The Graph Regularity Method

Jacob Fox
Stanford University

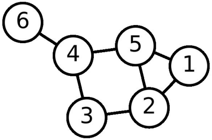
Marston Morse Lecture Series
Institute for Advanced Studies

October 24, 2016

Graphs

Definition (graph)

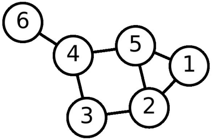
A *graph* $G = (V, E)$ has a vertex set V and a set E of edges, which are pairs of vertices.



Graphs

Definition (graph)

A *graph* $G = (V, E)$ has a vertex set V and a set E of edges, which are pairs of vertices.



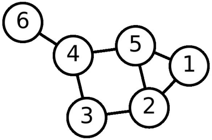
Examples:

- The Internet with computers connected by links

Graphs

Definition (graph)

A *graph* $G = (V, E)$ has a vertex set V and a set E of edges, which are pairs of vertices.



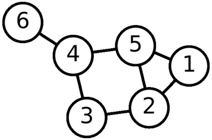
Examples:

- The Internet with computers connected by links
- The World Wide Web with webpages and hyperlinks

Graphs

Definition (graph)

A *graph* $G = (V, E)$ has a vertex set V and a set E of edges, which are pairs of vertices.



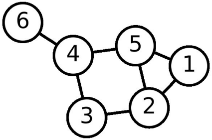
Examples:

- The Internet with computers connected by links
- The World Wide Web with webpages and hyperlinks
- Social networks like Facebook with users and friendships

Graphs

Definition (graph)

A *graph* $G = (V, E)$ has a vertex set V and a set E of edges, which are pairs of vertices.



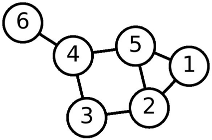
Examples:

- The Internet with computers connected by links
- The World Wide Web with webpages and hyperlinks
- Social networks like Facebook with users and friendships
- Chemical networks with atoms and chemical bonds

Graphs

Definition (graph)

A *graph* $G = (V, E)$ has a vertex set V and a set E of edges, which are pairs of vertices.



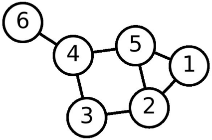
Examples:

- The Internet with computers connected by links
- The World Wide Web with webpages and hyperlinks
- Social networks like Facebook with users and friendships
- Chemical networks with atoms and chemical bonds
- Biological networks like the brain with neurons and synapses

Graphs

Definition (graph)

A *graph* $G = (V, E)$ has a vertex set V and a set E of edges, which are pairs of vertices.



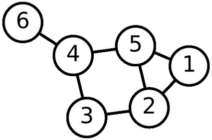
Examples:

- The Internet with computers connected by links
- The World Wide Web with webpages and hyperlinks
- Social networks like Facebook with users and friendships
- Chemical networks with atoms and chemical bonds
- Biological networks like the brain with neurons and synapses
- Engineered networks like chips with transistors and wires

Graphs

Definition (graph)

A *graph* $G = (V, E)$ has a vertex set V and a set E of edges, which are pairs of vertices.



Examples:

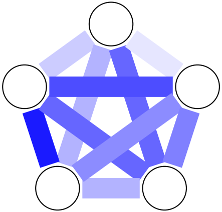
- The Internet with computers connected by links
- The World Wide Web with webpages and hyperlinks
- Social networks like Facebook with users and friendships
- Chemical networks with atoms and chemical bonds
- Biological networks like the brain with neurons and synapses
- Engineered networks like chips with transistors and wires

From *Large Networks and Graph Limits* by Lovász

Szemerédi's Regularity Lemma

Szemerédi's regularity lemma

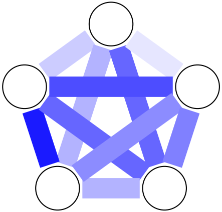
Roughly speaking, every graph can be partitioned into a bounded number of roughly equally-sized parts so that the graph is random-like between almost all pairs of parts.



Szemerédi's Regularity Lemma

Szemerédi's regularity lemma

Roughly speaking, every graph can be partitioned into a bounded number of roughly equally-sized parts so that the graph is random-like between almost all pairs of parts.

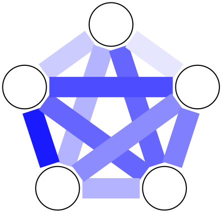


- Rough structural result for all graphs.

Szemerédi's Regularity Lemma

Szemerédi's regularity lemma

Roughly speaking, every graph can be partitioned into a bounded number of roughly equally-sized parts so that the graph is random-like between almost all pairs of parts.



- Rough structural result for all graphs.
- One of the most powerful tool in combinatorics.

Regularity

Let X and Y be vertex subsets of a graph G.

Regularity

Let X and Y be vertex subsets of a graph G.

$e(X, Y) =$ number of pairs in $X \times Y$ that are edges.

Regularity

Let X and Y be vertex subsets of a graph G.

$e(X, Y) =$ number of pairs in $X \times Y$ that are edges.

Definition (density)

$$d(X, Y) = \frac{e(X, Y)}{|X|\,|Y|}$$

Regularity

Let X and Y be vertex subsets of a graph G.

$e(X, Y)$ = number of pairs in $X \times Y$ that are edges.

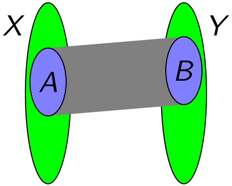
Definition (density)

$$d(X, Y) = \frac{e(X, Y)}{|X|\,|Y|}$$

Definition (irregularity)

$\mathrm{irreg}(X, Y)$ is the maximum over all $A \subset X$ and $B \subset Y$ of

$$\Big| e(A, B) - d(X, Y)|A||B| \Big|$$



Regularity

Let X and Y be vertex subsets of a graph G.

$e(X, Y)$ = number of pairs in $X \times Y$ that are edges.

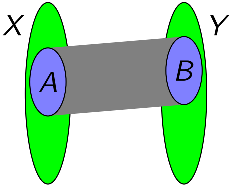
Definition (density)

$$d(X, Y) = \frac{e(X, Y)}{|X|\,|Y|}$$

Definition (irregularity)

$\operatorname{irreg}(X, Y)$ is the maximum over all $A \subset X$ and $B \subset Y$ of

$$\Big| e(A, B) - d(X, Y)|A||B| \Big|$$



(X, Y) is *ε-regular* if $\operatorname{irreg}(X, Y) \leq \varepsilon |X||Y|$.

Szemerédi's regularity lemma

Definition (Partition irregularity)

The *irregularity* of a vertex partition P of a graph $G = (V, E)$ is

$$\mathrm{irreg}(P) := \sum_{X,Y \in P} \mathrm{irreg}(X, Y)$$

Szemerédi's regularity lemma

Definition (Partition irregularity)

The *irregularity* of a vertex partition P of a graph $G = (V, E)$ is

$$\mathrm{irreg}(P) := \sum_{X,Y \in P} \mathrm{irreg}(X, Y)$$

Partition P is ε*-regular* if $\mathrm{irreg}(P) \leq \varepsilon |V|^2$.

Szemerédi's regularity lemma

Definition (Partition irregularity)

The *irregularity* of a vertex partition P of a graph $G = (V, E)$ is

$$\mathrm{irreg}(P) := \sum_{X,Y \in P} \mathrm{irreg}(X, Y)$$

Partition P is ε-*regular* if $\mathrm{irreg}(P) \leq \varepsilon |V|^2$.

Szemerédi's regularity lemma

For every $\varepsilon > 0$, there is an $M(\varepsilon)$ so that every graph G has an ε-regular vertex partition P with at most $M(\varepsilon)$ parts.

Number of parts

Szemerédi's regularity lemma

Every graph has an ε-regular partition into at most $M(\varepsilon)$ parts.

Number of parts

Szemerédi's regularity lemma

Every graph has an ε-regular partition into at most $M(\varepsilon)$ parts.

Question

How big is $M(\varepsilon)$?

Number of parts

Szemerédi's regularity lemma

Every graph has an ε-regular partition into at most $M(\varepsilon)$ parts.

Question

How big is $M(\varepsilon)$?

Tower function $T(n)$ is given by $T(1) = 2$ and $T(n) = 2^{T(n-1)}$.

Number of parts

Szemerédi's regularity lemma

Every graph has an ε-regular partition into at most $M(\varepsilon)$ parts.

Question

How big is $M(\varepsilon)$?

Tower function $T(n)$ is given by $T(1) = 2$ and $T(n) = 2^{T(n-1)}$.

The proof of the regularity lemma gives $M(\varepsilon) \leq T(\varepsilon^{-2})$.

Number of parts

Szemerédi's regularity lemma

Every graph has an ε-regular partition into at most $M(\varepsilon)$ parts.

Question

How big is $M(\varepsilon)$?

Tower function $T(n)$ is given by $T(1) = 2$ and $T(n) = 2^{T(n-1)}$.

The proof of the regularity lemma gives $M(\varepsilon) \leq T(\varepsilon^{-2})$.

Gowers proved $M(\varepsilon) \geq T(\varepsilon^{-c})$ for some constant $c > 0$.

Number of parts

Szemerédi's regularity lemma

Every graph has an ε-regular partition into at most $M(\varepsilon)$ parts.

Question

How big is $M(\varepsilon)$?

Tower function $T(n)$ is given by $T(1) = 2$ and $T(n) = 2^{T(n-1)}$.

The proof of the regularity lemma gives $M(\varepsilon) \leq T(\varepsilon^{-2})$.

Gowers proved $M(\varepsilon) \geq T(\varepsilon^{-c})$ for some constant $c > 0$.

Question (Gowers 1997)

Determine the order of the tower height of $M(\varepsilon)$.

Number of parts

Szemerédi's regularity lemma

Every graph has an ε-regular partition into at most $M(\varepsilon)$ parts.

Question

How big is $M(\varepsilon)$?

Tower function $T(n)$ is given by $T(1) = 2$ and $T(n) = 2^{T(n-1)}$.

The proof of the regularity lemma gives $M(\varepsilon) \leq T(\varepsilon^{-2})$.

Gowers proved $M(\varepsilon) \geq T(\varepsilon^{-c})$ for some constant $c > 0$.

Question (Gowers 1997)

Determine the order of the tower height of $M(\varepsilon)$.

Theorem (F.-L. M. Lovász 2016+)

$$M(\varepsilon) = T(\Theta(\varepsilon^{-2}))$$

Szemerédi's regularity lemma: proof idea

Definition (Mean square density)

For a vertex partition $P : V = V_1 \cup \ldots \cup V_k$,

$$q(P) = \sum_{i,j} p_i p_j d(V_i, V_j)^2$$

where $p_i = \frac{|V_i|}{|V|}$.

Szemerédi's regularity lemma: proof idea

Definition (Mean square density)

For a vertex partition $P : V = V_1 \cup \ldots \cup V_k$,

$$q(P) = \sum_{i,j} p_i p_j d(V_i, V_j)^2$$

where $p_i = \frac{|V_i|}{|V|}$.

Properties:

- $0 \leq q(P) \leq 1$.
- If P' is a refinement of P, then $q(P') \geq q(P)$.

Szemerédi's regularity lemma: proof idea

Definition (Mean square density)

For a vertex partition $P : V = V_1 \cup \ldots \cup V_k$,

$$q(P) = \sum_{i,j} p_i p_j d(V_i, V_j)^2$$

where $p_i = \frac{|V_i|}{|V|}$.

Properties:

- $0 \leq q(P) \leq 1$.
- If P' is a refinement of P, then $q(P') \geq q(P)$.

Claim

If P with $|P| = k$ is not ε-regular, then there is a refinement P' into at most $k2^{k+1}$ parts such that $q(P') \geq q(P) + \varepsilon^2$.

Szemerédi's regularity lemma: proof idea

Definition (Mean square density)

For a vertex partition $P : V = V_1 \cup \ldots \cup V_k$,

$$q(P) = \sum_{i,j} p_i p_j d(V_i, V_j)^2$$

where $p_i = \frac{|V_i|}{|V|}$.

Properties:

- $0 \leq q(P) \leq 1$.
- If P' is a refinement of P, then $q(P') \geq q(P)$.

Claim

If P with $|P| = k$ is not ε-regular, then there is a refinement P' into at most $k2^{k+1}$ parts such that $q(P') \geq q(P) + \varepsilon^2$.

At most ε^{-2} iterations before obtaining an ε-regular partition.

Regularity method

Regularity method

1. Apply Szemerédi's regularity lemma.
2. Use a counting lemma for embedding small graphs.

Regularity method

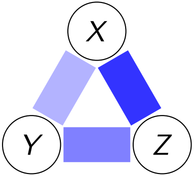
Regularity method

1. Apply Szemerédi's regularity lemma.
2. Use a counting lemma for embedding small graphs.

Triangle counting lemma

If each pair of parts is ε-regular, the number of triangles across the three parts is

$$\approx d(X,Y)d(X,Z)d(Y,Z)\,|X|\,|Y|\,|Z|\,.$$



Triangle removal lemma

Triangle removal lemma (Ruzsa-Szemerédi 1976)

For any $\varepsilon > 0$ there is a $\delta > 0$ such that every n-vertex graph with $\leq \delta n^3$ triangles can be made triangle-free by removing εn^2 edges.

Triangle removal lemma

Triangle removal lemma (Ruzsa-Szemerédi 1976)

For any $\varepsilon > 0$ there is a $\delta > 0$ such that every n-vertex graph with $\leq \delta n^3$ triangles can be made triangle-free by removing εn^2 edges.

Has many applications in extremal graph theory, additive number theory, theoretical computer science, and discrete geometry.

Triangle removal lemma

Triangle removal lemma (Ruzsa-Szemerédi 1976)

For any $\varepsilon > 0$ there is a $\delta > 0$ such that every n-vertex graph with $\leq \delta n^3$ triangles can be made triangle-free by removing εn^2 edges.

Has many applications in extremal graph theory, additive number theory, theoretical computer science, and discrete geometry.

Proof idea:

1. Apply Szemerédi's regularity lemma.

Triangle removal lemma

Triangle removal lemma (Ruzsa-Szemerédi 1976)

For any $\varepsilon > 0$ there is a $\delta > 0$ such that every n-vertex graph with $\leq \delta n^3$ triangles can be made triangle-free by removing εn^2 edges.

Has many applications in extremal graph theory, additive number theory, theoretical computer science, and discrete geometry.

Proof idea:

1. Apply Szemerédi's regularity lemma.
2. Delete edges between pairs which are irregular or sparse.

Triangle removal lemma

Triangle removal lemma (Ruzsa-Szemerédi 1976)

For any $\varepsilon > 0$ there is a $\delta > 0$ such that every n-vertex graph with $\leq \delta n^3$ triangles can be made triangle-free by removing εn^2 edges.

Has many applications in extremal graph theory, additive number theory, theoretical computer science, and discrete geometry.

Proof idea:

1. Apply Szemerédi's regularity lemma.
2. Delete edges between pairs which are irregular or sparse.
3. If there is a remaining triangle, then its edges go between pairs which are both dense and regular. The counting lemma then implies that there are more than δn^3 triangles.

Triangle removal lemma

Triangle removal lemma (Ruzsa-Szemerédi 1976)

For any $\varepsilon > 0$ there is a $\delta > 0$ such that every n-vertex graph with $\leq \delta n^3$ triangles can be made triangle-free by removing εn^2 edges.

Triangle removal lemma

Triangle removal lemma (Ruzsa-Szemerédi 1976)

For any $\varepsilon > 0$ there is a $\delta > 0$ such that every n-vertex graph with $\leq \delta n^3$ triangles can be made triangle-free by removing εn^2 edges.

Problem (Erdős, Alon, Gowers, Tao)

Find a new proof which gives a better bound.

Triangle removal lemma

Triangle removal lemma (Ruzsa-Szemerédi 1976)

For any $\varepsilon > 0$ there is a $\delta > 0$ such that every n-vertex graph with $\leq \delta n^3$ triangles can be made triangle-free by removing εn^2 edges.

Problem (Erdős, Alon, Gowers, Tao)

Find a new proof which gives a better bound.

Theorem (F. 2011)

We may take δ^{-1} to be a tower of twos of height $\log \varepsilon^{-1}$.

Triangle removal lemma

Triangle removal lemma (Ruzsa-Szemerédi 1976)

For any $\varepsilon > 0$ there is a $\delta > 0$ such that every n-vertex graph with $\leq \delta n^3$ triangles can be made triangle-free by removing εn^2 edges.

Problem (Erdős, Alon, Gowers, Tao)

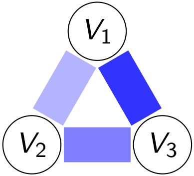
Find a new proof which gives a better bound.

Theorem (F. 2011)

We may take δ^{-1} to be a tower of twos of height $\log \varepsilon^{-1}$.

Best known lower bound on δ^{-1} is only $\varepsilon^{-c \log \varepsilon^{-1}}$.

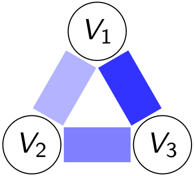
Triangle removal lemma: new proof idea



Key lemma

If there are at most $\alpha|V_1||V_2||V_3|$ triangles across V_1, V_2, V_3,

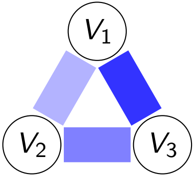
Triangle removal lemma: new proof idea



Key lemma

If there are at most $\alpha|V_1||V_2||V_3|$ triangles across V_1, V_2, V_3, then there are $1 \leq i < j \leq 3$ and equitable partitions Q_i of V_i and Q_j of V_j each with at most $2^{\alpha^{-O(1)}}$ parts such that there are at least $\frac{1}{10}|Q_i||Q_j|$ pairs $(X, Y) \in Q_i \times Q_j$ with $d(X, Y) < 2\alpha^{1/3}$.

Triangle removal lemma: new proof idea



Key lemma

If there are at most $\alpha|V_1||V_2||V_3|$ triangles across V_1, V_2, V_3, then there are $1 \leq i < j \leq 3$ and equitable partitions Q_i of V_i and Q_j of V_j each with at most $2^{\alpha^{-O(1)}}$ parts such that there are at least $\frac{1}{10}|Q_i||Q_j|$ pairs $(X, Y) \in Q_i \times Q_j$ with $d(X, Y) < 2\alpha^{1/3}$.

Definition: Mean entropy density

For a vertex partition $P : V = V_1 \cup \ldots \cup V_k$,

$$h(P) = \sum_{i,j} p_i p_j d(V_i, V_j) \log d(V_i, V_j)$$

where $p_i = \frac{|V_i|}{|V|}$.

Property Testing

Property Testing

Goal

Quickly distinguish between objects that have a property from objects that are far from having that property.

Property Testing

Goal

Quickly distinguish between objects that have a property from objects that are far from having that property.

Introduced by Rubinfeld and Sudan in 1996 and investigated by Goldreich, Goldwasser, and Ron.

Property Testing

Goal

Quickly distinguish between objects that have a property from objects that are far from having that property.

Introduced by Rubinfeld and Sudan in 1996 and investigated by Goldreich, Goldwasser, and Ron.

Testable graph properties

A graph property P is *testable* if for each $\varepsilon > 0$ there is a randomized algorithm which in constant time (depending on ε and P) distinguishes with probability at least $2/3$ between graphs having P and graphs which are ε-far from having P.

What is testable?

Claim

Triangle-freeness is testable.

What is testable?

Claim

Triangle-freeness is testable.

Proof: Pick $1/\delta$ random triples of vertices.

What is testable?

Claim

Triangle-freeness is testable.

Proof: Pick $1/\delta$ random triples of vertices.
Accept if no triple of vertices makes a triangle; reject otherwise.

What is testable?

Claim

Triangle-freeness is testable.

Proof: Pick $1/\delta$ random triples of vertices.
Accept if no triple of vertices makes a triangle; reject otherwise.
If the graph is triangle-free, then the algorithm accepts.

What is testable?

Claim

Triangle-freeness is testable.

Proof: Pick $1/\delta$ random triples of vertices.
Accept if no triple of vertices makes a triangle; reject otherwise.
If the graph is triangle-free, then the algorithm accepts.
If the graph is ε-far from being triangle-free, then it accepts with probability $\leq (1-6\delta)^{1/\delta} \leq e^{-6} < .01$. □

What is testable?

Claim

Triangle-freeness is testable.

Proof: Pick $1/\delta$ random triples of vertices.
Accept if no triple of vertices makes a triangle; reject otherwise.
If the graph is triangle-free, then the algorithm accepts.
If the graph is ε-far from being triangle-free, then it accepts with probability $\leq (1-6\delta)^{1/\delta} \leq e^{-6} < .01$. □

Further, H-freeness is testable (by the graph removal lemma).

What is testable?

Claim

Triangle-freeness is testable.

Proof: Pick $1/\delta$ random triples of vertices.
Accept if no triple of vertices makes a triangle; reject otherwise.
If the graph is triangle-free, then the algorithm accepts.
If the graph is ε-far from being triangle-free, then it accepts with probability $\leq (1-6\delta)^{1/\delta} \leq e^{-6} < .01$. □

Further, H-freeness is testable (by the graph removal lemma).

Theorem (Alon-Fischer-Krivelevich-Szegedy 2000)

Induced H-freeness is testable.

What is testable?

Claim

Triangle-freeness is testable.

Proof: Pick $1/\delta$ random triples of vertices.
Accept if no triple of vertices makes a triangle; reject otherwise.
If the graph is triangle-free, then the algorithm accepts.
If the graph is ε-far from being triangle-free, then it accepts with probability $\leq (1-6\delta)^{1/\delta} \leq e^{-6} < .01$. □

Further, H-freeness is testable (by the graph removal lemma).

Theorem (Alon-Fischer-Krivelevich-Szegedy 2000)

Induced H-freeness is testable.

Theorem (Alon-Shapira 2008)

Every hereditary graph property is testable.

Induced graph removal lemma

Lemma (Alon-Fischer-Krivelevich-Szegedy 2000)

For each $\varepsilon > 0$ and graph H on h vertices there is $\delta > 0$ such that every graph on n vertices with at most δn^h induced copies of H can be made induced H-free by adding or removing εn^2 edges.

Induced graph removal lemma

Lemma (Alon-Fischer-Krivelevich-Szegedy 2000)

For each $\varepsilon > 0$ and graph H on h vertices there is $\delta > 0$ such that every graph on n vertices with at most δn^h induced copies of H can be made induced H-free by adding or removing εn^2 edges.

Proof developed the *strong regularity lemma*,

Induced graph removal lemma

Lemma (Alon-Fischer-Krivelevich-Szegedy 2000)

For each $\varepsilon > 0$ and graph H on h vertices there is $\delta > 0$ such that every graph on n vertices with at most δn^h induced copies of H can be made induced H-free by adding or removing εn^2 edges.

Proof developed the *strong regularity lemma*, which is proved by repeated use of Szemerédi's regularity lemma and gives a wowzer-type bound.

Induced graph removal lemma

Lemma (Alon-Fischer-Krivelevich-Szegedy 2000)

For each $\varepsilon > 0$ and graph H on h vertices there is $\delta > 0$ such that every graph on n vertices with at most δn^h induced copies of H can be made induced H-free by adding or removing εn^2 edges.

Proof developed the *strong regularity lemma*, which is proved by repeated use of Szemerédi's regularity lemma and gives a wowzer-type bound.

Wowzer function is given by $W(1) = 2$ and $W(n) = T(W(n-1))$.

Induced graph removal lemma

Lemma (Alon-Fischer-Krivelevich-Szegedy 2000)

For each $\varepsilon > 0$ and graph H on h vertices there is $\delta > 0$ such that every graph on n vertices with at most δn^h induced copies of H can be made induced H-free by adding or removing εn^2 edges.

Proof developed the *strong regularity lemma*, which is proved by repeated use of Szemerédi's regularity lemma and gives a wowzer-type bound.

Wowzer function is given by $W(1) = 2$ and $W(n) = T(W(n-1))$.

Problem (Alon)

Estimate the dependence of δ on ε.

Induced graph removal lemma

Lemma (Alon-Fischer-Krivelevich-Szegedy 2000)

For each $\varepsilon > 0$ and graph H on h vertices there is $\delta > 0$ such that every graph on n vertices with at most δn^h induced copies of H can be made induced H-free by adding or removing εn^2 edges.

Proof developed the *strong regularity lemma*, which is proved by repeated use of Szemerédi's regularity lemma and gives a wowzer-type bound.

Wowzer function is given by $W(1) = 2$ and $W(n) = T(W(n-1))$.

Problem (Alon)

Estimate the dependence of δ on ε.

In particular, can the wowzer-type bound be improved?

Wowzer and Tower

Wowzer function is given by $W(1) = 2$ and $W(n) = T(W(n-1))$.

Wowzer and Tower

Wowzer function is given by $W(1) = 2$ and $W(n) = T(W(n-1))$.

Theorem (Conlon-F. 2012)

$W(\varepsilon^{-\Omega(1)})$ parts are needed in the strong regularity lemma.

Wowzer and Tower

Wowzer function is given by $W(1) = 2$ and $W(n) = T(W(n-1))$.

Theorem (Conlon-F. 2012)

$W(\varepsilon^{-\Omega(1)})$ parts are needed in the strong regularity lemma.

Lemma (Alon-Fischer-Krivelevich-Szegedy 2000)

For each $\varepsilon > 0$ and graph H on h vertices there is $\delta > 0$ such that every graph on n vertices with at most δn^h induced copies of H can be made H-free by adding or removing εn^2 edges.

Wowzer and Tower

Wowzer function is given by $W(1) = 2$ and $W(n) = T(W(n-1))$.

Theorem (Conlon-F. 2012)

$W(\varepsilon^{-\Omega(1)})$ parts are needed in the strong regularity lemma.

Lemma (Alon-Fischer-Krivelevich-Szegedy 2000)

For each $\varepsilon > 0$ and graph H on h vertices there is $\delta > 0$ such that every graph on n vertices with at most δn^h induced copies of H can be made H-free by adding or removing εn^2 edges.

Theorem (Conlon-F. 2012)

We can take $\delta^{-1} = T(\varepsilon^{-O(1)})$ in the induced graph removal lemma.

Wowzer and Tower

Wowzer function is given by $W(1) = 2$ and $W(n) = T(W(n-1))$.

Theorem (Conlon-F. 2012)

$W(\varepsilon^{-\Omega(1)})$ parts are needed in the strong regularity lemma.

Lemma (Alon-Fischer-Krivelevich-Szegedy 2000)

For each $\varepsilon > 0$ and graph H on h vertices there is $\delta > 0$ such that every graph on n vertices with at most δn^h induced copies of H can be made H-free by adding or removing εn^2 edges.

Theorem (Conlon-F. 2012)

We can take $\delta^{-1} = T(\varepsilon^{-O(1)})$ in the induced graph removal lemma.

Further, a tower-type bound holds for testing hereditary properties.

Algorithmic regularity lemma

Question

Can the regularity lemma be made algorithmic?

Algorithmic regularity lemma

Question

Can the regularity lemma be made algorithmic?

Importance lies in that it would give approximation algorithms for a large class of computationally hard problems.

Algorithmic regularity lemma

Question

Can the regularity lemma be made algorithmic?

Importance lies in that it would give approximation algorithms for a large class of computationally hard problems.

Bottleneck: Determine if a pair of parts X, Y is ε-regular, and, if not, find $X' \subset X$, $Y' \subset Y$ realizing this.

Algorithmic regularity lemma

Question

Can the regularity lemma be made algorithmic?

Importance lies in that it would give approximation algorithms for a large class of computationally hard problems.

Bottleneck: Determine if a pair of parts X, Y is ε-regular, and, if not, find $X' \subset X$, $Y' \subset Y$ realizing this.

Theorem (Alon-Duke-Lefmann-Rödl-Yuster)

Yes!

Algorithmic regularity lemma

Question

Can the regularity lemma be made algorithmic?

Importance lies in that it would give approximation algorithms for a large class of computationally hard problems.

Bottleneck: Determine if a pair of parts X, Y is ε-regular, and, if not, find $X' \subset X$, $Y' \subset Y$ realizing this.

Theorem (Alon-Duke-Lefmann-Rödl-Yuster)

Yes!

Up to changing ε, checking regularity is the same as showing the density of C_4 with edges across $X \times Y$ is $\approx d(X, Y)^4$.

Algorithmic regularity lemma

Question

Can the regularity lemma be made algorithmic?

Importance lies in that it would give approximation algorithms for a large class of computationally hard problems.
Bottleneck: Determine if a pair of parts X, Y is ε-regular, and, if not, find $X' \subset X$, $Y' \subset Y$ realizing this.

Theorem (Alon-Duke-Lefmann-Rödl-Yuster)

Yes!

Up to changing ε, checking regularity is the same as showing the density of C_4 with edges across $X \times Y$ is $\approx d(X, Y)^4$.

Theorem (Kohayakawa-Rödl-Thoma)

Can be done in $O(n^2)$ time.

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices.

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices.

Zuckerman showed NP-hard to approximate the size of the largest clique within a factor $n^{1-\varepsilon}$.

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices.

Zuckerman showed NP-hard to approximate the size of the largest clique within a factor $n^{1-\varepsilon}$.

How fast can we approximate the count within an additive εn^k?

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices within an additive εn^k.

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices within an additive εn^k.

A simple randomized algorithm gives 99% certainty:

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices within an additive εn^k.

A simple randomized algorithm gives 99% certainty:

Sample $10/\varepsilon^2$ random k-sets of vertices.

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices within an additive εn^k.

A simple randomized algorithm gives 99% certainty:

Sample $10/\varepsilon^2$ random k-sets of vertices.

What about deterministic algorithms?

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices within an additive εn^k.

A simple randomized algorithm gives 99% certainty:

Sample $10/\varepsilon^2$ random k-sets of vertices.

What about deterministic algorithms?

Can use the regularity method!

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices within an additive εn^k deterministically.

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices within an additive εn^k deterministically.

Duke, Lefmann, Rödl 1996: Can be done in time $2^{(k/\varepsilon)^{O(1)}} n^{\omega}$.

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices within an additive εn^k deterministically.

Duke, Lefmann, Rödl 1996: Can be done in time $2^{(k/\varepsilon)^{O(1)}} n^{\omega}$.

Theorem (F.-L. M. Lovász-Zhao 2016)

Can be done in time $O(\varepsilon^{-k^2} n + n^{\omega})$.

Counting cliques

Algorithmic problem

Count the number of cliques of order k in a graph on n vertices within an additive εn^k deterministically.

Duke, Lefmann, Rödl 1996: Can be done in time $2^{(k/\varepsilon)^{O(1)}} n^{\omega}$.

Theorem (F.-L. M. Lovász-Zhao 2016)

Can be done in time $O(\varepsilon^{-k^2} n + n^{\omega})$.

Corollary

We can approximate the count of K_{1000} in a graph on n vertices within an additive $n^{1000-10^{-6}}$ in time $O(n^{2.4})$.

Regularity method for sparse graphs?

The original regularity method is only useful for dense graphs.

Regularity method for sparse graphs?

The original regularity method is only useful for dense graphs.

Kohayakawa, Rödl, and Scott proved a sparse regularity lemma.

Regularity method for sparse graphs?

The original regularity method is only useful for dense graphs.

Kohayakawa, Rödl, and Scott proved a sparse regularity lemma.

Main Problem

Prove a counting lemma in sparse graphs.

Regularity method for sparse graphs?

The original regularity method is only useful for dense graphs.

Kohayakawa, Rödl, and Scott proved a sparse regularity lemma.

Main Problem

Prove a counting lemma in sparse graphs.

Caveat: A general counting lemma in sparse graphs cannot hold.

Regularity method for sparse graphs?

The original regularity method is only useful for dense graphs.

Kohayakawa, Rödl, and Scott proved a sparse regularity lemma.

Main Problem

Prove a counting lemma in sparse graphs.

Caveat: A general counting lemma in sparse graphs cannot hold.

How about within subgraphs of sparse pseudorandom graphs?

Regularity method for sparse graphs?

The original regularity method is only useful for dense graphs.

Kohayakawa, Rödl, and Scott proved a sparse regularity lemma.

Main Problem

Prove a counting lemma in sparse graphs.

Caveat: A general counting lemma in sparse graphs cannot hold.

How about within subgraphs of sparse pseudorandom graphs?

Regularity method for sparse graphs?

The original regularity method is only useful for dense graphs.

Kohayakawa, Rödl, and Scott proved a sparse regularity lemma.

Main Problem

Prove a counting lemma in sparse graphs.

Caveat: A general counting lemma in sparse graphs cannot hold.

How about within subgraphs of sparse pseudorandom graphs?

Theorem (Conlon-F.-Zhao)

A sparse counting lemma in graphs and hypergraphs.

Green–Tao Theorem (2008)

The primes contain arbitrarily long arithmetic progressions.

Examples:

- 3, 5, 7
- 5, 11, 17, 23, 29
- 7, 37, 67, 97, 127, 157
- Longest known: 26 terms

Green–Tao Theorem (2008)

The primes contain arbitrarily long arithmetic progressions (AP).

Szemerédi's Theorem (1975)

Every positive density subset of contains arbitrarily long APs.

(upper) density of $A \subset$ is $\limsup_{N\to\infty} \dfrac{|A \cap [N]|}{N}$

$[N] := \{1, 2, \ldots, N\}$

$P =$ prime numbers

Prime number theorem: $\dfrac{|P \cap [N]|}{N} \sim \dfrac{1}{\log N}$

Proof strategy of the Green–Tao theorem

Step 1:

Relative Szemerédi theorem (informally)

If $S \subset$ satisfies certain pseudorandomness conditions, then every subset of S of relative positive density contains long APs.

Proof strategy of the Green–Tao theorem

Step 1:

Relative Szemerédi theorem (informally)

If $S \subset$ satisfies certain pseudorandomness conditions, then every subset of S of relative positive density contains long APs.

Step 2: Construct a superset of the primes satisfying these conditions.

Proof strategy of the Green–Tao theorem

Step 1:

Relative Szemerédi theorem (informally)

If $S \subset$ satisfies certain pseudorandomness conditions, then every subset of S of relative positive density contains long APs.

Step 2: Construct a superset of the primes satisfying these conditions.

$P =$ prime numbers, $Q =$ "almost primes"

$P \subseteq Q$ with relative positive density, i.e., $\dfrac{|P \cap [N]|}{|Q \cap [N]|} > \delta$

Relative Szemerédi theorem (informally)

If $S \subset$ satisfies certain pseudorandomness conditions, then every subset of S of positive density contains long APs.

What pseudorandomness conditions?

Green–Tao:

1. Linear forms condition
2. Correlation condition

Relative Szemerédi theorem

Relative Szemerédi theorem (informally)

If $S \subset$ satisfies certain pseudorandomness conditions, then every subset of S of positive density contains long APs.

What pseudorandomness conditions?

Green–Tao:

1. Linear forms condition
2. Correlation condition

A natural question (asked by Green, Gowers, . . .)

Does a relative Szemerédi theorem hold with weaker and more natural hypotheses?

Relative Szemerédi theorem

Relative Szemerédi theorem (informally)

If $S \subset$ satisfies certain pseudorandomness conditions, then every subset of S of positive density contains long APs.

What pseudorandomness conditions?

Green–Tao:

1. Linear forms condition
2. Correlation condition $\leftarrow$ no longer needed

A natural question (asked by Green, Gowers, . . .)

Does a relative Szemerédi theorem hold with weaker and more natural hypotheses?

Our main result

Yes! A weak linear forms condition suffices.

Roth's theorem

Roth's theorem (1952)

If $A \subseteq [N]$ is 3-AP-free, then $|A| = o(N)$.

$[N] := \{1, 2, \ldots, N\}$

3-AP = 3-term arithmetic progression

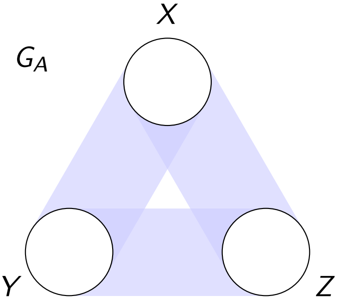
It'll be easier (and equivalent) to work in $\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}$.

Proof of Roth's theorem

Roth's theorem (1952)

If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Given A, construct tripartite graph G_A with vertex sets $X = Y = Z = \mathbb{Z}_N$.

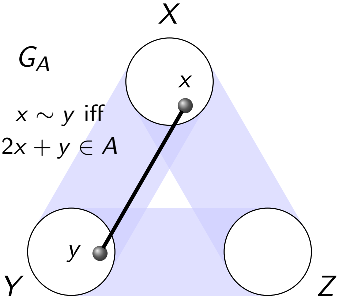


Proof of Roth's theorem

Roth's theorem (1952)

If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Given A, construct tripartite graph G_A with vertex sets $X = Y = Z = \mathbb{Z}_N$.

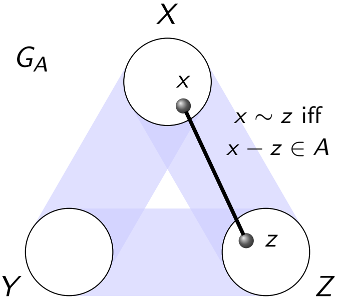


Proof of Roth's theorem

Roth's theorem (1952)

If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

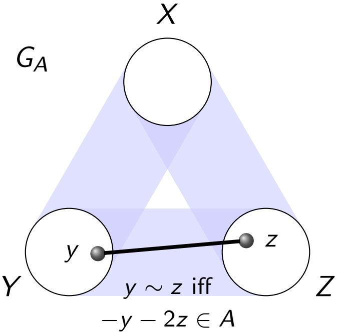
Given A, construct tripartite graph G_A with vertex sets $X = Y = Z = \mathbb{Z}_N$.



Roth's theorem (1952)

If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Given A, construct tripartite graph G_A with vertex sets $X = Y = Z = \mathbb{Z}_N$.

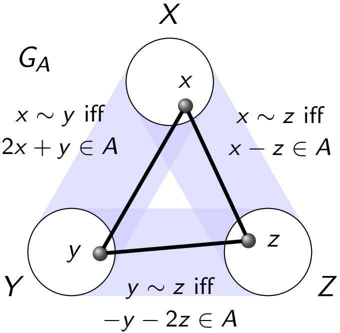


Proof of Roth's theorem

Roth's theorem (1952)

If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Given A, construct tripartite graph G_A with vertex sets $X = Y = Z = \mathbb{Z}_N$.



Proof of Roth's theorem

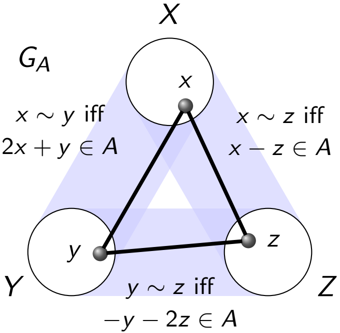
Roth's theorem (1952)

If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Given A, construct tripartite graph G_A with vertex sets $X = Y = Z = \mathbb{Z}_N$.

Triangle xyz in $G_A \iff$

$2x + y,\ x - z,\ -y - 2z\ \in A$



Roth's theorem (1952)

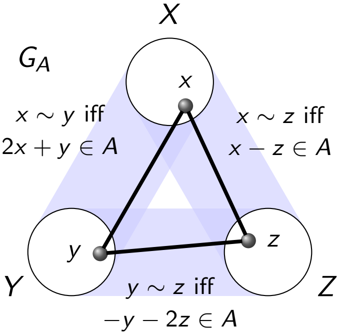
If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Given A, construct tripartite graph G_A with vertex sets $X = Y = Z = \mathbb{Z}_N$.

Triangle xyz in G_A $\iff$

$2x + y,\ x - z,\ -y - 2z \in A$

It's a 3-AP with diff $-x - y - z$



Proof of Roth's theorem

Roth's theorem (1952)

If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

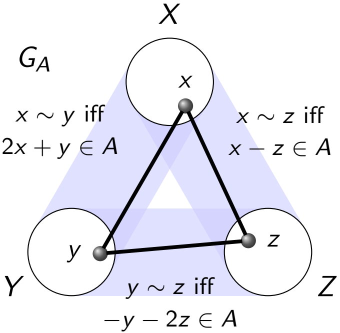
Given A, construct tripartite graph G_A with vertex sets $X = Y = Z = \mathbb{Z}_N$.

Triangle xyz in $G_A \iff$

$2x + y,\ x - z,\ -y - 2z \in A$

It's a 3-AP with diff $-x - y - z$

No triangles?



Proof of Roth's theorem

Roth's theorem (1952)

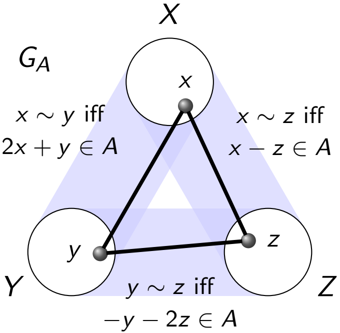
If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Given A, construct tripartite graph G_A with vertex sets $X = Y = Z = \mathbb{Z}_N$.

Triangle xyz in G_A $\iff$

$2x + y,\ x - z,\ -y - 2z \in A$

It's a 3-AP with diff $-x - y - z$



No triangles? Only triangles $\longleftrightarrow$ trivial 3-APs with diff 0.

Proof of Roth's theorem

Roth's theorem (1952)

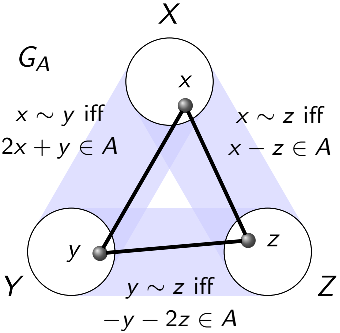
If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Given A, construct tripartite graph G_A with vertex sets $X = Y = Z = \mathbb{Z}_N$.

Triangle xyz in $G_A \iff$

$2x + y,\ x - z,\ -y - 2z \in A$

It's a 3-AP with diff $-x - y - z$



No triangles? Only triangles $\longleftrightarrow$ trivial 3-APs with diff 0.
Every edge of the graph is contained in exactly one triangle (the one with $x + y + z = 0$).

Proof of Roth's theorem

Roth's theorem (1952)

If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Constructed a graph with

- $3N$ vertices
- $3N|A|$ edges
- every edge in exactly one triangle

Proof of Roth's theorem

Roth's theorem (1952)

If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Constructed a graph with

- $3N$ vertices
- $3N|A|$ edges
- every edge in exactly one triangle

Theorem (Ruzsa & Szemerédi '76)

If every edge in a graph $G = (V, E)$ is contained in exactly one triangle, then $|E| = o(|V|^2)$.

(a consequence of the *triangle removal lemma*)

So $3N|A| = o(N^2)$. Thus $|A| = o(N)$.

Relative Roth theorem

Roth's theorem (1952)

If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Relative Roth theorem (Conlon, F., Zhao)

If $S \subseteq \mathbb{Z}_N$ satisfies the 3-linear forms condition,
and $A \subseteq S$ is 3-AP-free, then $|A| = o(|S|)$.

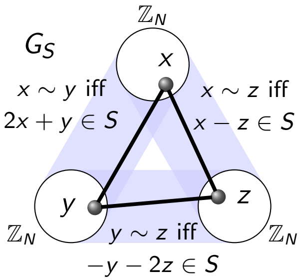
Relative Roth theorem

Roth's theorem (1952)

If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Relative Roth theorem (Conlon, F., Zhao)

If $S \subseteq \mathbb{Z}_N$ satisfies the 3-linear forms condition,
and $A \subseteq S$ is 3-AP-free, then $|A| = o(|S|)$.



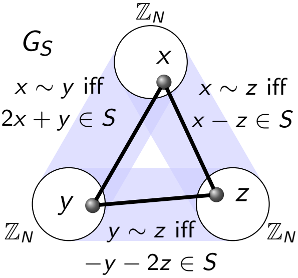
Relative Roth theorem

Roth's theorem (1952)

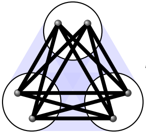
If $A \subseteq \mathbb{Z}_N$ is 3-AP-free, then $|A| = o(N)$.

Relative Roth theorem (Conlon, F., Zhao)

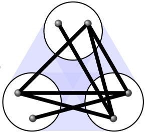
If $S \subseteq \mathbb{Z}_N$ satisfies the 3-linear forms condition, and $A \subseteq S$ is 3-AP-free, then $|A| = o(|S|)$.



3-linear forms condition:
G_S has asymp. the expected number of embeddings of $K_{2,2,2}$ & its subgraphs (compared to random graph of same density)



$K_{2,2,2}$ & subgraphs, e.g.,



Relative Szemerédi theorem

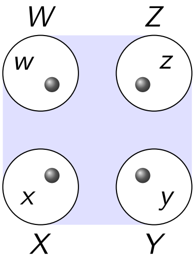
Relative Szemerédi theorem (Conlon, F., Zhao)

Fix $k \geq 3$. If $S \subseteq \mathbb{Z}_N$ satisfies the k-linear forms condition, and $A \subseteq S$ is k-AP-free, then $|A| = o(|S|)$.

$k = 4$: build a weighted 4-partite 3-uniform hypergraph

Vertex sets $W = X = Y = Z = \mathbb{Z}_N$

- $xyz \in E \iff 3w + 2x + y \in S$
- $wxz \in E \iff 2w + x - z \in S$
- $wyz \in E \iff w - y - 2z \in S$
- $xyz \in E \iff -x - 2y - 3z \in S$



common diff: $-w - x - y - z$

Relative Szemerédi theorem

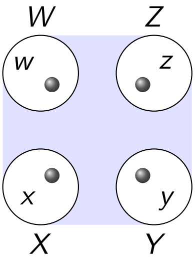
Relative Szemerédi theorem (Conlon, F., Zhao)

Fix $k \geq 3$. If $S \subseteq \mathbb{Z}_N$ satisfies the k-linear forms condition, and $A \subseteq S$ is k-AP-free, then $|A| = o(|S|)$.

$k = 4$: build a weighted 4-partite 3-uniform hypergraph

Vertex sets $W = X = Y = Z = \mathbb{Z}_N$

- $xyz \in E \iff 3w + 2x + y \in S$
- $wxz \in E \iff 2w + x - z \in S$
- $wyz \in E \iff w - y - 2z \in S$
- $xyz \in E \iff -x - 2y - 3z \in S$



common diff: $-w - x - y - z$

4-linear forms condition: correct count of the 2-blow-up of the simplex $K_4^{(3)}$ (as well as its subgraphs)

Avoiding regularity

A major drawback of the regularity lemma is the tower dependence.

Problem

Find alternative proofs of the applications which avoid using the regularity lemma and give improved bounds.

Preview of the next lecture

Preview of the next lecture

Green developed an arithmetic regularity lemma and used it to prove the following two extensions of Roth's theorem:

Preview of the next lecture

Green developed an arithmetic regularity lemma and used it to prove the following two extensions of Roth's theorem:

Green's Arithmetic Removal Lemma

For each $\varepsilon > 0$ there is $\delta > 0$ such that if G is an abelian group and $A, B, C \subset G$ with at most $\delta|G|^2$ triples $(a, b, c) \in A \times B \times C$ with $a + b + c = 0$, then we can delete $\varepsilon|G|$ elements from A, B, C and get rid of all solutions.

Preview of the next lecture

Green developed an arithmetic regularity lemma and used it to prove the following two extensions of Roth's theorem:

Green's Arithmetic Removal Lemma

For each $\varepsilon > 0$ there is $\delta > 0$ such that if G is an abelian group and $A, B, C \subset G$ with at most $\delta|G|^2$ triples $(a, b, c) \in A \times B \times C$ with $a + b + c = 0$, then we can delete $\varepsilon|G|$ elements from A, B, C and get rid of all solutions.

Green's Roth theorem with popular differences

For each $\varepsilon > 0$ there is $N(\varepsilon)$ such that if G is an abelian group with $|G| \geq N(\varepsilon)$ and $A \subset G$ with $|A| = \alpha|G|$, then there is a nonzero $d \in G$ such that the density of three-term arithmetic progressions with common difference d in A is at least $\alpha^3 - \varepsilon$.

Preview of the next lecture

Green developed an arithmetic regularity lemma and used it to prove the following two extensions of Roth's theorem:

Green's Arithmetic Removal Lemma

For each $\varepsilon > 0$ there is $\delta > 0$ such that if G is an abelian group and $A, B, C \subset G$ with at most $\delta |G|^2$ triples $(a, b, c) \in A \times B \times C$ with $a + b + c = 0$, then we can delete $\varepsilon |G|$ elements from A, B, C and get rid of all solutions.

Green's Roth theorem with popular differences

For each $\varepsilon > 0$ there is $N(\varepsilon)$ such that if G is an abelian group with $|G| \geq N(\varepsilon)$ and $A \subset G$ with $|A| = \alpha |G|$, then there is a nonzero $d \in G$ such that the density of three-term arithmetic progressions with common difference d in A is at least $\alpha^3 - \varepsilon$.

Next lecture: tight quantitative bounds in vector spaces.