Removal Lemmas: Summer School 2025

2 When is the removal lemma polynomial?

For which graphs H does it hold that the parameters in the H-removal lemma satisfy $\delta_H(\varepsilon) = \text{poly}(\varepsilon)$? A classical result in extremal graph theory, namely the Kővári-Sós-Turán theorem, shows that this is the case if H is bipartite.

Theorem 2.1 (Kővári-Sós-Turán theorem, supersaturation form)**.** *An n-vertex graph with εn^2 edges contains at least poly(ε)n^{s+t} copies of $K_{s,t}$.*

Returning to the H-removal lemma for a bipartite H, observe that if G is ε-far from H-free then G (trivially) contains at least εn^2 edges, hence G contains $\text{poly}(\varepsilon)n^{v(H)}$ copies of H by Theorem 2.1. Thus, if H is bipartite then the H-removal lemma is polynomial. Alon proved that the converse also holds, i.e., that bipartite graphs are the only ones which admit a polynomial removal lemma.

Theorem 2.2 (Alon 2002)**.** *For a graph H, $\delta_H(\varepsilon) = \text{poly}(\varepsilon)$ if and only if H is bipartite.*

We will first prove Theorem 2.2 in the case that H is an odd cycle. For this, we need a number-theoretic construction.

Theorem 2.3. *Let $k \geq 3$. There is a set $S \subseteq [n]$ with $|S| \geq n^{1-o(1)}$, such that for every $x_1, \dots, x_k \in S$, if $x_1 + \cdots + x_{k-1} = (k-1)x_k$, then $x_1 = \cdots = x_k$.*

The case $k = 3$ is Behrend's construction of a large set with no 3-term arithmetic progressions. The general case is a straightforward generalization.

Proof of Theorem 2.3. Write $n = d^t$ for d, t to be chosen later. Represent the numbers $1, \dots, n$ in base d. I.e., for $x \in [n]$, write

$$x = \sum_{i=0}^{t-1} a_i d^i,$$

where $0 \leq a_i \leq d-1$. Write $v(x) := (a_0, \dots, a_{t-1})$. Let U be the set of all x for which $a_0, \dots, a_{t-1} \leq \frac{d-1}{k-1}$. This property guarantees that for $x_1, \dots, x_{k-1} \in U$, we have

$$v(x_1 + \cdots + x_{k-1}) = v(x_1) + \cdots + v(x_{k-1}).$$

I.e., there is no carry when summing $x_1, \dots, x_{k-1}$. Similarly, $v((k-1) \cdot x_k) = (k-1) \cdot v(x_k)$ for every $x_k \in U$.

Now fix $r \geq 1$, to be chosen later, and take S to be the set of all $x \in U$ with $\|v(x)\| = r$, where $\|\cdot\|$ is the Euclidean norm. Suppose that $x_1, \dots, x_k \in S$ satisfy $x_1 + \cdots + x_{k-1} = (k-1)x_k$. Putting $v_i = v(x_i)$, we get $v_1 + \cdots + v_{k-1} = (k-1)v_k$. Now we take norms. The norm of the RHS is $(k-1)r$.

For the LHS, by Cauchy-Schwarz we have $\|v_1 + \cdots + v_{k-1}\| \le \sqrt{\sum_{i=1}^{k-1} \|v_i\|^2} \cdot \sqrt{k-1} = (k-1)r$, with equality if and only if $v_1 = \cdots = v_{k-1}$. So we must have $v_1 = \cdots = v_k$ and hence $x_1 = \cdots = x_k$.[1]

Now we estimate the size of S. For every $x \in [n]$, we have $\|x\|^2 \le td^2$, so the number of choices for r is at most td^2. By pigeonhole, there exists r such that

$$|S| \ge \frac{|U|}{td^2} \ge \frac{(d/k)^t}{td^2} = \frac{n}{k^t t d^2}.$$

Choose t, d such that $k^t = d$. As $d^t = n$, this gives $t = \sqrt{\frac{\log(n)}{\log(k)}}$, $d = e^{\sqrt{\log(k)\log(n)}}$. So

$$|S| \ge \frac{n}{e^{O_k(\sqrt{\log n})}} = n^{1-o(1)}.$$

■

Now we prove Theorem 2.2 for odd cycles.

Theorem 2.4. *For every odd $k \ge 3$, there exists an n-vertex graph G with εn^2 edge-disjoint copies of C_k, but only $\varepsilon^{\omega(1)} n^k$ copies of C_k in total.*

Proof. Let $\varepsilon > 0$. Let $S \subseteq [n]$ be the set given by Theorem 2.3. Choose n such that $|S| = \varepsilon n$. As $|S| = n^{1-o(1)}$, this means that $n = (1/\varepsilon)^{\omega(1)}$. Define a graph with k parts $V_1, \ldots, V_k$, each of size kn and identified with $[kn]$.[2] For each $y \in [n]$ and $x \in S$, add a copy of C_k on the vertices $v_1 = y, v_2 = y + x, v_3 = y + 2x, \ldots, v_k = y + (k-1)x$ (so $v_i = y + (i-1)x$) such that $v_i \in V_i$.[3] Denote this copy by $C_{x,y}$. We claim that the copies $C_{x,y}$ are edge-disjoint. Indeed, even stronger, any two such copies share at most one vertex, because if $C_{x,y}$ and $C_{x',y'}$ have the same vertex in V_i and V_j, then $y + (i-1)x = y' + (i-1)x'$ and $y + (j-1)x = y' + (j-1)x'$, and solving this system of equations gives $x = x', y = y'$. The number of copies $C_{x,y}$ is $n|S| \ge \varepsilon n^2$. Thus, the graph has a collection of εn^2 edge-disjoint copies of C_k.

Now we bound the total number of copies of C_k. Crucially, as k is odd, we can only have copies of C_k of the form $(v_1, \ldots, v_k, v_1)$ with $v_i \in V_i$.[4] Now consider such a copy $v_1, \ldots, v_k$. Then for each $1 \le i \le k-1$ there are y_i, x_i with $v_i, v_{i+1} \in C_{x_i, y_i}$, and there are y_k, x_k with $v_k, v_1 \in C_{x_k, y_k}$. Then

$$x_1 + \cdots + x_{k-1} = v_k - v_1 = (k-1)x_k.$$

By the property of the set S, we get $x_1 = \cdots = x_k =: x$ (from which we can also deduce that $y_1 = \cdots = y_k$). So $(v_1, \ldots, v_k) \in C_{x,y_1}$. This shows that any copy of C_k in the graph is one of the "original" copies $C_{x,y}$ we put in. Their number is

$$n|S| \le n^2 \le \frac{|V(G)|^k}{n} \le \varepsilon^{\omega(1)} |V(G)|^k.$$

■

Remarks:

---

[1] What we are using here is that S is a sphere, and a sphere has no point in the convex hull of other points (unless all points are equal.

[2] Thus, we are actually defining a graph on $k^2 n$ vertices, but we can of course adjust the parameters.

[3] Note that we choose each V_i to be $[kn]$ so that the numbers $v_i = y + (i-1)x$ "fit" in V_i.

[4] What we are using here is that C_k is not homomorphic to any of its proper subgraphs.

- We can take blowups of the graph defined in the proof of Theorem 2.4 to get constructions of any (large enough) size.
- The above proof gives a connection between the triangle removal lemma and the problem of estimating the largest possible size $r_3(n)$ of a subset of $[n]$ with no 3-term arithmetic progression. Indeed, in the proof, we use a lower bound on $r_3(n)$ (via Theorem 2.3) to show that the triangle removal lemma is not polynomial. In the other direction, one can use the triangle removal lemma to show that $r_3(n) = o(n)$, which is the statement of Roth's theorem. We note, however, that this gives a very poor quantitative bound of roughly $r_3(n) \le n/\log_*(n)$. Much better bounds are known.

To prove Theorem 2.2 for a general non-bipartite H, we would like to use the same strategy as in Theorem 2.4. Namely, if $V(H) = \{1, \dots, h\}$, we construct an H-partite graph with parts $V_1, \dots, V_h$ and put a copy of H on $y, y + x, \dots, y + (h-1)x$ for $y \in [n], x \in S$. We will also use that H has an odd cycle. The issue is that we want to make sure that every copy of H is of the form $v_1, \dots, v_h$ with $v_i \in V_i$ (and v_i plays the role of $i \in [h]$). Note that the construction is homomorphic to H via the homomorphism $V_i \mapsto i$.[5] Thus, what we want is that H has no homomorphism to a proper subgraph of itself. This might not be true of H itself, but there is a maximal subgraph of H which has this property, and we will exploit this for our construction. Let us now define this subgraph.

Definition 2.5. *The* core *of H is the minimal subgraph K of H (in terms of the number of vertices) such that there is a homomorphism from H to K.*

We will show soon that the core is well defined, in the sense that K is unique up to isomorphism. Observe that K is not homomorphic to any of its proper subgraphs. Indeed, if there is a homomorphism $\psi : K \to J$ for J with $V(J) \subsetneq V(K)$, then by taking a homomorphism $\varphi : H \to K$, we get a homomorphism $\psi \circ \varphi$ from H to J, contradicting the minimality of K. Thus, every homomorphism from K to itself is injective and hence an isomorphism. Similarly, we can show that the core is unique up to isomorphism: If K_1, K_2 are both cores of H, then there are homomorphisms $\varphi_1 : K_2 \to K_1$ and $\varphi_2 : K_1 \to K_2$ (we obtain φ_i by taking a homomorphism from H to K_i and restricting it to K_{3-i}). Now, $\varphi_1 \circ \varphi_2$ is a homomorphism from K_1 to itself and hence an isomorphism, and similarly for $\varphi_2 \circ \varphi_1$. It follows that φ_1, φ_2 are bijective and hence isomorphisms.

Note that if H is bipartite (and has at least one edge), then the core of H is an edge. On the other hand, if H is not bipartite then neither is its core. Using cores, we can now prove Theorem 2.2. The idea is to do the construction for the core of H, and then blow it up by a constant factor to get a construction for H.

Proof of Theorem 2.2. Let K be the core of H. Then K is also not bipartite. Write $V(K) = \{1, \dots, k\}$, where $(1, \dots, \ell, 1)$ is an odd cycle. Take $S \subseteq [n]$ from Theorem 2.3 (with parameter ℓ), and define a graph G with sides $V_1, \dots, V_k$ by doing the following: For each $y \in [n]$ and $x \in S$, put a copy $K_{x,y}$ of K on $v_1, \dots, v_k$, where $v_i = y + (i-1)x \in V_i$ (in this copy, v_i plays the role of i). As in the proof of Theorem 2.4, the copies $K_{x,y}$ are edge-disjoint, and hence G has εn^2 edge-disjoint copies of K.

On the other hand, since K is a core, every copy of K in G is of the form $v_1, \dots, v_k$ with $v_i \in V_i$ playing the role of i. Hence, for each such copy $v_1, \dots, v_k$, the vertices $v_1, \dots, v_\ell$ makes an odd cycle. By the same argument as in the proof of Theorem 2.4, each such odd cycle is of the form

[5] A *homomorphism* from a graph G to a graph H is a mapping $\varphi : V(G) \to V(H)$ such that $\varphi(x)\varphi(y) \in E(H)$ for every $xy \in E(G)$.

$(y, y+x, \ldots, y+(\ell-1)x)$ for some $y \in [n], x \in S$, and hence the number of such odd cycles is at most n^2. Thus, the total number of copies of K in G is at most $n^2 \cdot n^{k-\ell} \leq n^k/n \leq \varepsilon^{\omega(1)} n^k$.

To obtain a construction for H, take the above construction for K and blow it up by a factor of $h := |V(H)|$. Then each copy $K_{x,y}$ of K gives rise to a copy of H (because H is homomorphic to K, i.e., contained in a blowup of K). Hence, the resulting graph (which has $O(n)$ vertices) has εn^2 edge-disjoint copies of H. On the other hand, each copy of H must contain a copy of K (because K is a subgraph of H). Also, the blown-up graph has $O(\varepsilon^{\omega(1)} n^k)$ copies of K (the only way to get copies of K is from blowups of copies of K in G, as K is a core). Thus, the total number of copies of H is $O(\varepsilon^{\omega(1)} n^k) \cdot n^{h-k} = O(\varepsilon^{\omega(1)} n^h)$, as required. ■