MA 5109: Extremal Graph Theory

Amit Rajaraman

Last updated September 2, 2021

Contents

§0. Notation

We use $[n]$ to represent $\{1, 2, \ldots, n\}$.

For integers a and b, $[a, b]$ means $\{a, a+1, \ldots, b\}$.

A graph G_n is a graph with n vertices.

Given a graph G, $e(G)$ is the number of edges G has.

For a vertex v, denote by $N(v)$ or $\Gamma(v)$ the set of *neighbours* of v – all the vertices that have an edge to v.

For a vertex v, denote by $d_G(v) = |\Gamma(v)|$ the *degree* of v – the number of edges incident on it. If the graph G is clear from context, we write simply $d(V)$.

For $v \in V$ and $K \subseteq V$, $d(v, K)$ is the number of edges

$$\big|\{u \in K : uv \in E\}\big|$$

from v into K.

Given a graph $G = (V, E)$, denote by $\delta(G)$ and $\Delta(G)$ the minimum and maximum degree in G respectively. That is,

$$\delta(G) = \min_{v \in V} d(v) \text{ and } \Delta(G) = \max_{v \in V} d(v).$$

We shorten "there exists n_0 such that for all $n > n_0$" to "for all n sufficiently large", which in turn is shortened to "for all $n \gg 0$".

We say that something occurs "with high probability" if the probability of the complement occurring converges to 0 as the input size grows to ∞.

§1. Introduction

1.1. Basic Definitions

Definition 1.1. A (simple undirected) **graph** G is an ordered pair (V, E) where V is a finite set called the *vertex set* and E, called the *edge set*, is a subset of $\binom{V}{2}$, where $\binom{S}{k}$ represents the set of all k-element subsets of S.

We typically represent graphs pictorially, showing vertices as dots and edges as arcs joining the vertices present in the corresponding subset.
A few important graphs are:

- the *null graph* with vertex set V, where $E = \varnothing$.
- the *complete graph* K_n, where $V = [n]$ and $E = \binom{[n]}{2}$.
- the *complete bipartite graph* $K_{m,n}$, where $V = A \cup B$ with $|A| = m$, $|B| = n$, and A, B are disjoint, and $E = \{\{a, b\} : a \in A, b \in B\}$.
- the *path graph* P_{n+1} of length n, where $V = [n+1]$ and $E = \{\{m, m+1\} : m \in [n]\}$.
- the *cycle* of length n, where $V = [n]$ and $E = \{\{l, m\} : l, m \in [n], (m - l) \equiv 1 \pmod{n}\}$.

Now, consider the graph G with vertex set $[4]$ and edge set $\{\{1, 3\}, \{3, 2\}, \{2, 4\}\}$. This graph appears to be the same as the path graph of length 3, but how do we make this correspondence more concrete?
Relabeling vertices doesn't create a "new" graph.

Definition 1.2 (Graph Isomorphism)**.** Two graphs $G = (V, E)$ and $G' = (V', E')$ are said to be **isomorphic** and we write $G \simeq G'$ if there exists a bijection $f : V \to V'$ such that there is an edge between two vertices u and v in G if and only if there is an edge between $f(u)$ and $f(v)$ in G'.

If two graphs are isomorphic, they are identical for our purposes (we only care about graphs up to isomorphism).
We now give a few more definitions that are useful.

Definition 1.3 (Subgraph)**.** Given a graph $G = (V, E)$, a **subgraph** $H = (V', E')$ is a graph such that $V' \subseteq V$ and $E' \subseteq E$. Given $V' \subseteq V$, the subgraph *induced* by V' on G is that with vertex set V' and edge set $\binom{V'}{2} \cap E$.

Definition 1.4 (r-partite Graph)**.** A graph $G = (V, E)$ is said to be r**-partite** if there exists a partition $V_1, V_2, \ldots, V_r$ of V such that for any edge $e = uv \in E$, u and v are in distinct V_i. That is, there are no edges within any of the V_i.
In particular, a 2-partite graph is said to be **bipartite**.

Definition 1.5 (Independent Set)**.** Given a graph $G = (V, E)$, $I \subseteq V$ is said to be **independent** if no two vertices of I are adjacent (the subgraph induced by I is null).
$\alpha(G)$, the *independence number* of G, denotes the size of the largest independent set in G.

Definition 1.6 (Clique)**.** Given a graph $G = (V, E)$, $K \subseteq V$ is said to be a **clique** if any two vertices of K are adjacent (the subgraph induced by I is complete). $\omega(G)$, the *clique number* of G, denotes the size of the largest clique in G.

Definition 1.7 (Complement Graph)**.** Given a graph $G = (V, E)$, the **complement graph** of G is $\bar{G} = (V, \binom{V}{2} \setminus E)$.

Observe that $S \subseteq V$ is independent in G if and only if S is a clique in $\bar{G}$. In particular, $\alpha(G) = \omega(\bar{G})$.

Definition 1.8 (Connectedness)**.** A graph G is said to be **connected** if for any pair of vertices u, v, there is a sequence $u = v_0, v_1, \ldots, v_r = v$ for some r such that $v_{i-1}v_i$ is an edge for each $i \in [r]$.

Definition 1.9 (Girth)**.** The **girth** of a graph G is the smallest k (> 2) for which C_k is isomorphic to a subgraph of G.

If G has no cycles, it is said to have infinite girth.

1.2. K_{r+1}-free graphs

Extremal graph theory is motivated by the following simple problem:

> At most how many edges can a graph G_n have if it contains no triangles?

More precisely, what is

$$\max_{\substack{\text{no subgraph of } G_n \\ \text{is isomorphic to } K_3}} e(G_n)?$$

Clearly, this number is well-defined since a graph on n vertices cannot have more than $\binom{n}{2}$ vertices.
A simple observation is that any complete bipartite graph has no triangles: if there were a triangle, then two vertices would be in the same "part", which contradicts the existence of edges only between the two parts.
As a consequence, for any $1 \leq m \leq n$, it is possible to construct $m \times (n - m)$ edges (with this bound being attained for $K_{m,n-m}$). In particular, it is possible to construct a graph with $\lfloor n^2/4 \rfloor$ edges.

Theorem 1.1 (Mantel's Theorem)**.** If G_n has no triangle, then

$$e(G_n) \leq \left\lfloor \frac{n^2}{4} \right\rfloor.$$

Further, equality is attained iff $G_n \simeq K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

Proof. Suppose G_n has no triangles. Saying that G_n has no triangles is equivalent to saying that for distinct adjacent u, v, $\Gamma(u) \cap \Gamma(v) = \varnothing$.

So, $d(u) + d(v) \leq n$. Therefore,

$$\begin{aligned}
ne(G_n) &\overset{(1)}{\geq} \sum_{uv \in E} d(u) + d(v) \\
&= \sum_{uv \in E} |\Gamma(u) \cup \Gamma(v)| \\
&= \left|(e, w) : e = uv \in E, w \in \Gamma(u) \cup \Gamma(v)\right| \\
&= \sum_{u \in V} |\{(e, w) : w \in \Gamma(u), e = uv \in E\}| \\
&= \sum_{u \in V} |\{(v, w) : v, w \in \Gamma(u)\}| \\
&\overset{(2)}{=} \sum_{u \in V} d(u)^2 \\
&\overset{(3)}{\geq} \frac{1}{n} \left(\sum_{u \in V} d(u) \right)^2 \\
&\overset{(4)}{=} \frac{4e(G_n)^2}{n},
\end{aligned}$$

where (2) follows from the changing the main thing being summed over to u, the "middle" vertex in the L-like structure, (3) follows from the Cauchy-Schwarz inequality, and (3) follows from the handshaking lemma.

What happens when equality is attained? Let us look at the case where n is even.
(1) is only tight when $d(u) + d(v) = n$ for all edges uv and (3) is only tight when $d(u)$ is a constant (independent of u). This implies that $d(u) = \frac{n}{2}$ for every $u \in V$. Now, if uv is an edge, $\Gamma(u) \cap \Gamma(v) = \varnothing$ implies that $\Gamma(u) \cup \Gamma(v) = V$, and so $G_n = K_{\frac{n}{2},\frac{n}{2}}$.
The case where n is odd is analyzed similarly, with slight nuances in (3) since exact equality is not attained. ■

While the above is one of the early results in extremal graph theory, the subject was only really born due to Turán in the following result.

Theorem 1.2 (Turán's Thoerem)**.** If G_n has no K_{r+1} $(r \geq 2)$, then $e(G_n) \leq t_r(n)$, with equality attained iff $G \simeq T_r(n)$.

The version for $r = 2$ is just a triangle-free graph and is the same as Mantel's Theorem. In the proof of this, we split the vertex set into two parts and dumped all the edges between these parts.
If we want to avoid K_4 $(r = 3)$, then perhaps we could split the vertex set into three parts and dump all the edges between these parts.
In general, we want to partition V of size n into r "almost equal" parts and set only those edges between vertices in distinct parts – such a graph is known as the **Turán graph** $T_r(n)$ and the number of edges $e(T_r(n))$ is the **Turán number** $t_r(n)$.
In particular, when $r \mid n$,

$$t_r(n) = \binom{r}{2} \left(\frac{n}{r}\right)^2 = \frac{n^2}{2}\left(1 - \frac{1}{r}\right).$$

Here, we give three proofs of Turán's Theorem.

Proof of Turán's Theorem. We perform strong induction on $n + r$. We have already proved the result for $r = 2$.
Suppose $e(G_n) \geq t_r(n)$ and G_n is K_{r+1}-free, where $r > 2$. We wish to prove that $G \simeq T_r(n)$.
Since $t_r(n) > t_{r-1}(n)$[1], the inductive hypothesis implies that G has a copy $K \subseteq V$ of K_r. Observe that for $v \notin K$,

[1] To show this, use the fact that any $(r-1)$-partite graph can be thought of as r-partite graph with one of the pieces being empty, and then that the Turán graph has the most edges among r-partite graphs.

$d(v, K) \le r - 1$ – otherwise, there would be a copy of K_{r+1} in G.
As a result, $e(V \setminus K, K) \le (r-1)(n-r)$. By the induction hypothesis, $e(V \setminus K, V \setminus K) \le t_r(n-r)$. Therefore,

$$t_r(n) \le e(G_n) \le t_r(n-r) + (r-1)(n-r) + \binom{r}{2}.$$

However, as can be checked manually, $t_r(n-r) + (r-1)(n-r) + \binom{r}{2} = t_r(n)$!
It follows that equality holds everywhere – $e(G_n) = t_r(n)$, $e(V \setminus K) = t_r(n-r)$, and $d(v, K) = r-1$ for all $v \in V \setminus K$.
This graph is then isomorphic to $T_r(n)$ – for each $v \in V \setminus K$, we can put the vertex in K that is not adjacent to v in the same bucket as v. Then, the only edges are those between distinct buckets (Why?), so $G_n \simeq T_r(n)$. ■

Erdős' Proof of Turán's Theorem. Erdős proves a slightly more general claim: given a K_{r+1}-free graph G_n, there exists an r-partite graph H on V such that $d_G(v) \le d_H(v)$ for all $v \in V$.
It is then a simple task to check that among the r-partite graphs on n vertices, the Turán graph $T_r(n)$ has the most edges.

To prove our claim, we perform induction on r.
The claim is trivial for the base case $r = 1$.
Now, suppose the claim holds for values less than r. Let $v_0 \in V$ such that $d_G(v_0) = \max_{v \in V} d_G(v)$ (the vertex of maximum degree in G) and $W = \Gamma(z)$. Since G is K_{r+1}-free, W is K_r-free. Inductively, there is an $(r-1)$-partite graph H' on W such that for all $v \in W$, $d_{H'}(v) \ge d_W(v)$.
Let $U = V \setminus W$. For each $u \in U$, remove all its edges in G and set its new neighbour set as W.
Our desired graph H is that with these edges along with those in H' and the edges from v_0 to W. That is, the rth part is $U \cup \{v_0\}$ and the remaining $(r-1)$ parts are those formed by H'. The graph is clearly r-partite by definition. What about the degree inequality?

- $d_G(v_0) = d_H(v_0)$ trivially.
- For $u \in U$, $d_H(u) = d_G(v_0) \ge d_G(u)$.
- For $w \in W$,

$$d_H(w) = |U| + 1 + d_{H'}(w) \ge |U| + 1 + d_W(w) \ge d_U(w) + 1 + d_W(w) = d_G(w).$$ ■

(Why does equality imply that the graph is isomorphic to $T_r(n)$?)

Theorem 1.3 (Turán's Theorem, reformulation)**.** If $d = e(G_n)/n$ is the average degree of the vertices of G_n, then G_n has an independent set of size at least $n/(d+1)$.

Proof. Why is this equivalent to Turán's Theorem?
If G_n has no K_{r+1}, then $\alpha(\bar{G}) \le r$. If $\bar{G_n}$ has average degree d, the above result would imply that $r \ge n/(d+1)$, that is, $d \le (n/r) - 1$. The total number of edges in G_n is then

$$\binom{n}{2} - \frac{nd}{2} \le \binom{n}{2} - \frac{n}{2}\left(\frac{n}{r} - 1\right) = \frac{n^2}{2}\left(1 - \frac{1}{r}\right),$$

which gives Turán's bound!

Let us now get to the proof of the above reformulation. First, consider the following algorithm to come up with *some* independent set in G:

1. Order V to get $\{v_1, \ldots, v_n\}$ and initialize $S = \varnothing$.
2. Add v_1 to S.

3. Having processed vertices v_1 through v_i, add v_{i+1} to S iff there is no vertex in S that is adjacent to v_{i+1}.

It is clear that this always produces an independent set, but the size of the independent set depends on the ordering we choose at the beginning.
For a given ordering σ, denote by $\mathcal{A}(\sigma)$ the independent set produced by the algorithm.
How do we choose a "good" ordering?
Enter the probabilistic method. Define the random variable π to be uniformly random on the set of all permutations of V. Then,

$$\begin{aligned}\mathbf{E}[|\mathcal{A}(\pi)|] &= \mathbf{E}\left[\sum_{v\in V} \mathbb{1}_{v\in\mathcal{A}(\pi)}\right] \\ &= \sum_{v\in V} \mathbf{E}\left[\mathbb{1}_{v\in\mathcal{A}(\pi)}\right] \\ &= \sum_{v\in V} \Pr\left[v \in \mathcal{A}(\pi)\right].\end{aligned}$$

Fix some $v \in V$ and permutation σ. What is the probability that $v \in \mathcal{A}(\sigma)$?
If at the time of processing v for the ordering σ, $\Gamma(v)\cap S \neq \varnothing$, then v is not picked. In particular, if v is the first element of $\Gamma(v)\cup\{v\}$ in the ordering σ, then it is definitely chosen by the algorithm. The probability of this occurring is $\frac{1}{d(v)+1}$.
So,

$$\begin{aligned}\mathbf{E}[|\mathcal{A}(\pi)|] &= \sum_{v\in V} \Pr\left[\mathbb{1}_{v\in\mathcal{A}(\pi)}\right] \\ &\geq \sum_{v\in V} \frac{1}{d(v)+1} \\ &\overset{(*)}{\geq} \frac{n^2}{\sum_{v\in V}(d(v)+1)} = \frac{n}{d+1},\end{aligned}$$

where $(*)$ follows from the AM-HM inequality.
Since the expectation of $|\mathcal{A}(\pi)|$ is at least $n/(d+1)$, there must exist some permutation σ such that $|\mathcal{A}(\sigma)| \geq n/(d+1)$, proving the result. ■

1.3. The Zarankiewicz Problem

Turán's Theorem is the primary result that birthed Extremal Graph Theory. To generalize the problem studied in the previous section, define the following.

Definition 1.10 (Extremal Function)**.** Given a graph H, define the **extremal function**

$$\mathrm{ex}(n; H) = \max_{\substack{\text{no subgraph of } G_n \\ \text{is isomorphic to } H}} e(G_n) \tag{1.1}$$

as the maximum number of edges in a graph on n vertices without H as a subgraph.
More generally, given graphs $H_1, \ldots, H_m$, define

$$\mathrm{ex}(n; H_1, \ldots, H_m) = \max_{\substack{\text{no subgraph of } G_n \\ \text{is isomorphic to a } H_i}} e(G_n).$$

With this notation, Turán's Theorem then says that $\mathrm{ex}(n; K_{r+1}) = t_r(n)$, with the corresponding maximum in (1.1) being attained iff $G_n \simeq T_r(n)$.

Definition 1.11 (Zarankiewicz Function)**.** Fix $s, t \in \mathbb{N}$ with $t \geq s \geq 2$ and $m, n \in \mathbb{N}$. The **Zarankiewicz function** $z(m, n; s, t)$ is the maximum number of edges in a bipartite graph $G = (A \sqcup B, E)$ such that

- the two components A and B of the graph are of cardinality m and n respectively[2], and
- there exist no $S \subseteq A, T \subseteq B$ with $|S| = s$, $|T| = t$, and all edges between S and T present in E.[3]

For ease of writing, we refer to the above described criterion as the *Zarankiewicz condition*.
That is, we "forbid" the subgraph $K_{s,t}$ with the components of size s and t on the side of A and B respectively.

The **Zarankiewicz problem** asks for a closed form representation of $z(m, n; s, t)$. Failing this, for fixed t, it asks for a tight asymptotic bound on $z(n, n; t, t)$ as n grows large.
Perhaps surprisingly, this problem remains unsolved! (as of the time of writing these notes)

Theorem 1.4 (Kővári-Sós-Turán Theorem)**.** For $t \geq s \geq 2$ and $m, n \in \mathbb{N}$,

$$z(m, n; s, t) \leq (s-1)^{1/t}(n - t + 1)m^{1-1/t} + (t-1)m.$$

Proof. Let G be bipartite with vertex set $A \sqcup B$ and satisfy the Zarankiewicz condition.
Fix $a \in A$. By definition, $|\Gamma(a)| = d(a)$. Now, $\binom{d(a)}{t}$ is the number of t-element subsets of $\Gamma(a)$. We may assume that for all $a \in A$, $d(a) \geq t - 1$. Indeed, otherwise, we may add arbitrary edges to a to make its degree $t - 1$; any vertex from A in a subgraph isomorphic to $K_{s,t}$ must have at least degree t so a cannot be part of it.
We have that

$$\begin{aligned}
\sum_{x \in A} \binom{d(x)}{t} &= \left|\{(x, T) : x \in A, T \subseteq \Gamma(x), |T| = t\}\right| \\
&= \sum_{\substack{T \subseteq B \\ |T| = t}} |\{x \in A : T \subseteq \Gamma(x)\}|.
\end{aligned}$$

If we fix a T, then the number of such x for that T is at most $s - 1$, due to our assumption. As a result,

$$\sum_{x \in A} \binom{d(x)}{t} \leq \binom{n}{t}(s-1).$$

Now, observe that the function

$$f(x) = \frac{x(x-1)\cdots(x-t+1)}{t!}$$

is convex on $[t-1, \infty)$. Using Jensen's inequality together with our assumption that $d(x) \geq t - 1$ for all $x \in A$,

$$\begin{aligned}
\binom{n}{t}(s-1) &\geq \sum_{x \in A} \binom{d(x)}{t} \\
&\geq m \binom{\frac{1}{m}\sum_{x \in A} d(x)}{t} \\
&= m \binom{e(G)/m}{t}.
\end{aligned} \tag{1.2}$$

[1]by "components" of the bipartite graph we mean that for any edge uv in the graph, $u \in A$ and $v \in B$ or $u \in B$ and $v \in A$.
[2]we make the tacit assumption that $s \leq m$ and $t \leq n$.

Let $d = e(G)/m$, the average degree of the vertices in A. Simplifying the above expression,

$$\frac{s-1}{m} \geq \frac{d(d-1)\cdots(d-t+1)}{n(n-1)\cdots(n-t+1)} \geq \left(\frac{d-t+1}{n-t+1}\right)^t,$$

where the second inequality follows from the fact that $d \leq n$. Therefore,

$$e(G) \leq m\left(\left(\frac{s-1}{m}\right)^{1/t}(n-t+1)+(t-1)\right) = (s-1)^{1/t}(n-t+1)m^{1-1/t}+(t-1)m,$$

completing the proof. ■

Next, let us look at some consequences of the above bound.

1.3.1. The Zarankiewicz problem and the extremal function for complete bipartite graphs

We get a bound on $\mathrm{ex}(n; K_{s,t})$. We claim that for $n \in \mathbb{N}$ and $s, t \geq 2$,

$$\mathrm{ex}(n; K_{s,t}) \leq \frac{1}{2} z(n, n; s, t). \tag{1.3}$$

Indeed, if $G_n = (V, E)$ has no $K_{s,t}$, make a bipartite graph G' that has vertex set $V \times \{0, 1\}$ such that uv is an edge in G iff $\{(u, 0), (v, 1)\}$ is an edge in G'.
G' satisfies the Zarankiewicz condition. If there do exist $S \subseteq V \times \{0\}$ and $T' \subseteq V \times \{1\}$ such that all S-T edges are in G, then $\pi_1(S) \cap \pi_1(T') = \varnothing$ (otherwise, a vertex would be adjacent to itself in G, which is clearly false). This implies that $K_{s,t} \subseteq G$, which is a contradiction.
Since $e(G') = 2e(G)$, the claim follows.

1.3.2. The case where $s = t = 2$

When $s = t = 2$, we get

$$z(m, n; 2, 2) \leq (n-1)m^{1/2} + m \text{ and } z(n, n; 2, 2) \leq (n-1)n^{1/2} + n.$$

Therefore,

$$\mathrm{ex}(n; K_{2,2}) \leq \frac{1}{2}(n + (n-1)\sqrt{n}).$$

Note that $K_{2,2} \simeq C_4$. Therefore, a square-free graph on n vertices has $\mathcal{O}(n^{3/2})$ edges.

In fact, this bound is tight! We give an algebraic construction of a suitable graph with no $K_{2,2}$, known as the *Levi graph for the projective plane*.
Let q be a prime power and consider the 3-dimensional vector space $\mathcal{V} = \mathbb{F}_q^3$ (over $\mathbb{F}_q$). Let $\mathcal{P}$ and $\mathcal{L}$ be the set of all 1- and 2-dimensional subspaces of $\mathcal{V}$ respectively.
Define the graph $G = (\mathcal{P} \sqcup \mathcal{L}, E)$ as follows. For $x \in \mathcal{P}$ and $L \in \mathcal{L}$, let x be adjacent to L in G iff $x \subseteq L$.
We claim that G has no $K_{2,2}$. Suppose otherwise and let $x_1, x_2 \in \mathcal{P}$ and $L_1, L_2 \in \mathcal{L}$ such that the x_i are adjacent to the L_j. If $x_1 = \langle u \rangle$ and $x_2 = \langle v \rangle$, then u and v are linearly independent, which implies that $L_i = \langle u, v \rangle$. This contradicts the fact that the L_i are distinct!
What are the cardinalities of $\mathcal{P}$ and $\mathcal{L}$?

- To get a 1-dimensional subspace, we pick a non-zero u and consider $\langle u \rangle$. In $\mathcal{V}$, there are $q^3 - 1$ non-zero u. We must now divide by $q-1$ to account for the fact that linearly dependent vectors generate the same 1-dimensional subspace. Any non-zero u has precisely $q - 1$ non-zero multiples. Therefore,

$$|\mathcal{P}| = \frac{q^3-1}{q-1} = q^2 + q + 1.$$

- It turns out that the number of 2-dimensional subspace is equal to the number of 1-dimensional subspaces (more generally, the number of d-dimensional subspaces is equal to the number of $(n-d)$-dimensional subspaces of $\mathbb{F}_q^n$), so
$$|\mathcal{L}| = q^2 + q + 1.$$

How many edges does G have?
Fix $x = \langle u \rangle \in \mathcal{P}$. We wish to determine how many $L \in \mathcal{L}$ are adjacent to x in G. Such an L can be created by choosing $v \notin x$ and letting $L = \langle u, v \rangle$.
The number of choices of v is $q^3 - q$, but each such subspace is repeated for $q^2 - q$ choices of v since the cardinality of $\langle u, v \rangle$ is q^2. Therefore,
$$d(x) = \frac{q^3 - q}{q^2 - q} = q + 1.$$
The total number of edges is
$$|\mathcal{P}|(q+1) = (q+1)(q^2+q+1) = q^3 + 2q^2 + 2q + 1,$$
which is $\mathcal{O}(|\mathcal{P}|^{3/2})$! Therefore, our $\mathcal{O}(n^{3/2})$ bound is tight.
In fact, the Levi graph is optimal in the case where $n = 2(q^2+q+1)$, as seen in Corollary 1.5.

Corollary 1.5. For $n \in \mathbb{N}$,
$$z(n, n; 2, 2) \leq \frac{n(1 + \sqrt{4n-3})}{2}. \tag{1.4}$$
Consequently,
$$\mathrm{ex}(n; C_4) \leq \frac{1}{4} n (1 + \sqrt{4n-3}). \tag{1.5}$$

Proof. Equation (1.5) clearly follows from Equations (1.3) and (1.4), so it suffices to show the first equation. Equation (1.2) in the proof of the Kővári-Sós-Turán Theorem for the case where $s = t = 2$ just says that
$$\binom{n}{2} \geq n \binom{d}{2},$$
where $d = e(G)/m$. That is, $d^2 - d - (n-1) \leq 0$. Then,
$$d \leq \frac{1 + \sqrt{1 + 4(n-1)}}{2} = \frac{1 + \sqrt{4n-3}}{2},$$
which is exactly the bound we want. This bound is tight in the case where $n = 2(q^2+q+1)$, as seen in the Levi graph. ∎

Before we move on to the next section, let us build a tiny bit of intuition for why the construction of the Levi graph works.
The projective plane is chosen to ensure that any two distinct points determine a unique line (which holds even in the non-projective plane setting), and any two distinct lines determine a unique point. This corresponds to the absence of $K_{2,2}$ – if it *was* present as a subgraph, then there would be two lines (points) that determine two points (lines), which cannot happen!

1.3.3. The case where $s = t = 3$

We next look at $\mathrm{ex}(n; K_{3,3})$.

The Kővári-Sós-Turán Theorem applied here gives
$$z(n, n; 3, 3) \leq (2)^{1/3}(n-2)n^{2/3} + 2n,$$
which is $\mathcal{O}(n^{5/3})$.

Similar to the Levi graph, we construct an (algebraic) extremal example. Let p be a prime and fix some $r \in \mathbb{F}_p$. Consider the graph G that has vertex set $\mathcal{V} = \mathbb{F}_p^3$ where (x, y, z) is adjacent to (u, v, w) iff

$$(x-u)^2 + (y-v)^2 + (z-w)^2 = r. \tag{1.6}$$

Before moving on to explaining why this works, let us try to impart some intuition. (1.6) resembles the equation of a sphere in $\mathbb{R}^3$. If we take three spheres of the same radius, the points of intersection of all three must lie on two circles, corresponding to the circles of intersection of two pairs of spheres. Since any two circles meet at atmost two points, the absence of $K_{3,3}$ follows.
It may be shown that even over $\mathbb{F}_p$, two "spheres" intersect on a "circle" and any two circles meet at at most 2 points (Check this!).
So, if we have a $K_{3,3}$ in the described graph, we have three spheres (centered at each of the three points) that intersect at three points, which is not possible.

It remains to count the number of edges in this graph. To do so, let us estimate the degree of $(0, 0, 0)$, since all vertices have the same degree (Why?). That is, we want to determine

$$|\{(x, y, z) \in \mathbb{F}_p^3 : x^2 + y^2 + z^2 = r\}|.$$

Letting z be arbitrary, we want to find

$$N(\xi) = |\{(x, y) \in \mathbb{F}_p^2 : x^2 + y^2 = \xi\}|$$

for any arbitrary $\xi \in \mathbb{F}_p$.

Definition 1.12 (Legendre Symbol)**.** For $a \in \mathbb{F}_p$, define the **Legendre symbol**

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & a \in \mathbb{F}_p^\times \text{ is a square,} \\ -1, & a \in \mathbb{F}_p^\times \text{ is not a square,} \\ 0, & a = 0. \end{cases}$$

With the above, it is not too difficult to see that

$$N(\xi) = \sum_{(a,b):a+b=\xi} \left(1 + \left(\frac{a}{p}\right)\right)\left(1 + \left(\frac{b}{p}\right)\right).$$

Let us now compute the above quantity.

- First of all,
$$\sum_{(a,b):a+b=\xi} 1 = p.$$
- Observe that
$$\sum_{a \in \mathbb{F}_p} \left(\frac{a}{p}\right) = 0.$$
Indeed, the number of squares and non-squares in $\mathbb{F}_p^\times$ is the same. As a result,
$$\sum_{(a,b):a+b=\xi} \left(\frac{a}{p}\right) = \sum_{a \in \mathbb{F}_p} \left(\frac{a}{p}\right) = 0.$$

Therefore,

$$N(\xi) = p + \sum_{(a,b):a+b=\xi} \left(\frac{a}{p}\right)\left(\frac{b}{p}\right). \tag{1.7}$$

Notice that the map $\mathbb{F}_p^\times \to \{-1, 1\}$ given by $x \mapsto \left(\frac{x}{p}\right)$ is a group homomorphism.

Lemma 1.6. If $\xi \neq 0$, $N(\xi) = N(1)$.

Proof. Using (1.7) and letting $a' = a/\xi$ and $b' = b/\xi$,

$$\begin{aligned}
N(\xi) - p &= \sum_{a+b=\xi} \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \\
&= \sum_{a'+b'=1} \left(\frac{a'\xi}{p}\right)\left(\frac{b'\xi}{p}\right) \\
&= \sum_{a'+b'=1} \left(\frac{a'}{p}\right)\left(\frac{b'}{p}\right)\left(\frac{\xi^2}{p}\right) \\
&= \sum_{a'+b'=1} \left(\frac{a'}{p}\right)\left(\frac{b'}{p}\right) = N(1) - p.
\end{aligned}$$

∎

So, we have

$$\begin{aligned}
(p-1)(N(1)-p) &= \sum_{\xi \in \mathbb{F}_p^\times} \sum_{(a,b):a+b=\xi} \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \\
&= \sum_{\xi \in \mathbb{F}_p^\times} \sum_{a \in \mathbb{F}_p} \left(\frac{a}{p}\right)\left(\frac{\xi - a}{p}\right) \\
&= \sum_{a \in \mathbb{F}_p} \left(\left(\frac{a}{p}\right) \sum_{\xi \in \mathbb{F}_p^\times} \left(\frac{\xi - a}{p}\right) \right) \\
&= \sum_{a \in \mathbb{F}_p} \left(\frac{a}{p}\right) \left(0 - \left(\frac{-a}{p}\right)\right) \\
&= -\left(\frac{-1}{p}\right) \sum_{a \in \mathbb{F}_p} \left(\frac{a^2}{p}\right) && \text{(using the group homomorphism property)} \\
&= -(p-1)\left(\frac{-1}{p}\right)
\end{aligned}$$

and therefore, $N(1) = p - \left(\frac{-1}{p}\right)$.

If we choose r to be a non-square, then $r - z^2 \neq 0$ for any $z \in \mathbb{F}_p$ and $N(r - z^2) = p - \left(\frac{-1}{p}\right)$. In this case, the degree of $(0,0,0)$ is $\Theta(p^2)$. The size n of the vertex set $\mathcal{V}$ is p^3, so the number of edges is $\Theta(p^5)$ which is $\Theta(n^{5/3})$ and thus, the bound given by the Kővári-Sós-Turán Theorem is (asymptotically) tight.

For $s \leq t$, the Kővári-Sós-Turán Theorem gives

$$\mathrm{ex}(n; K_{s,t}) < C_{s,t} n^{2-1/s}$$

for some $c_{s,t}$ depending only on s and t. However, it is not known if the above bound is tight for most cases.
Lower bounds of the form

$$\mathrm{ex}(n; K_{s,t}) \geq c_{s,t} n^{2-1/t}$$

are known only if $t > (s-1)!$. See [KRS96] for more details.

1.4. P_{k+1}-free graphs

Next, we study $\mathrm{ex}(n; P_{k+1})$.

First of all, note that K_k is P_{k+1}-free. So, if we split the n vertices up into blocks of k vertices and add all edges within each block, the resulting graph will be P_{k+1}-free as well. That is, G_n is a disjoint union of K_ks (and possibly one K_r for $r < k$). For this particular graph G_n,

$$e(G_n) \leq \frac{(k-1)n}{2}.$$

It turns out that we cannot do better than this.

Theorem 1.7. $\mathrm{ex}(n; P_{k+1}) \leq (k-1)n/2$ with equality if and only if G_n is a disjoint union of K_ks.

The above theorem just says that equality is not attained for connected graphs without P_{k+1}.
To prove this, we prove the following (seemingly) more general statement.

Lemma 1.8. Let $G_n = (V, E)$ be connected and suppose $d(v) \geq k$ for all $v \in V$. If $n \geq 2k$, G_n contains a path of length $2k$. Otherwise, G_n contains an n-cycle.

Proof. We prove the result for the case where $n \geq 2k$ first. Consider the longest path $u = v_1, v_2, \ldots, v_r = v$ in G_n and let $U = \{v_i : i \in [r]\}$. We must show that $r \geq 2k$. Suppose otherwise and let $r < 2k$.
First of all, we must have $\Gamma(u) \subseteq U$ – otherwise, the path can be extended by adding another vertex from $\Gamma(u) \setminus U$. Similarly, $\Gamma(v) \subseteq U$.
Next, v_1 and v_r cannot be adjacent. If they are, then we can obtain a longer path by cycling through and choosing some edge from a vertex in U to one outside of U (such a vertex must exist since the graph is connected and has at least $2k$ vertices). More generally, if there exists i such that v_1, v_{i+1} and v_i, v_r are adjacent, then we arrive at a contradiction (for the same reason).
Let $S = \{i : v_i v_r \in E\}$ and $T = \{i : v_1 v_{i+1} \in E\}$. By the above observation, $S \cap T = \varnothing$. However, by our first observation, $|S| = d(v_r) \geq k$ and $|T| = d(v_1) \geq k$. Therefore, $r \geq |S \cup T| \geq 2k$.
The result for $n < 2k$ is shown using nearly the same proof. ∎

Proof of Theorem 1.7. We perform strong induction on n. We may assume that $n > k$. Suppose G_n has no P_{k+1}. If G_n is *not* connected, then it consists of a disjoint union of connected subgraphs. We may then apply the inductive hypothesis to each of these smaller pieces.
So, let G_n be connected. By Lemma 1.8, there is some vertex v such that $d(v) \leq (k-1)/2$ (otherwise, there must be a path of length k). Additionally, observe that G_n does not have any subgraph isomorphic to K_k – using the connectedness assumption gives a path of length k otherwise.
The graph $G \setminus \{x\}$[4] has no P_{k+1} and further, it has no K_k either. Therefore, $G \setminus \{x\}$ is *not* extremal and $e(G \setminus \{x\}) < (k-1)(n-1)/2$. Therefore,

$$\begin{aligned} e(G) &= d(x) + e(G \setminus x) \\ &< \frac{k-1}{2} + \frac{(k-1)(n-1)}{2} = \frac{(k-1)n}{2}. \end{aligned}$$ ∎

[4]This is the subgraph induced by G on the vertex set $V \setminus \{x\}$

§2. Fundamental Results in Extremal Graph Theory

2.1. The Erdős-Stone-Simonovits Theorem

2.1.1. Motivation

Using the folklore result that a graph is bipartite iff it has no odd cycle, we have

$$\mathrm{ex}(n; C_{2k+1}) \geq \left\lfloor \frac{n^2}{4} \right\rfloor.$$

Remark. It in fact turns out that for $n \gg 0$, $\mathrm{ex}(n; C_{2k+1}) = \lfloor n^2/4 \rfloor$, as we shall see later..
Is there any more general relationship between forbidden subgraphs and r-partite graphs?

Definition 2.1 (Chromatic Number)**.** Given a graph G, its **chromatic number** is given by

$$\chi(G) = \min\{r : G \text{ is } r\text{-partite}\}.$$

Alternatively, we can define the above using the following.

Definition 2.2. Given a graph $G = (V, E)$, an r**-coloring** of G is a function $f : V \to [r]$ such that for any $uv \in E$, $f(u) \neq f(v)$.
The chromatic number of a graph is the least r such that it is r-colorable.

Suppose H is an arbitrary graph such that $\chi(H) = r + 1$. Then, no r-partite graph contains H. As a result,

$$\mathrm{ex}(n; H) > t_r(n).$$

Our earlier observation on $\mathrm{ex}(n; C_{2k+1})$ is then just a consequence of the fact that C_{2k+1} is 3-colorable.
Let us give a more concrete example. Suppose we want to find $\mathrm{ex}(n; \text{Petersen})$, where Petersen is the Petersen graph.
It may be checked that Petersen has chromatic number 3. So,

$$\mathrm{ex}(n; \text{Petersen}) > t_2(n) = \left\lfloor \frac{n^2}{4} \right\rfloor.$$

Can we do better?
This is answered by the Erdős-Stone-Simonovits Theorem.

2.1.2. The result

Theorem 2.1 (Erdős-Stone-Simonovits Theorem, Version 1)**.** Let $r \geq 1$ and $0 < \varepsilon < 1/2$. Then, for $n \gg 0$ and any graph G_n, if

$$\delta(G_n) \geq \left(1 - \frac{1}{r} + \varepsilon\right) n, \tag{2.1}$$

there exist pairwise disjoint subsets $V_1, \ldots, V_{r+1}$ of V such that for each i,

$$|V_i| = t \geq \frac{\varepsilon \log n}{2^{r-1}(r-1)!}$$

and the complete $(r+1)$-partite graph on these subsets is contained in G_n.

Theorem 2.2 (Erdős-Stone-Simonovits Theorem, Version 2)**.** Let $r \geq 1$ and $0 < \varepsilon < 1/2$. Then, for $n \gg 0$ and any graph G_n, if

$$e(G_n) \geq \left(1 - \frac{1}{r} + \varepsilon\right)\binom{n}{2}, \tag{2.2}$$

there exist pairwise disjoint subsets $V_1, \dots, V_{r+1}$ of V such that for each i,

$$|V_i| = t \geq \frac{\varepsilon \log n}{2^{r+1}(r-1)!}$$

and the complete $(r+1)$-partite graph on these subsets is contained in G_n.

Observe the difference in the exponent of 2 in the denominator of t in the two versions.
For example, if $n \geq e^{32/\varepsilon}$ and Equation (2.2) is satisfied, then $T_3(12)$, and thus Petersen, is a subgraph of G_n.
Therefore, (replacing ε with 2ε)

$$\mathrm{ex}(n; \mathsf{Petersen}) \leq \left(\frac{1}{2} + 2\varepsilon\right)\binom{n}{2} \leq \left(\frac{1}{4} + \varepsilon\right)n^2.$$

for all $\varepsilon > 0$.

Corollary 2.3. If $\chi(H) = r + 1$, then for any $\varepsilon > 0$, for $n \gg 0$,

$$\left(1 - \frac{1}{r}\right)\frac{n^2}{2} < \mathrm{ex}(n; H) < \left(1 - \frac{1}{r} + \varepsilon\right)n^2 - \mathcal{O}(n).$$

Proof of Erdős-Stone-Simonovits Theorem, Version 1. We first show the result for $r = 1$. Let G_n be a graph such that $\delta(G_n) \geq \varepsilon n$. We want a "large" t such that $K_{t,t}$ is isomorphic to a subgraph G_n.
Suppose that G_n is $K_{t,t}$-free (we shall fix t later). Now, we have that

$$t\binom{n}{t} \overset{(1)}{>} |\{(v, T) : v \in V, T \subseteq \Gamma(v), |T| = t\}| = \sum_{v \in V}\binom{d(v)}{t} \overset{(2)}{\geq} n\binom{\varepsilon n}{t},$$

where (2) follows from the hypothesis on $\delta(G_n)$ and (1) follows due to our assumption that G_n is $K_{t,t}$-free.
Therefore,

$$n\binom{\varepsilon n}{t} < t\binom{n}{t}.$$

So,

$$\begin{aligned} 1 &< \frac{t}{n} \cdot \frac{n(n-1)\cdot(n-t+1)}{\varepsilon n(\varepsilon n - 1)\cdot(\varepsilon n - t + 1)} \\ &< \frac{t}{n} \cdot \left(\frac{n}{\varepsilon n - t + 1}\right)^t \\ &= \frac{t}{n}\frac{1}{\varepsilon^t}\left(\frac{1}{1 - \frac{t-1}{\varepsilon n}}\right)^t. \end{aligned}$$

We desire a large t such that the above results in a contradiction. Let us look at the last term. Suppose we want

$$\left(1 - \frac{t-1}{\varepsilon n}\right)^t \to 1.$$

Equivalently,

$$t \log\left(1 - \frac{t-1}{\varepsilon n}\right) \to 0$$

$$t\left(\frac{t-1}{\varepsilon n} + \frac{1}{2}\left(\frac{t-1}{2n}\right)^2 + \frac{1}{3}\left(\frac{t-1}{\varepsilon n}\right)^3 + \cdots\right) \to 0.$$

If $\varepsilon n > t^4$, the above holds (Why?).
That is, if $t < (\varepsilon n)^{1/4}$, then for sufficiently large n, t,

$$\left(1 - \frac{t-1}{\varepsilon n}\right)^{-t} < 2$$

and

$$1 < \frac{t}{n} \cdot \frac{1}{\varepsilon^t}\left(1 - \frac{t-1}{\varepsilon n}\right)^{-t} < \frac{2t}{\varepsilon^t n}.$$

If the above is ≤ 1, we shall arrive at a contradiction. That is, if

$$\log(2t) + t\log\left(\frac{1}{\varepsilon}\right) < \log n,$$

then we arrive at a contradiction. In particular,

$$t = \lceil \varepsilon \log n \rceil$$

is a suitable choice (Why?), completing the proof of the theorem for $r = 1$.

Now, let us prove the general case by performing induction on r. Let G_n be a graph with

$$\delta(G_n) \geq \left(1 - \frac{1}{r} + \varepsilon\right) n.$$

Now,

$$1 - \frac{1}{r} + \varepsilon > 1 - \frac{1}{r-1} + \frac{1}{r(r-1)}.$$

Let $\varepsilon' = 1/r(r-1)$. By induction, there are $V_1', \ldots, V_r'$ such that

$$|V_i'| \geq t' = \frac{\log n}{r(r-1)2^{r-2}(r-2)!} = \frac{\log n}{2^{r-2}r!} \tag{2.3}$$

for each i and the complete r-partite graph on these sets is a subgraph of G_n. Let $K = \bigcup V_i'$.
We obviously have $\varepsilon < 1/r$, since the claim is vacuously true otherwise.
We shall find $V_i \subseteq V_i'$ for each $1 \leq i \leq r$ and some $V_{r+1} \subseteq V \setminus K$ such that the complete bipartite graph on $(V_1, \ldots, V_{r+1})$ is the required subgraph of G_n.
Now, define

$$U = \left\{x \in V \setminus K : d(x, K) \geq \left(1 - \frac{1}{r} + \lambda\right)|K|\right\}$$

for some λ we shall fix later. We shall bound $e(K, V \setminus K)$ in two different ways. From the perspective of K,

$$e(K, V \setminus K) \geq |K|\left(\left(1 - \frac{1}{r} + \varepsilon\right)n - |K|\right).$$

From the perspective of $V \setminus K$,

$$\begin{aligned} e(K, V \setminus K) &= e(V \setminus (K \sqcup U), K) + e(U, K) \\ &\leq \big(n - |U| - |K|\big)\left(1 - \frac{1}{r} + \lambda\right)|K| + |U||K|. \end{aligned}$$

Putting the two together,

$$|K|\left(\left(1-\frac{1}{r}+\varepsilon\right)n-|K|\right)\leq(n-|U|-|K|)\left(1-\frac{1}{r}+\lambda\right)|K|+|U||K|$$
$$\left(1-\frac{1}{r}+\varepsilon\right)n-|K|\leq(n-|U|-|K|)\left(1-\frac{1}{r}+\lambda\right)+|U|.$$

Set $\lambda=\varepsilon/2$. We then have

$$\frac{r}{2}\varepsilon n\leq\left(1-\frac{r\varepsilon}{2}\right)(|U|+|K|).$$

So,

$$|U|\geq\frac{r\varepsilon}{2-r\varepsilon}n-|K|\overset{(*)}{\geq}\frac{\varepsilon}{1-\varepsilon}n-rt',$$

where $(*)$ follows from the fact that $r\varepsilon/(2-r\varepsilon)$ is increasing in r and $r\geq 2$.
Since t', and thus rt', is of the order of $\log n$, the first term in the expression dominates for sufficiently large n. So, for $n\gg 0$, $|U|\geq\varepsilon n$.

Now,

$$\left(1-\frac{1}{r}+\frac{\varepsilon}{2}\right)|K|\geq\left(1-\frac{1}{r}+\frac{\varepsilon}{2}\right)rt'$$
$$\geq(r-1)t'+\frac{\varepsilon r}{2}t'.$$

This implies that each $u\in U$ has at least $(\varepsilon r/2)t'$ neighbours in each V_i'.
We can now use a pigeonhole argument to choose a subset of U whose vertices are all adjacent to some common set of vertices in each V_i'. To do so, consider

$$|\{(u,W_1,\ldots,W_r):W_i\subseteq V_i',|W_i|=(\varepsilon r/2)t',\text{ and } u \text{ is adjacent to all the vertices of each } W_i\}|.$$

By our earlier observation, this must be at least $|U|\geq\varepsilon n$.
On the other hand, it is at most the number of ways of choosing the W_i, which is $\binom{t'}{(\varepsilon r/2)t'}^r$.
In particular, using a pigeonhole argument, there exist $V_1,\ldots,V_r$ and a $V_{r+1}\subseteq U$ such that $V_i\subseteq V_i'$ for each $1\leq i\leq r$ and for all $u\in V_{r+1}$, $(u,V_1,\ldots,V_r)$ is in the set whose cardinality we just considered, and $|V_{r+1}|\geq\varepsilon n/\binom{t'}{(\varepsilon r/2)t'}^r$. Let us now bound this expression.

$$\begin{aligned}|V_{r+1}|&\geq\frac{\varepsilon n}{\binom{t'}{(\varepsilon r/2)t'}^r}\\&\geq\frac{\varepsilon n}{(2e/\varepsilon r)^{t'\varepsilon r^2/2}}&&\text{(see here for the bound used)}\\&\geq\varepsilon n\cdot\left(\frac{\varepsilon}{e}\right)^{t'\varepsilon r^2/2}&&\text{(since } r\geq 2,\ 2/r\leq 1).\end{aligned}$$

Setting

$$t=\frac{\varepsilon\log n}{2^{r-1}(r-1)!},$$

we see that $t=\varepsilon rt'/2$.

Keeping in mind that the bound we want is $|V_{r+1}| \geq t$,

$$\begin{aligned}
\log\left(n\left(\frac{\varepsilon}{e}\right)^{t'\varepsilon r^2/2}\right) &\geq \log n + \frac{\varepsilon r^2 t'}{2}\log\left(\frac{\varepsilon}{e}\right) \\
&\geq t'\left(2^{r-2}r! - \log\left(\frac{e}{\varepsilon}\right)\cdot\frac{\varepsilon r^2}{2}\right) && \text{(using the expression for } t' \text{ in (2.3))} \\
&\geq t'\left(2^{r-2}r! - \log(2e)\cdot\frac{r^2}{4}\right) && (-\varepsilon\log(e/\varepsilon) \text{ is decreasing in } \varepsilon) \\
&\geq \log(rt'/2).
\end{aligned}$$

Therefore, $V_{r+1} \geq t$. Since $|V_i| = \varepsilon r t'/2 = t$ by definition, the proof is complete. ■

Whew. Let us now give a simple corollary of the above result.

Porism 2.4. Suppose $H_1, \ldots, H_m$ are graphs. Then, for any $\varepsilon > 0$ and $n \gg 0$,

$$\mathrm{ex}(n; H_1, \ldots, H_m) \leq \left(1 - \frac{1}{r} + \varepsilon\right)\frac{n^2}{2} - \mathcal{O}(n^2),$$

where $r + 1 = \max\{\chi(H_1), \ldots, \chi(H_m)\}$.

Before we move onto the proof of the second version of the Erdős-Stone-Simonovits Theorem, we give a lemma that will assist in its proof.

Lemma 2.5. For $n \gg 0$, if $e(G_n) \geq (c+\varepsilon)\binom{n}{2}$ for some $c > 0$ and $\varepsilon > 0$, then there exists $H \subseteq G_n$ such that

1. $|H| \geq \sqrt{\varepsilon} n$.
2. $\delta(H) \geq c|H|$.

Proof. Assume that G_n itself does not satisfy the conclusions. Then, there exists some $x_n \in V(G_n)$ such that $d(x_n) < cn$.
Let $H_{n-1} = G \setminus \{x_n\}$. If H_{n-1} fails the second conclusion of the theorem, there exists $x_{n-1} \in V(H_{n-1})$ such that $d(x_{n-1}) < c(n-1)$.
Repeating the above, we get a sequence of graphs $G_n = H_n \supsetneq H_{n-1} \supsetneq \cdots \supsetneq H_\ell$, where $V(H_{n-r}) \setminus V(H_{n-r-1}) = \{x_{n-r}\}$ for each r and $\ell \geq \sqrt{\varepsilon} n$. Further, for each i, $d_{H_i}(x_i) < ci$ for $i = \ell, \ldots, n$.
Now,

$$\begin{aligned}
e(H_\ell) &> e(H_n) - \big(cn + c(n-1) + \cdots + c(\ell+1)\big) \\
&= (c+\varepsilon)\binom{n}{2} - c\left(\frac{n(n+1)}{2} - \frac{\ell(\ell+1)}{2}\right) \\
&= (c+\varepsilon)\binom{n}{2} - c\left(\binom{n}{2} + n - \binom{\ell+1}{2}\right) \\
&= \varepsilon\binom{n}{2} - cn + c\binom{\ell+1}{2}.
\end{aligned}$$

Now, the final expression must be at most $\binom{\ell}{2}$ (since H_ℓ has ℓ vertices).
As a result, it would suffice to show that the above expression on taking $\ell = \lfloor \varepsilon n \rfloor$ is greater than $\binom{\ell}{2}$ for sufficiently

large n (this implies that the sequence must stop before reaching this ℓ due to one of the graphs satisfying the conclusions). Indeed, this is seen to be true as

$$\begin{aligned}
\varepsilon\binom{n}{2} - cn + c\binom{\ell+1}{2} &\geq \varepsilon\binom{n}{2} + c\left(\frac{(\lfloor\sqrt{\varepsilon}n\rfloor+1)\lfloor\sqrt{\varepsilon}n\rfloor}{2} - n\right) \\
&\geq \varepsilon\binom{n}{2} && \text{(for sufficiently large } n) \\
&\geq \binom{\ell}{2},
\end{aligned}$$

completing the proof. ■

2.2. An Introduction to Random Graphs

Definition 2.3 (Erdős-Rényi Model)**.** Fix $0 \leq p \leq 1$. The **Erdős-Rényi random graph model**, denoted $G_{n,p}$ is the random variable which is a graph with vertex set $[n]$, such that for each $\{i,j\} \in \binom{[n]}{2}$, $\{i,j\}$ is an edge with probability p, independently across distinct pairs.

So, for any graph H on vertex set $[n]$,

$$\Pr(G_{n,p} = H) = p^{e(H)}(1-p)^{\binom{n}{2}-e(H)}.$$

Remark. $G_{n,1/2}$ is the uniform distribution on the set of graphs on $[n]$.

A recurring theme in probability theory is that random objects tend to behave very nicely given a large number of samples (along the lines of the laws of large numbers and the central limit theorem).

2.2.1. A motivating extremal problem (bounding $\mathrm{ex}(n; C_{2k})$)

To understand why random graphs are important, let us look at $\mathrm{ex}(n; C_{2k})$.

Theorem 2.6. For any k, there exists a constant c such that for $n \gg 0$, there is a C_k-free G_n with

$$e(G_n) \geq c \cdot n^{1+1/(k-1)}.$$

The above result does not yield anything useful for k odd.

Proof. Consider $G_{n,p}$ for some p we shall fix later. Let $N(G)$ be the number of copies of C_k in a given graph G. Given a cycle $(v_1, \ldots, v_k)$, observe that the sequences $(v_2, v_3, \ldots, v_k, v_1)$ and $(v_k, v_{k-1}, \ldots, v_1)$ determine the same cycle. That is, performing cyclic shifts of a sequence of vertices or reversing their order around gives the same cycle. Let $\mathcal{C}$ be the set of all these cycles.[5] By our observation,

$$|\mathcal{C}| = \frac{n(n-1)\cdots(n-k+1)}{2k} = \frac{n!}{2k \cdot k!}.$$

[5]This can be made more formal by taking all length k sequences of $[n]$ consisting of distinct elements and considering the equivalence classes formed by the equivalence relation defined on the previous line.

Now,

$$\begin{aligned}
\mathbf{E}[N(G_{n,p})] &= \mathbf{E}\left[\sum_{(v_1,\dots,v_k)\in\mathcal{C}} \mathbb{1}_{v_1v_2,v_2v_3,\dots,v_{k-1}v_k,v_kv_1 \text{ are edges}}\right] \\
&= \sum_{(v_1,\dots,v_k)\in\mathcal{C}} \Pr\left[\mathbb{1}_{v_1v_2,v_2v_3,\dots,v_{k-1}v_k,v_kv_1 \text{ are edges}}\right] && \text{(linearity of expectation)} \\
&= \sum_{(v_1,\dots,v_k)\in\mathcal{C}} p^k \\
&= \frac{n!}{2k\cdot k!}p^k.
\end{aligned}$$

This quantity is obviously less than $(np)^k/2k$. If p, and thus the expectation is small, we expect to not see many C_ks. On the other hand,

$$\begin{aligned}
\mathbf{E}[e(G_{n,p})] &= \mathbf{E}\left[\sum_{\{i,j\}\in\binom{[n]}{2}} \mathbb{1}_{ij \text{ is an edge}}\right] \\
&= \binom{n}{2}p.
\end{aligned}$$

Given a graph, if we delete an (arbitrary) edge from each copy of C_k in it, we will be left with no cycles. That is, given any graph G, there is a graph on the same vertex set with $e(G) - N(G)$ edges that is C_k-free. Inspired by this, by the linearity of expectation,

$$\mathbf{E}[e(G_{n,p}) - N(G_{n,p})] \geq \binom{n}{2}p - \frac{(np)^k}{2k}.$$

If we set $p = \left(\frac{k}{2}\right)^{1/(k-1)} n^{-1+1/(k-1)}$, then the above quantity is at least $n(n-1)p/4 \geq c\cdot n^{1+1/(k-1)}$ for an appropriate constant c and $n \gg 0$, completing the proof. ■

2.2.2. Digression: A coloring result of Erdős

The question we consider in this section is:

> Are there C_3-free graphs with large chromatic number?

More generally,

> Are there graphs with large chromatic number and large girth?

For example, if the girth of a graph is 7, then there cannot be adjacent vertices v, w such that $\Gamma(v)\cap\Gamma(u)$ and $\Gamma(w)\cap\Gamma(u)$ are non-empty for some u distinct from v, w. As a result, we can draw a "2-step tree" rooted at any u, which has $\Gamma(u)$ at the first level and the neighbours (other than u) of vertices of $\Gamma(u)$ at the second.
This seems to suggest some level of sparseness in the graph, due to which there are not too many edges and as a result, the chromatic number is low. However, it turns out that this intuition is not true, as proved by Erdős in [Erd59].

Theorem 2.7. There exist graphs with arbitrarily large girth and chromatic number. That is, given a $g, k \geq 3$, there exists a graph G_n such that $\operatorname{girth}(G) > g$ and $\chi(G) > k$.

Proof. Consider $G_{n,p}$ for some p we fix later. Further assume that $np \geq 1$.
Given a graph G, let $N_i(G)$ (for $3 \leq i \leq g$) be the number of cycles of size i in G. As we saw in the proof of Theorem 2.6,

$$\mathbf{E}[N_i(G_{n,p})] = \frac{n!}{2i \cdot i!} p^i < \frac{(np)^i}{6}.$$

Let $N = \sum_{i=3}^{g} N_i$. Then

$$\begin{aligned}
\mathbf{E}\left[N(G_{n,p})\right] &= \mathbf{E}\left[\sum_{i=3}^{g} N_i(G_{n,p})\right] \\
&< \frac{(np)^3}{6}\left(\frac{(np)^{g-2}-1}{np-1}\right) \\
&< \frac{(np)^g}{3}. && \text{(since } np \geq 1\text{)}
\end{aligned}$$

Using Markov's inequality,

$$\Pr\left[N(G_{n,p}) > \frac{2}{3}(np)^g\right] < \frac{1}{2} \tag{2.4}$$

This takes care of the girth (we want the above probability to be small). On the other hand, we need to make the chromatic number large. Towards this, observe that $\chi(G) \geq n/\alpha(G)$ (Why?). We have

$$\begin{aligned}
\Pr\left[\alpha(G_{n,p}) \geq r\right] &= \Pr\left[\bigcup_{X \subseteq [n] : |X| = r} \{X \text{ is independent}\}\right] \\
&\leq \sum_{X \subseteq [n] : |X| = r} \Pr[X \text{ is independent}] \\
&= \sum_{X \subseteq [n] : |X| = r} (1-p)^{\binom{r}{2}} \\
&\leq \binom{n}{r} \cdot e^{-pr(r-1)/2} \\
&\leq \left(\frac{en}{r}\right)^r \cdot e^{-pr^2/3} \\
&= \left(\frac{e^{1-pr/3} n}{r}\right)^r.
\end{aligned}$$

We shall choose r and p such that with positive probability, $N(G_{n,p}) \leq 2(np)^g/3$ *and* $\alpha(G_{n,p}) < r$. This implies the existence of a graph G_n such that both of the above hold.
We cannot use the tactic of removing edges we did in the earlier proof since that might increase α. Deleting vertices on the other hand works, since this can increase neither α nor N.
If we delete a single vertex from each cycle involved in N, the resulting graph will have girth greater than g. That is, given a graph G_n, there exists a graph with at least $n - N(G_n)$ vertices that has girth greater than g. Denote this corresponding graph as G'_n.

Set $p = n^{1/(g+1)-1}$. In this case, $2(np)^g/3 < n/2$ for $n \gg 0$ and using (2.4),

$$\Pr\left[n - N(G_{n,p}) > \frac{n}{2}\right] \geq \frac{1}{2}. \tag{2.5}$$

Set $r = 4 \log n/p = 4n^{1-1/(g+1)} \log n$. For these values of p and r,

$$\Pr\left[\alpha(G_{n,p}) \geq r\right] \xrightarrow{n \to \infty} 0. \tag{2.6}$$

Using (2.5), our construction of G', and (2.7), it is true with with positive probability that

$$|G'_{n,p}| \geq \frac{n}{2}, \operatorname{girth}(G'_{n,p}) > g,$$

and

$$\chi(G'_{n,p}) \geq \frac{n}{\alpha(G'_{n,p})} \geq \frac{n}{r} = \frac{n^{1/(g+1)}}{4\log n} \xrightarrow{n\to\infty} \infty.$$

Therefore, for $g, k \geq 3$, there exists $n \gg 0$ and graph G on n vertices such that $\operatorname{girth}(G) > g$ and $\chi(G) > k$. ■

2.3. Szemerédi's Regularity Lemma

The second of our powerful results in extremal graph theory (after the Erdős-Stone-Simonovits Theorem) is Szemerédi's Regularity Lemma, which says that any sufficiently large graph behaves in some way like a random graph.

2.3.1. Motivation

First, let us give a bound from probability theory that will be useful.

Lemma 2.8 (Chernoff Bound)**.** Suppose $X \sim B(n,p)$, the binomial distribution with parameters n, p. Then for any $t \geq 0$,

$$\Pr\left[X - \mathbf{E}[X] \geq t\right] \leq \exp\left(-\frac{t^2}{2\left(\mathbf{E}[X] + t/3\right)}\right)$$

and

$$\Pr\left[X - \mathbf{E}[X] \leq -t\right] \leq \exp\left(-\frac{t^2}{2\mathbf{E}[X]}\right).$$

Fix disjoint $A, B \subseteq [n]$ and let $|A| = a$, $|B| = b$. Then given a graph G on $[n]$,

$$e(A,B) = \sum_{x \in A, y \in B} \mathbb{1}_{xy \in E(G)}.$$

Fix $0 < p < 1$. Then if $G \sim G_{n,p}$,

$$e(A,B) \sim B\left(ab, p\right).$$

By the Chernoff Bound, for some fixed constant c,

$$\Pr\left[|e(A,B) - pab| > c\left(b\sqrt{pa\log\left(\frac{2n}{b}\right)}\right)\right] \xrightarrow{n\to\infty} 0.$$

In particular, if $a = b = \alpha n$ for some $0 < \alpha < 1/3$, then with high probability,

$$|e(A,B) - pab| = \mathcal{O}\left(b\sqrt{ap\log\left(\frac{2n}{b}\right)}\right)$$

for any sets A, B of sizes a and b respectively.
This seems to say that the actual number of edges between two sets of the given size does not deviate very much from the expected number of edges between the two sets. The expression on the right is of the order of $\mathcal{O}(n\sqrt{n})$, which is asymptotically less than the expectation $pab = \mathcal{O}(n^2)$.

2.3.2. The Result

The regularity lemma gives a qualitative version of the above observation. Before we move to the actual result, let us provide some notation.

Definition 2.4 (Density)**.** Given a graph $G = (V, E)$ andr $U, W \subseteq V$, the **density** $d(U, W)$ is equal to $e(U, W)/|U||W|$.

Remark. Here, $e(U, W)$ is $\{(u, w) \in U \times W : \{u, w\} \in E\}$. If U and W are disjoint, this is the same as our earlier definition of $e(\cdot, \cdot)$. If they are not disjoint however, edges within the intersection are counted *twice* in our current definition.
This does not matter all that much since we usually apply the regularity lemma on disjoint sets.

Definition 2.5 (ε-regular pair)**.** Suppose $0 < \varepsilon < 1$. A pair of subsets (U, W) is said to be ε**-regular** if for any $A \subseteq U$, $B \subseteq W$ with $|A| \geq \varepsilon|U|$ and $|B| \geq \varepsilon|W|$, we have

$$|d(A, B) - d(U, W)| \leq \varepsilon$$

This corresponds to some sort of uniform behaviour throughout the sets, where subsets behave similarly to their parent sets in terms of density. If (U, W) is ε-regular, all sufficiently large subsets of U, W have roughly the same edge density as (U, W).

Theorem 2.9 (Szemerédi's Regularity Lemma)**.** Given $0 < \varepsilon < 1$, there exists M such that for $n \gg 0$, any graph G_n admits a vertex partition $\mathcal{P} = (V_0, V_1, \ldots, V_k)$, where

- $k \leq M$,
- $|V_0| \leq \varepsilon n$ (V_0 is known as an "exceptional set"),
- all the V_i for $1 \leq i \leq k$ are of equal size, and
- the number of ε-*ir*regular pairs (V_i, V_j) $(1 \leq i, j \leq k)$ is at most εk^2.

Such a partition where the number of ε-regular pairs is at most εk^2 is often referred to as an ε**-regular partition**.

We present the proof of the above, that uses an "energy increment" argument, over a series of lemmas.

Definition 2.6 (Energy)**.** Given a graph G_n with vertex set V, for disjoint $U, W \subseteq V$, define the **energy** of the pair (U, W) by

$$q(U, W) = \frac{|U||W|}{n^2} d^2(U, W).$$

If $\mathcal{U} = \{U_1, \ldots, U_m\}$ and $\mathcal{W} = \{W_1, \ldots, W_\ell\}$ are partitions of U and W respectively, then the energy of the pair $(\mathcal{U}, \mathcal{W})$ is

$$q(\mathcal{U}, \mathcal{W}) = \sum_{\substack{1 \leq i \leq m \\ 1 \leq j \leq \ell}} q(U_i, W_j).$$

For a partition $\mathcal{P}$ of V, let

$$q(\mathcal{P}) = q(\mathcal{P}, \mathcal{P}) = \sum_{U,W \in \mathcal{P}} q(U, W).$$

Lemma 2.10. If U,W are disjoint subsets of V and $\mathcal{U}$,$\mathcal{W}$ are partitions of U,W respectively, then

$$q(\mathcal{U}, \mathcal{W}) \geq q(U, W).$$

Proof. Independently pick u,w uniformly randomly from V. Define the random variable

$$Z(u, w) = \begin{cases} d(U', W'), & u \in U' \in \mathcal{U}, w \in W' \in \mathcal{W}, \\ 0, & \text{otherwise.} \end{cases}$$

Then,

$$\begin{aligned}
\mathbf{E}[Z]^2 &= \left(\sum_{\substack{U' \in \mathcal{U} \\ W' \in \mathcal{W}}} d(U', W') \left(\frac{|U'||W'|}{|U||W|} \right) \right)^2 \\
&= \left(\frac{1}{|U||W|} \sum_{\substack{U' \in \mathcal{U} \\ W' \in \mathcal{W}}} e(U', W') \right)^2 \\
&= d(U, W)^2 \\
&= \frac{n^2}{|U||W|} q(U, W) \text{ and} \\
\mathbf{E}[Z^2] &= \sum_{\substack{U' \in \mathcal{U} \\ W' \in \mathcal{W}}} d^2(U', W') \left(\frac{|U'||W'|}{|U||W|} \right) \\
&= \frac{n^2}{|U||W|} q(\mathcal{U}, \mathcal{W}).
\end{aligned}$$

Since variance is always non-negative, $\mathbf{E}[Z^2] \geq \mathbf{E}[Z]^2$ and thus,

$$q(\mathcal{U}, \mathcal{W}) \geq q(U, W).$$ ■

Partitioning increases energy! Also observe that since the random variable Z is at most 1 $(1/2$, in fact), so is the energy.
This also implies that if the partition $\mathcal{Q}$ is a refinement of $\mathcal{P}$, then

$$q(\mathcal{Q}) \geq q(\mathcal{P}).$$

Lemma 2.11. If (U, W) is not ε-regular, there is a partition $\mathcal{U} = \{U_1, U_2\}$ and $\mathcal{W} = \{W_1, W_2\}$ of U,W respectively such that

$$q(\mathcal{U}, \mathcal{W}) > q(U, W) + \varepsilon^4 \frac{|U||W|}{n^2}.$$

Proof. By the ε-irregularity, there exist $U_1 \subseteq U$ and $W_1 \subseteq W$ with $|U_1| \geq \varepsilon|U|$, $|W_1| \geq \varepsilon|W|$, and

$$|d(U_1, W_1) - d(U, W)| > \varepsilon.$$

Consider the partitions $\mathcal{U} = \{U_1, U \setminus U_1\}$ and $\mathcal{W} = \{W_1, W \setminus W_1\}$. With the same Z as in the proof of the previous lemma,

$$\begin{aligned}\mathbf{E}[Z^2] - \mathbf{E}[Z]^2 &= \mathbf{E}[(Z - \mathbf{E}[Z])^2] \\ &\geq \mathbf{E}[(Z - \mathbf{E}[Z])^2 \mathbb{1}_{u \in U', w \in W'}] \\ &= \varepsilon^4.\end{aligned}$$

The claim follows. ■

Lemma 2.12. Suppose $0 < \varepsilon < 1/4$ and let $\mathcal{P} = (V_0, V_1, \dots, V_k)$ be a partition of V such that

- $\mathcal{P} \setminus \{V_0\}$ is not ε-regular (there are at least εk^2 irregular pairs),
- $\mathcal{P} \setminus \{V_0\}$ is equitable[6], and
- $|V_0| \leq \varepsilon n$.

Then, there is a refinement $\mathcal{Q} = (V_0', V_1', \dots, V_\ell')$ of $\mathcal{P}$ such that

- $\ell \leq k4^k$,
- $\mathcal{Q} \setminus \{V_0'\}$ is equitable,
- $|V_0'| \leq |V_0| + n/2^k$, and
- $q(\mathcal{Q}) \geq q(\mathcal{P}) + \varepsilon^5/2$.

Proof. Let $|V_i| = t$ for all $1 \leq i \leq k$.
Suppose (V_i, V_j) is an ε-irregular pair in $\mathcal{P}$ and let $(V_{i,1}, V_{i,2})$ and $(V_{j,1}, V_{j,2})$ be the partitions of V_i, V_j respectively described in Lemma 2.11. If $\mathcal{Q}_1$ is this particular refinement

$$\mathcal{P} \cup \{V_{i,1}, V_{i,2}, V_{j,1}, V_{j,2}\} \setminus \{V_i, V_j\}$$

of $\mathcal{P}$, then

$$q(\mathcal{Q}_1) \geq q(\mathcal{P}) + \varepsilon^4 \frac{|V_i||V_j|}{n^2}$$

Let us similarly produce refinements corresponding to all irregular pairs of $\mathcal{P}$. Let $\mathcal{Q}'$ be the "common" refinement of all these partitions $(\mathcal{Q}_i)$. That is, for each $v \in V$, v is placed in the subset $\bigcap_{W \in \mathcal{Q}_i : v \in W} W$.
We then have

$$\begin{aligned}q(\mathcal{Q}') &\geq q(\mathcal{P}) + \varepsilon^4 \cdot \frac{t^2}{n^2}(\varepsilon k^2) \\ &= q(\mathcal{P}) + \varepsilon^5 \frac{(tk)^2}{n^2} \\ &\geq q(\mathcal{P}) + \frac{\varepsilon^5}{2}. \qquad (tk = n - |V_0| \geq (1-\varepsilon)n \geq 3n/4 \text{ and } 9/16 \geq 1/2) \end{aligned} \tag{2.7}$$

To make the partition $\mathcal{Q}'$ equitable, create the partition $\mathcal{Q}$ using it as follows.
Suppose we partition V_i into (V_{ij}), where V_{ij} is formed due to the irregularity of (V_i, V_j). Partition each of these parts of $\mathcal{Q}'$ into sets of size $b := \lfloor t/4^k \rfloor$. Whatever residual part cannot be cut out in this manner, we merge with V_0.
Since partitioning can only increase energy, this operation will only strengthen (2.7), if anything.

[6]All the blocks of the partition are of equal size.

Since the size of any block of $\mathcal{Q}$ is b, the number of blocks is at most $n/b = k4^k$.
Finally, what is the size of $|V_0'|$?
$\mathcal{Q}'$ has at most $k \cdot 2^{k-1}$ parts (there are at most $(k-1)$ 2-part partitions of each block of $\mathcal{P}$, and together they give a partition of size at most 2^{k-1}). So,

$$\begin{aligned}|V_0'| &\le |V_0| + |\mathcal{Q}'|b \\ &\le |V_0| + k2^{k-1}\frac{t}{4^k} \\ &\le |V_0| + \frac{n}{2^k},\end{aligned}$$

as desired. ■

Observe that with the above lemma, Szemerédi's Regularity Lemma follows without too much difficulty.
Begin with a partition $\mathcal{P}_0$ with k_0 parts, where $2^{k_0} \ge 2/\varepsilon$.
Given $\mathcal{P}_k$, let $\mathcal{P}_{k+1}$ be the partition defined by Lemma 2.12. Since the energy of any partition is bounded above by 1, this process must terminate after at most $2/\varepsilon^5$ steps.
Further, since the size of the refined partition is bounded between quantities dependent solely on the old partition, the size of the final partition after termination of the above process is at most some quantity dependent only on ε (this quantity might be massive[7], but that is besides the point).

2.3.3. Interesting corollaries

In this section, we cover several interesting corollaries of Szemerédi's Regularity Lemma.

Before we begin however, how do we process the regularity lemma? We typically begin with a "cleaning" of the graph as follows.

- Given $\varepsilon > 0$, get a partition as described in the lemma.
- First, delete all edges between irregular pairs. This loses at most $(\varepsilon k^2)t^2 < \varepsilon n^2$ edges.
- Delete all edges between "sparse" pairs, where we say that a pair (V_i, V_j) is sparse if $d(V_i, V_j) < \varepsilon$ (say). This loses at most $\binom{k}{2}(\eta t^2) < (\varepsilon/2)n^2$ edges.
- Delete all edges inside the exceptional part V_0. This loses at most $(\varepsilon^2/2)n^2$ edges.

All of these deletions cost at most $2\varepsilon n^2$ edges. In the remaining graph, *all* pairs (V_i, V_j) for $1 \le i, j \le k$ are ε-regular. Further, if $e(V_i, V_j) > 0$ (for $i \ne j$), then $d(V_i, V_j) \ge \varepsilon$.

Based on this, let us give a corollary of Szemerédi's Lemma.

Theorem 2.13 (Triangle Counting Lemma)**.** Suppose V_1,V_2,V_3 forms a pairwise ε-regular partition of the vertex set of graph G_n, and that $d(V_i, V_j) \ge d$ (for some $d \ge 2\varepsilon$). Then, there are at least

$$\varepsilon^2(1-2\varepsilon)(d-\varepsilon)|V_1||V_2||V_3|$$

triangles xyz with $x \in V_1$, $y \in V_2$, $z \in V_3$.

[7] Unbelievably so. For example, $\varepsilon = 1/8$ gives a bound of the order of $4 \uparrow\uparrow 2^{15}$ (using Knuth's up-arrow notation). It further turns out that such a massive tetration-type bound is necessary, as proved in [Gow97].

Proof. Let

$$X_1 = \{x \in V_1 : d(x, V_2) \leq \varepsilon|V_2|\}.$$

We claim that $|X_1| < \varepsilon|V_1|$. Suppose otherwise. Then $e(X_1, V_2)/|X_1||V_2| < \varepsilon$ and further, ε-regularity implies that $d(X_1, V_2) \geq (d - \varepsilon)$, which leads to a contradiction.
Similarly, $X_2 := \{x \in V_1 : d(x, V_3) \leq \varepsilon|V_3|\}$ has size less than $\varepsilon|V_1|$ too.
For all $x \in V_1 \setminus (X_1 \cup X_2)$, $d(x, V_2) \geq \varepsilon|V_2|$ and $d(x, V_3) \geq \varepsilon|V_3|$. Fix such an x and let $X_i' = \Gamma(x) \cap V_i$ (for $i = 2, 3$).
By the ε-regularity of (V_2, V_3), $d(V_2, V_3) - d(X_2', X_3') \leq \varepsilon$ so $d(X_2', X_3') \geq d - \varepsilon$. As a result,

$$e(X_2', X_3') \geq (d - \varepsilon)|X_2'||X_3'| \geq (d - \varepsilon)\varepsilon^2|V_2||V_3|.$$

Therefore, the number of triangles of the desired form is at least

$$\underbrace{(1 - 2\varepsilon)|V_1|}_{\substack{\text{the number of} \\ \text{such } x}} \cdot \underbrace{(d - \varepsilon)\varepsilon^2|V_2||V_3|}_{\substack{\text{the number of triangles} \\ \text{corresponding to each } x}},$$

completing the proof. ■

Along similar lines is the following result.

Theorem 2.14 (Triangle Removal Lemma)**.** Given $\varepsilon > 0$, there exists δ (depending only on ε) such that for $n \gg 0$, any graph G_n with at most δn^3 triangles can be made triangle-free by deleting at most εn^2 edges.

Proof. Start with an $(\varepsilon/4)$ regular partition of G_n using Szemerédi's Regularity Lemma and the cleaning process from earlier. Delete all edges within each of the V_i. This costs at most $k(t^2/2) < (kt)^2/2k \leq \varepsilon n^2/4$ edges.
So, we have lost $\leq \varepsilon n^2$ edges in all. If there is a triangle remaining in the graph, it must come from a triple (V_i, V_j, V_k) with all three pairs being $(\varepsilon/4)$-regular and density at least $\varepsilon/2$. We can then use the triangle counting lemma to conclude that there are at least $t^3(\varepsilon/4)^3(1 - \varepsilon/2) > n^3(\varepsilon/8)^3/M(\varepsilon)^3$. Letting δ to be the $1/6$ this quantity, we are done. ■

Next, we describe Roth's Theorem. The result deals with a conjecture of Erdős and Turán:

> Given $\varepsilon > 0$ and $r \in \mathbb{N}$, there exists N_0 such that for all $N \geq N_0$, the following holds. If $A \subseteq [N]$ with $|A| \geq \varepsilon N$, then A contains an arithmetic progression of length r.

Roth proved in the early 50s that the conjecture holds for $r = 3$. Szemerédi proved in the early 70s that it holds for all r. In fact, Szemerédi had originally come up with the regularity lemma to prove this result (this is why it is a *lemma*).

Lemma 2.15. Suppose that every edge of G_n is in exactly one triangle. Then, $e(G_n) = o(n^2)$.

How is this related to Roth's Theorem? Given $N \gg 0$ and $A \subseteq [N]$, suppose A is 3-AP free. Let $M = 2N + 1$ and construct a 3-partite graph G whose three parts X, Y, and Z are copies of $\mathbb{Z}/M$.
For $x \in X$, $y \in Y$, and $z \in Z$, keep an edge from x to y iff $y - x \in A$, y to z iff $z - y \in A$, and z to x iff $(z - x)/2 \in A$.
The final part is well-defined since 2 is invertible in $\mathbb{Z}/M$ (Why?).
Observe that if xyz is a triangle with $a = y - x$, $b = z - y$, and $2c = z - x$, then $a + b = 2c$, so a,c,b are in AP.
Since a, b, and c are all in A, then by our assumption we must have $a = b = c$, $y - x = z - y = (z - x)/2$, and so x, y, z must be in AP in $\mathbb{Z}/M$.
This implies that any edge must be in a unique triangle.
By Lemma 2.15, $e(G) = o((3M)^2) = o(N^2)$. On the other hand, $e(G) = 3M|A|$. So, $3M|A| = o(N^2)$ and $|A| = o(N)$.

Proof of Lemma 2.15. The number of triangles in G_n is exactly $e(G_n)/3 = o(n^3)$. By the Triangle Removal Lemma, G_n can be made triangle-free by removing $o(n^2)$ edges. However, we must remove at least $e(G_n)/3$ edges to make the graph triangle-free, so $e(G_n) = o(n^2)$. ■

Inspired by this problem, a natural question to ask is: what is the maximum sized $A \subseteq [N]$ that is 3-AP free? Picking A greedily gives $O(\sqrt{n})$ elements.

It is possible to do better, as shown by *Behrend's construction*.
The basic idea is that on a sphere, the midpoint of any two points does not lie on the sphere.
Consider the set $S = [k]^d \subseteq \mathbb{R}^d$ for some k, d. Observe that $\|x\|^2 \leq k^2 d$ for any point in S. By the pigeonhole principle, there is a (centered) sphere with at least k^{d-2}/d of these points.
To get a subset A from these points, project $x = (a_1, \ldots, a_d)$ to $\sum_{j=1}^{d} a_j (2k+1)^{j-1}$, thus converting it to an integer in base $(2k+1)$. Observe that if $x_1 \mapsto m_1$ and $x_2 \mapsto m_2$, then $(x_1 + x_2)/2 \mapsto (m_1 + m_2)/2$. As a result, if X is a subset of S on a single sphere, its projection by this map gives a 3-AP free subset.
The maximum element by this projection is $k \sum_{i=0}^{d-1} (2k+1)^i = ((2k+1)^d - 1)/2$. Set $(2k+1)^d = 2N + 1$. By the argument from before, there is a 3-AP free set of size k^{d-2}/d.
This achieves a bound of

$$\frac{N}{e^{c\sqrt{\log n}}}$$

for some constant c. This is $\Omega(N^{1-\delta})$ for any (fixed) $\delta > 0$!

Definition 2.7 (Corner)**.** An $A \subseteq [N]^2$ is said to have a **corner** if the points (x, y), $(x + d, y)$, and $(x, y + d)$ are in A for some $d > 0$ and x,$y \in [N]$.

Theorem 2.16 (No Corners Theorem)**.** Suppose $A \subseteq [N]^2$ has no corners. Then $|A| = o(N^2)$.

Proof. First, let us get rid of the (slightly unnatural) $d > 0$ clause.
Denote $A + A$ as $\{a + b : a, b \in A\}$ (the Minkowski sum of the two sets) and $x + A = \{x + a : a \in A\}$.
Obviously, $A + A \subseteq [2N]^2$. By the pigeonhole principle, there exists a $z \in [2N]^2$ such that $z = a + b$ in at least $|A|^2/4N^2$ ways. Consider $A' = A \cap (z - A)$ for such a z. Then

$$|A'| \geq \frac{|A|^2}{4N^2}$$

by the definition of z. Further, A' is *symmetric* about $z/2$, that is, $A' = z - A'$ (Why?). As a result, there is a correspondence between corners with positive d and negative d in A'. Further, if A is corner-free, so is A'. By the cardinality bound on A', it suffices to show that $|A'| = o(N^2)$.
So, let us drop the $d > 0$ condition on corners and work with a corner-free A.

We construct a 3-partite graph $G = (V, E)$. Let the three parts of V be H, the set of horizontal lines in $[N]^2$, V, the set of vertical lines in $[N]^2$, and D, the set of line segments with slope -1 in $[2N]^2$.
Given ℓ, ℓ' in distinct parts of the vertex set, let ℓ and ℓ' be adjacent if $\ell \cap \ell' \in A$. Triangles in this graph correspond to either corners or three lines all passing through a single vertex in A. Due to our assumption on A, each edge is in a unique triangle. The result then follows from Lemma 2.15. ■

Let us go back to the triangle counting lemma for a moment. Let (V_1, V_2), (V_2, V_3), and (V_3, V_1) all be ε-regular pairs with densities d_{12},d_{23},$d_{13} > 2\varepsilon$. In a random graph on $V_1 \sqcup V_2 \sqcup V_3$ vertices with edge probabilities d_{ij} between V_i and V_j, the expected number of triangles is $|V_1||V_2||V_3|d_{12}d_{23}d_{31}$.
Ideally, we would have a result that the number of triangles is indeed close to this quantity (so the graph behaves almost randomly, in a sense similar to Szemerédi's Regularity Lemma).

Theorem 2.17 (Graph Counting Lemma)**.** Let G be a graph on n vertices and H a graph on $[k]$. Suppose $V_1, \dots, V_k \subseteq V(G)$ (with the V_i pairwise-disjoint) such that (V_i, V_j) is ε-regular whenever $ij \in E(H)$. Then,

$$\left|\{(v_1, \dots, v_k) : v_i \in V_i \text{ and } \{v_1, \dots, v_k\} \text{ form a copy of } H \text{ in } G\}\right|$$

is within $\varepsilon e(H)|V_1||V_2| \cdots |V_k|$ of the (expected) quantity

$$\prod_{i=1}^{k} |V_i| \prod_{ij \in E(H)} d(V_i, V_j),$$

assuming that $\prod_{ij \in E(H)} d(V_i, V_j) > \varepsilon e(H)$.

Proof. We shall prove this by inducting on $e(H)$. If $e(H) = 0$, the result is trivial.
Let us rephrase the problem probabilistically. Pick $v_i \in V_i$ independently and uniformly. Then, we wish to prove that

$$\left| \Pr\left[v_i v_j \in E(G) \text{ for all } ij \in E(H)\right] - \prod_{ij \in E(H)} d(V_i, V_j) \right| \leq \varepsilon e(H). \tag{2.8}$$

Suppose $\{1, 2\} \in E(H)$. It suffices to show that

$$\left| \Pr\left[v_i v_j \in E(G) \text{ for all } ij \in E(H)\right] - d(V_1, V_2) \Pr\left[v_i v_j \in E(G) \text{ for all } E(H) \setminus \{1, 2\}\right] \right| \leq \varepsilon. \tag{2.9}$$

Indeed, by induction,

$$\left| \Pr\left[v_i v_j \in E(G) \text{ for all } E(H) \setminus \{1, 2\}\right] - \prod_{\substack{ij \in H \\ ij \neq \{1,2\}}} d(V_i, V_j) \right| < \varepsilon(e(H) - 1), \tag{2.10}$$

and (2.8) follows on using the union bound. We shall prove that (2.9) holds when we condition on the choices v_i for $i > 2$. Let

$$A_j = \{v_j \in V_j : \{v_j, v_i\} \in E(G) \text{ when } \{j, i\}\} \in E(H).$$

for $j = 1, 2$. Equation (2.9) is then equivalent to

$$\left| \frac{e(A_1, A_2)}{|V_1||V_2|} - d(V_1, V_2) \frac{|A_1||A_2|}{|V_1||V_2|} \right| \leq \varepsilon. \tag{2.11}$$

We claim that (2.11) holds for all A_1,A_2. If $|A_1| \geq \varepsilon |V_1|$ and $|A_2| \geq \varepsilon |V_2|$, then ε-regularity implies that $|d(A_1, A_2) - d(V_1, V_2)| \leq \varepsilon$, that is,

$$\left| \frac{e(A_1, A_2)}{|A_1||A_2|} - d(V_1, V_2) \right| \leq \varepsilon.$$

So,

$$\frac{|V_1||V_2|}{|A_1||A_2|} \left| \frac{e(A_1, A_2)}{|V_1||V_2|} - d(V_1, V_2) \frac{|A_1||A_2|}{|V_1||V_2|} d(V_1, V_2) \right| \leq \varepsilon,$$

and (2.11) easily follows. If $|A_1| < \varepsilon |V_1|$, then the above follows immediately anyway, so we are done. ■

References

[Erd59] P. Erdös. Graph theory and probability. *Canadian Journal of Mathematics*, 11:34–38, 1959.

[Gow97] W. T. Gowers. Lower bounds of tower type for szemerédi's uniformity lemma. *Geometric & Functional Analysis GAFA*, 7(2):322–337, May 1997.

[KRS96] János Kollár, Lajos Rónyai, and Tibor Szabó. Norm-graphs and bipartite turán numbers. *Combinatorica*, 16(3):399–406, 1996.