Minimal characteristic factors for linear systems - a quantitative approach

Julia Wolf

Rutgers University

joint work with Tim Gowers, University of Cambridge

May 13, 2009

Furstenberg's proof of Szemerédi's Theorem

Theorem (Szemerédi's Theorem)

Let $A \subset \{1, \ldots, N\}$ be a set of density α, and suppose that A contains no arithmetic progressions of length k. Then

$$\alpha = o_k(1).$$

Furstenberg's proof of Szemerédi's Theorem

Theorem (Szemerédi's Theorem)

Let $A \subset \{1, \ldots, N\}$ be a set of density α, and suppose that A contains no arithmetic progressions of length k. Then

$$\alpha = o_k(1).$$

Every set of positive upper density contains an arithmetic progression of length k.

Furstenberg's proof of Szemerédi's Theorem

Szemerédi's Theorem can be deduced from a recurrence statement in ergodic measure preserving systems.

Furstenberg's proof of Szemerédi's Theorem

Szemerédi's Theorem can be deduced from a recurrence statement in ergodic measure preserving systems.

Theorem (Furstenberg's Correspondence Principle, 1977)

Let A be a set of integers of positive upper density. Then there exist an ergodic measure-preserving system $(X, \mathcal{X}, \mu, T)$ and a set $E \in \mathcal{X}$ with $\mu(E) = d^*(A)$ such that

$$\mu(T^n E \cap \cdots \cap T^{kn} E) \leq d^*((A+n) \cap \cdots \cap (A+kn))$$

for all integers $k \geq 1$ and all $n \in \mathbb{Z}$.

Furstenberg's proof of Szemerédi's Theorem

Theorem (Furstenberg Multiple Recurrence, 1977)

Let $(X, \mathcal{X}, \mu, T)$ *be an ergodic measure-preserving system, and let* $f \in L^\infty(\mu)$ *Then*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int f(x) T^n f(x)\ T^{2n} f(x)\ \ldots\ T^{kn} f(x) d\mu(x)$$

is strictly greater than zero.

Furstenberg's proof of Szemerédi's Theorem

Theorem (Furstenberg Multiple Recurrence, 1977)

Let $(X, \mathcal{X}, \mu, T)$ be an ergodic measure-preserving system, and let $f \in L^\infty(\mu)$ Then

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int f(x) T^n f(x)\ T^{2n} f(x)\ \ldots\ T^{kn} f(x) d\mu(x)$$

is strictly greater than zero.

In order to understand multiple ergodic averages, look at so-called "characteristic factors".

Characteristic factors

Definition

We say a factor Y of X is *characteristic* for the average

$$\frac{1}{N^d} \sum_{n_1,\ldots,n_d=1}^{N} T^{L_1(n)} f(x) \ \ldots \ T^{L_m(n)} f(x) d\mu(x)$$

if and only if the difference with

$$\frac{1}{N^d} \sum_{n_1,\ldots,n_d=1}^{N} T^{L_1(n)} \mathbb{E}(f|Y)(x) \ \ldots \ T^{L_m(n)} \mathbb{E}(f|Y)(x) d\mu(x)$$

tends to 0 in $L^2(\mu)$.

Host-Kra norms

Definition

For $f \in L^\infty(\mu)$ and $k \in \mathbb{N}$., we define the *Host-Kra semi-norms* as

$$|||f|||_k := \left(\int_{X^{2^k}} f \otimes \cdots \otimes f d\mu^{[k]}\right)^{1/2^k}.$$

Host-Kra norms

Definition

For $f \in L^\infty(\mu)$ and $k \in \mathbb{N}.$, we define the *Host-Kra semi-norms* as

$$|||f|||_k := \left(\int_{X^{2^k}} f \otimes \cdots \otimes f d\mu^{[k]} \right)^{1/2^k} .$$

These seminorms are nested, in the sense that they satisfy

$$|||f|||_1 \leq |||f|||_2 \leq \cdots \leq |||f|||_k \leq \cdots \leq \|f\|_\infty,$$

Host-Kra norms

Definition

For $f \in L^\infty(\mu)$ and $k \in \mathbb{N}$., we define the *Host-Kra semi-norms* as

$$|||f|||_k := \left(\int_{X^{2^k}} f \otimes \cdots \otimes f d\mu^{[k]} \right)^{1/2^k} .$$

These seminorms are nested, in the sense that they satisfy

$$|||f|||_1 \leq |||f|||_2 \leq \cdots \leq |||f|||_k \leq \cdots \leq \|f\|_\infty,$$

and can also be defined inductively via the formula

$$|||f|||_{k+1}^{2^{k+1}} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |||f \cdot T^n f|||^{2^k}.$$

Characteristic factors

With this definition, every ergodic m.p.s has a natural sequence of characteristic factors $\mathcal{Z}_k$, defined via this sequence of semi-norms.

Characteristic factors

With this definition, every ergodic m.p.s has a natural sequence of characteristic factors $\mathcal{Z}_k$, defined via this sequence of semi-norms.

Definition

Given a measure-preserving system $(X, \mathcal{X}, \mu, T)$, there is a nested sequence of factors $\mathcal{Z}_k$ of X such that for any bounded function f on X

$$|||f|||_{k+1} = 0 \text{ if and only if } \mathbb{E}(f|\mathcal{Z}_k) = 0.$$

Characteristic factors

With this definition, every ergodic m.p.s has a natural sequence of characteristic factors $\mathcal{Z}_k$, defined via this sequence of semi-norms.

Definition

Given a measure-preserving system $(X, \mathcal{X}, \mu, T)$, there is a nested sequence of factors $\mathcal{Z}_k$ of X such that for any bounded function f on X

$$\|f\|_{k+1} = 0 \text{ if and only if } \mathbb{E}(f|\mathcal{Z}_k) = 0.$$

- $\mathcal{Z}_1$ corresponds to the classical Kronecker factor

Characteristic factors

With this definition, every ergodic m.p.s has a natural sequence of characteristic factors $\mathcal{Z}_k$, defined via this sequence of semi-norms.

Definition

Given a measure-preserving system $(X, \mathcal{X}, \mu, T)$, there is a nested sequence of factors $\mathcal{Z}_k$ of X such that for any bounded function f on X

$$\|\!|f|\!\|_{k+1} = 0 \text{ if and only if } \mathbb{E}(f|\mathcal{Z}_k) = 0.$$

- $\mathcal{Z}_1$ corresponds to the classical Kronecker factor
- $\mathcal{Z}_2$ corresponds to the (quadratic) Conze-Lesigne factor

Characteristic factors

With this definition, every ergodic m.p.s has a natural sequence of characteristic factors $\mathcal{Z}_k$, defined via this sequence of semi-norms.

Definition

Given a measure-preserving system $(X, \mathcal{X}, \mu, T)$, there is a nested sequence of factors $\mathcal{Z}_k$ of X such that for any bounded function f on X

$$|||f|||_{k+1} = 0 \text{ if and only if } \mathbb{E}(f|\mathcal{Z}_k) = 0.$$

- $\mathcal{Z}_1$ corresponds to the classical Kronecker factor
- $\mathcal{Z}_2$ corresponds to the (quadratic) Conze-Lesigne factor
- $\mathcal{Z}_k$ ...

Structure theorem for characteristic factors

A deep theorem by Host and Kra characterizes the structure of $\mathcal{Z}_k$ for general k.

Structure theorem for characteristic factors

A deep theorem by Host and Kra characterizes the structure of $\mathcal{Z}_k$ for general k.

Theorem (Host-Kra, 2006)

For each integer k, the factor $\mathcal{Z}_k$ is isomorphic to an inverse limit of k-step nilsystems.

Example of a 2-step nilsystems

Example

$$G := \begin{pmatrix} 1 & \mathbb{R} & \mathbb{R} \\ 0 & 1 & \mathbb{R} \\ 0 & 0 & 1 \end{pmatrix} \qquad \Gamma := \begin{pmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ 0 & 1 & \mathbb{Z} \\ 0 & 0 & 1 \end{pmatrix}$$

Example of a 2-step nilsystems

Example

$$G := \begin{pmatrix} 1 & \mathbb{R} & \mathbb{R} \\ 0 & 1 & \mathbb{R} \\ 0 & 0 & 1 \end{pmatrix} \qquad \Gamma := \begin{pmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ 0 & 1 & \mathbb{Z} \\ 0 & 0 & 1 \end{pmatrix}$$

Then $X = G/\Gamma$ is a 2-step nilmanifold. Then the transformation T defined as translation by

$$g := \begin{pmatrix} 1 & \gamma & \beta \\ 0 & 1 & \alpha \\ 0 & 0 & 1 \end{pmatrix}$$

Example of a 2-step nilsystems

Example

$$G := \begin{pmatrix} 1 & \mathbb{R} & \mathbb{R} \\ 0 & 1 & \mathbb{R} \\ 0 & 0 & 1 \end{pmatrix} \qquad \Gamma := \begin{pmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ 0 & 1 & \mathbb{Z} \\ 0 & 0 & 1 \end{pmatrix}$$

Then $X = G/\Gamma$ is a 2-step nilmanifold. Then the transformation T defined as translation by

$$g := \begin{pmatrix} 1 & \gamma & \beta \\ 0 & 1 & \alpha \\ 0 & 0 & 1 \end{pmatrix}$$

together with the Borel σ-algebra $\mathcal{X}$ and Haar measure μ defines a 2-step nilsystem.

Example of a 2-step nilsystems

Example

For the Heisenberg nilsystem, we have

$$T(x,y,z) = (x+\alpha, y+\beta+\gamma x, z+\gamma)$$

Example of a 2-step nilsystems

Example

For the Heisenberg nilsystem, we have

$$T(x, y, z) = (x + \alpha, y + \beta + \gamma x, z + \gamma)$$

and

$$T^n(x, y, z) = (x + n\alpha, y + n\beta + \frac{1}{2}n(n+1)\alpha, z + n\gamma)$$

Example of a 2-step nilsystems

Example

For the Heisenberg nilsystem, we have

$$T(x, y, z) = (x + \alpha, y + \beta + \gamma x, z + \gamma)$$

and

$$T^n(x, y, z) = (x + n\alpha, y + n\beta + \frac{1}{2}n(n+1)\alpha, z + n\gamma)$$

Note that the behaviour is quadratic in n.

Multiple linear ergodic averages

Proposition

Assume that $(X, \mathcal{X}, \mu, T)$ *is ergodic and let* $d, k, m \in \mathbb{N}$*. If* $f \in L^\infty(\mu)$*, then*

$$\limsup_{N\to\infty} \left\| \frac{1}{N^d} \sum_{n_1,n_2,\ldots,n_d=0}^{N-1} T^{L_1(n)} f(x) \ \ldots \ T^{L_m(n)} f(x) \right\|_2 \ll \| f \|_m.$$

Multiple linear ergodic averages

Proposition

Assume that $(X, \mathcal{X}, \mu, T)$ is ergodic and let $d, k, m \in \mathbb{N}$. If $f \in L^\infty(\mu)$, then

$$\limsup_{N\to\infty} \left\| \frac{1}{N^d} \sum_{n_1, n_2, \ldots, n_d = 0}^{N-1} T^{L_1(n)} f(x) \ \ldots \ T^{L_m(n)} f(x) \right\|_2 \ll \| f \|_m .$$

The proof is a simple application of Van der Corput's Lemma.

Multiple linear ergodic averages

Proposition

Assume that $(X, \mathcal{X}, \mu, T)$ is ergodic and let $d, k, m \in \mathbb{N}$. If $f \in L^\infty(\mu)$, then

$$\limsup_{N\to\infty} \left\| \frac{1}{N^d} \sum_{n_1, n_2, \ldots, n_d = 0}^{N-1} T^{L_1(n)} f(x) \ \ldots \ T^{L_m(n)} f(x) \right\|_2 \ll \| f \|_m.$$

The proof is a simple application of Van der Corput's Lemma.

Question

What is the degree of the minimal characteristic factor for a given multiple ergodic average?

Minimal characteristic factors

Theorem (Leibman, 2007)

The factor $\mathcal{Z}_1$ is minimal characteristic if and only if the linear forms are square-independent.

Minimal characteristic factors

Theorem (Leibman, 2007)

The factor $\mathcal{Z}_1$ is minimal characteristic if and only if the linear forms are square-independent.

The linear forms are said to be square-independent if $L_1^2, L_2^2, \ldots, L_m^2$ are linearly independent (over $\mathbb{Z}$).

Minimal characteristic factors

Theorem (Leibman, 2007)

The factor $\mathcal{Z}_1$ is minimal characteristic if and only if the linear forms are square-independent.

The linear forms are said to be square-independent if $L_1^2, L_2^2, \ldots, L_m^2$ are linearly independent (over $\mathbb{Z}$).

Example

Consider the linear forms $n+m$, $2n+4m$, $3n+9m$, $4n+16m$, $5n+25m$, $6n+37m$.

Minimal characteristic factors

Theorem (Leibman, 2007)

More generally, the factor $\mathcal{Z}_k$ is minimal characteristic if and only if k is the least integer such that $L_1^{k+1}, L_2^{k+1}, \ldots, L_m^{k+1}$ are linearly independent.

Minimal characteristic factors

Theorem (Leibman, 2007)

More generally, the factor $\mathcal{Z}_k$ is minimal characteristic if and only if k is the least integer such that $L_1^{k+1}, L_2^{k+1}, \ldots, L_m^{k+1}$ are linearly independent.

Idea of the proof:

Minimal characteristic factors

Theorem (Leibman, 2007)

More generally, the factor $\mathcal{Z}_k$ is minimal characteristic if and only if k is the least integer such that $L_1^{k+1}, L_2^{k+1}, \ldots, L_m^{k+1}$ are linearly independent.

Idea of the proof:

- Use van der Corput to control the multiple ergodic average by $\|\!|\cdot|\!\|_k$ for some k.

Minimal characteristic factors

Theorem (Leibman, 2007)

More generally, the factor $\mathcal{Z}_k$ is minimal characteristic if and only if k is the least integer such that $L_1^{k+1}, L_2^{k+1}, \ldots, L_m^{k+1}$ are linearly independent.

Idea of the proof:

- Use van der Corput to control the multiple ergodic average by $\|\cdot\|_k$ for some k.
- This tells us that $\mathcal{Z}_k$ is a characteristic factor. By the theorem of Host and Kra, $\mathcal{Z}_k$ has the structure of a k-step nilmanifold.

Minimal characteristic factors

Theorem (Leibman, 2007)

More generally, the factor $\mathcal{Z}_k$ is minimal characteristic if and only if k is the least integer such that $L_1^{k+1}, L_2^{k+1}, \ldots, L_m^{k+1}$ are linearly independent.

Idea of the proof:

- Use van der Corput to control the multiple ergodic average by $\|\!|\cdot|\!\|_k$ for some k.
- This tells us that $\mathcal{Z}_k$ is a characteristic factor. By the theorem of Host and Kra, $\mathcal{Z}_k$ has the structure of a k-step nilmanifold.
- Explicitly compute the orbit of the diagonal of the nilmanifold under the linear actions.

Minimal characteristic factors

Remarks:

Minimal characteristic factors

Remarks:

- The linear case is fully resolved.

Minimal characteristic factors

Remarks:

- The linear case is fully resolved.
- In the polynomial case, the orbit can be explicitly described when the nilmanifold is connected, as well as when the nilmanifold is a torus or a Weyl system.

Minimal characteristic factors

Remarks:

- The linear case is fully resolved.
- In the polynomial case, the orbit can be explicitly described when the nilmanifold is connected, as well as when the nilmanifold is a torus or a Weyl system.
- For the general polynomial case, Leibman produces an algorithm which yields an upper bound on the degree of the minimal characteristic factor.

The discrete Fourier transform

Let us start off by considering a classical Fourier decomposition.

The discrete Fourier transform

Let us start off by considering a classical Fourier decomposition.

- Fourier transform: $\widehat{f}(\gamma) := \mathbb{E}_{x \in G} f(x)\gamma(x)$

The discrete Fourier transform

Let us start off by considering a classical Fourier decomposition.

- Fourier transform: $\widehat{f}(\gamma) := \mathbb{E}_{x \in G} f(x)\gamma(x)$
- Fourier inversion: $f(x) = \sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma)\gamma(-x)$

The discrete Fourier transform

Let us start off by considering a classical Fourier decomposition.

- Fourier transform: $\widehat{f}(\gamma) := \mathbb{E}_{x \in G} f(x)\gamma(x)$
- Fourier inversion: $f(x) = \sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma)\gamma(-x)$
- Parseval's identity: $\mathbb{E}_{x \in G}|f(x)|^2 = \sum_{\gamma \in \widehat{G}} |\widehat{f}(\gamma)|^2$

The discrete Fourier transform

Let us start off by considering a classical Fourier decomposition.

- Fourier transform: $\widehat{f}(\gamma) := \mathbb{E}_{x \in G} f(x)\gamma(x)$
- Fourier inversion: $f(x) = \sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma)\gamma(-x)$
- Parseval's identity: $\mathbb{E}_{x \in G}|f(x)|^2 = \sum_{\gamma \in \widehat{G}} |\widehat{f}(\gamma)|^2$

Throughout this talk, we will consider $G = \mathbb{Z}/N\mathbb{Z}$ or $G = \mathbb{F}_p^n$.

The discrete Fourier transform

Let us start off by considering a classical Fourier decomposition.

- Fourier transform: $\widehat{f}(\gamma) := \mathbb{E}_{x \in G} f(x)\gamma(x)$
- Fourier inversion: $f(x) = \sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma)\gamma(-x)$
- Parseval's identity: $\mathbb{E}_{x \in G}|f(x)|^2 = \sum_{\gamma \in \widehat{G}} |\widehat{f}(\gamma)|^2$

Throughout this talk, we will consider $G = \mathbb{Z}/N\mathbb{Z}$ or $G = \mathbb{F}_p^n$.

Definition

We say a set $A \subseteq G$ is uniform if the largest non-trivial Fourier coefficient of its characteristic function is small.

Counting arithmetic progressions

Fact

If a subset A of G of density α is uniform, then it contains the expected number α^3 of 3-term progressions.

Counting arithmetic progressions

Fact

If a subset A of G of density α is uniform, then it contains the expected number α^3 of 3-term progressions.

This corresponds to the statement that the Kronecker factor is characteristic for the ergodic average along 3-term arithmetic progressions.

Counting arithmetic progressions

Fact

If a subset A of G of density α is uniform, then it contains the expected number α^3 of 3-term progressions.

This corresponds to the statement that the Kronecker factor is characteristic for the ergodic average along 3-term arithmetic progressions.

$$\mathbb{E}_{x,d\in G} 1_A(x)1_A(x+d)1_A(x+2d)$$

Counting arithmetic progressions

Fact

If a subset A of G of density α is uniform, then it contains the expected number α^3 of 3-term progressions.

This corresponds to the statement that the Kronecker factor is characteristic for the ergodic average along 3-term arithmetic progressions.

$$\mathbb{E}_{x,d\in G} 1_A(x) 1_A(x+d) 1_A(x+2d)$$

$$= \sum_{\gamma \in \widehat{G}} \widehat{1_A}(\gamma)^2 \widehat{1_A}(-2\gamma)$$

Counting arithmetic progressions

Fact

If a subset A of G of density α is uniform, then it contains the expected number α^3 of 3-term progressions.

This corresponds to the statement that the Kronecker factor is characteristic for the ergodic average along 3-term arithmetic progressions.

$$\mathbb{E}_{x,d\in G} 1_A(x) 1_A(x+d) 1_A(x+2d)$$

$$= \sum_{\gamma\in\widehat{G}} \widehat{1_A}(\gamma)^2 \widehat{1_A}(-2\gamma)$$

$$= \alpha^3 + o(1)$$

Counting arithmetic progressions

Fact

Fourier analysis is not sufficient for counting longer progressions

Counting arithmetic progressions

Fact

Fourier analysis is not sufficient for counting longer progressions

For example, the following set is uniform but contains many more than the expected number of 4-APs.

$$A = \{x \in \mathbb{Z}_N : |x^2| \quad \text{small}\}$$

Counting arithmetic progressions

Fact

Fourier analysis is not sufficient for counting longer progressions

For example, the following set is uniform but contains many more than the expected number of 4-APs.

$$A = \{x \in \mathbb{Z}_N : |x^2| \quad \text{small}\}$$

$$x^2 - 3(x+d)^2 + 3(x+2d)^2 - (x+3d)^2 = 0$$

Counting arithmetic progressions

Fact

Fourier analysis is not sufficient for counting longer progressions

For example, the following set is uniform but contains many more than the expected number of 4-APs.

$$A = \{x \in \mathbb{Z}_N : |x^2| \quad \text{small}\}$$

$$x^2 - 3(x+d)^2 + 3(x+2d)^2 - (x+3d)^2 = 0$$

This corresponds to the Furstenberg-Weiss example in ergodic theory.

Uniformity norms

Definition (Gowers, 2001)

For any positive integer $k \geq 2$, and any function $f : G \to [-1, 1]$, define the U^k*-norm* by the formula

$$\|f\|_{U^k}^{2^k} := \mathbb{E}_{x,h_1,\ldots,h_k \in G} \prod_{\omega \in \{0,1\}^k} f(x + \sum_i \omega_i h_i).$$

Uniformity norms

Definition (Gowers, 2001)

For any positive integer $k \geq 2$, and any function $f : G \to [-1, 1]$, define the U^k*-norm* by the formula

$$\|f\|_{U^k}^{2^k} := \mathbb{E}_{x,h_1,\ldots,h_k \in G} \prod_{\omega \in \{0,1\}^k} f(x + \sum_i \omega_i h_i).$$

This definition corresponds to that of the semi-norms $\| \cdot \|_k$ in ergodic theory.

Uniformity norms

Definition (Gowers, 2001)

For any positive integer $k \geq 2$, and any function $f : G \to [-1, 1]$, define the U^k-*norm* by the formula

$$\|f\|_{U^k}^{2^k} := \mathbb{E}_{x,h_1,\ldots,h_k \in G} \prod_{\omega \in \{0,1\}^k} f(x + \sum_i \omega_i h_i).$$

This definition corresponds to that of the semi-norms $|\!|\!| \cdot |\!|\!|_k$ in ergodic theory.
In particular,

$$\|f\|_{U^2}^4 = \mathbb{E}_{x,a,b \in G} f(x) f(x+a) f(x+b) f(x+a+b) = \|\widehat{f}\|_4^4.$$

Counting arithmetic progressions, II

Definition

A set $A \subseteq G$ is said to be quadratically uniform if its balanced function $f_A = 1_A - \alpha$ is small in U^3.

Counting arithmetic progressions, II

Definition

A set $A \subseteq G$ is said to be quadratically uniform if its balanced function $f_A = 1_A - \alpha$ is small in U^3.

If a subset A of G is quadratically uniform, then it contains the expected number of 4-term progressions.

Counting arithmetic progressions, II

Definition

A set $A \subseteq G$ is said to be quadratically uniform if its balanced function $f_A = 1_A - \alpha$ is small in U^3.

If a subset A of G is quadratically uniform, then it contains the expected number of 4-term progressions.

Proposition (Gowers, 2001)

Let $f : G \to [-1, 1]$*. Then*

$$|\mathbb{E}_{x,d \in G} f(x)f(x+d)f(x+2d)f(x+3d)| \leq \|f\|_{U^3}.$$

Counting arithmetic progressions, II

Definition

A set $A \subseteq G$ is said to be quadratically uniform if its balanced function $f_A = 1_A - \alpha$ is small in U^3.

If a subset A of G is quadratically uniform, then it contains the expected number of 4-term progressions.

Proposition (Gowers, 2001)

Let $f : G \to [-1, 1]$*. Then*

$$|\mathbb{E}_{x,d \in G} f(x)f(x+d)f(x+2d)f(x+3d)| \leq \|f\|_{U^3}.$$

This is analogous to the statement that the Conze-Lesigne factor is characteristic for averages along 4-term progressions.

Counting arithmetic progressions, II

More generally, k-APs are controlled by the U^{k-1} norm.

Counting arithmetic progressions, II

More generally, k-APs are controlled by the U^{k-1} norm.

Proposition (Gowers, 2001)

Let $f : G \to [-1,1]$. Then

$$|\mathbb{E}_{x,d \in G} f(x)f(x+d)f(x+2d)\dots f(x+kd)| \leq \|f\|_{U^k}.$$

Counting arithmetic progressions, II

More generally, k-APs are controlled by the U^{k-1} norm.

Proposition (Gowers, 2001)

Let $f : G \to [-1, 1]$*. Then*

$$|\mathbb{E}_{x,d \in G} f(x)f(x+d)f(x+2d)\dots f(x+kd)| \leq \|f\|_{U^k}.$$

This is proved using multiple applications of the Cauchy-Schwarz inequality.

Counting arithmetic progressions, II

More generally, k-APs are controlled by the U^{k-1} norm.

Proposition (Gowers, 2001)

Let $f : G \to [-1,1]$. Then

$$|\mathbb{E}_{x,d\in G} f(x)f(x+d)f(x+2d)\dots f(x+kd)| \leq \|f\|_{U^k}.$$

This is proved using multiple applications of the Cauchy-Schwarz inequality.

Question

What replaces the Host-Kra structure theorem in the finite world?

An inverse theorem for U^3

Theorem (Green-Tao, 2005, Samorodnitsky, 2007)

Suppose $\|f\|_\infty \leq 1$ is such that $\|f\|_{U^3} \geq \delta$. Then there exists a quadratic phase function ϕ such that

$$|\mathbb{E}_x f(x)\phi(x)| \geq c(\delta).$$

An inverse theorem for U^3

Theorem (Green-Tao, 2005, Samorodnitsky, 2007)

Suppose $\|f\|_\infty \leq 1$ *is such that* $\|f\|_{U^3} \geq \delta$. *Then there exists a quadratic phase function* ϕ *such that*

$$|\mathbb{E}_x f(x)\phi(x)| \geq c(\delta).$$

We shall be deliberately vague here about what we mean by a quadratic phase function.

A classical Fourier decomposition

We can write

$$f(x) = \sum_{\gamma \in R} \widehat{f}(\gamma)\gamma(x) + \sum_{t \notin R} \widehat{f}(\gamma)\gamma(x),$$

where R denotes the set of frequencies where the Fourier transform of f is large. Here

A classical Fourier decomposition

We can write

$$f(x) = \sum_{\gamma \in R} \widehat{f}(\gamma)\gamma(x) + \sum_{t \notin R} \widehat{f}(\gamma)\gamma(x),$$

where R denotes the set of frequencies where the Fourier transform of f is large. Here

f_1 has linear structure
and bounded complexity ($\rightarrow |R|$ is small)

A classical Fourier decomposition

We can write

$$f(x) = \sum_{\gamma \in R} \widehat{f}(\gamma)\gamma(x) + \sum_{t \notin R} \widehat{f}(\gamma)\gamma(x),$$

where R denotes the set of frequencies where the Fourier transform of f is large. Here

f_1 has linear structure
and bounded complexity ($\rightarrow |R|$ is small)

and

f_2 is uniform in the classical sense ($\rightarrow$ small in U^2).

A decomposition into quadratic phases

Many proofs in arithmetic combinatorics proceed via a dichotomy: a decomposition theorem can often be used to encode such a dichotomy.

A decomposition into quadratic phases

Many proofs in arithmetic combinatorics proceed via a dichotomy: a decomposition theorem can often be used to encode such a dichotomy.

However, there does not seem to be a canonical way to decompose a function into quadratic phases.

A decomposition into quadratic phases

We aim for a decomposition of the form

$$f = f_1 + f_2,$$

where

A decomposition into quadratic phases

We aim for a decomposition of the form

$$f = f_1 + f_2,$$

where

f_1 has some sort of quadratic structure
and bounded complexity

A decomposition into quadratic phases

We aim for a decomposition of the form

$$f = f_1 + f_2,$$

where

f_1 has some sort of quadratic structure
and bounded complexity

and

f_2 is quadratically uniform ($\rightarrow$ small in U^3).

A first decomposition theorem

Theorem (Green-Tao, 2005)

Let $\delta > 0$. *Given* $f : \mathbb{F}_p^n \to [-1, 1]$, *there exists* $d(\delta)$ *and a quadratic factor* $(\mathcal{B}_1, \mathcal{B}_2)$ *of complexity* d *with together with a decomposition*

$$f = f_1 + f_2,$$

where $f_1 := \mathbb{E}(f|\mathcal{B}_2)$ *and* $\|f_2\|_{U^3} \leq \delta$.

A first decomposition theorem

Theorem (Green-Tao, 2005)

Let $\delta > 0$. Given $f : \mathbb{F}_p^n \to [-1, 1]$, there exists $d(\delta)$ and a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ of complexity d with together with a decomposition

$$f = f_1 + f_2,$$

where $f_1 := \mathbb{E}(f|\mathcal{B}_2)$ and $\|f_2\|_{U^3} \leq \delta$.

A quadratic factor is a partition of $\mathbb{F}_p^n$ into simultaneous level sets of at most d linear and quadratic phases.

A simple decomposition into quadratic phases

Theorem (Gowers-W., 2008)

Let $f : \mathbb{F}_p^n \to \mathbb{C}$ be a function such that $\|f\|_2 \leq 1$. Then for every $\delta > 0$ there exists $M(\delta)$ such that f has a decomposition of the form

$$f(x) = \sum_i \lambda_i \omega^{q_i(x)} + g(x) + h(x),$$

where the q_i are quadratic forms on $\mathbb{F}_p^n$, $\|g\|_{U^3} \leq \delta$, $\|h\|_1 \leq \delta$ and $\sum_i |\lambda_i| \leq M$.

A simple decomposition into quadratic phases

Theorem (Gowers-W., 2008)

Let $f : \mathbb{F}_p^n \to \mathbb{C}$ be a function such that $\|f\|_2 \leq 1$. Then for every $\delta > 0$ there exists $M(\delta)$ such that f has a decomposition of the form

$$f(x) = \sum_i \lambda_i \omega^{q_i(x)} + g(x) + h(x),$$

where the q_i are quadratic forms on $\mathbb{F}_p^n$, $\|g\|_{U^3} \leq \delta$, $\|h\|_1 \leq \delta$ and $\sum_i |\lambda_i| \leq M$.

Here M can be taken to be $\exp(C(\delta^2)^{-C})$.

A simple decomposition into quadratic phases

Theorem (Gowers-W., 2008)

Let $f : \mathbb{F}_p^n \to \mathbb{C}$ be a function such that $\|f\|_2 \leq 1$. Then for every $\delta > 0$ there exists $M(\delta)$ such that f has a decomposition of the form

$$f(x) = \sum_i \lambda_i \omega^{q_i(x)} + g(x) + h(x),$$

where the q_i are quadratic forms on $\mathbb{F}_p^n$, $\|g\|_{U^3} \leq \delta$, $\|h\|_1 \leq \delta$ and $\sum_i |\lambda_i| \leq M$.

Here M can be taken to be $\exp(C(\delta^2)^{-C})$.

This is deduced from the inverse theorem via the Hahn-Banach theorem from functional analysis.

A higher-order inverse theorem

Conjecture

Let $0 < \delta \leq 1$ and let p be a prime. Let $f : \mathbb{F}_p^n \to \mathbb{C}$ be a function with $\|f\|_\infty \leq 1$ and $\|f\|_{U^{k+1}} \geq \delta$. Then there exists a polynomial $q : \mathbb{F}_p^n \to \mathbb{F}_p$ of degree k and a constant $c(\delta)$ such that

$$|\mathbb{E}_x f(x)\omega^{q(x)}| \geq c(\delta).$$

A higher-order inverse theorem

Conjecture

Let $0 < \delta \leq 1$ and let p be a prime. Let $f : \mathbb{F}_p^n \to \mathbb{C}$ be a function with $\|f\|_\infty \leq 1$ and $\|f\|_{U^{k+1}} \geq \delta$. Then there exists a polynomial $q : \mathbb{F}_p^n \to \mathbb{F}_p$ of degree k and a constant $c(\delta)$ such that

$$|\mathbb{E}_x f(x)\omega^{q(x)}| \geq c(\delta).$$

Recently proved in the case when $p \geq k$ by Bergelson-Tao-Ziegler using ergodic theory and a correspondence principle.

A higher-order inverse theorem

Conjecture

Let $0 < \delta \leq 1$ and let p be a prime. Let $f : \mathbb{F}_p^n \to \mathbb{C}$ be a function with $\|f\|_\infty \leq 1$ and $\|f\|_{U^{k+1}} \geq \delta$. Then there exists a polynomial $q : \mathbb{F}_p^n \to \mathbb{F}_p$ of degree k and a constant $c(\delta)$ such that

$$|\mathbb{E}_x f(x)\omega^{q(x)}| \geq c(\delta).$$

Recently proved in the case when $p \geq k$ by Bergelson-Tao-Ziegler using ergodic theory and a correspondence principle.

The conjecture when $p < k$ had been disproved by Green-Tao and Lovett-Meshulam-Samorodnitsky (2008).

Cauchy-Schwarz complexity

Definition (Green-Tao, 2006)

Let $\mathcal{L} = (L_1, ..., L_m)$ be a system of m linear forms in d variables. $\mathcal{L}$ is said to have *Cauchy-Schwarz complexity* k if, after an appropriate reparametrization, there exists a set of $k+1$ variables that are simultaneously used by only one of the linear forms.

Cauchy-Schwarz complexity

Definition (Green-Tao, 2006)

Let $\mathcal{L} = (L_1, ..., L_m)$ be a system of m linear forms in d variables. $\mathcal{L}$ is said to have *Cauchy-Schwarz complexity* k if, after an appropriate reparametrization, there exists a set of $k+1$ variables that are simultaneously used by only one of the linear forms.

Example

A 4-term progression can be expressed as

$$(y + 2z + 3w, -x + z + 2w, -2x - y + w, -3x - 2y - z),$$

and thus has Cauchy-Schwarz complexity 2.

Cauchy-Schwarz complexity

Definition (Green-Tao, 2006)

Let $\mathcal{L} = (L_1, \dots, L_m)$ be a system of m linear forms in d variables. $\mathcal{L}$ is said to have *Cauchy-Schwarz complexity* k if, after an appropriate reparametrization, there exists a set of $k+1$ variables that are simultaneously used by only one of the linear forms.

Example

A 4-term progression can be expressed as

$$(y + 2z + 3w, -x + z + 2w, -2x - y + w, -3x - 2y - z),$$

and thus has Cauchy-Schwarz complexity 2. More generally, CSC(k-AP)$= k - 2$ and CSC(cube of dimension d)$= d - 1$.

Cauchy-Schwarz complexity

Proposition (Green-Tao, 2006)

Let $f : G \to [-1,1]$, and let $L_1, \ldots, L_m$ be a system of Cauchy-Schwarz complexity k. Then

$$|\mathbb{E}_{x_1,\ldots,x_d \in G} \prod_{i=1}^{m} f(L_i(x_1,\ldots,x_d))| \leq \|f\|_{U^{k+1}}.$$

Cauchy-Schwarz complexity

Proposition (Green-Tao, 2006)

Let $f : G \to [-1,1]$, and let $L_1, \ldots, L_m$ be a system of Cauchy-Schwarz complexity k. Then

$$|\mathbb{E}_{x_1,\ldots,x_d \in G} \prod_{i=1}^{m} f(L_i(x_1,\ldots,x_d))| \leq \|f\|_{U^{k+1}}.$$

Question

What is the minimal k such that U^{k+1} controls the average over the linear system?

True complexity

Recall the example of a uniform set containing too many 4-APs:

$$A = \{x \in \mathbb{Z}_N : |x^2| \quad \text{small}\}$$

True complexity

Recall the example of a uniform set containing too many 4-APs:

$$A = \{x \in \mathbb{Z}_N : |x^2| \quad \text{small}\}$$

$$x^2 - 3(x+d)^2 + 3(x+2d)^2 - (x+3d)^2 = 0$$

True complexity

Recall the example of a uniform set containing too many 4-APs:

$$A = \{x \in \mathbb{Z}_N : |x^2| \quad \text{small}\}$$

$$x^2 - 3(x+d)^2 + 3(x+2d)^2 - (x+3d)^2 = 0$$

Conjecture (Gowers-W., 2007)

A linear system is controlled by U^2 if and only if the functions L_i^2 are linearly independent.

True complexity

Recall the example of a uniform set containing too many 4-APs:

$$A = \{x \in \mathbb{Z}_N : |x^2| \quad \text{small}\}$$

$$x^2 - 3(x+d)^2 + 3(x+2d)^2 - (x+3d)^2 = 0$$

Conjecture (Gowers-W., 2007)

A linear system is controlled by U^2 if and only if the functions L_i^2 are linearly independent.

More generally, the linear system is controlled by U^k if and only if k is the least integer such that the functions L_i^{k+1} are linearly independent.

Square independence matters

Theorem (Gowers-W., 2007)

Let $G = \mathbb{F}_p^n$, and let $\mathcal{L}$ be a linear system of Cauchy-Schwarz complexity 2. Then $\mathcal{L}$ is controlled by U^2 if and only if the $L_i^T L_i$ are linearly independent.

Square independence matters

Theorem (Gowers-W., 2007)

Let $G = \mathbb{F}_p^n$, and let $\mathcal{L}$ be a linear system of Cauchy-Schwarz complexity 2. Then $\mathcal{L}$ is controlled by U^2 if and only if the $L_i^T L_i$ are linearly independent.

Idea of proof:

Square independence matters

Theorem (Gowers-W., 2007)

Let $G = \mathbb{F}_p^n$, and let $\mathcal{L}$ be a linear system of Cauchy-Schwarz complexity 2. Then $\mathcal{L}$ is controlled by U^2 if and only if the $L_i^T L_i$ are linearly independent.

Idea of proof:

- Decompose the balanced function f into a quadratically structured part f_1 and a quadratically uniform part f_2.

Square independence matters

Theorem (Gowers-W., 2007)

Let $G = \mathbb{F}_p^n$, and let $\mathcal{L}$ be a linear system of Cauchy-Schwarz complexity 2. Then $\mathcal{L}$ is controlled by U^2 if and only if the $L_i^T L_i$ are linearly independent.

Idea of proof:

- Decompose the balanced function f into a quadratically structured part f_1 and a quadratically uniform part f_2.
- By Green and Tao's Cauchy-Schwarz argument, the contribution from f_2 is negligible since $\mathcal{L}$ has CSC 2.

Square independence matters

Theorem (Gowers-W., 2007)

Let $G = \mathbb{F}_p^n$, and let $\mathcal{L}$ be a linear system of Cauchy-Schwarz complexity 2. Then $\mathcal{L}$ is controlled by U^2 if and only if the $L_i^T L_i$ are linearly independent.

Idea of proof:

- Decompose the balanced function f into a quadratically structured part f_1 and a quadratically uniform part f_2.
- By Green and Tao's Cauchy-Schwarz argument, the contribution from f_2 is negligible since $\mathcal{L}$ has CSC 2.
- By explicit computation the contribution from f_1 is negligible since $\mathcal{L}$ is square-independent and f highly uniform.

Square independence matters

Some remarks:

Square independence matters

Some remarks:

- This gives an easy proof of Szemerédi-type theorems for translation invariant square-independent systems.

Square independence matters

Some remarks:

- This gives an easy proof of Szemerédi-type theorems for translation invariant square-independent systems.
- What is the correct dependence on the U^2 norm?

Square independence matters

Some remarks:

- This gives an easy proof of Szemerédi-type theorems for translation invariant square-independent systems.
- What is the correct dependence on the U^2 norm?
- The case $\mathbb{Z}/N\mathbb{Z}$ is technically much more difficult since we lack global quadratic correlation in the inverse theorem.

Square independence matters

Some remarks:

- This gives an easy proof of Szemerédi-type theorems for translation invariant square-independent systems.
- What is the correct dependence on the U^2 norm?
- The case $\mathbb{Z}/N\mathbb{Z}$ is technically much more difficult since we lack global quadratic correlation in the inverse theorem.
- What about systems of higher Cauchy-Schwarz complexity? What about cube-independent systems?

Square independence matters

Some remarks:

- This gives an easy proof of Szemerédi-type theorems for translation invariant square-independent systems.
- What is the correct dependence on the U^2 norm?
- The case $\mathbb{Z}/N\mathbb{Z}$ is technically much more difficult since we lack global quadratic correlation in the inverse theorem.
- What about systems of higher Cauchy-Schwarz complexity? What about cube-independent systems?
- What about polynomial averages?

Bibliography

- B. Host and B. Kra. *Nonconventional ergodic averages and nilmanifolds*, Annals of Math. 161, 1 (2005) 397-488.
- W.T. Gowers and J. Wolf. *The true complexity of a system of linear equations*, 2007. Available at http://arxiv.org/abs/0711.0185.
- W.T. Gowers and J. Wolf. *Linear forms and uniformity for functions on* $\mathbb{F}_p^n$, *2009*. In preparation.
- B.J. Green. *Montreal lecture notes on quadratic Fourier analysis*, 2006. Available at http://arxiv.org/abs/math/0604089.
- A. Leibman. *Orbit of the diagonal of a power of a nilmanifold*, 2007. Available at http://www.math.ohio-state.edu/∼Leibman.