Extensions of Removal Lemmas from Triangles and Arithmetic Analogues

Leo Hsu

December 22, 2022

1 Introduction and Background

In this paper, we will give a summary of a paper co-authored by Jacob Fox and Yufei Zhao (see [2]). We begin by stating a result due to Ruzsa and Szemerédi that is most well-known as the Triangle Removal Lemma. In particular, we will survey a close cousin of the Triangle Removal Lemma, which is called the Triangle-free Lemma, in which we bound the optimal constant in the statemtent of the lemma, along with introducing the notion of an "ε-approximate homomorphism". In this paper, we will also see how statements that extend triangles to arbitrary graphs are made, as well as arithmetic analogues of triangles over finite fields. In light of the famous Graph Regularity Lemma ([5]), we see that it provides us with a solid background to the problems of interest.

Theorem 1.1 (Triangle Removal Lemma [4]). *For every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that if G is a graph on n vertices having at most δn^3 triangles, then it can be made triangle-free by deleting at most εn^2 edges.*

As it turns out, the Regularity Lemma, when used as a means to give a proof of Theorem 1.1, has implications about the structure of a triangle-free spanning subgraph (subgraph obtained by only edge deletions) of the graph G as described in the statement of the Triangle Removal Lemma. The following theorem gives a precise statement about the aforementioned structure.

Theorem 1.2 (Triangle Removal Lemma with Bounded Complexity [2]). *For every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ and M such that if G is a graph on n vertices having at most δn^3 triangles, then there is a partition $V_1, ..., V_M$ of $V(G)$ and a triangle-free graph G' on $V(G)$ that is either complete or empty between each pair of parts (V_i, V_j) satisfying $|E(G) \setminus E(G')| \leq \varepsilon n^2$.*

Now, in order to give a more concise statement of Theorem 1.2, we recall the following definition of a graph homomorphism.

Definition 1.3. Let G and F be graphs. A *graph homomorphism* is a map (equivalently, a function)

$$\phi : V(G) \to V(F)$$

such that if $uv \in E(G)$, then $\phi(u)\phi(v) \in E(F)$.

Suppose we are given two graphs G and G' on the same vertex set, and we have the identity map

$$\begin{aligned} f : V(G) &\to V(G') \\ v &\mapsto v. \end{aligned}$$

If $E(G) = E(G')$, then it is clear that f is a graph homomorphism. However, one immediately sees that if G and G' are as described in Theorem 1.2, then f is simply a function from $V(G)$ to itself without the adjacency preservation property. In that setting, since $|E(G) \setminus E(G')| \leq \varepsilon |V(G)|^2$, the following definition of a function that is close to being a graph homomorhism is then motivated.

Definition 1.4 ([2])**.** Let G and F be graphs. An *ε-approximate homomorphism* is a map

$$\phi : V(G) \to V(F)$$

such that it maps all but at most $\varepsilon|V(G)|^2$ edges of G to edges of F.

With the notion of an ε-approximate homomorphism between graphs, a restatement of Theorem 1.2 can then be made. The equivalence of the statement of Theorem 1.2 and that of what follows might not readily be clear to the reader, we therefore present a proof of the following restatement of Theorem 1.2.

Theorem 1.5 (Triangle Removal Lemma with Bounded Complexity, Rephrased[2])**.** *For each $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ and M such that if G is a graph on n vertices having at most δn^3 triangles, then G has an ε-approximate homomorphism into a triangle-free graph with at most M vertices.*

Proof. Let $\varepsilon > 0$, and let δ and M be as in Theorem 1.2. Suppose G is a graph on n vertices with fewer than δn^3 triangles, we may then let $V_1, ..., V_M$ be a partition of $V(G)$, and let G' be a triangle-free spanning subgraph of G that is either complete or empty between each pair of parts (V_i, V_j) with $|E(G) \setminus E(G')| \leq \varepsilon n^2$. Note that we may assume that there are no edges within each part V_i, and that is since for each edge within any part, we may put each endpoint of it in a part on its own, and so it is trivially complete between this pair of parts. To this end, the identity map

$$\begin{aligned} f : V(G) &\to V(G') \\ v &\mapsto v \end{aligned}$$

is clearly an ε-approximate homomorphism between G and G'.
By the structure of G', it can be realized as the graph G^* with the vertex set $V(G^*) =$

$\{V_1, ..., V_M\}$ such that $V_iV_j \in E(G^*)$ if and only if it is complete between the pair (V_i, V_j) in G'. Since G' is triangle-free, it forbids the existence of a complete tripartite subgraph, which corresponds to a triangle in G^*. Thus, the graph G^* is also a triangle-free graph (on M vertices). Now, the map

$$\begin{aligned} g : V(G') &\to V(G^*) \\ v &\mapsto V_i \end{aligned}$$

where $v \in V_i$ in G', is well-defined since each $v \in V(G')$ is in exactly one part. We claim the composition $g \circ f$ is an ε-approximate homomorphism between G and G^*. Indeed, since the collection of edges of G that map to non-edges of G' under f are precisely $E(G) \setminus E(G')$, and since $|E(G) \setminus E(G')| \leq \varepsilon n^2$, it will suffice to show that each edge in $E(G) \setminus E(G')$ remains a non-edge and that each edge in $E(G')$ remains an edge in G^* under g. But if $xy \in E(G) \setminus E(G')$, then $x \in V_i$ and $y \in V_j$ for some $i, j \in [M]$, and either $i = j$ or $i \neq j$. In the case when $i = j$, we are done. Suppose therefore that $i \neq j$, but then it is empty between X_i and X_j in G' by the structure of G', and so $X_iX_j \notin E(G^*)$. Now, let $xy \in E(G')$, then $x \in V_i$ and $y \in V_j$ for some $i, j \in [M]$ with $i \neq j$ (there are no edges within each part in G'). Again, by the structure of G^*, it follows that it is complete between X_i and X_j, and so $X_iX_j \in E(G^*)$. This completes the proof. □

In particular, if we restrict ourselves to only looking at triangle-free graphs, this following corollary of the previous theorem gives us a rather nice result (in the case when G contains no triangles in Theorem 1.5).

Corollary 1.6 (Triangle-free Lemma [2])**.** *For every $\varepsilon > 0$, there exists M such that every triangle-free graph has an ε-approximate homomorphism into some triangle-free graph on at most M vertices.*

We now make the following definitions in order to bound the optimal constant M in the above theorem.

Definitions 1.7 ([2])**.** Let $\varepsilon > 0$.

(a) $\delta_{TRL}(\varepsilon)$ is defined to be the largest possible δ in the Triangle Removal Lemma (Theorem 1.1);

(b) $M_{TRL}(\varepsilon)$ is defined to be the smallest possible M in the Triangle Removal Lemma with Bouned Complexity, Rephrased (Theorem 1.5);

(c) $M_{TFL}(\varepsilon)$ is defined to be the smallest possible M in the Triangle-free Lemma (Corollary 1.6).

Given a graph G, Szemerédi's Regularity Lemma has a famously large upper bound on the number of sets in a partition of $V(G)$ which satisfies the statement(s) of the lemma (i.e., a regular partition). In particular, this bound can go as far as an exponential tower of 2's of height $\varepsilon^{-O(1)}$.

Let us now denote the exponential tower of 2's of height m by tower(m). Noting that the usual proofs of the Triangle Removal Lemmas (Theorems 1.1, 1.2, and 1.5) rely on the result of the Regularity Lemma (the partition of $V(G)$, in particular), M in Corollary 1.6, and hence $M_{TFL}(\varepsilon)$, is then bounded above by tower$(\varepsilon^{-O(1)})$. One sees as well, in the standard proof of the Triangle Removal Lemma (Theorem 1.1) using the Regularity Lemma, that δ is defined to be a fraction with the aforementioned upper bound on the number of sets in a regular partition of $V(G)$ (to some power) in the denominator. Thus, δ_{TRL}^{-1} is then also bounded above by tower$(\varepsilon^{-O(1)})$, but by a result due to Fox [1], the upper bound of δ_{TRL}^{-1} has been improved to tower$(O(\log(\varepsilon^{-1})))$.

Moreover, the (usual) bound for M_{TFL} of tower$(\varepsilon^{-O(1)})$ is also improved [2]:

$$M_{TFL} \leq e^{O(\delta_{TRL}(\varepsilon/C)^{-2})}, \tag{1.8}$$

and we see that this is a better bound since $\delta_{TRL} \leq$ tower$(O(\log(\varepsilon^{-1})))$.

As a consequence of the Triangle Removal Lemma, the following theorem was left as an exercise in Math 492/529 at UVic [[3], Exercise 5.8], and provides us a lower bound to M_{TFL}.

Theorem 1.9 (Dimond-free Lemma [2])**.** *For every $\varepsilon > 0$, there exists some N such that if $n \geq N$, every n-vertex graph where each edge lies in a unique trianlge has at most $n\varepsilon^2$ edges.*

Definition 1.10 ([2])**.** $N_{DFL}(\varepsilon)$ is defined to be the smallest possible N in Theorem 1.9.

Theorem 1.11 ([2])**.** *For every $\varepsilon > 0$, there exists a constant C such that*

$$M_{TFL}(\varepsilon) \geq e^{\varepsilon N_{DFL}(C\varepsilon)/C}.$$

As a consequence of the above theorem, using the best known lower bound on $N_{DFL}(\varepsilon)$ [2], we have the following:

Corollary 1.12 ([2])**.** *For every $0 < \varepsilon < 1/2$, there exists a constant c such that*

$$M_{TFL}(\varepsilon) \geq e^{(\varepsilon^{-1})^{clog(\varepsilon^{-1})}}.$$

As it turns out, N_{DFL} and δ_{TRL}^{-1} have similar growth [2], and so together with the inequality (1.8), M_{TFL} can then be bounded quite nicely. In section 3, where we establish an arithmetic analogue of the above setting, we will see a result that the best upper bound and lower bound coincide. Moreover, in the below section, we establish an extension from triangles to arbitrary graphs and state the main theorem for graphs.

2 Extensions to Arbitrary Graphs from Triangles

In this section, we make statements about how Theorems 1.1, 1.6, and 1.9 may be extended to an aribitrary graph H from a triangle, and instead of looking at actual copies of H as we did with triangles, we look at "homomorphic copies of H" which is defined in what follows. Furthermore, we will see that in the main theorem for graphs has bounds similar in spirit to equation (1.8) and Theorem 1.11.

Definitions 2.1 ([2])**.** Let G and H be graphs.

(a) A graph G is said to be *H-homomorhism-free* if there does not exist a graph homomorhism from H to G.

(b) A *homomorphic copy of H in G* is a subgraph of G that is the image of a graph homomorhism from H (to G).

(c) The *core* of H, denoted core(H), is the smallest subgraph of H that can arise as the image of a graph homomorphism from H to itself (i.e., the minimal homomorphic copy of H in H).

Remark 2.2 ([2])**.** By the above definition(s), if G is H-homomorphism-free, then G does not contain any homomorphic copies of H. Moreover, the core of H is unique (up to graph isomorphism), and to see this, suppose that ϕ and ψ are graph homomorphisms of H such that the images $\phi(H)$ and $\psi(H)$ are both cores of H. It will then suffice to argue that there is a surjective graph homomorphism from $\phi(H)$ to $\psi(H)$ and a surjective graph homomorphism from $\psi(H)$ to $\phi(H)$. But it is clear that the restrictions

$$\phi|_{\psi(H)} : \psi(H) \to \phi(H)$$

and

$$\psi|_{\phi(H)} : \phi(H) \to \psi(H)$$

are graph homomorphisms. It now only remains to show that they are indeed surjective. Suppose for the sake of contradiction, that $\phi|_{\psi(H)}$ is not surjective. Then the image $\mathrm{Im}(\phi|_{\psi(H)}) \subset \phi(H)$. Moreover, since composition of homomorphisms is again a homomorphism, the composition

$$\phi|_{\psi(H)} \circ \psi(H) : H \to \mathrm{Im}(\phi|_{\psi(H)})$$

is a graph homomorphism. But then $\mathrm{Im}(\phi|_{\psi(H)} \circ \psi(H)) \subseteq \mathrm{Im}(\phi|_{\psi(H)}) \subset \phi(H) \subseteq H$, and so the above restriction is a graph homomorphim of H with the image strictly smaller than $\phi(H)$ which is a core of H, a contradiction. Similarly, the other restriction $\psi|_{\phi(H)}$ is also surjective. Whence, $\phi(H) \cong \psi(H)$.

With the above definitions in mind, the following theorem gives us a rather satisfying analogue to earlier theorems that involve triangles.

Theorem 2.3 ([2])**.** *Let H be a graph, and let $\varepsilon > 0$.*

(a) There exists $\delta = \delta(\varepsilon) > 0$ such that if G is a graph on n vertices with fewer than $\delta n^{|V(H)|}$ homomorphic copies of H, then G can be made H-homomorphism-free by removing at most εn^2 edges.

(b) There exists some M such that every H-homomorphism-free graph has an ε-approximate homomorphism to an H-homomorphism-free graph on at most M vertices.

(c) Suppose further that H is connected and non-bipartite. There exists some N such that if $n \geq N$ and G is a graph on n vertices such that each edge of G lies in a unique homomorphic copy of core(H), then G has at most εn^2 edges.

Remark 2.4. Theorem 2.3 (a), (b), and (c) correspond to Theorems 1.1, 1.6, and 1.9, respectively. It is not hard to see that if H is a bipartite graph having at least one edge, then $core(H)$ is just an edge, and so it makes sense that we consider non-bipartite graphs in (c).

Definition 2.5 ([2])**.** $\delta_H(\varepsilon), M_H(\varepsilon)$, and $N_H(\varepsilon)$ are defined to be the optimal constants δ, M, and N, respectively in Theorem 2.3 (a), (b), and (c).

To this end, we state here the main result for graphs.

Theorem 2.6 (Main Theorem for Graphs [2])**.** *For every $0 < \varepsilon < 1$ and every connected non-bipartite graph H, there is some constant $C = C_H > 0$ such that*

$$e^{N_H(C\varepsilon)/C} \leq M_H(\varepsilon) \leq e^{C\delta_H(\varepsilon/C)^{-2}}.$$

Since being H-homomorphism-free is equivalent to being core(H)-homomorphism-free, we consider our graph $H = \text{core}(H)$. Let G be a graph on n vertices in which there are some homomorphic copies of H such that every edge is contained in a unique homomorphic copy of H. The lower bound of the above theorem is obtained by first constructing an H-homomorphism-free G' from G with a process called the "partial binary blow-up" (see [2]) in which the edge set of each homomorphic copy is arbitrarily partitioned into non-empty sets, and in turn being in position to apply Theorem 2.3 and together with a novel entropy argument. The upper bound, however, is obtained by using what is called the "Weak Regularity Lemma" (see [2]) as well as the "Weighted Graph Removal Lemma" (see [2]) with edge weights in $[0, 1]$.

3 An Arithmetic Analogue

We begin this section with the following definition.

Definition 3.1 ([2])**.** Let G be an abelian group. Given $X, Y, Z \subseteq G$, a *triangle* in $X \times Y \times Z$ is a triple $(x, y, z) \in X \times Y \times Z$ with $x + y + z = 0$.

Bearing in mind the above definition of an “arithmetic triangle”, arithmetic analogues of the theorems from the previous section are presented in what follows.

Theorem 3.2 (Arithmetic Triangle Removal Lemma[2])**.** *For every $\varepsilon > 0$ there exists $\delta > 0$ such that for every finite abelian group G, and subsets $X, Y, Z \subseteq G$ with fewer than $\delta|G|^2$ triangles in $X \times Y \times Z$, we can remove all triangles by deleting at most $\varepsilon|G|$ elements from each of $X, Y,$ and Z.*

Theorem 3.3 (Arithmetic Dimond-free Lemma [2])**.** *For every $\varepsilon > 0$, there exists N such that for every finite abelian group G with $|G| \geq N$, and $x_1, ..., x_l, y_1, ..., y_l, z_1, ..., z_l \in G$ satisfying $x_i + y_j + z_k = 0$ if and only if $i = j = k$, then $l \leq \varepsilon|G|$.*

Definition 3.4. Let p be a prime, and let $n \in \mathbb{N}$. $\mathbb{F}_p^n$ is the (abelian) group of n-tuples (n-dimensional vectors) with entries in $\mathbb{F}_p$ where $\mathbb{F}_p$ is a field with p elements under the usual binary operation of the additive group of $\mathbb{F}_p$.

Consider now only the case where the group we regard is of the form F_p^n for a fixed prime p. We also note here that if $G = \mathbb{F}_p^n$, then $|G| = p^n$, so that (the analogues to the theorems from the previous sections of) the following theorems become more clear to the reader.

Theorem 3.5 (Arithmetic Triangle Removal Lemma with Bounded Complexity[2])**.** *For every $\varepsilon > 0$ and a prime p, there exists $\delta > 0$ and a positive integer m such that if $X, Y, Z \subseteq \mathbb{F}_p^n$ are such that $X \times Y \times Z$ has at most δp^{2n} triangles, then there exist $X', Y', Z' \subseteq \mathbb{F}_p^m$ with $X' \times Y' \times Z'$ being triangle-free, and a linear map $\phi : \mathbb{F}_p^n \to \mathbb{F}_p^m$ such that at most εp^n elements from each of $X, Y,$ and Z do not get mapped to $X', Y',$ and Z', respectively.*

Theorem 3.6 (Arithmetic Triangle-free Lemma[2])**.** *For every $\varepsilon > 0$ and prime p, there exists a positive integer m such that if $X, Y, Z \subseteq \mathbb{F}_p^n$ are such that $X \times Y \times Z$ is triangle-free, then there exist $X', Y', Z' \subseteq \mathbb{F}_p^m$ with $X' \times Y' \times Z'$ being triangle-free, and a linear map $\phi : \mathbb{F}_p^n \to \mathbb{F}_p^m$ such that at most εp^n elements from each of $X, Y,$ and Z do not get mapped to $X', Y',$ and Z', respectively.*

Definitions 3.7 ([2])**.** For every $\varepsilon > 0$ and a fixed prime p,

(a) $\delta_p(\varepsilon)$ is defined to be the largest possible constant δ in the Arithmetic Triangle Removal Lemma (Theorem 3.2) when restricted to groups of the form $\mathbb{F}_p^n$.

(b) $N_p(\varepsilon)$ is defined to be the smallest possible positive integer N in the Arithmetic Dimond-free Lemma (Theorem 3.3) when restricted to groups of the form $\mathbb{F}_p^n$ with $N_p(\varepsilon) \leq p^n$.

(c) $m_p(\varepsilon)$ is defined to be the smallest constant m in the Arithmetic Triangle-free Lemma (Theorem 3.6). Moreover, we define $M_p(\varepsilon) = p^{m_p(\varepsilon)}$.

With the above definitions, we can now state the main result, analogous to the main result for graphs (Theorem 2.6), in the arithmetic setting.

Theorem 3.8 (Main Theorem, Arithmetic Analogue [2]). *For every* $0 < \varepsilon < 1$ *and a fixed prime* p*,*

$$p^{\varepsilon N_p(5\varepsilon)/p} \le M_p(\varepsilon) \le p^{27\delta_p(\varepsilon/4)^{-2}}.$$

Let p be a prime, let $G = \mathbb{F}_p^n$. Let $X, Y, Z \subseteq G$ such that $X \times Y \times Z$ is dimond-free (i.e., satisfies the hypothesis of Theorem 3.3). The derivation of the lower bound of the above theorem is similar in spirit to that of the Main Theorem for Graphs (Theorem 2.6), in that a "blow-up" construction from $X \times Y \times Z$ is executed except, here, in the arithmetic setting, we partition into cosets instead of arbitrary partitions, to obtain a triangle-free $X' \times Y' \times Z'$, and hence has a cleaner proof (see [2]). The derivation of the upper bound of the above theorem, again, fairly similar to that of Theorem 2.6, but in the arithmetic setting, uses the Weak Regularity Lemma, and the Weighted Arithmetic Triangle Removal Lemma (see [2]).

4 Conclusion

It is already impressive, in the sense of extremal combinatorics, that a thorough (if not comprehensive) analogue in abstract algrbra of problems in graph theory can be drawn. Furhtermore, in this paper, we see that either in the setting of graphs or the arithmetic setting, bounds for the optimal constant of the corresponding Triangle-free Lemma can be bounded in terms of that of its close variants. Lastly, as one of the most profound discoveries in graph theory, Szemerédi's Graph Regularity Lemma (and its weak variant) grants a powerful means to obtain satisfying bounds for the aforementioned constants.

References

[1] Jacob Fox. A new proof of the graph removal lemma. *Ann. of Math. (2)*, 174(1):561–579, 2011.

[2] Jacob Fox and Yufei Zhao. Removal lemmas and approximate homomorphisms. *Combin. Probab. Comput.*, 31(4):721–736, 2022.

[3] J. Noel. Math 492/529: Extremal Combinatorics, University of Victoria. `https://www.math.uvic.ca/~noelj/ExtremalCombinatorics202209.pdf`, Accessed on November 30, 2022.

[4] I. Z. Ruzsa and E. Szemerédi. Triple systems with no six points carrying three triangles. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, volume 18 of *Colloq. Math. Soc. János Bolyai*, pages 939–945. North-Holland, Amsterdam-New York, 1978.

[5] Endre Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.