Green's Conjecture and Testing Linear Invariant Properties

Asaf Shapira
Georgia Inst. of Tech.

Triangle Removal Lemma

[Ruzsa-Szemeredi '76]: If a graph G has $o(n^3)$ triangles, then one can remove $o(n^2)$ edges and make G triangle-free.

Main application: A short proof of *Roth's Theorem*: every $S \subseteq [n]$ of size $\Omega(n)$, contains a 3-term arithmetic progression .

Original motivation: Bounding the number of edges in a *3-uniform* hypergraph without *6* vertices carrying *3* edges.

Main tool: Szemeredi's regularity lemma.

Green's Removal Lemma

[Green '03]: If $S \subseteq [n]$ contains $o(n^2)$ solutions of $E : x+y=z$, then S can be made *E-free* by the removal of $o(n)$ elements.

Main tool: An analytic regularity lemma for *abelian groups*.

[Green '03]: If $S \subseteq [n]$ contains $o(n^{k-1})$ solutions of a homogenous linear equation E, in k unknowns, then S can be made *E-free* by the removal of $o(n)$ elements.

Green's Conjecture

Definition: Let $Mx=b$ be a set of t linear equations in p unknowns. A set $S \subseteq [n]$ is (M,b)-*free* if it contains no solution to $Mx=b$. That is, if there is no $v \in S^p$, satisfying $Mv=b$.

Conjecture [Green '03]: If $S \subseteq [n]$ has $o(n^{p-t})$ solutions to $Mx=0$, then S can be made $(M,0)$-*free* by the removal of $o(n)$ elements.

Previous Results

[*Green '03*]: If $S \subseteq [n]$ contains $o(n^2)$ solutions to $E : x+y=z$, then S can be made *E-free* by the removal of $o(n)$ elements.

Main tool: An analytic regularity lemma for *abelian groups*.

Holds for every *single homogenous* linear equation over an *abelian group*.

[*Kral, Serra and Vena '07*]*:* A simple proof of *Green's result*. Holds over *any group* and for *non-homogenous* linear equations.

* Avoids the *analytic* regularity lemma by using the removal lemma for *directed graphs* [*Alon-S '05*].

Main Idea: Variant of *Ruzsa-Szemeredi* proof of *Roth's Theo*.

Independent Results

Conj. [Green '03]: If $S \subseteq [n]$ has $o(n^{p-t})$ solutions to $Mx=0$, then S can be made $(M,0)$-*free* be the removal of $o(n)$ elements.

[Candela '08, Kral-Serra-Vena '08]: Conjecture holds if every t columns of M are linearly independent.

A Computational Angle

[*Green '03*]: If $S \subseteq [n]$ contains $o(n^2)$ solutions to $E : x+y=z$, then S can be made *E-free* by the removal of $o(n)$ elements.

Question 1: Getting an $O(|S|^{2-c})$ algorithm for deciding if S is *E-free* is a major open problem in Comp Geometry.

Question 2: How fast can we decide whether S is *E-free*, or if one should remove εn elements from S to make it *E-free*?

Answer: If this is the case then S has $\delta(\varepsilon)n^2$ solutions to E.

* Therefore, in this case a sample of $O(1/\delta(\varepsilon))$ elements from S contains a solution to E whp.

* If S is *E-free*, then the sample will be *E-free* with prob 1.

Testing Boolean Functions

Bhattacharyya, Chen, Sudan and Xie '08:

Conjectured that a certain family of properties of Boolean functions can all be tested with a constant number of queries.

* Motivated by previous work of [*Kaufman-Sudan '08*]

* Their conjecture is "equivalent" to Green's conjecture.

[*BCSX '08*] Verified Green's conjecture for some special sets of equations $Mx=0$.

Our Main Result

Conjecture [Green '03 and Bhattacharyya et al. '08]:

Let $Mx=0$ be t linear equations in p unknowns over F. If $S \subseteq F$ has $o(n^{p-t})$ solutions to $Mx=0$, then we can remove $o(n)$ elements from S and make it $(M,0)$-free.

Theorem [S-'08]: Above conjecture holds for every set of linear equations $Mx=b$.

Corollary: Gives a testing algorithm for the properties of boolean functions studied in [*BCSX '08*].

Theorem [Kral-Serra-Vena '08]: Independently obtained the same result .

Proof Overview

Main idea: Apply results on extremal hypergraphs.

$*$ More reminiscent of questions on dense graphs *[GGR '96]*

[Ruzsa-Szemeredi '76]: If G has $o(n^3)$ triangles then G can be made triangle free by the removal of $o(n^2)$ edges.

Main application: A short proof of *Roth's Theorem*: every $S \subseteq [n]$ of size $\Omega(n)$, contains a 3-term arithmetic progression .

"We can represent the solutions of a single equation using a graph"

Hypergraph Removal Lemma

Graph Removal lemma $\Rightarrow$ Roth's Theorem

Szemeredi's Theorem: Every $S \subseteq [n]$ of size $\Omega(n)$, contains a *k-term arithmetic progression*.

Folklore belief: Cannot be proved using *removal-lemma* in *graphs*.

[*Frankl-Rodl '02*]*:* Szemeredi's theorem would follow from a removal lemma for *hypergraphs*.

"We can represent some sets of equations using hypergraphs"

The Overall Strategy

[FR '02]: Removal lemma for *hypergraphs* $\Rightarrow$ Szemeredi's theorem.

Hypergraph Removal Lemma: If a *k-uniform* hypergraph *G* contains $o(n^h)$ copies of *H*, then we can remove $o(n^k)$ edges and thus make it *H-free*.

Obtained recently by [*Gowers '07, Rodl et al. '06, Tao '06*].

$\Downarrow$ (strategy)

Theorem [S 08]: Let $Mx=b$ be t linear equations in p unknowns over F. If $S \subseteq F$ contains $o(n^{p-t})$ solutions to $Mx=b$, then we can remove $o(n)$ elements from S and thus make it *(M,b)-free*.

The Overall Strategy

Hypergraph Removal Lemma: If a *k-uniform* hypergraph *G* contains $o(n^h)$ copies of *H*, then we can remove $o(n^k)$ edges and thus make it *H-free*.

$\Downarrow$ (strategy)

Theorem [S 08]: Let $Mx=b$ be t linear equations in p unknowns over F. If $S \subseteq F$ contains $o(n^{p-t})$ solutions to $Mx=b$, then we can remove $o(n)$ elements from S and thus make it (M,b)*-free*.

"Extending" the case of a single equation

[Candela '08, Kral-Serra-Vena '08]: Conjecture holds if every t columns of M are linearly independent.

Proof Overview

Theo [S 08]: Let $Mx=b$ be t linear equations in p unknowns over F. If $S \subseteq F$ contains δn^{p-t} solutions to $Mx=b$, we can remove $\varepsilon(\delta)n$ elements from S and make it (M,b)-*free*.

1st step: Prove a stronger result by induction. Each variable x_i may have its own subset $S_i \subseteq [n]$.

2nd step: Use hypergraphs with *larger* (than expected) edges. Simplifies many technicalities.

3rd step: Change M into an *easy to handle* set of equations M'.

4th step: Find a "small" hypergraph H based on M', which will represent the solutions to $M'x=b$.

5th step: Just do it!

Open Problems

1. Extend our result to *non-monotone* variants. E.g., given *S*, decide: *are there* $x, y \in S$ *and* $z \notin S$ *satisfying* $x + y = z$?

   * Analogous to being induced *H-free* a-la [*AFKS '00*].

2. We get *astronomical* bounds that apply to every set of equations. For which sets can we get *civilized* bounds?

   * Even the case $x + y = z$ is open [*Green '03*].

   * Analogous question for graphs were answered in [*Alon '00, Alon-S '04*].

3. [*Alon-S '05*] obtained a graph removal lemma for *infinite* sets of forbidden subgraphs. Is there a similar removal lemma for infinite sets of linear equations?

Thank You

Open Problems

1. Extend our result to *non-monotone* variants. E.g., given *S*, decide: *are there* $x, y \in S$ *and* $z \notin S$ *satisfying* $x + y = z$?

   * Analogous to being induced *H-free* a-la [*AFKS '00*].

2. We get *astronomical* bounds that apply to every set of equations. For which sets can we get *civilized* bounds?

   * Even the case $x + y = z$ is open [*Green '03*].

   * Analogous question for graphs were answered in [*Alon '00, Alon-S '04*].

3. [*Alon-S '05*] obtained a graph removal lemma for *infinite* sets of forbiden subgraphs. Is there a similar removal lemma for infinite sets of linear equations?