On the unreasonable effectiveness of ergodic theory in combinatorial number theory

Emilio Corso

ETH Zürich

May 31, 2020

Introduction

The fruitful interaction between ergodic theory and number theory can be traced back to the early days of the former.

Introduction

The fruitful interaction between ergodic theory and number theory can be traced back to the early days of the former.

Figure: L. Boltzmann (1844-1906)

Figure: G.D. Birkhoff (1884-1944)

Introduction

The fruitful interaction between ergodic theory and number theory can be traced back to the early days of the former.

Figure: L. Boltzmann (1844-1906)

Figure: G.D. Birkhoff (1884-1944)

What is so unreasonable about this long-standing connection?

Ergodic vs. number theory

Here is a tentative answer:

Ergodic vs. number theory

Here is a tentative answer:

- *Ergodic theory* investigates the long-term behaviour of evolving systems from a statistical point of view; it provides meaningful information about the time-evolution of **typical** initial data.

Ergodic vs. number theory

Here is a tentative answer:

- *Ergodic theory* investigates the long-term behaviour of evolving systems from a statistical point of view; it provides meaningful information about the time-evolution of **typical** initial data.
- *Number theory* is primarily concerned with questions involving **specific** points or subsets of arithmetically defined objects. Even if some of these questions can be dynamically formulated, ergodic tools may not be sufficiently powerful to give a full understanding of the resulting dynamics.

Ergodic vs. number theory

Here is a tentative answer:

- *Ergodic theory* investigates the long-term behaviour of evolving systems from a statistical point of view; it provides meaningful information about the time-evolution of **typical** initial data.
- *Number theory* is primarily concerned with questions involving **specific** points or subsets of arithmetically defined objects. Even if some of these questions can be dynamically formulated, ergodic tools may not be sufficiently powerful to give a full understanding of the resulting dynamics.

Notwithstanding their inherently different purposes, a wealth of number-theoretical results have been established via ergodic-theoretical methods. For many of these, no other proofs are known to date.

A list of results

Below is a small sample of successful applications of ergodic theory to long-standing problems in number theory:

A list of results

Below is a small sample of successful applications of ergodic theory to long-standing problems in number theory:

1. Furstenberg's alternative proof of Szemerédi's theorem (1977);

A list of results

Below is a small sample of successful applications of ergodic theory to long-standing problems in number theory:

1. Furstenberg's alternative proof of Szemerédi's theorem (1977);
2. Margulis' full proof of the Oppenheim conjecture, on density of values of irrational quadratic forms at integral points (1985);

A list of results

Below is a small sample of successful applications of ergodic theory to long-standing problems in number theory:

1. Furstenberg's alternative proof of Szemerédi's theorem (1977);
2. Margulis' full proof of the Oppenheim conjecture, on density of values of irrational quadratic forms at integral points (1985);
3. Green-Tao's theorem on the existence of arbitrary long arithmetic progressions in primes (2004);

A list of results

Below is a small sample of successful applications of ergodic theory to long-standing problems in number theory:

1. Furstenberg's alternative proof of Szemerédi's theorem (1977);
2. Margulis' full proof of the Oppenheim conjecture, on density of values of irrational quadratic forms at integral points (1985);
3. Green-Tao's theorem on the existence of arbitrary long arithmetic progressions in primes (2004);
4. Einsiedler-Katok-Lindenstrauss' major advancement in Littlewood's conjecture on simultaneous Diophantine approximation (2006);

A list of results

Below is a small sample of successful applications of ergodic theory to long-standing problems in number theory:

1. Furstenberg's alternative proof of Szemerédi's theorem (1977);
2. Margulis' full proof of the Oppenheim conjecture, on density of values of irrational quadratic forms at integral points (1985);
3. Green-Tao's theorem on the existence of arbitrary long arithmetic progressions in primes (2004);
4. Einsiedler-Katok-Lindenstrauss' major advancement in Littlewood's conjecture on simultaneous Diophantine approximation (2006);
5. Ellenberg-Venkatesh's local-global principle for representability of integral quadratic forms (2008);

A list of results

Below is a small sample of successful applications of ergodic theory to long-standing problems in number theory:

1. Furstenberg's alternative proof of Szemerédi's theorem (1977);
2. Margulis' full proof of the Oppenheim conjecture, on density of values of irrational quadratic forms at integral points (1985);
3. Green-Tao's theorem on the existence of arbitrary long arithmetic progressions in primes (2004);
4. Einsiedler-Katok-Lindenstrauss' major advancement in Littlewood's conjecture on simultaneous Diophantine approximation (2006);
5. Ellenberg-Venkatesh's local-global principle for representability of integral quadratic forms (2008);
6. Venkatesh's subconvexity bounds for a class of standard and Rankin-Selberg L-functions (2010).

The Abel prize for 2020

The most recent testimony to the mathematical relevance of the connection between the two areas is 2020's Abel prize, awarded to Furstenberg and Margulis

Figure: H. Furstenberg

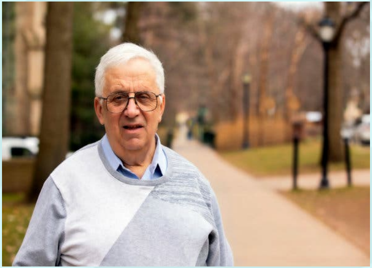

Figure: G.A. Margulis

for having pioneered the use of ergodic theory, and more generally probability theory, in several other domains of mathematics.

Arithmetic progressions in various sets

In this presentation, we focus on the application of recurrence theorems in dynamics to questions pertaining to the additive structure of notable sets of integers.

Arithmetic progressions in various sets

In this presentation, we focus on the application of recurrence theorems in dynamics to questions pertaining to the additive structure of notable sets of integers.

A k-term arithmetic progression is a finite set of natural numbers, of the form $\{a, a+d, a+2d, \dots, a+(k-1)d\}$, where $a, d \geq 1$. We shall revolve around the following:

Arithmetic progressions in various sets

In this presentation, we focus on the application of recurrence theorems in dynamics to questions pertaining to the additive structure of notable sets of integers.

A k-term arithmetic progression is a finite set of natural numbers, of the form $\{a, a+d, a+2d, \dots, a+(k-1)d\}$, where $a, d \geq 1$. We shall revolve around the following:

Question

Which (infinite) subsets of the natural numbers contain arbitrarily long arithmetic progressions? How "large" do they need to be?

Arithmetic progressions in various sets

In this presentation, we focus on the application of recurrence theorems in dynamics to questions pertaining to the additive structure of notable sets of integers.

A k-term arithmetic progression is a finite set of natural numbers, of the form $\{a, a+d, a+2d, \ldots, a+(k-1)d\}$, where $a, d \geq 1$. We shall revolve around the following:

Question

Which (infinite) subsets of the natural numbers contain arbitrarily long arithmetic progressions? How "large" do they need to be?

For instance, the set of even numbers $\{2, 4, 6, \ldots, \}$, the set of odd numbers $\{1, 3, 5, \ldots\}$, and more generally the set of multiples of a given natural number $a\mathbb{N} = \{na : n \geq 1\}$ (and translates thereof) all have this property.

Arithmetic progressions in various sets

In this presentation, we focus on the application of recurrence theorems in dynamics to questions pertaining to the additive structure of notable sets of integers.

A k-term arithmetic progression is a finite set of natural numbers, of the form $\{a, a+d, a+2d, \ldots, a+(k-1)d\}$, where $a, d \geq 1$. We shall revolve around the following:

Question

Which (infinite) subsets of the natural numbers contain arbitrarily long arithmetic progressions? How "large" do they need to be?

For instance, the set of even numbers $\{2, 4, 6, \ldots, \}$, the set of odd numbers $\{1, 3, 5, \ldots\}$, and more generally the set of multiples of a given natural number $a\mathbb{N} = \{na : n \geq 1\}$ (and translates thereof) all have this property.

But this is essentially tautological!

Arithmetic progressions in various sets

What about "sparser" sets, and in particular multiplicatively-defined ones (squares, cubes, geometric progressions)? The question is an instance of the much broader, long-pursued effort to understand how the additive and the multiplicative structure of the integers intertwine.

Arithmetic progressions in various sets

What about "sparser" sets, and in particular multiplicatively-defined ones (squares, cubes, geometric progressions)? The question is an instance of the much broader, long-pursued effort to understand how the additive and the multiplicative structure of the integers intertwine.

Notation: $\mathbb{N} = \{1, 2, 3, \dots\}$

Arithmetic progressions in various sets

What about "sparser" sets, and in particular multiplicatively-defined ones (squares, cubes, geometric progressions)? The question is an instance of the much broader, long-pursued effort to understand how the additive and the multiplicative structure of the integers intertwine.

Notation: $\mathbb{N} = \{1, 2, 3, \dots\}$

- It is obvious that a geometric progression $\{ab^n : n \in \mathbb{N}\}$, $b > 1$, cannot contain a 3-term arithmetic progression.

Arithmetic progressions in various sets

What about "sparser" sets, and in particular multiplicatively-defined ones (squares, cubes, geometric progressions)? The question is an instance of the much broader, long-pursued effort to understand how the additive and the multiplicative structure of the integers intertwine.

Notation: $\mathbb{N} = \{1, 2, 3, \dots\}$

- It is obvious that a geometric progression $\{ab^n : n \in \mathbb{N}\}$, $b > 1$, cannot contain a 3-term arithmetic progression.
- Less obvious is that the set $\mathbb{N}^2 = \{n^2 : n \in \mathbb{N}\}$ does not contain any 4-term arithmetic progression; this was claimed by Fermat in 1640, but first rigorously proved by Euler in 1780.

Arithmetic progressions in various sets

What about "sparser" sets, and in particular multiplicatively-defined ones (squares, cubes, geometric progressions)? The question is an instance of the much broader, long-pursued effort to understand how the additive and the multiplicative structure of the integers intertwine.

Notation: $\mathbb{N} = \{1, 2, 3, \dots\}$

- It is obvious that a geometric progression $\{ab^n : n \in \mathbb{N}\}$, $b > 1$, cannot contain a 3-term arithmetic progression.
- Less obvious is that the set $\mathbb{N}^2 = \{n^2 : n \in \mathbb{N}\}$ does not contain any 4-term arithmetic progression; this was claimed by Fermat in 1640, but first rigorously proved by Euler in 1780.
- Tremendously hard is to prove that $\mathbb{N}^m = \{n^m : n \in \mathbb{N}\}$ does not contain any 3-term arithmetic progression, except possibly for a finite number of trivial ones, whenever $m \geq 3$ (Darmon-Merel, 1997); admittedly, it is almost as difficult as proving Fermat's last theorem.

The conjectures of Erdös and Turán

On the other hand, it has always been widely believed that:

Conjecture (Folklore)

The set of primes $\{2, 3, 5, 7, 11, \ldots\}$ *contains arbitrarily long arithmetic progressions.*

The conjectures of Erdös and Turán

On the other hand, it has always been widely believed that:

Conjecture (Folklore)

The set of primes $\{2, 3, 5, 7, 11, \dots\}$ contains arbitrarily long arithmetic progressions.

More generally, Erdös had postulated the following:

Conjecture (Erdös, 1972)

If $A \subset \mathbb{N}$ satisfies

$$\sum_{n \in A} \frac{1}{n} = \infty,$$

then A contains arbitrarily long arithmetic progressions.

The conjectures of Erdös and Turán

On the other hand, it has always been widely believed that:

Conjecture (Folklore)

The set of primes $\{2, 3, 5, 7, 11, \dots\}$ *contains arbitrarily long arithmetic progressions.*

More generally, Erdös had postulated the following:

Conjecture (Erdös, 1972)

If $A \subset \mathbb{N}$ *satisfies*

$$\sum_{n \in A} \frac{1}{n} = \infty,$$

then A *contains arbitrarily long arithmetic progressions.*

At present, no progress whatsoever has been made on this conjecture in its full generality; it is not even known whether such sets contain a 3-term arithmetic progression.

The conjectures of Erdös and Turán

Actually, Erdös and Turán formulated in 1936 a weaker conjecture, asserting that a sufficient condition for a set $A \subset \mathbb{N}$ to contain arbitrarily long AP's is to have strictly positive *upper density*, the latter being defined by

$$\overline{d}(A) = \limsup_{N \to \infty} \frac{|A \cap [1, N]|}{N} \in [0, 1],$$

where $|B|$ denotes the cardinality of a finite set B, and $[1, N] = \{x \in \mathbb{R} : 1 \leq x \leq N\}$.

The conjectures of Erdös and Turán

Actually, Erdös and Turán formulated in 1936 a weaker conjecture, asserting that a sufficient condition for a set $A \subset \mathbb{N}$ to contain arbitrarily long AP's is to have strictly positive *upper density*, the latter being defined by

$$\overline{d}(A) = \limsup_{N \to \infty} \frac{|A \cap [1, N]|}{N} \in [0, 1],$$

where $|B|$ denotes the cardinality of a finite set B, and $[1, N] = \{x \in \mathbb{R} : 1 \leq x \leq N\}$.

Notice: Erdös' conjecture implies the assertion about primes, while the weaker formulation of Erdös and Turán does not, as the Prime Number Theorem gives

$$\pi(N) \sim \frac{N}{\log N}, \quad \text{where } \pi(N) = |\{\text{primes in } \{1, \ldots, N\}\}|,$$

so that primes have zero density.

The theorems of van der Waerden and Szemerédi

The first result identifying non-trivial classes of sets containing arbitrarily long AP's is due to van der Waerden:

Theorem (Van der Waerden, 1927)

If the natural numbers are coloured using a finite set of colors, then there are monochromatic arithmetic progressions of arbitrary length.

The theorems of van der Waerden and Szemerédi

The first result identifying non-trivial classes of sets containing arbitrarily long AP's is due to van der Waerden:

Theorem (Van der Waerden, 1927)

If the natural numbers are coloured using a finite set of colors, then there are monochromatic arithmetic progressions of arbitrary length.

It took half a century for the next major progress towards Erdös-Turán's conjectures:

Theorem (Szemerédi, 1975)

Let $A \subset \mathbb{N}$ be a set of positive upper density. Then A contains arbitrarily long arithmetic progressions.

The theorems of van der Waerden and Szemerédi

The first result identifying non-trivial classes of sets containing arbitrarily long AP's is due to van der Waerden:

Theorem (Van der Waerden, 1927)

If the natural numbers are coloured using a finite set of colors, then there are monochromatic arithmetic progressions of arbitrary length.

It took half a century for the next major progress towards Erdös-Turán's conjectures:

Theorem (Szemerédi, 1975)

Let $A \subset \mathbb{N}$ be a set of positive upper density. Then A contains arbitrarily long arithmetic progressions.

Szemerédi's proof relies on an intricate combinatorial argument.

The ergodic approach and the theorem of Green and Tao

Two years after Szemerédi's result, Furstenberg realized that the theorem would follow from a refinement of Poincaré's recurrence theorem in ergodic theory, and went on to prove this refinement.

The ergodic approach and the theorem of Green and Tao

Two years after Szemerédi's result, Furstenberg realized that the theorem would follow from a refinement of Poincaré's recurrence theorem in ergodic theory, and went on to prove this refinement.

His investigation spawned intensive research by ergodic and number theorists alike, aimed at establishing quantitative analogues of his results.

The ergodic approach and the theorem of Green and Tao

Two years after Szemerédi's result, Furstenberg realized that the theorem would follow from a refinement of Poincaré's recurrence theorem in ergodic theory, and went on to prove this refinement.

His investigation spawned intensive research by ergodic and number theorists alike, aimed at establishing quantitative analogues of his results.

These efforts culminated in the crowning achievement of Ben Green and Terence Tao:

Theorem (Green-Tao, 2004)

The set of primes contains arbitrarily long arithmetic progressions.

Green-Tao and Dirichlet

Observe that reversing the statement produces a well-known theorem, dating back to the first half of the 19th century:

Green-Tao and Dirichlet

Observe that reversing the statement produces a well-known theorem, dating back to the first half of the 19th century:

Theorem (Dirichlet, 1837)

Let $a, d \in \mathbb{N}$ be coprime. Then the infinite arithmetic progression $a + d\mathbb{N} = \{a, a + d, \ldots, a + nd, \ldots\}$ contains infinitely many primes.

Green-Tao and Dirichlet

Observe that reversing the statement produces a well-known theorem, dating back to the first half of the 19th century:

Theorem (Dirichlet, 1837)

Let $a, d \in \mathbb{N}$ be coprime. Then the infinite arithmetic progression $a + d\mathbb{N} = \{a, a + d, \dots, a + nd, \dots\}$ contains infinitely many primes.

However, the exact transposition of Dirichlet's theorem fails to hold: there are no infinite arithmetic progressions consisting of prime numbers!

Green-Tao and Dirichlet

Observe that reversing the statement produces a well-known theorem, dating back to the first half of the 19th century:

Theorem (Dirichlet, 1837)

Let $a, d \in \mathbb{N}$ be coprime. Then the infinite arithmetic progression $a + d\mathbb{N} = \{a, a+d, \ldots, a+nd, \ldots\}$ contains infinitely many primes.

However, the exact transposition of Dirichlet's theorem fails to hold: there are no infinite arithmetic progressions consisting of prime numbers!

As of 2020, the longest known arithmetic progression of primes has 27 terms, and it starts with

$$224584605939537911\,.$$

A primer in ergodic theory

Definition

A *probability measure preserving system* is a quadruple $(X, \mathcal{A}, \mu, T)$, where $(X, \mathcal{A}, \mu)$ is a probability measure space and $T\colon X \to X$ is a measurable map such that $T_*(\mu) = \mu$, that is

$$\mu(T^{-1}(A)) = \mu(A) \text{ for any } A \in \mathcal{A}.$$

A primer in ergodic theory

Definition

A *probability measure preserving system* is a quadruple $(X, \mathcal{A}, \mu, T)$, where $(X, \mathcal{A}, \mu)$ is a probability measure space and $T\colon X \to X$ is a measurable map such that $T_*(\mu) = \mu$, that is

$$\mu(T^{-1}(A)) = \mu(A) \text{ for any } A \in \mathcal{A}.$$

We shall only be interested in the case where X is endowed with a compact, metrizable topology, $\mathcal{A}$ is the Borel σ-algebra associated to it, and T is a continuous map.

A primer in ergodic theory

Definition

A *probability measure preserving system* is a quadruple $(X, \mathcal{A}, \mu, T)$, where $(X, \mathcal{A}, \mu)$ is a probability measure space and $T\colon X \to X$ is a measurable map such that $T_*(\mu) = \mu$, that is

$$\mu(T^{-1}(A)) = \mu(A) \text{ for any } A \in \mathcal{A}.$$

We shall only be interested in the case where X is endowed with a compact, metrizable topology, $\mathcal{A}$ is the Borel σ-algebra associated to it, and T is a continuous map.

Example: $X = \{0,1\}^{\mathbb{N}}$ with the topology generated by *cylinders*

$$C^{i_1,\dots,i_r} = \{x = (x_n)_n \in X : x_1 = i_1, \dots, x_r = i_r\},\ i_1, \dots, i_r \in \{0,1\},$$

$$T\colon X \to X \text{ defined by } T((x_n)_n) = (x_{n+1})_n,\ \mu = \tfrac{1}{2}(\delta_0 + \delta_1)^{\otimes \mathbb{N}}.$$

Poincaré recurrence theorem

The main focus of ergodic theory lies in describing qualitatively the long-term behaviour of *typical* orbits $\{T^n(x) : n \in \mathbb{N}\}$.

Poincaré recurrence theorem

The main focus of ergodic theory lies in describing qualitatively the long-term behaviour of *typical* orbits $\{T^n(x) : n \in \mathbb{N}\}$.
Finiteness of the measure entails remarkable recurrence phenomena:

Theorem (Poincaré recurrence)

Let $(X, \mathcal{A}, \mu, T)$ be a probability measure preserving system, $E \subset X$ a measurable subset. Then μ-almost every $x \in E$ returns to E infinitely often, that is there exists $E' \subset E$ such that $\mu(E \setminus E') = 0$ and

$$\{n \in \mathbb{N} : T^n(x) \in E\} \text{ is infinite for all } x \in E'.$$

Poincaré recurrence theorem

The main focus of ergodic theory lies in describing qualitatively the long-term behaviour of *typical* orbits $\{T^n(x) : n \in \mathbb{N}\}$.
Finiteness of the measure entails remarkable recurrence phenomena:

Theorem (Poincaré recurrence)

Let $(X, \mathcal{A}, \mu, T)$ be a probability measure preserving system, $E \subset X$ a measurable subset. Then μ-almost every $x \in E$ returns to E infinitely often, that is there exists $E' \subset E$ such that $\mu(E \setminus E') = 0$ and

$$\{n \in \mathbb{N} : T^n(x) \in E\} \text{ is infinite for all } x \in E'.$$

A short proof: define

$$F = \limsup_{n\to\infty} T^{-n}(E), \; F_n = \bigcup_{k \geq n} T^{-k}(E) \text{ for all } n \geq 0.$$

Then the assumptions on μ imply $\mu(F_0 \setminus F) = 0$. Hence

$$\mu(E) = \mu(E \cap F) + \mu(E \setminus F) \leq \mu(E \cap F) + \mu(F_0 \setminus F) = \mu(E \cap F).$$

More structure in recurrence

Does $\{n \in \mathbb{N} : T^n(x) \in E\}$ enjoy some additional arithmetic structure?

More structure in recurrence

Does $\{n \in \mathbb{N} : T^n(x) \in E\}$ enjoy some additional arithmetic structure?

Theorem (Furstenberg's multiple recurrence theorem, 1977)

Let $(X, \mathcal{A}, \mu, T)$ be a probability measure preserving system, $E \subset X$ a measurable subset with $\mu(E) > 0$. For every $k \in \mathbb{N}$ there exists $n \in \mathbb{N}$ such that

$$\mu(E \cap T^{-n}(E) \cap T^{-2n}(E) \cap \cdots \cap T^{-(k-1)n}(E)) > 0 \, .$$

More structure in recurrence

Does $\{n \in \mathbb{N} : T^n(x) \in E\}$ enjoy some additional arithmetic structure?

Theorem (Furstenberg's multiple recurrence theorem, 1977)

Let $(X, \mathcal{A}, \mu, T)$ be a probability measure preserving system, $E \subset X$ a measurable subset with $\mu(E) > 0$. For every $k \in \mathbb{N}$ there exists $n \in \mathbb{N}$ such that

$$\mu(E \cap T^{-n}(E) \cap T^{-2n}(E) \cap \cdots \cap T^{-(k-1)n}(E)) > 0 \,.$$

A note on the proof: if the sets $E, T^{-n}(E), \ldots, T^{-(k-1)n}(E)$ *decorrelate* as n tends to infinity, the assertion is intuitively clear (and obvious if they become independent). This is the case for *weak mixing* systems.

More structure in recurrence

Does $\{n \in \mathbb{N} : T^n(x) \in E\}$ enjoy some additional arithmetic structure?

Theorem (Furstenberg's multiple recurrence theorem, 1977)

Let $(X, \mathcal{A}, \mu, T)$ be a probability measure preserving system, $E \subset X$ a measurable subset with $\mu(E) > 0$. For every $k \in \mathbb{N}$ there exists $n \in \mathbb{N}$ such that

$$\mu(E \cap T^{-n}(E) \cap T^{-2n}(E) \cap \cdots \cap T^{-(k-1)n}(E)) > 0 \, .$$

A note on the proof: if the sets $E, T^{-n}(E), \ldots, T^{-(k-1)n}(E)$ *decorrelate* as n tends to infinity, the assertion is intuitively clear (and obvious if they become independent). This is the case for *weak mixing* systems. The theorem is also obvious for a *periodic* dynamics ($T^m = T$ for some $m > 1$, for instance $T : x \mapsto x + \alpha$ on $\mathbb{R}/\mathbb{Z}$ with $\alpha \in \mathbb{Q}$).

More structure in recurrence

Does $\{n \in \mathbb{N} : T^n(x) \in E\}$ enjoy some additional arithmetic structure?

Theorem (Furstenberg's multiple recurrence theorem, 1977)

Let $(X, \mathcal{A}, \mu, T)$ be a probability measure preserving system, $E \subset X$ a measurable subset with $\mu(E) > 0$. For every $k \in \mathbb{N}$ there exists $n \in \mathbb{N}$ such that

$$\mu(E \cap T^{-n}(E) \cap T^{-2n}(E) \cap \cdots \cap T^{-(k-1)n}(E)) > 0 .$$

A note on the proof: if the sets $E, T^{-n}(E), \ldots, T^{-(k-1)n}(E)$ *decorrelate* as n tends to infinity, the assertion is intuitively clear (and obvious if they become independent). This is the case for *weak mixing* systems. The theorem is also obvious for a *periodic* dynamics ($T^m = T$ for some $m > 1$, for instance $T : x \mapsto x + \alpha$ on $\mathbb{R}/\mathbb{Z}$ with $\alpha \in \mathbb{Q}$).
The general case follows from Furstenberg's *structure theorem*, which allows to decompose an arbitrary system into several "layers" consisting either of weak mixing or of *almost periodic* systems.

Proof of Szemerédi's theorem

Assume $A \subset \mathbb{N}$ satisfies $\overline{d}(A) > 0$, choose a subsequence $(N_j)_{j \geq 1}$ so that

$$\lim_{j \to \infty} \frac{|A \cap [1, N_j]|}{N_j} > 0 \, .$$

Proof of Szemerédi's theorem

Assume $A \subset \mathbb{N}$ satisfies $\overline{d}(A) > 0$, choose a subsequence $(N_j)_{j\geq 1}$ so that

$$\lim_{j\to\infty} \frac{|A \cap [1, N_j]|}{N_j} > 0 \, .$$

Define $\omega = (\omega_n)_n \in X = \{0,1\}^{\mathbb{N}}$ by $\omega_n = \mathbb{1}_A(n)$ for every $n \geq 1$. Let $E = C^1 = \{x = (x_n)_n \in X : x_1 = 1\}$.

Proof of Szemerédi's theorem

Assume $A \subset \mathbb{N}$ satisfies $\overline{d}(A) > 0$, choose a subsequence $(N_j)_{j\geq 1}$ so that

$$\lim_{j\to\infty} \frac{|A \cap [1, N_j]|}{N_j} > 0 \, .$$

Define $\omega = (\omega_n)_n \in X = \{0,1\}^{\mathbb{N}}$ by $\omega_n = \mathbb{1}_A(n)$ for every $n \geq 1$. Let $E = C^1 = \{x = (x_n)_n \in X : x_1 = 1\}$.
Pretend for a moment that the probability measure

$$\mu_{N_j} = \frac{1}{N_j} \sum_{i=0}^{N_j - 1} \delta_{T^i(\omega)}$$

is T-invariant and verifies $\mu_{N_j}(E) > 0$ for some N_j.

Proof of Szemerédi's theorem

Applying Furstenberg's multiple recurrence, and unravelling it in our context, we get: for every $k \in N$ there exists $n \in \mathbb{N}$, $i \in \{0, \ldots, N_j - 1\}$ such that $T^i(\omega) \in E \cap T^{-n}(E) \cap \cdots \cap T^{-(k-1)n}(E)$, that is, $\omega_{i+1} = \omega_{i+1+n} = \cdots = \omega_{i+1+(k-1)n} = 1$, or in other words

$$\{i+1, i+1+n, \ldots, i+1+(k-1)n\} \subset A\,.$$

Proof of Szemerédi's theorem

Applying Furstenberg's multiple recurrence, and unravelling it in our context, we get: for every $k \in N$ there exists $n \in \mathbb{N}$, $i \in \{0, \ldots, N_j - 1\}$ such that $T^i(\omega) \in E \cap T^{-n}(E) \cap \cdots \cap T^{-(k-1)n}(E)$, that is, $\omega_{i+1} = \omega_{i+1+n} = \cdots = \omega_{i+1+(k-1)n} = 1$, or in other words

$$\{i+1, i+1+n, \ldots, i+1+(k-1)n\} \subset A\,.$$

However, there is no reason for μ_{N_j} to be T-invariant nor, *a priori*, to give E positive measure.

Proof of Szemerédi's theorem

Applying Furstenberg's multiple recurrence, and unravelling it in our context, we get: for every $k \in N$ there exists $n \in \mathbb{N}$, $i \in \{0, \dots, N_j - 1\}$ such that $T^i(\omega) \in E \cap T^{-n}(E) \cap \cdots \cap T^{-(k-1)n}(E)$, that is, $\omega_{i+1} = \omega_{i+1+n} = \cdots = \omega_{i+1+(k-1)n} = 1$, or in other words

$$\{i+1, i+1+n, \dots, i+1+(k-1)n\} \subset A\,.$$

However, there is no reason for μ_{N_j} to be T-invariant nor, *a priori*, to give E positive measure.

Fortunately, functional analysis rescues the argument: the set of T-invariant Borel probability measures on X is sequentially compact for the weak* topology, being a closed bounded subset of the topological dual of the separable Banach space $C(X) = \{f : X \to \mathbb{C} \text{ continuous}\}$.

Proof of Szemerédi's theorem

If $\mu = \lim_{l \to \infty} \mu_{N_{j_l}}$ is a weak* limit of $(\mu_{N_j})_j$, then

Proof of Szemerédi's theorem

If $\mu = \lim_{l \to \infty} \mu_{N_{j_l}}$ is a weak* limit of $(\mu_{N_j})_j$, then

- μ is T-invariant: indeed, if $f : X \to \mathbb{C}$ is continuous,

$$\int_X f \, dT_* \mu = \int_X f \circ T \, d\mu = \lim_{l \to \infty} \frac{1}{N_{j_l}} \sum_{i=0}^{N_{j_l}-1} (f \circ T)(T^i(\omega))$$

$$= \lim_{l \to \infty} \frac{1}{N_{j_l}} \sum_{i=1}^{N_{j_l}} f(T^i(\omega)) = \lim_{l \to \infty} \frac{1}{N_{j_l}} \sum_{i=0}^{N_{j_l}-1} f(T^i(\omega)) = \int_X f \, d\mu \, ;$$

Proof of Szemerédi's theorem

If $\mu = \lim_{l\to\infty} \mu_{N_{j_l}}$ is a weak* limit of $(\mu_{N_j})_j$, then

- μ is T-invariant: indeed, if $f\colon X \to \mathbb{C}$ is continuous,

$$\int_X f\, dT_*\mu = \int_X f \circ T\, d\mu = \lim_{l\to\infty} \frac{1}{N_{j_l}} \sum_{i=0}^{N_{j_l}-1} (f \circ T)(T^i(\omega))$$

$$= \lim_{l\to\infty} \frac{1}{N_{j_l}} \sum_{i=1}^{N_{j_l}} f(T^i(\omega)) = \lim_{l\to\infty} \frac{1}{N_{j_l}} \sum_{i=0}^{N_{j_l}-1} f(T^i(\omega)) = \int_X f\, d\mu\ ;$$

- E has positive μ-measure:

$$\mu(E) = \lim_{l\to\infty} \frac{1}{N_{j_l}} \sum_{i=0}^{N_{j_l}-1} \delta_{T^i(\omega)}(E) = \lim_{l\to\infty} \frac{|A \cap [1, N_{j_l}]|}{N_{j_l}} > 0\ .$$

A "finitary" version of Szemerédi's theorem

Rigorously, we thus apply Furstenberg's multiple recurrence to μ, and conclude as before using that

$$\mu(E \cap \cdots \cap T^{-(k-1)n}(E)) > 0 \implies \mu_{N_j}(E \cap \cdots \cap T^{-(k-1)n}(E)) > 0$$

for all j sufficiently large.

A "finitary" version of Szemerédi's theorem

Rigorously, we thus apply Furstenberg's multiple recurrence to μ, and conclude as before using that

$$\mu(E \cap \cdots \cap T^{-(k-1)n}(E)) > 0 \implies \mu_{N_j}(E \cap \cdots \cap T^{-(k-1)n}(E)) > 0$$

for all j sufficiently large.

Furstenberg's proof relies crucially on a compactness argument, thus failing to provide any quantitative refinement.

A "finitary" version of Szemerédi's theorem

Rigorously, we thus apply Furstenberg's multiple recurrence to μ, and conclude as before using that

$$\mu(E \cap \cdots \cap T^{-(k-1)n}(E)) > 0 \implies \mu_{N_j}(E \cap \cdots \cap T^{-(k-1)n}(E)) > 0$$

for all j sufficiently large.

Furstenberg's proof relies crucially on a compactness argument, thus failing to provide any quantitative refinement. However, elementary combinatorial reasoning allows to deduce the following statement:

Theorem (Quantitative Szemerédi)

Let $k \in \mathbb{N}, 0 < \delta \leq 1$. There is $N_0 = N_0(k, \delta)$ such that, for any $N \geq N_0$, any set $A \subset \{1, \ldots, N\}$ with $|A| \geq \delta N$ contains a k-term arithmetic progression.

Further quantitative refinements

Heuristics derived by choosing A randomly among density-δ subsets of $\{1, \ldots, N\}$ suggests that such an A should contain $\sim \delta^k N^2$ k-term arithmetic progressions.

Further quantitative refinements

Heuristics derived by choosing A randomly among density-δ subsets of $\{1, \ldots, N\}$ suggests that such an A should contain $\sim \delta^k N^2$ k-term arithmetic progressions.

The intuition is confirmed by the following, apparently stronger result:

Further quantitative refinements

Heuristics derived by choosing A randomly among density-δ subsets of $\{1, \ldots, N\}$ suggests that such an A should contain $\sim \delta^k N^2$ k-term arithmetic progressions.
The intuition is confirmed by the following, apparently stronger result:

Notation: for every $N \in \mathbb{N}$, we denote $\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}$. If A is a finite set, $f : A \to \mathbb{C}$ is a function, we denote

$$\mathbb{E}(f) = \mathbb{E}(f(x)|x \in A) = \frac{1}{|A|} \sum_{a \in A} f(a) .$$

Further quantitative refinements

Heuristics derived by choosing A randomly among density-δ subsets of $\{1, \ldots, N\}$ suggests that such an A should contain $\sim \delta^k N^2$ k-term arithmetic progressions.
The intuition is confirmed by the following, apparently stronger result:

Notation: for every $N \in \mathbb{N}$, we denote $\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}$. If A is a finite set, $f : A \to \mathbb{C}$ is a function, we denote

$$\mathbb{E}(f) = \mathbb{E}(f(x)|x \in A) = \frac{1}{|A|} \sum_{a \in A} f(a) \, .$$

Proposition

Fix $k \in \mathbb{N}$, $0 < \delta \leq 1$. There exist $N_0 = N_0(k, \delta) \in \mathbb{N}, c_{k,\delta} > 0$ such that, if $N \geq N_0$ and $f : \mathbb{Z}_N \to \mathbb{R}$ satisfies both $0 \leq f(n) \leq 1$ for every $n \in \mathbb{Z}_N$ and $\mathbb{E}(f) \geq \delta$, then

$$\mathbb{E}(f(n)f(n+r)\cdots f(n+(k-1)r) \mid (n, r) \in \mathbb{Z}_N \times \mathbb{Z}_N) \geq c_{k,\delta} \, .$$

Von Neumann ergodic theorem

There is an abstract ergodic-theoretic analogue to this proposition, which greatly inspired Green-Tao's approach for dealing with prime numbers.

Von Neumann ergodic theorem

There is an abstract ergodic-theoretic analogue to this proposition, which greatly inspired Green-Tao's approach for dealing with prime numbers. It is a generalization of the celebrated:

Theorem (von Neumann ergodic theorem)

Let $(X, \mathcal{A}, \mu, T)$ be a probability measure preserving system, $f \in L^2(X, \mu)$. Then, there exists $\tilde{f} \in L^2(X, \mu)$ satisfying

- $\tilde{f} \circ T = \tilde{f}$ *(in $L^2(X, \mu)$)*
- $\mathbb{E}_\mu(\tilde{f}) = \mathbb{E}_\mu(f)$

such that the sequence

$$\frac{1}{N} \sum_{n=0}^{N-1} f \circ T^n \, , \; N \in \mathbb{N},$$

converges towards $\tilde{f}$ in the L^2-norm.

A generalized von Neumann ergodic theorem

The generalization reads as follows:

Theorem (Host-Kra, 2005)

Under the assumptions of von Neumann's theorem, supposing in addition that $f \in L^\infty(X, \mu)$, it holds that, for every $k \in \mathbb{N}$, there exists $f_k \in L^2(X, \mu)$ such that the sequence

$$\frac{1}{N} \sum_{n=0}^{N-1} f \circ T^n \cdot f \circ T^{2n} \cdots f \circ T^{kn}, \quad N \in \mathbb{N},$$

converges towards f_k in the L^2-norm.

A generalized von Neumann ergodic theorem

The generalization reads as follows:

Theorem (Host-Kra, 2005)

Under the assumptions of von Neumann's theorem, supposing in addition that $f \in L^\infty(X, \mu)$, it holds that, for every $k \in \mathbb{N}$, there exists $f_k \in L^2(X, \mu)$ such that the sequence

$$\frac{1}{N} \sum_{n=0}^{N-1} f \circ T^n \cdot f \circ T^{2n} \cdots f \circ T^{kn}, \quad N \in \mathbb{N},$$

converges towards f_k in the L^2-norm.

The result is influenced by the stronger version of Furstenberg's recurrence theorem: if $\mu(A) > 0$, then

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(A \cap T^{-n}(A) \cap \cdots \cap T^{-(k-1)n}(A)) > 0 \text{ for every } k \in \mathbb{N}.$$

A generalized von Neumann ergodic theorem

Since convergence in the theorem of Host and Kra occurs with respect to the L^2-norm, it entails convergence of the expectations:

$$\mathbb{E}_\mu\left(\frac{1}{N}\sum_{n=0}^{N-1} f\circ T^n \cdot f\circ T^{2n}\cdots f\circ T^{kn}\right) \xrightarrow{N\to\infty} \mathbb{E}_\mu(f_k)\,.$$

A generalized von Neumann ergodic theorem

Since convergence in the theorem of Host and Kra occurs with respect to the L^2-norm, it entails convergence of the expectations:

$$\mathbb{E}_\mu\left(\frac{1}{N}\sum_{n=0}^{N-1} f\circ T^n\cdot f\circ T^{2n}\cdots f\circ T^{kn}\right)\xrightarrow{N\to\infty}\mathbb{E}_\mu(f_k)\,.$$

The analogy with the proposition is thus explained in that

$$\mathbb{E}(f(n)f(n+r)\cdots f(n+(k-1)r)\mid (n,r)\in\mathbb{Z}_N\times\mathbb{Z}_N)$$
$$=\mathbb{E}\left(\frac{1}{N}\sum_{r=0}^{N-1} f(n)\cdot f(T^r(n))\cdots f(T^{(k-1)r}(n))\,\middle|\, n\in\mathbb{Z}_N\right)$$

where $T\colon \mathbb{Z}_N\ni x\mapsto x+1\in\mathbb{Z}_N$ preserves the uniform measure on $\mathbb{Z}_N$.

Adapting to primes: pseudorandomness

If, in the proposition, we could take f to be the restriction to $\mathbb{Z}_N$ of a function supported on the primes (that is, such that $f(n) = 0$ for every composite $n \in \mathbb{N}$), then Green-Tao's theorem would follow readily, or at least up to some "wraparound" issues in $\mathbb{Z}_N$ (easy to circumvent).

Adapting to primes: pseudorandomness

If, in the proposition, we could take f to be the restriction to $\mathbb{Z}_N$ of a function supported on the primes (that is, such that $f(n) = 0$ for every composite $n \in \mathbb{N}$), then Green-Tao's theorem would follow readily, or at least up to some "wraparound" issues in $\mathbb{Z}_N$ (easy to circumvent).

However, this is nothing but wishful thinking, as $f(p)$ should grow at least logarithmically in p for $\mathbb{E}(f) \geq \delta$ to be satisfied asymptotically!

Adapting to primes: pseudorandomness

If, in the proposition, we could take f to be the restriction to $\mathbb{Z}_N$ of a function supported on the primes (that is, such that $f(n) = 0$ for every composite $n \in \mathbb{N}$), then Green-Tao's theorem would follow readily, or at least up to some "wraparound" issues in $\mathbb{Z}_N$ (easy to circumvent).

However, this is nothing but wishful thinking, as $f(p)$ should grow at least logarithmically in p for $\mathbb{E}(f) \geq \delta$ to be satisfied asymptotically!

What if we could loosen the upper bound $f(n) \leq 1$ in the assumptions of the proposition, and retain a similar conclusion?

Adapting to primes: pseudorandomness

If, in the proposition, we could take f to be the restriction to $\mathbb{Z}_N$ of a function supported on the primes (that is, such that $f(n) = 0$ for every composite $n \in \mathbb{N}$), then Green-Tao's theorem would follow readily, or at least up to some "wraparound" issues in $\mathbb{Z}_N$ (easy to circumvent).

However, this is nothing but wishful thinking, as $f(p)$ should grow at least logarithmically in p for $\mathbb{E}(f) \geq \delta$ to be satisfied asymptotically!

What if we could loosen the upper bound $f(n) \leq 1$ in the assumptions of the proposition, and retain a similar conclusion?

This is the major insight of Green and Tao: Szemerédi's theorem should hold not just for positive-proportion subsets of $\mathbb{N}$, but also of sufficiently "random" (from an additive perspective) subsets of $\mathbb{N}$.

Adapting to primes: pseudorandomness

Their major contribution lies in replacing the constant function $\nu \equiv 1$ by a *pseudorandom measure* $(\nu_N)_{N \in \mathbb{N}}$, where $\nu_N \colon \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ is called a *measure* if it satisfies $\mathbb{E}(\nu_N) = 1 + o(1)$, with $o(1) \to 0$ as $N \to \infty$.

Adapting to primes: pseudorandomness

Their major contribution lies in replacing the constant function $\nu \equiv 1$ by a *pseudorandom measure* $(\nu_N)_{N \in \mathbb{N}}$, where $\nu_N : \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ is called a *measure* if it satisfies $\mathbb{E}(\nu_N) = 1 + o(1)$, with $o(1) \to 0$ as $N \to \infty$.

The formal notion of pseudorandomness is rather involved, and draws upon previous work of T. Gowers (1998), who devised yet another proof of Szemerédi's theorem via Fourier-analytic methods.

Adapting to primes: pseudorandomness

Their major contribution lies in replacing the constant function $\nu \equiv 1$ by a *pseudorandom measure* $(\nu_N)_{N\in\mathbb{N}}$, where $\nu_N \colon \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ is called a *measure* if it satisfies $\mathbb{E}(\nu_N) = 1 + o(1)$, with $o(1) \to 0$ as $N \to \infty$.

The formal notion of pseudorandomness is rather involved, and draws upon previous work of T. Gowers (1998), who devised yet another proof of Szemerédi's theorem via Fourier-analytic methods.

Vaguely it amounts to requiring that, for any collection of "$\mathbb{Q}$-linearly independent" affine forms $\psi_1, \ldots, \psi_m \colon \mathbb{Z}_N^t \to \mathbb{Z}_N$, where m and t are small integral parameters, the *random variables* $\nu_N(\psi_1(\boldsymbol{x})), \ldots, \nu_N(\psi_m(\boldsymbol{x})), \boldsymbol{x} \in \mathbb{Z}_N^t$ are independent.

Adapting to primes: pseudorandomness

Their major contribution lies in replacing the constant function $\nu \equiv 1$ by a *pseudorandom measure* $(\nu_N)_{N \in \mathbb{N}}$, where $\nu_N \colon \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ is called a *measure* if it satisfies $\mathbb{E}(\nu_N) = 1 + o(1)$, with $o(1) \to 0$ as $N \to \infty$.

The formal notion of pseudorandomness is rather involved, and draws upon previous work of T. Gowers (1998), who devised yet another proof of Szemerédi's theorem via Fourier-analytic methods.

Vaguely it amounts to requiring that, for any collection of "$\mathbb{Q}$-linearly independent" affine forms $\psi_1, \ldots, \psi_m \colon \mathbb{Z}_N^t \to \mathbb{Z}_N$, where m and t are small integral parameters, the *random variables* $\nu_N(\psi_1(\boldsymbol{x})), \ldots, \nu_N(\psi_m(\boldsymbol{x})), \boldsymbol{x} \in \mathbb{Z}_N^t$ are independent.

The driving principle is that there should be functions ν_N supported on (almost-) primes enjoying this property, namely the events "$\psi_j(\boldsymbol{x})$ is almost prime" are independent of each other as j varies.
That's what is meant by random additive behaviour of primes.

Adapting to primes: pseudorandomness

Quantitative Szemerédi holds in the context of pseudorandom measures:

Theorem (Green, Tao)

Fix $k \in \mathbb{N}$, $0 < \delta \leq 1$. *There exist* $N_0 = N_0(k, \delta) \in \mathbb{N}, c'_{k,\delta} > 0$ *such that, for any* k*-pseudorandom measure* $(\nu_N)_N$, *any* $N \geq N_0$ *and any* $f : \mathbb{Z}_N \to \mathbb{R}$ *satisfying both* $0 \leq f(n) \leq \nu_N(n)$ *for every* $n \in \mathbb{Z}_N$ *and* $\mathbb{E}(f) \geq \delta$, *we have*

$$\mathbb{E}(f(n)f(n+r)\cdots f(n+(k-1)r) \mid (n,r) \in \mathbb{Z}_N \times \mathbb{Z}_N) \geq c'_{k,\delta} \,.$$

Adapting to primes: pseudorandomness

Quantitative Szemerédi holds in the context of pseudorandom measures:

Theorem (Green, Tao)

Fix $k \in \mathbb{N}$, $0 < \delta \leq 1$. There exist $N_0 = N_0(k, \delta) \in \mathbb{N}, c'_{k,\delta} > 0$ such that, for any k-pseudorandom measure $(\nu_N)_N$, any $N \geq N_0$ and any $f : \mathbb{Z}_N \to \mathbb{R}$ satisfying both $0 \leq f(n) \leq \nu_N(n)$ for every $n \in \mathbb{Z}_N$ and $\mathbb{E}(f) \geq \delta$, we have

$$\mathbb{E}(f(n)f(n+r)\cdots f(n+(k-1)r) \mid (n,r) \in \mathbb{Z}_N \times \mathbb{Z}_N) \geq c'_{k,\delta} \,.$$

Assuming this, existence of a k-term arithmetic progression of primes is inferred taking as f the restriction to $\mathbb{Z}_N$ of (a modified version of) the von Mangoldt function

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^j \text{ for some prime } p \text{ and } m \in \mathbb{N} \\ 0 & \text{otherwise .} \end{cases}$$

Concluding remarks

- The proof of Green-Tao's main quantitative theorem closely resembles, in the overarching strategy, Furstenberg's proof of the structure theorem for measure preserving systems (decomposition into a tower of compact extensions of weak-mixing systems). The function f is split into two components: a "uniform" part (in the sense of Gowers uniformity norms), whose contribution to the expectation is controlled via a von Neumann-type estimate, and an "anti-uniform" part, which is bounded by a constant and thus taken care of by quantitative Szemerédi.

Concluding remarks

- The proof of Green-Tao's main quantitative theorem closely resembles, in the overarching strategy, Furstenberg's proof of the structure theorem for measure preserving systems (decomposition into a tower of compact extensions of weak-mixing systems). The function f is split into two components: a "uniform" part (in the sense of Gowers uniformity norms), whose contribution to the expectation is controlled via a von Neumann-type estimate, and an "anti-uniform" part, which is bounded by a constant and thus taken care of by quantitative Szemerédi.
- The foremost limitation of an ergodic-theoretic approach to number theory is that proofs tend to be non-quantitative and non-effective. However, the pursuit of quantitative analogues of classical ergodic theorems may lead, as in Green-Tao's example, to the disclosure of surprising qualitative results.