The primes contain arbitrarily long arithmetic progressions

Antonios Manoussos

Universität Bielefeld
Fakultät für Mathematik

Bielefeld 2010

The Green-Tao theorem

Theorem

The prime numbers contain infinitely many arithmetic progressions of length k for all $k \geq 1$.

The Green-Tao theorem

Theorem

The prime numbers contain infinitely many arithmetic progressions of length k for all $k \geq 1$.

Related problems

- (Erdös conjecture) Let A be a subset of $\mathbb{N}$. If $\sum_{a \in A} \frac{1}{a} = \infty$ then A contains arithmetic progressions of any length.

The Green-Tao theorem

Theorem

The prime numbers contain infinitely many arithmetic progressions of length k for all $k \geq 1$.

Related problems

- (Erdös conjecture) Let A be a subset of $\mathbb{N}$. If $\displaystyle\sum_{a \in A} \frac{1}{a} = \infty$ then A contains arithmetic progressions of any length.
- (Szemerédi's theorem) Let A be a subset of the positive integers with positive upper density; i.e.

$$\limsup_{N \to \infty} \frac{|A \cap [1, N]|}{N} > 0.$$

Then for each $k \geq 1$, the set A contains at least one arithmetic progression of length k.

The Green-Tao theorem

Theorem (The Green-Tao theorem is Szemerédi's theorem in the primes)

Let A be any subset of the prime numbers of positive relative upper density; i.e.

$$\limsup_{N\to\infty} \frac{|A \cap [1, N]|}{\pi(N)} > 0,$$

where $\pi(N)$ denotes the number of primes less than or equal to N. Then A contains infinitely many arithmetic progressions of length k for all k.

The Green-Tao theorem

History

- Van der Corput (1939) proved that the primes contain infinitely many arithmetic progressions of length 3.

The Green-Tao theorem

History

- Van der Corput (1939) proved that the primes contain infinitely many arithmetic progressions of length 3.
- Heath-Brown (1981) showed that there are infinitely many arithmetic progressions of length 4 consisting of three primes and a semiprime (: a product of at most two primes).

Basic definitions and notation

Definitions

- Let $N \in \mathbb{N}$ be a prime. We denote $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$.

Basic definitions and notation

Definitions

- Let $N \in \mathbb{N}$ be a prime. We denote $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$.
- We will write $o(1)$ for a quantity that tends to zero as $N \to \infty$ and we write $O(1)$ for a bounded quantity.

Basic definitions and notation

Definitions

- Let $N \in \mathbb{N}$ be a prime. We denote $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$.
- We will write $o(1)$ for a quantity that tends to zero as $N \to \infty$ and we write $O(1)$ for a bounded quantity.
- (Expected value) Let $n \geq 1$ and $f : A \to \mathbb{R}$ where $A \subseteq \mathbb{Z}_N^n$ with $A \neq \emptyset$. Then we define

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

If f is defined on all $\mathbb{Z}_N^n$ we write $\mathbb{E}(f) = \mathbb{E}(f(x)|x \in \mathbb{Z}_N^n)$.

Basic definitions and notation

Definitions

- Let $N \in \mathbb{N}$ be a prime. We denote $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$.
- We will write $o(1)$ for a quantity that tends to zero as $N \to \infty$ and we write $O(1)$ for a bounded quantity.
- (Expected value) Let $n \geq 1$ and $f : A \to \mathbb{R}$ where $A \subseteq \mathbb{Z}_N^n$ with $A \neq \emptyset$. Then we define

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

  If f is defined on all $\mathbb{Z}_N^n$ we write $\mathbb{E}(f) = \mathbb{E}(f(x)|x \in \mathbb{Z}_N^n)$.
- A function $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ is called a measure if $\mathbb{E}(\nu) = 1 + o(1)$.

Pseudorandom measures

Definition (Linear forms condition)

Let $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ be a measure. Let m_0, t_0 and L_0 be small positive integers. Then we say that ν satisfies the (m_0, t_0, L_0)-linear forms condition if the following holds. Let $m \leq m_0$ and $t \leq t_0$ and let (L_{ij}), $1 \leq i \leq m$, $1 \leq j \leq t$ are arbitrary rational numbers with numerator and denominator at most L_0 in absolute value. Suppose that as i ranges over $1, \ldots, m$ the t-tuples $(L_{ij})_{1 \leq j \leq t} \in \mathbb{Q}^t$ are nonzero, and no t-tuple is a rational multiple of any other. Let $b_1, \ldots, b_m \in \mathbb{Z}_N$ and define for $1, \ldots, m$ linear forms $\psi_i : \mathbb{Z}_N^t \to \mathbb{Z}_N$ by $\psi_i(x) = \sum_{j=1}^t L_{ij} x_j + b_i$, where $x = (x_1, \ldots, x_t)$. Then

$$\mathbb{E}(\nu(\psi_1(x)) \ldots \nu(\psi_m(x)) | x \in \mathbb{Z}_N^t) = 1 + o_{L_0, m_0, t_0}(1).$$

Pseudorandom measures

Remarks

1. The rational numbers L_{ij} are consider as elements of $\mathbb{Z}_N$ assuming that N is a prime larger than L_0. E.g. a/b is the unique solution of the equation $bx \equiv a \pmod{N}$ in $\mathbb{Z}_N$ (unique since $|a|, |b| < N$).

Pseudorandom measures

Remarks

1. The rational numbers L_{ij} are consider as elements of $\mathbb{Z}_N$ assuming that N is a prime larger than L_0. E.g. a/b is the unique solution of the equation $bx \equiv a \pmod{N}$ in $\mathbb{Z}_N$ (unique since $|a|, |b| < N$).
2. The use of linear forms comes naturally. For instance, if we look for arithmetic progressions of length k in the primes we can define the linear forms $\psi_i : \mathbb{Z}^2 \to \mathbb{Z}$ with $\psi_i(n, r) = n + (i-1)r$, $n, r \in \mathbb{Z}$, $i = 1, \ldots, k$ and ask if there exist positive integers n and r such that $\psi_i(n, r)$ are simultaneously primes.

Pseudorandom measures

Definition (Correlation condition)

Let $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ be a measure and let m_0 be positive integer parameter. We say that ν satisfies the m_0-correlation condition if for every $1 < m \leq m_0$ there is $\tau = \tau_m : \mathbb{Z}_N \to \mathbb{R}^+$ such that $\mathbb{E}(\tau^q) = O_{m,q}(1)$ for all $q \geq 1$ and such that

$$\mathbb{E}(\nu(x+h_1)\ldots\nu(x+h_m)|x \in \mathbb{Z}_N^t) \leq \sum_{1\leq i<j\leq m} \tau(h_i - h_j)$$

for all $h_1, \ldots, h_m \in \mathbb{Z}_N$ (not necessarily distinct).

Pseudorandom measures

Definition (Correlation condition)

Let $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ be a measure and let m_0 be positive integer parameter. We say that ν satisfies the m_0-correlation condition if for every $1 < m \leq m_0$ there is $\tau = \tau_m : \mathbb{Z}_N \to \mathbb{R}^+$ such that $\mathbb{E}(\tau^q) = O_{m,q}(1)$ for all $q \geq 1$ and such that

$$\mathbb{E}(\nu(x+h_1)\ldots\nu(x+h_m)|x \in \mathbb{Z}_N^t) \leq \sum_{1\leq i<j\leq m} \tau(h_i - h_j)$$

for all $h_1, \ldots, h_m \in \mathbb{Z}_N$ (not necessarily distinct).

Remark

The correlation condition arises in Goldston and Yildirim's work and is used for specific estimates applied to the prime numbers.

Pseudorandom measures

Definition (Pseudorandom measure)

A measure $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ is called k-pseudorandom if it satisfies the $(k2^{k-1}, 3k-4, k)$-linear forms condition and also the 2^{k-1}-correlation condition.

Pseudorandom measures

Definition (Pseudorandom measure)

A measure $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ is called k-pseudorandom if it satisfies the $(k2^{k-1}, 3k-4, k)$-linear forms condition and also the 2^{k-1}-correlation condition.

Examples

- The constant measure $\nu_{const} : \mathbb{Z}_N \to \mathbb{R}^+$ defined by $\nu_{const}(x) = 1$, for all $x \in \mathbb{Z}_N$ is k-pseudorandom for any k.

Pseudorandom measures

Definition (Pseudorandom measure)

A measure $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ is called k-pseudorandom if it satisfies the $(k2^{k-1}, 3k-4, k)$-linear forms condition and also the 2^{k-1}-correlation condition.

Examples

- The constant measure $\nu_{const} : \mathbb{Z}_N \to \mathbb{R}^+$ defined by $\nu_{const}(x) = 1$, for all $x \in \mathbb{Z}_N$ is k-pseudorandom for any k.
- Let ν be a k-pseudorandom measure, then

$$\nu_{1/2} := (\nu + \nu_{const})/2 = (\nu + 1)/2$$

is also a k-pseudorandom measure. So, pseudorandom measures are star shaped around 1.

The strategy of the proof

Theorem (Szemerédi's theorem again)

For any positive integer k and any real number $0 < \delta \leq 1$ there exists a positive integer $N_0(\delta, k)$ such that for every $N \geq N_0$, every set $A \subseteq \{1, \ldots, N\}$ of cardinality $|A| \geq \delta N$ contains at least one arithmetic progression of length k.

The strategy of the proof

Theorem (Szemerédi's theorem again)

For any positive integer k and any real number $0 < \delta \leq 1$ there exists a positive integer $N_0(\delta, k)$ such that for every $N \geq N_0$, every set $A \subseteq \{1, \ldots, N\}$ of cardinality $|A| \geq \delta N$ contains at least one arithmetic progression of length k.

Theorem (Reformulated Szemerédi's theorem, Varnavides (1959))

Let $0 < \delta \leq 1$ and $k \geq 1$ be fixed. Let $f : \mathbb{Z}_N \to \mathbb{R}^+$ such that $0 \leq f(x) \leq 1 = \nu_{const}(x)$, for all $x \in \mathbb{Z}_N$, and $\mathbb{E}(f) \geq \delta$. Then

$$\mathbb{E}(f(x)f(x+r)\cdots f(x+(k-1)r) | x, r \in \mathbb{Z}_N) \geq c(k, \delta)$$

for some constant $c(k, \delta) > 0$, not depending on f nor N.

The strategy of the proof

Theorem (Szemerédi's theorem relative to a pseudorandom measure)

let $0 < \delta \leq 1$ and $k \geq 3$ be fixed. Suppose that $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ is k-pseudorandom. Let $f : \mathbb{Z}_N \to \mathbb{R}^+$ such that $0 \leq f(x) \leq \nu(x)$, for all $x \in \mathbb{Z}_N$, and $\mathbb{E}(f) \geq \delta$. Then

$$\mathbb{E}(f(x)f(x+r)\cdots f(x+(k-1)r) \,|\, x, r \in \mathbb{Z}_N) \geq c(k,\delta) - o_{k,\delta}(1)$$

for some constant $c(k,\delta) > 0$, not depending on f nor N.

The strategy of the proof

Construction of f

Let W be the product of primes up to $\log\log N$. Now define the modified von Mangoldt function $\widetilde{\Lambda} : \mathbb{Z}^+ \to \mathbb{R}^+$ by

$$\widetilde{\Lambda}(n) := \begin{cases} \frac{\varphi(W)}{W} \log(Wn+1) & \text{when } Wn+1 \text{ is prime} \\ 0 & \text{otherwise,} \end{cases}$$

where φ denotes the Euler totient function (i.e. $\varphi(n)$ is the number of positive integers less than or equal to n that are coprime to n).

The strategy of the proof

Construction of f

Let W be the product of primes up to $\log\log N$. Now define the modified von Mangoldt function $\widetilde{\Lambda} : \mathbb{Z}^+ \to \mathbb{R}^+$ by

$$\widetilde{\Lambda}(n) := \begin{cases} \frac{\varphi(W)}{W} \log(Wn+1) & \text{when } Wn+1 \text{ is prime} \\ 0 & \text{otherwise,} \end{cases}$$

where φ denotes the Euler totient function (i.e. $\varphi(n)$ is the number of positive integers less than or equal to n that are coprime to n). Fix $\varepsilon_k > 0$. Now define $f : \mathbb{Z}_N \to \mathbb{R}^+$ by

$$f(n) := \begin{cases} \frac{1}{k2^{k+5}} \widetilde{\Lambda}(n) & \text{if } \varepsilon_k N \leq n \leq 2\varepsilon_k N \\ 0 & \text{otherwise,} \end{cases}$$

where we identify $\{0, \ldots, N-1\}$ with $\mathbb{Z}_N$ in the usual manner.

The strategy of the proof

Construction of ν

Let R be a parameter (it will be a small power of N). Define

$$\Lambda_R(n) := \sum_{d|n,\, d\leq R} \mu(d)\log(R/d) = \sum_{d|n} \mu(d)\log(R/d)_+$$

The strategy of the proof

Construction of ν

Let R be a parameter (it will be a small power of N). Define

$$\Lambda_R(n) := \sum_{d|n,\, d\leq R} \mu(d) \log(R/d) = \sum_{d|n} \mu(d) \log(R/d)_+$$

where $\log(x)_+ := \max(\log(x), 0)$ and μ is the Möbius function, i.e

- $\mu(n) = 1$, if n is a square-free positive integer with an even number of distinct prime factors.
- $\mu(n) = -1$, if n is a square-free positive integer with an odd number of distinct prime factors.
- $\mu(n) = 0$, if n is not square-free.

The strategy of the proof

Let $R := N^{k^{-1}2^{-k-4}}$, and let $\varepsilon_k = \frac{1}{2^k(k+4)!}$. We define a k-pseudorandom measure $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ by

$$\nu(n) := \begin{cases} \frac{\varphi(W)}{W} \frac{\Lambda_R(Wn+1)^2}{\log R} & \text{if } \varepsilon_k N \leq n \leq 2\varepsilon_k N \\ 1 & \text{otherwise.} \end{cases}$$

The strategy of the proof

Let $R := N^{k^{-1}2^{-k-4}}$, and let $\varepsilon_k = \frac{1}{2^k(k+4)!}$. We define a k-pseudorandom measure $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ by

$$\nu(n) := \begin{cases} \frac{\varphi(W)}{W} \frac{\Lambda_R(Wn+1)^2}{\log R} & \text{if } \varepsilon_k N \le n \le 2\varepsilon_k N \\ 1 & \text{otherwise.} \end{cases}$$

Proposition

Let $\varepsilon_k = \frac{1}{2^k(k+4)!}$ and let N a sufficiently large prime number. Then $0 \le f(x) \le \nu(x)$, for all $x \in \mathbb{Z}_N$.

The proof of the Green-Tao theorem

The main ingredient of the proof of the Green-Tao theorem is Szemerédi's theorem for pseudorandom measures.

The proof of the Green-Tao theorem

The main ingredient of the proof of the Green-Tao theorem is Szemerédi's theorem for pseudorandom measures.
After making an estimate to the generalized von Mangoldt function $\sum_{n\leq N}\widetilde{\Lambda}(n) = N(1+o(1))$ we get that $\mathbb{E}(f) = \frac{1}{k2^{k+5}}\varepsilon_k(1+o(1))$.

The proof of the Green-Tao theorem

The main ingredient of the proof of the Green-Tao theorem is Szemerédi's theorem for pseudorandom measures.
After making an estimate to the generalized von Mangoldt function $\sum_{n \leq N} \widetilde{\Lambda}(n) = N(1 + o(1))$ we get that $\mathbb{E}(f) = \frac{1}{k2^{k+5}} \varepsilon_k (1 + o(1))$.
For sufficiently large k we have $\frac{1}{k2^{k+5}} \varepsilon_k \leq 1$. So, if we let $\delta < \frac{1}{k2^{k+5}} \varepsilon_k$ the Szemerédi's theorem for pseudorandom measures gives

$$\mathbb{E}(f(x)f(x+r)\cdots f(x+(k-1)r)|x, r \in \mathbb{Z}_N) \geq c(k,\delta) - o_{k,\delta}(1)$$

for some constant $c(k,\delta) > 0$, not depending on f nor N.
Note that the right-hand side goes to $c(k,\delta) > 0$, as N goes to infinity.

The proof of the Green-Tao theorem

The product in the average on the left-hand is positive whenever f is positive on all elements in a configuration $\{x, \ldots, x+(k-1)r\}$. Therefore, must be infinitely many such configurations such that f is positive on all elements. Now notice that $f(x) > 0$ if and only if $Wx+1$ is prime and $\varepsilon_k N \leq x \leq 2\varepsilon_k N$. So, if $x+jr$, $j=0,\ldots,k-1$ is an arithmetic progression then so is $W(x+jr)$, since $W(x+jr)+1=(Wx+1)+j(Wr)$. So, by definition of f, each of the configurations gives us an arithmetic progression consisting of primes.

The proof of the Green-Tao theorem

Problems

There are two problems remaining. First of all the configuration is not an arithmetic progression of length k if $r = 0$, and secondly we must show that the arithmetic progressions we found in $\mathbb{Z}_N$ are genuine arithmetic progressions in $\mathbb{Z}$.

The proof of the Green-Tao theorem

1. The contribution to the average on the left-hand side for the degenerate case $r = 0$ is

$$\frac{1}{N^2} \sum_{x \in \mathbb{Z}_N} f(x)^k.$$

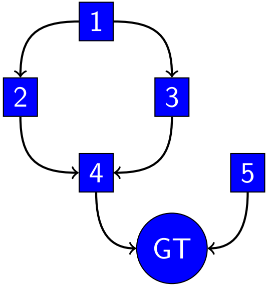
Note that $f(x) = O(\log N)$ since $\frac{\varphi(W)}{W} \leq 1$ and $\log(Wn + 1) \leq 2 \log N$ for large N. So the average for $r = 0$ is $O(\frac{(\log N)^k}{N}) = o(1)$ and thus can be discarded.

The proof of the Green-Tao theorem

2 Now we need to prove that the arithmetic progressions we find in $\mathbb{Z}_N$ are also progressions in $\mathbb{Z}$. Recall that in the configurations $\{x, \ldots, x+(k-1)r\}$ we found, we have that $\varepsilon_k N \leq x + jr \leq 2\varepsilon_k N$, for all $j = 0, \ldots, k-1$ because f has its support in this interval. So, if we assume that the arithmetic progression exceeds N the step r of the progression must be greater or equal to $N - 2\varepsilon_k N$. Therefore, $\frac{1}{k} > \varepsilon_k \geq \frac{N-r}{2N} \geq \frac{N-1}{2N}$. Hence, $N < \frac{k}{k-2}$ which is a contradiction since N is large enough.

The ingredients of the proof

1. A Furstenberg-type decomposition of a function into a Gowers uniform part and a bounded Gowers anti-uniform part.
2. Gowers uniform part is negligible by a generalized von Neumann theorem.
3. Gowers anti-uniform part is bounded from below by the reformulated Szemerédi's theorem.
4. Szemerédi's theorem relative to a pseudorandom measure.
5. Construction of f and ν.



The Gowers uniformity norm

Definition

Let $d \geq 0$ be a dimension (in practise, we will have $d = k - 1$, where k is the length of the arithmetic progressions under consideration). We let $\{0,1\}^d$ be the standard d-dimensional cube, consisting of d-tuples $\omega = (\omega_1, \ldots, \omega_d)$ where $\omega_j \in \{0,1\}$ for $j = 1, \ldots, d$. If $h = (h_1, \ldots, h_d) \in \mathbb{Z}_N^d$ we define $\omega \cdot h := \omega_1 h_1 + \ldots + \omega_d h_d$. If $(f_\omega)_{\omega \in \{0,1\}^d}$ is $\{0,1\}^d$-tuple of functions in $L^\infty(\mathbb{Z}_N)$, we define the d-dimensional Gowers inner product $< (f_\omega)_{\omega \in \{0,1\}^d} >_{U^d}$ by the formula

$$< (f_\omega)_{\omega \in \{0,1\}^d} >_{U^d} := \mathbb{E}\left(\prod_{\omega \in \{0,1\}^d} f_\omega(x + \omega \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right).$$

The Gowers uniformity norm

Definition

The Gowers uniformity norm $\|f\|_{U^d}$ of a function $f : \mathbb{Z}_N \to \mathbb{R}$ is defined by the formula

$$\|f\|_{U^d} := < (f)_{\omega \in \{0,1\}^d} >_{U^d}^{1/2^d} =$$

$$= \mathbb{E}\left(\prod_{\omega \in \{0,1\}^d} f(x + \omega \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right)^{1/2^d}.$$

The Gowers uniformity norm

- The Gowers uniformity norm is a shift invariant, i.e. $\|f(x)\|_{U^d} = \|f(x+h)\|_{U^d}$ for any $h \in \mathbb{Z}_N$ and they are genuinely norms for $d \geq 2$.

The Gowers uniformity norm

- The Gowers uniformity norm is a shift invariant, i.e. $\|f(x)\|_{U^d} = \|f(x+h)\|_{U^d}$ for any $h \in \mathbb{Z}_N$ and they are genuinely norms for $d \geq 2$.
- Moreover, $\|f\|_{U^1} \leq \|f\|_{U^2} \leq \ldots \leq \|f\|_{L^\infty}$.

The Gowers uniformity norm

- The Gowers uniformity norm is a shift invariant, i.e. $\|f(x)\|_{U^d} = \|f(x+h)\|_{U^d}$ for any $h \in \mathbb{Z}_N$ and they are genuinely norms for $d \geq 2$.
- Moreover, $\|f\|_{U^1} \leq \|f\|_{U^2} \leq \ldots \leq \|f\|_{L^\infty}$.

The pseudorandom measures ν are close to the constant measure ν_{const} in the U^d norms.

Suppose that ν is a k-pseudorandom measure. Then

$$\|\nu - \nu_{const}\|_{U^d} = \|\nu - 1\|_{U^d} = o(1)$$

for all $1 \leq d \leq k-1$.

Decomposition theorem

Proposition

Let ν be a k-pseudorandom measure, and let $f \in L^1(\mathbb{Z}_N)$ be a non-negative function satisfying $0 \leq f(x) \leq \nu(x)$ for all $x \in \mathbb{Z}_N$. Let $0 < \varepsilon \ll 1$ be a small parameter, and assume $N > N_0(\varepsilon)$ is sufficiently large. Then there exists a σ-algebra $\mathcal{B}$ and an exceptional set $\Omega \in \mathcal{B}$ such that

- *(smallness condition)*

$$\mathbb{E}(\nu \mathbf{1}_\Omega) = o_\varepsilon(1);$$

- *(ν is uniformly distributed outside of Ω)*

$$\|(1 - \mathbf{1}_\Omega)\mathbb{E}(\nu - 1|\mathcal{B})\|_{L^\infty} = o_\varepsilon(1)$$

and

Decomposition theorem

Proposition

- *(Gowers uniformity estimate)*

$$\|(1-\mathbf{1}_\Omega)(f-\mathbb{E}(f|\mathcal{B}))\|_{U^{k-1}} \leq \varepsilon^{1/2^k}.$$

Decomposition theorem

Definitions (The Gowers anti-uniformity)

We introduce the dual $(U^{k-1})^*$ norm, defined as usual by

$$\|g\|_{(U^{k-1})^*} = \sup\{| < f, g > | \,:\, f \in U^{k-1}(\mathbb{Z}_N),\ \|f\|_{U^{k-1}} \leq 1\}.$$

We say that g is Gowers anti-uniform if $\|g\|_{(U^{k-1})^*} = O(1)$ and $\|g\|_{L^\infty} = O(1)$. The dual function $\mathcal{D}F$ of F is defined by

$$\mathcal{D}F(x) := \mathbb{E}\left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0^{k-1}}} F(x + \omega \cdot h) \,\middle|\, h \in \mathbb{Z}_N^{k-1} \right)$$

where 0^{k-1} denotes the element of $\{0,1\}^{k-1}$ consisting entirely by zeroes.

Decomposition theorem

Lemma (Lack of Gowers uniformity implies correlation)

Let ν be a k-pseudorandom measure, and let $F \in L^1(\mathbb{Z}_N)$ be any function. Then the following hold

- $< F, \mathcal{D}F >= \|F\|_{U^{k-1}}^{2^{k-1}}$ *and*
- $\|\mathcal{D}F\|_{(U^{k-1})^*} = \|F\|_{U^{k-1}}^{2^{k-1}-1}$.

If furthermore we assume that $|F(x)| \leq \nu(x) + 1$ for all $x \in \mathbb{Z}_N$, then we have the estimate $\|\mathcal{D}F\|_{L^\infty} \leq 2^{2^{k-1}-1} + o(1)$.

Decomposition theorem

Sketch of the proof of the Decomposition Theorem

To construct the σ-algebra $\mathcal{B}$ required in the proposition, we will use the philosophy laid out by Furstenberg in his ergodic structure theorem, which decomposes any measure-preserving system into a weakly mixing extension of a tower of compact extensions.

Decomposition theorem

Sketch of the proof of the Decomposition Theorem

To construct the σ-algebra $\mathcal{B}$ required in the proposition, we will use the philosophy laid out by Furstenberg in his ergodic structure theorem, which decomposes any measure-preserving system into a weakly mixing extension of a tower of compact extensions.

In our setting, the idea is roughly speaking as follows. We initialize $\mathcal{B}$ to be the trivial σ-algebra $\mathcal{B} = \{\emptyset, \mathbb{Z}_N\}$. If the function $f - \mathbb{E}(f|\mathcal{B})$ is already Gowers uniform (in the sense of the Gowers uniformity estimate), then we can terminate the algorithm. Otherwise, we use the machinery of dual functions to locate a Gowers anti-uniform function $\mathcal{D}F_1$ which has some non-trivial correlation with f, and add the level sets of $\mathcal{D}F_1$ to the σ-algebra $\mathcal{B}$.

Sketch of the proof of the Decomposition Theorem

Sketch of the proof of the Decomposition Theorem

The nontrivial correlation property will ensure that the L^2 norm of $\mathbb{E}(f|\mathcal{B})$ increases by a non-trivial amount during the procedure and the pseudorandomness of ν will ensure that $\mathbb{E}(f|\mathcal{B})$ remains uniformly bounded.

Sketch of the proof of the Decomposition Theorem

Sketch of the proof of the Decomposition Theorem

The nontrivial correlation property will ensure that the L^2 norm of $\mathbb{E}(f|\mathcal{B})$ increases by a non-trivial amount during the procedure and the pseudorandomness of ν will ensure that $\mathbb{E}(f|\mathcal{B})$ remains uniformly bounded.
We then repeat the above algorithm until $f - \mathbb{E}(f|\mathcal{B})$ becomes sufficiently Gowers uniform, at which point we terminate the algorithm.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

The idea is to decompose an arbitrary function f into a Gowers uniform part and a bounded Gowers anti-uniform part.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

The idea is to decompose an arbitrary function f into a Gowers uniform part and a bounded Gowers anti-uniform part.

- The contribution of the Gowers-uniform part to the expression in Szemerédi's theorem relative to a pseudorandom measure, which counts k-term arithmetic progressions, will be negligible by a generalized von Neumann theorem.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

The idea is to decompose an arbitrary function f into a Gowers uniform part and a bounded Gowers anti-uniform part.

- The contribution of the Gowers-uniform part to the expression in Szemerédi's theorem relative to a pseudorandom measure, which counts k-term arithmetic progressions, will be negligible by a generalized von Neumann theorem.
- The contribution from the Gowers anti-uniform component will be bounded from below by Szemerédi's theorem in its traditional form.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Proposition (A generalized von Neumann theorem)

Suppose that ν is k-pseudorandom. Let $f_0, \ldots, f_{k-1} \in L^1(\mathbb{Z}_N)$ be functions which are pointwise bounded by $\nu + \nu_{const}$, or in other words

$$|f_j(x)| \leq \nu(x) + 1 \text{ for all } x \in \mathbb{Z}_N,\ 0 \leq j \leq k-1.$$

Let $c_0, \ldots, c_{k-1}$ be a permutation of k consecutive elements of $\{-k+1, \ldots, -1, 0, 1, \ldots, k-1\}$ (in practise we will take $c_j := j$). Then

$$\mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x + c_j r) \,\middle|\, x, r \in \mathbb{Z}_N \right) = O\Big(\inf_{0 \leq j \leq k-1} \|f_j\|_{U^{k-1}}\Big) + o(1).$$

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Remark

The reason we have an upper bound of $\nu(x) + 1$ instead of ν is because we shall apply this proposition to functions f_j which roughly have the form $f_j = f - \mathbb{E}(f \,|\, \mathcal{B})$, where f is some function bounded pointwise by ν, and $\mathcal{B}$ is a σ-algebra such that $\mathbb{E}(\nu \,|\, \mathcal{B})$ is essentially bounded (up to $o(1)$ errors) by 1, so that we can essentially bound $|f_j|$ by $\nu(x) + 1$.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Remark

The reason we have an upper bound of $\nu(x)+1$ instead of ν is because we shall apply this proposition to functions f_j which roughly have the form $f_j = f - \mathbb{E}(f \mid \mathcal{B})$, where f is some function bounded pointwise by ν, and $\mathcal{B}$ is a σ-algebra such that $\mathbb{E}(\nu \mid \mathcal{B})$ is essentially bounded (up to $o(1)$ errors) by 1, so that we can essentially bound $|f_j|$ by $\nu(x)+1$.

Note that $O(\inf_{0\le j\le k-1} \|f_j\|_{U^{k-1}})$ is small if at least one of the norms $\|f_j\|_{U^{k-1}}$ is small.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Definition

Let $f \in L^1(\mathbb{Z}_N)$ be a non-negative function satisfying $0 \leq f(x) \leq \nu(x)$ for all $x \in \mathbb{Z}_N$ and let $\mathcal{B}$ be the σ-algebra described in the Decomposition theorem. Let $\widetilde{f} := f_U + f_{U^\perp} = (1 - \mathbf{1}_\Omega)f$, where $f_U := (1 - \mathbf{1}_\Omega)(f - \mathbb{E}(f|\mathcal{B}))$ and $f_{U^\perp} := (1 - \mathbf{1}_\Omega)\mathbb{E}(f|\mathcal{B})$ (the subscript U stands for Gowers uniform, and $U^\perp$ for Gowers anti-uniform). Hence $\widetilde{f} = f$ outside a small set Ω and $0 \leq \widetilde{f} \leq f$.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Let f, δ be as in the Szemeredi's theorem relative to a pseudorandom measure, and let $0 < \varepsilon \ll \delta$ be a parameter to be chosen later.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Let f, δ be as in the Szemeredi's theorem relative to a pseudorandom measure, and let $0 < \varepsilon \ll \delta$ be a parameter to be chosen later.

Let $\mathcal{B}$ be as in the Decomposition theorem , and write $f_U := (1 - \mathbf{1}_\Omega)(f - \mathbb{E}(f|\mathcal{B}))$ and $f_{U^\perp} := (1 - \mathbf{1}_\Omega)\mathbb{E}(f|\mathcal{B})$ (the subscript U stands for Gowers uniform, and $U^\perp$ for Gowers anti-uniform).

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Let f, δ be as in the Szemeredi's theorem relative to a pseudorandom measure, and let $0 < \varepsilon \ll \delta$ be a parameter to be chosen later.
Let $\mathcal{B}$ be as in the Decomposition theorem , and write $f_U := (1 - \mathbf{1}_\Omega)(f - \mathbb{E}(f|\mathcal{B}))$ and $f_{U^\perp} := (1 - \mathbf{1}_\Omega)\mathbb{E}(f|\mathcal{B})$ (the subscript U stands for Gowers uniform, and $U^\perp$ for Gowers anti-uniform).
Observe that from the smallness condition, the assumptions of our theorem and the measurability of Ω we have

$$\begin{aligned}\mathbb{E}(f_{U^\perp}) &= \mathbb{E}((1 - \mathbf{1}_\Omega)\mathbb{E}(f|\mathcal{B})) = \mathbb{E}((1 - \mathbf{1}_\Omega)f) \geq \\ &\mathbb{E}(f) - \mathbb{E}(\nu\mathbf{1}_\Omega) \geq \delta - o_\varepsilon(1).\end{aligned}$$

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Also, since ν is uniformly distributed outside of Ω, we see that $f_{U^\perp}$ is bounded above by $1 + o_\varepsilon(1)$. Since f is non-negative, $f_{U^\perp}$ is also.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Also, since ν is uniformly distributed outside of Ω, we see that $f_{U^\perp}$ is bounded above by $1 + o_\varepsilon(1)$. Since f is non-negative, $f_{U^\perp}$ is also.

We may thus apply the reformulated Szemerédi's theorem to obtain

$$\mathbb{E}(f_{U^\perp}(x)f_{U^\perp}(x+r)\cdots f_{U^\perp}(x+(k-1)r)\,|\,x, r \in \mathbb{Z}_N) \\ \geq c(k,\delta) - o_\varepsilon(1) - o_{k,\delta}(1).$$

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Also, since ν is uniformly distributed outside of Ω, we see that $f_{U^\perp}$ is bounded above by $1+o_\varepsilon(1)$. Since f is non-negative, $f_{U^\perp}$ is also.
We may thus apply the reformulated Szemerédi's theorem to obtain

$$\mathbb{E}(f_{U^\perp}(x)f_{U^\perp}(x+r)\cdots f_{U^\perp}(x+(k-1)r)\,|\,x,r\in\mathbb{Z}_N)$$
$$\geq c(k,\delta)-o_\varepsilon(1)-o_{k,\delta}(1).$$

On the other hand, since $(1-\mathbf{1}_\Omega)f$ is bounded by ν and $f_{U^\perp}$ is bounded by $1+o_\varepsilon(1)$ then $f_U=(1-\mathbf{1}_\Omega)f-f_{U^\perp}$ is pointwise bounded by $\nu+1+o_\varepsilon(1)$. Note that from the Gowers uniformity estimate, we have that $\|f_U\|_{U^{k-1}}\leq\varepsilon^{1/2^k}$.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Applying the generalized von Neumann theorem we thus see that

$$\mathbb{E}(f_0(x)f_1(x+r)\cdots f_{k-1}(x+(k-1)r)\,|\,x,r\in\mathbb{Z}_N) = O(\varepsilon^{1/2^k}) + o_\varepsilon(1)$$

whenever each f_j is equal to f_U or $f_{U^\perp}$ with at least one f_j equals to f_U.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

Applying the generalized von Neumann theorem we thus see that

$$\mathbb{E}(f_0(x)f_1(x+r)\cdots f_{k-1}(x+(k-1)r) \mid x, r \in \mathbb{Z}_N) = O(\varepsilon^{1/2^k}) + o_\varepsilon(1)$$

whenever each f_j is equal to f_U or $f_{U^\perp}$ with at least one f_j equals to f_U.
Adding 2^k estimates of this kind we obtain

$$\begin{aligned}\mathbb{E}(\widetilde{f}(x)\widetilde{f}(x+r)\cdots\widetilde{f}(x+(k-1)r) \mid x, r \in \mathbb{Z}_N)\\ \geq c(k,\delta) - O(\varepsilon^{1/2^k}) - o_\varepsilon(1) - o_{k,\delta}(1),\end{aligned}$$

where $\widetilde{f} := f_U + f_{U^\perp} = (1 - \mathbf{1}_\Omega)f$.

The proof of the Szemeredi's theorem relative to a pseudorandom measure

But since $0 \leq \widetilde{f} = (1 - \mathbf{1}_\Omega) f \leq f$ we obtain that

$$\begin{aligned} &\mathbb{E}(f(x)f(x+r)\cdots f(x+(k-1)r) \,|\, x, r \in \mathbb{Z}_N) \\ &\qquad \geq c(k,\delta) - O(\varepsilon^{1/2^k}) - o_\varepsilon(1) - o_{k,\delta}(1). \end{aligned}$$

The proof of the Szemeredi's theorem relative to a pseudorandom measure

But since $0 \leq \widetilde{f} = (1 - \mathbf{1}_\Omega)f \leq f$ we obtain that

$$\begin{aligned}\mathbb{E}(f(x)f(x+r)\cdots f(x+(k-1)r) \,|\, x, r \in \mathbb{Z}_N) \\ \geq c(k,\delta) - O(\varepsilon^{1/2^k}) - o_\varepsilon(1) - o_{k,\delta}(1).\end{aligned}$$

Since ε can be made arbitrarily small (as long N is taken sufficiently large), the error terms on the right-hand side can be taken to be arbitrarily small by choosing N sufficiently large depending on k and δ and the proof is finished.

The work of H. Furstenberg (1977)

Theorem (Multiple recurrence)

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving probability system and let $k \geq 1$ be an integer. For any subset $E \in \mathcal{B}$ with $\mu(E) > 0$

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(E \cap T^{-n}E \cap T^{-2n}E \cap \ldots \cap T^{-(k-1)n}E) > 0.$$

The work of H. Furstenberg (1977)

Theorem (Multiple recurrence)

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving probability system and let $k \geq 1$ be an integer. For any subset $E \in \mathcal{B}$ with $\mu(E) > 0$

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(E \cap T^{-n}E \cap T^{-2n}E \cap \ldots \cap T^{-(k-1)n}E) > 0.$$

Corollary

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving probability system and let $k \geq 1$ be an integer. For any subset $E \in \mathcal{B}$ with $\mu(E) > 0$, there exists an integer $n \geq 1$ such that

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

The work of H. Furstenberg (1977)

Furstenberg then made the beautiful connection to combinatorics, showing that regularity properties of integers with positive upper density correspond to multiple recurrence results:

The work of H. Furstenberg (1977)

Furstenberg then made the beautiful connection to combinatorics, showing that regularity properties of integers with positive upper density correspond to multiple recurrence results:

Theorem (Correspondence principle)

Assume that A is a subset of integers with positive upper density. There exist a measure preserving probability system $(X, \mathcal{B}, \mu, T)$ and a measurable set $E \in \mathcal{B}$ with $\mu(E) = d^*(A)$, where $d^*(A)$ denotes the upper density of A, such that for all integers $k \geq 1$ and all integers $m_1, \ldots, m_{k-1} \geq 1$

$$d^*(A \cap (A+m_1) \cap \ldots \cap (A+m_{k-1})) \geq \mu(E \cap T^{-m_1} E \cap \ldots \cap T^{-m_{k-1}} E).$$

Taking $m_1 = n, m_2 = 2n, \ldots, m_{k-1} = (k-1)n$, Szemerédi's theorem follows from the Corollary.

The work of H. Furstenberg (1977)

To prove the Multiple Recurrence Theorem, Furstenberg showed that in any measure preserving system, one of two distinct phenomena occurs to make the measure of this intersection positive. The first is weak mixing. The system $(X, \mathcal{B}, \mu, T)$ is weakly mixing if for all $A, B \in \mathcal{B}$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mu(T^{-n}A \cap B) - \mu(A)\mu(B)| = 0.$$

The work of H. Furstenberg (1977)

To prove the Multiple Recurrence Theorem, Furstenberg showed that in any measure preserving system, one of two distinct phenomena occurs to make the measure of this intersection positive. The first is weak mixing. The system $(X, \mathcal{B}, \mu, T)$ is weakly mixing if for all $A, B \in \mathcal{B}$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mu(T^{-n}A \cap B) - \mu(A)\mu(B)| = 0.$$

So, for any set E, $\mu(E \cap T^{-n}E)$ is approximately $\mu(E)^2$ for most choices of the integer n. Then it can be shown that

$$\mu(E \cap T^{-n}E \cap T^{-2n}E \cap \ldots \cap T^{-(k-1)n}E)$$

is approximately $\mu(E)^k$ for most choices of n, which is clearly positive when E is a set of positive measure.

The work of H. Furstenberg (1977)

The opposite situation is rigidity, when for appropriately chosen n, T^n is very close to the identity. Then $T^{jn}E$ is very closed to E and

$$\mu(E \cap T^{-n}E \cap T^{-2n}E \cap \ldots \cap T^{-(k-1)n}E)$$

is very close to $\mu(E)$, again giving positive intersection for a set E of positive measure.

The work of H. Furstenberg (1977)

The opposite situation is rigidity, when for appropriately chosen n, T^n is very close to the identity. Then $T^{jn}E$ is very closed to E and

$$\mu(E \cap T^{-n}E \cap T^{-2n}E \cap \ldots \cap T^{-(k-1)n}E)$$

is very close to $\mu(E)$, again giving positive intersection for a set E of positive measure.
One then has to show that the average along arithmetic progressions for any function can be decomposed into two pieces, one which exhibits a generalized weak mixing property and another that exhibits a generalized rigidity property.