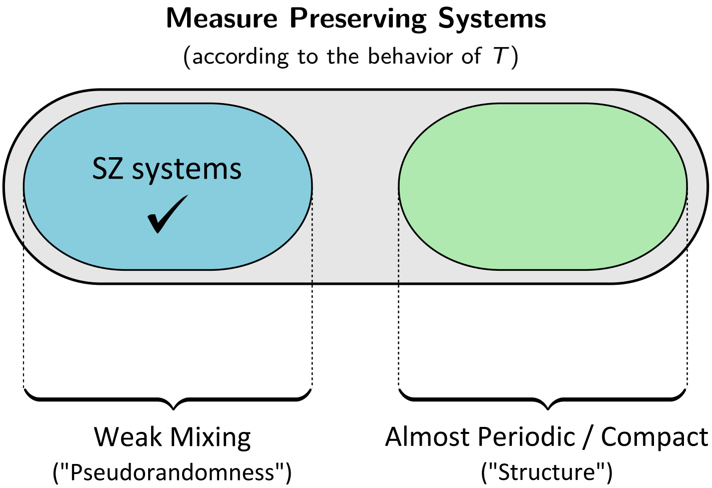
Compact Systems & Almost Periodic Functions

Definition. A function $f \in L^2(X)$ is **almost periodic** if for every $\varepsilon > 0$, the set $S_\varepsilon = \{n \in \mathbb{Z} \mid \|f - T^n f\|_2 < \varepsilon\}$ has bounded gaps, which means $\exists N > 0 \colon S_\varepsilon \cap [m, m+N] \neq \emptyset$ for all $m \in \mathbb{Z}$.

Definition. A system $(X, \mathcal{X}, \mu, T)$ is called **compact** if every $f \in L^2(X)$ is almost periodic.

Theorem. Every compact system is SZ.

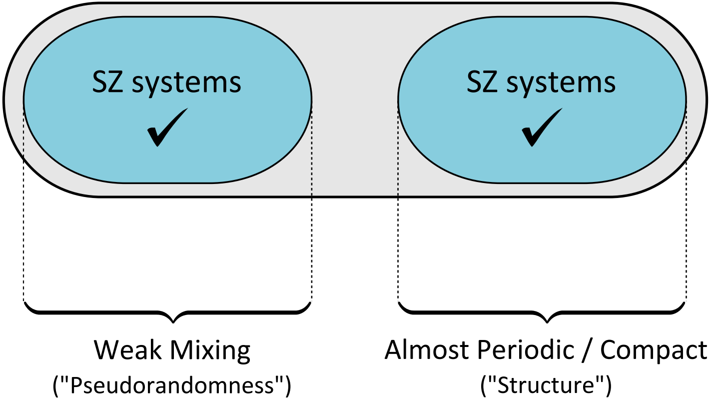
Ideas of the Proof – Types of Systems



Ideas of the Proof – Types of Systems

Measure Preserving Systems

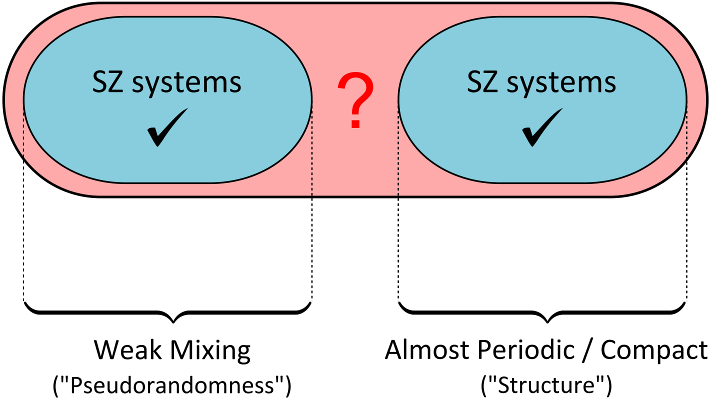
(according to the behavior of T)



Ideas of the Proof – Types of Systems

Measure Preserving Systems

(according to the behavior of T)



Weak Mixing & Almost Periodic Components

Notation. $WM(X) := \{f \in L^2(X) \mid f \text{ is weak mixing}\}$
$AP(X) := \{f \in L^2(X) \mid f \text{ is almost periodic}\}$

Weak Mixing & Almost Periodic Components

Notation. $WM(X) := \{f \in L^2(X) \mid f \text{ is weak mixing}\}$
$AP(X) := \{f \in L^2(X) \mid f \text{ is almost periodic}\}$

Key Proposition. For any system, $\boxed{L^2(X) = WM(X) \oplus AP(X)}$ as an orthogonal direct sum of Hilbert spaces.

Weak Mixing & Almost Periodic Components

Notation. $WM(X) := \{f \in L^2(X) \mid f \text{ is weak mixing}\}$
$AP(X) := \{f \in L^2(X) \mid f \text{ is almost periodic}\}$

Key Proposition. For any system, $\boxed{L^2(X) = WM(X) \oplus AP(X)}$
as an orthogonal direct sum of Hilbert spaces.

Proof ingredients.

- $AP(X) \subseteq L^2(X)$ is a closed T-invariant subspace.
- $f \in WM(X) \iff \langle f, g \rangle = 0$ for all $g \in AP(X)$.

Weak Mixing & Almost Periodic Components

Notation. $WM(X) := \{f \in L^2(X) \mid f \text{ is weak mixing}\}$
$AP(X) := \{f \in L^2(X) \mid f \text{ is almost periodic}\}$

Key Proposition. For any system, $\boxed{L^2(X) = WM(X) \oplus AP(X)}$
as an orthogonal direct sum of Hilbert spaces.

Proof ingredients.

- $AP(X) \subseteq L^2(X)$ is a closed T-invariant subspace.
- $f \in WM(X) \iff \langle f, g \rangle = 0$ for all $g \in AP(X)$.

Message. Unless a system is completely "pseudorandom" ($\leftrightsquigarrow$ WM), it must contain some "structured" ($\leftrightsquigarrow$ AP) piece.

Weak Mixing & Almost Periodic Components

Notation. $WM(X) := \{f \in L^2(X) \mid f \text{ is weak mixing}\}$
$AP(X) := \{f \in L^2(X) \mid f \text{ is almost periodic}\}$

Key Proposition. For any system, $\boxed{L^2(X) = WM(X) \oplus AP(X)}$ as an orthogonal direct sum of Hilbert spaces.

Proof ingredients.

- $AP(X) \subseteq L^2(X)$ is a closed T-invariant subspace.
- $f \in WM(X) \iff \langle f, g \rangle = 0$ for all $g \in AP(X)$.

Message. Unless a system is completely "pseudorandom" ($\leftrightsquigarrow WM$), it must contain some "structured" ($\leftrightsquigarrow AP$) piece.

Now we are ready to prove **Roth's Theorem**.

Roth's Theorem – Statement of the Theorem

Theorem (Roth)**.** Every subset A of $\mathbb{Z}$ with $\overline{\delta}(A) > 0$ contains a 3-term arithmetic progression.

Roth's Theorem – Statement of the Theorem

Theorem (Roth). Every subset A of $\mathbb{Z}$ with $\overline{\delta}(A) > 0$ contains a 3-term arithmetic progression. $\iff$

Theorem (Roth). Every system is SZ of level 3. In other words, let $(X, \mathcal{X}, \mu, T)$ be a system. Then for every $f \in L^\infty(X)$ with $f \geq 0$ and $\mathbb{E}(f) > 0$, we have

$$\boxed{\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f \cdot T^n f \cdot T^{2n} f \, \mathrm{d}\mu > 0.}$$

Roth's Theorem – Statement of the Theorem

Theorem (Roth). Every subset A of $\mathbb{Z}$ with $\overline{\delta}(A) > 0$ contains a 3-term arithmetic progression. $\Longleftrightarrow$

Theorem (Roth). Every system is SZ of level 3. In other words, let $(X, \mathcal{X}, \mu, T)$ be a system. Then for every $f \in L^\infty(X)$ with $f \geq 0$ and $\mathbb{E}(f) > 0$, we have

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f \cdot T^n f \cdot T^{2n} f \, d\mu > 0.$$

Intuition. We get rid of the weak mixing part of the system, and project it onto its almost periodic piece, which is known to be SZ, as it is a compact system.

Roth's Theorem – Key Ingredients

Proposition (♠). $L^2(X) = WM(X) \oplus AP(X)$ as an orthogonal direct sum of Hilbert spaces. Therefore each $f \in L^2(X)$ can be written as $f = f_{WM} + f_{AP} \in WM(X) \oplus AP(X)$.

Proposition (♣). Let $(X, \mathcal{X}, \mu, T)$ be an ergodic system. Then for any $f, g \in L^\infty(X)$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \left(T^n f \, T^{2n} g - T^n f_{AP} T^{2n} g_{AP} \right) = 0 \text{ in } L^2.$$

[The proof relies on **von Neumann's mean Ergodic Theorem**.]

Theorem (■). Every compact system is SZ.

Roth's Theorem – Proof

Let $f \in L^\infty(X)$ with $f \geq 0$ and $\mathbb{E}(f) > 0$.

Roth's Theorem – Proof

Let $f \in L^\infty(X)$ with $f \geq 0$ and $\mathbb{E}(f) > 0$.

Via **ergodic decomposition** we may assume ergodicity of X.

Roth's Theorem – Proof

Let $f \in L^\infty(X)$ with $f \geq 0$ and $\mathbb{E}(f) > 0$.

Via **ergodic decomposition** we may assume ergodicity of X.

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f \cdot T^n f \cdot T^{2n} f \, \mathrm{d}\mu$$

Roth's Theorem – Proof

Let $f \in L^\infty(X)$ with $f \geq 0$ and $\mathbb{E}(f) > 0$.

Via **ergodic decomposition** we may assume ergodicity of X.

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f \cdot T^n f \cdot T^{2n} f \, \mathrm{d}\mu \overset{\clubsuit}{=}$$

$$= \liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f \cdot T^n f_{AP} \cdot T^{2n} f_{AP} \, \mathrm{d}\mu$$

Roth's Theorem – Proof

Let $f \in L^\infty(X)$ with $f \geq 0$ and $\mathbb{E}(f) > 0$.

Via **ergodic decomposition** we may assume ergodicity of X.

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f \cdot T^n f \cdot T^{2n} f \, d\mu \stackrel{\clubsuit}{=}$$

$$= \liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f \cdot T^n f_{AP} \cdot T^{2n} f_{AP} \, d\mu \stackrel{\spadesuit}{=}$$

$$= \liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X (f_{WM} + f_{AP}) \cdot T^n f_{AP} \cdot T^{2n} f_{AP} \, d\mu$$

Roth's Theorem – Proof

Let $f \in L^\infty(X)$ with $f \geq 0$ and $\mathbb{E}(f) > 0$.

Via **ergodic decomposition** we may assume ergodicity of X.

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f \cdot T^n f \cdot T^{2n} f \, d\mu \overset{\clubsuit}{=}$$

$$= \liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f \cdot T^n f_{AP} \cdot T^{2n} f_{AP} \, d\mu \overset{\spadesuit}{=}$$

$$= \liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X (f_{WM} + f_{AP}) \cdot T^n f_{AP} \cdot T^{2n} f_{AP} \, d\mu \overset{\spadesuit}{=}$$

$$= \liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f_{AP} \cdot T^n f_{AP} \cdot T^{2n} f_{AP} \, d\mu \overset{\blacksquare}{>} 0,$$

as $\mathbb{E}(f_{AP}) = \mathbb{E}(\mathbb{E}(f|\mathcal{X}_{AP})) = \mathbb{E}(f) > 0$ (see Appendix). □

References

Klaus F. Roth:
On Certain Sets of Integers.
J. London Math. Soc. **28**(1953), 104–109.

Endre Szemerédi:
On Sets of Integers Containing no k Elements in Arithmetic Progression.
Acta Arith. **27**(1975), 199–245.

Peter Walters:
An Introduction to Ergodic Theory (GTM **79**).
Springer, New York, 1982.

Yufei Zhao:
Szemerédi's Theorem via Ergodic Theory.
Cambridge University, 2011.

What About the Set of Primes?

Let $\mathcal{P} := \{\text{nonnegative primes}\} \subset \mathbb{Z}$, and $\pi(x) := \mathcal{P} \cap [0, x]$.

Claim (elementary). $\lim_{x\to\infty} \pi(x)/x = 0. \implies \boldsymbol{\overline{\delta}(\mathcal{P}) = 0}$.

Prime Number Theorem. $\pi(x)/x \sim \log(x)^{-1} \quad (\text{as } x \to \infty)$.

Question. Longest arithmetic progression containing only primes?

Theorem (B. Green and T. Tao, 2004)

For any $k \in \mathbb{N}_+$ there exists a k-term arithmetic progression in $\mathcal{P}$.

Szemerédi's Theorem is a key ingredient of the proof.

Example (B. Perichon, J. Wróblewski, and G. Reynolds, 2010)

$43,142,746,595,714,191 + 23,681,770 \cdot 223,092,870 \cdot n$, for $n = 0$ to 25. $\implies$ 26-term arithmetic progression of primes.

Ergodic Decomposition: An Example

Goal. Decompose an arbitrary system into **ergodic components**.

Why? Ergodic theorems have **simple** forms for ergodic systems.

Example (a special finite case, but works in general)

$X = \{1,2,3,4,5,6\}$, $\mathcal{X} = 2^X$, μ: uniform, $T = (2\,3)(4\,5\,6) \in S_X$.

$\Longrightarrow T(1) = 1, T(2) = 3, T(3) = 2, T(4) = 5, T(5) = 6, T(6) = 4.$

NOT ergodic: $E = \{1\}$ is T-invariant, but $\mu(E) = 1/6 \notin \{0,1\}$.

BUT consider: $\mu_1 = \mathbb{1}_{\{1\}}$, $\mu_2 = (1/2) \cdot \mathbb{1}_{\{2,3\}}$, $\mu_3 = (1/3) \cdot \mathbb{1}_{\{4,5,6\}}$.

Note: $(X, \mathcal{X}, \mu_i, T)$ **IS** ergodic for $i = 1,2,3$. Moreover, we have

$$\boxed{\mu = \frac{1}{6}\mu_1 + \frac{1}{3}\mu_2 + \frac{1}{2}\mu_3} \Longrightarrow$$

weighted average of ergodic measures.

Szemerédi $\Longrightarrow$ Furstenberg

Lemma. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $E \in \mathcal{X}$ with $\mu(E) > 0$.
$\Longrightarrow \exists F \in \mathcal{X}$, $\mu(F) > 0$, s. t. $\forall x \in F \colon \overline{\delta}\left(\{n \in \mathbb{Z} \mid T^n x \in E\}\right) > 0$.

Szemerédi $\implies$ Furstenberg

Lemma. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $E \in \mathcal{X}$ with $\mu(E) > 0$.
$\implies \exists F \in \mathcal{X}$, $\mu(F) > 0$, s. t. $\forall x \in F \colon \overline{\delta}\left(\{n \in \mathbb{Z} \mid T^n x \in E\}\right) > 0$.

Fix a system $(X, \mathcal{X}, \mu, T)$, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$.

Szemerédi $\Longrightarrow$ Furstenberg

Lemma. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $E \in \mathcal{X}$ with $\mu(E) > 0$.
$\Longrightarrow \exists F \in \mathcal{X}$, $\mu(F) > 0$, s. t. $\forall x \in F \colon \overline{\delta}\left(\{n \in \mathbb{Z} \mid T^n x \in E\}\right) > 0$.

Fix a system $(X, \mathcal{X}, \mu, T)$, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$.

- $\mathcal{AP}_k := \{\boldsymbol{a} = (a_1, \ldots, a_k) \subset \mathbb{Z} \mid \boldsymbol{a}$ is arithmetic progression$\}$.

Szemerédi $\Longrightarrow$ Furstenberg

Lemma. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $E \in \mathcal{X}$ with $\mu(E) > 0$.
$\Longrightarrow \exists F \in \mathcal{X}$, $\mu(F) > 0$, s. t. $\forall x \in F \colon \overline{\delta}\left(\{n \in \mathbb{Z} \mid T^n x \in E\}\right) > 0$.

Fix a system $(X, \mathcal{X}, \mu, T)$, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$.

- $\mathcal{AP}_k := \{\boldsymbol{a} = (a_1, \ldots, a_k) \subset \mathbb{Z} \mid \boldsymbol{a}$ is arithmetic progression$\}$.
- $B_{\boldsymbol{a}} := \{x \in X \mid T^{a_i} x \in E,\ 1 \leq i \leq k\}$; $B_k := \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$.

Szemerédi $\Longrightarrow$ Furstenberg

Lemma. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $E \in \mathcal{X}$ with $\mu(E) > 0$.
$\Longrightarrow \exists F \in \mathcal{X}$, $\mu(F) > 0$, s. t. $\forall x \in F\colon \overline{\delta}\left(\{n \in \mathbb{Z} \mid T^n x \in E\}\right) > 0$.

Fix a system $(X, \mathcal{X}, \mu, T)$, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$.

- $\mathcal{AP}_k := \{\boldsymbol{a} = (a_1, \ldots, a_k) \subset \mathbb{Z} \mid \boldsymbol{a}$ is arithmetic progression$\}$.
- $B_{\boldsymbol{a}} := \{x \in X \mid T^{a_i} x \in E,\ 1 \le i \le k\}$; $B_k := \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$.

Note: $B_k = \left\{x \in X \mid \{n \in \mathbb{Z} \mid T^n x \in E\}\right.$ contains some $\left.\boldsymbol{a} \in \mathcal{AP}_k\right\}$.

Szemerédi $\implies$ Furstenberg

Lemma. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $E \in \mathcal{X}$ with $\mu(E) > 0$.
$\implies \exists F \in \mathcal{X}$, $\mu(F) > 0$, s. t. $\forall x \in F \colon \overline{\delta}\left(\{n \in \mathbb{Z} \mid T^n x \in E\}\right) > 0$.

Fix a system $(X, \mathcal{X}, \mu, T)$, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$.

- $\mathcal{AP}_k := \{\boldsymbol{a} = (a_1, \ldots, a_k) \subset \mathbb{Z} \mid \boldsymbol{a}$ is arithmetic progression$\}$.
- $B_{\boldsymbol{a}} := \{x \in X \mid T^{a_i} x \in E,\ 1 \le i \le k\}$; $B_k := \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$.

Note: $B_k = \left\{x \in X \mid \{n \in \mathbb{Z} \mid T^n x \in E\}\ \text{contains some}\ \boldsymbol{a} \in \mathcal{AP}_k\right\}$.

Choose $F \in \mathcal{X}$ as in the **Lemma**.

Szemerédi $\Longrightarrow$ Furstenberg

Lemma. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $E \in \mathcal{X}$ with $\mu(E) > 0$.
$\Longrightarrow \exists F \in \mathcal{X}$, $\mu(F) > 0$, s. t. $\forall x \in F \colon \overline{\delta}\left(\{n \in \mathbb{Z} \mid T^n x \in E\}\right) > 0$.

Fix a system $(X, \mathcal{X}, \mu, T)$, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$.

- $\mathcal{AP}_k := \{\boldsymbol{a} = (a_1, \ldots, a_k) \subset \mathbb{Z} \mid \boldsymbol{a} \text{ is arithmetic progression}\}$.
- $B_{\boldsymbol{a}} := \{x \in X \mid T^{a_i} x \in E,\ 1 \leq i \leq k\}$; $B_k := \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$.

Note: $B_k = \left\{x \in X \mid \{n \in \mathbb{Z} \mid T^n x \in E\} \text{ contains some } \boldsymbol{a} \in \mathcal{AP}_k\right\}$.

Choose $F \in \mathcal{X}$ as in the **Lemma**. By **Szemerédi**, $\{n \in \mathbb{Z} \mid T^n x \in E\}$ contains some $\boldsymbol{a} \in \mathcal{AP}_k$ for each $x \in F$.

Szemerédi $\implies$ Furstenberg

Lemma. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $E \in \mathcal{X}$ with $\mu(E) > 0$.
$\implies \exists F \in \mathcal{X}$, $\mu(F) > 0$, s. t. $\forall x \in F$: $\overline{\delta}\left(\{n \in \mathbb{Z} \mid T^n x \in E\}\right) > 0$.

Fix a system $(X, \mathcal{X}, \mu, T)$, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$.

- $\mathcal{AP}_k := \{\boldsymbol{a} = (a_1, \ldots, a_k) \subset \mathbb{Z} \mid \boldsymbol{a}$ is arithmetic progression$\}$.
- $B_{\boldsymbol{a}} := \{x \in X \mid T^{a_i} x \in E,\ 1 \leq i \leq k\}$; $B_k := \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$.

Note: $B_k = \left\{x \in X \mid \{n \in \mathbb{Z} \mid T^n x \in E\}\right.$ contains some $\left.\boldsymbol{a} \in \mathcal{AP}_k\right\}$.

Choose $F \in \mathcal{X}$ as in the **Lemma**. By **Szemerédi**, $\{n \in \mathbb{Z} \mid T^n x \in E\}$ contains some $\boldsymbol{a} \in \mathcal{AP}_k$ for each $x \in F$. $\implies F \subseteq B_k = \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$, a countable union.

Szemerédi $\implies$ Furstenberg

Lemma. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $E \in \mathcal{X}$ with $\mu(E) > 0$.
$\implies \exists F \in \mathcal{X}$, $\mu(F) > 0$, s. t. $\forall x \in F$: $\overline{\delta}\left(\{n \in \mathbb{Z} \mid T^n x \in E\}\right) > 0$.

Fix a system $(X, \mathcal{X}, \mu, T)$, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$.

- $\mathcal{AP}_k := \{\boldsymbol{a} = (a_1, \ldots, a_k) \subset \mathbb{Z} \mid \boldsymbol{a}$ is arithmetic progression$\}$.
- $B_{\boldsymbol{a}} := \{x \in X \mid T^{a_i} x \in E,\ 1 \le i \le k\}$; $B_k := \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$.

Note: $B_k = \left\{x \in X \mid \{n \in \mathbb{Z} \mid T^n x \in E\}\text{ contains some } \boldsymbol{a} \in \mathcal{AP}_k\right\}$.

Choose $F \in \mathcal{X}$ as in the **Lemma**. By **Szemerédi**, $\{n \in \mathbb{Z} \mid T^n x \in E\}$ contains some $\boldsymbol{a} \in \mathcal{AP}_k$ for each $x \in F$. $\implies F \subseteq B_k = \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$, a countable union. $\implies$ **Since** $\mu(F) > 0$, $\exists \boldsymbol{b} \in \mathcal{AP}_k$: $\mu(B_{\boldsymbol{b}}) > 0$.

Szemerédi $\Longrightarrow$ Furstenberg

Lemma. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $E \in \mathcal{X}$ with $\mu(E) > 0$.
$\Longrightarrow \exists F \in \mathcal{X}$, $\mu(F) > 0$, s. t. $\forall x \in F$: $\overline{\delta}\left(\{n \in \mathbb{Z} \mid T^n x \in E\}\right) > 0$.

Fix a system $(X, \mathcal{X}, \mu, T)$, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$.

- $\mathcal{AP}_k := \{\boldsymbol{a} = (a_1, \ldots, a_k) \subset \mathbb{Z} \mid \boldsymbol{a}$ is arithmetic progression$\}$.
- $B_{\boldsymbol{a}} := \{x \in X \mid T^{a_i} x \in E,\ 1 \leq i \leq k\}$; $B_k := \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$.

Note: $B_k = \left\{x \in X \mid \{n \in \mathbb{Z} \mid T^n x \in E\}\right.$ contains some $\left.\boldsymbol{a} \in \mathcal{AP}_k\right\}$.

Choose $F \in \mathcal{X}$ as in the **Lemma**. By **Szemerédi**, $\{n \in \mathbb{Z} \mid T^n x \in E\}$ contains some $\boldsymbol{a} \in \mathcal{AP}_k$ for each $x \in F$. $\Longrightarrow F \subseteq B_k = \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$, a countable union. $\Longrightarrow$ **Since** $\mu(F) > 0$, $\exists \boldsymbol{b} \in \mathcal{AP}_k$: $\mu(B_{\boldsymbol{b}}) > 0$.
$\Longrightarrow T^c B_{\boldsymbol{b}} \subseteq E \cap T^n E \cap \cdots \cap T^{(k-1)n} E$ for some $c \in \mathbb{Z}$ and $n \in \mathbb{N}_+$.

Szemerédi $\Longrightarrow$ Furstenberg

Lemma. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $E \in \mathcal{X}$ with $\mu(E) > 0$.
$\Longrightarrow \exists F \in \mathcal{X},\ \mu(F) > 0$, s. t. $\forall x \in F\colon \overline{\delta}\left(\{n \in \mathbb{Z} \mid T^n x \in E\}\right) > 0$.

Fix a system $(X, \mathcal{X}, \mu, T)$, $E \in \mathcal{X}$ with $\mu(E) > 0$, and $k \in \mathbb{N}_+$.

- $\mathcal{AP}_k := \{\boldsymbol{a} = (a_1, \dots, a_k) \subset \mathbb{Z} \mid \boldsymbol{a}$ is arithmetic progression$\}$.
- $B_{\boldsymbol{a}} := \{x \in X \mid T^{a_i} x \in E,\ 1 \le i \le k\}$; $B_k := \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$.

Note: $B_k = \left\{x \in X \mid \{n \in \mathbb{Z} \mid T^n x \in E\}\ \text{contains some}\ \boldsymbol{a} \in \mathcal{AP}_k\right\}$.

Choose $F \in \mathcal{X}$ as in the **Lemma**. By **Szemerédi**, $\{n \in \mathbb{Z} \mid T^n x \in E\}$ contains some $\boldsymbol{a} \in \mathcal{AP}_k$ for each $x \in F$. $\Longrightarrow F \subseteq B_k = \bigcup_{\boldsymbol{a} \in \mathcal{AP}_k} B_{\boldsymbol{a}}$, a countable union. $\Longrightarrow$ **Since** $\mu(F) > 0$, $\exists \boldsymbol{b} \in \mathcal{AP}_k\colon \mu(B_{\boldsymbol{b}}) > 0$.
$\Longrightarrow T^c B_{\boldsymbol{b}} \subseteq E \cap T^n E \cap \dots \cap T^{(k-1)n} E$ for some $c \in \mathbb{Z}$ and $n \in \mathbb{N}_+$.
$\Longrightarrow \mu(E \cap T^n E \cap \dots \cap T^{(k-1)n} E) \ge \mu(T^c B_{\boldsymbol{b}}) = \mu(B_{\boldsymbol{b}}) > 0$. □

Compact Factors

Definition. $\mathcal{Y}$ is ***T*-invariant**, if $TE, T^{-1}E \in \mathcal{Y}$ for any $E \in \mathcal{Y}$.

Definition. If $X = (X, \mathcal{X}, \mu, T)$ is a system, $X' = (X, \mathcal{X}', \mu, T)$ is called a **factor** of X if $\mathcal{X}'$ is a ***T*-invariant** sub-σ-algebra of $\mathcal{X}$.

A factor X' is **trivial**, if $\mu(E) \in \{0, 1\}$ for all $E \in \mathcal{X}'$.

It is **compact**, if X' is a compact measure preserving system.

Theorem

Let X be a system. Exactly one of the followings is true:

1. X is weak mixing ("**pseudorandomness**");
2. X has a nontrivial compact factor ("**structure**").

Kronecker Factors

Notation. If $X = (X, \mathcal{X}, \mu, T)$, $\mathcal{X}_{AP} := \{A \in X \mid \mathbb{1}_A \in AP(X)\}$.

Claim. $\mathcal{X}_{AP}$ is a T-invariant sub-σ-algebra of $\mathcal{X}$. Therefore $(X, \mathcal{X}_{AP}, \mu, T)$ is a factor of X, called **Kronecker factor**.

Remark. The Kronecker is the maximal compact factor of X.

Proposition. Let $(X, \mathcal{X}, \mu, T)$ be a system, and $f \in L^2(X, \mathcal{X}, \mu)$.

1. $f \in AP(X)$ iff f is $\mathcal{X}_{AP}$-measurable: $AP(X) = L^2(X, \mathcal{X}_{AP}, \mu)$.
2. $f \in WM(X)$ iff $\mathbb{E}(f|\mathcal{X}_{AP}) = 0$ a. e.
3. **(♠)** We can write $f = f_{AP} + f_{WM}$, where $f_{AP} := \mathbb{E}(f|\mathcal{X}_{AP}) \in AP(X)$, and $f_{WM} := f - f_{AP} \in WM(X)$.