1. Are arbitrarily long arithmetic progressions the same as infinite arithmetic progressions?
   (a) Construct a subset $S \subseteq \mathbb{N}$ with arbitrarily long arithmetic progressions, but no infinite arithmetic progressions.
   (b) For some $r \geq 2$, construct an r-coloring of $\mathbb{N}$ with no infinite monochromatic arithmetic progressions.

$\oplus$ 2. In class, I motivated the Erdős–Turán conjecture (and thus Szemerédi's theorem) from van der Waerden's theorem, plus the intuition that the "largest" set should be the one with the k-AP. In this problem, you'll see that this intuition is not always right. Along the way, you'll see another cool connection between graph theory and number theory.
   (a) For every integer $r \geq 1$, prove the following. If $N \geq 3r!$ and we color the edges of the complete graph K_N with r colors, then there is a monochromatic triangle. *Hint:* Induction on r.
   (b) Using part (a), prove the following. If $N \geq 3r!$ and we color $[N]$ with r colors, then there is a monochromatic solution to $x + y = z$ (i.e. there exist three numbers, not necessarily distinct, such that $x + y = z$ and x, y, z receive the same color).
   (c) For every $r \geq 2$ and every N, construct an r-coloring of $[N]$ so that the $r - 1$ *largest* color classes have no solution to $x + y = z$. In other words, in the result from part (b), it is possible that the monochromatic solution is in the *smallest* color.
   $\star$ (d) Can you improve the bound in part (a) from $3r!$ to some smaller quantity?
   $\oplus\star$ (e) Using the result in part (b), prove that the modular form of Fermat's last theorem is false. Namely, prove that if $k \geq 2$ is fixed and p is a sufficiently large prime, then there *do* exist non-zero $a, b, c \in \mathbb{Z}/p\mathbb{Z}$ with

$$a^k + b^k \equiv c^k \pmod{p}.$$

$\star$ 3. We stated two equivalent forms of Roth's theorem. One says that if $T \subseteq \mathbb{N}$ has positive density, then T contains a three-term arithmetic progression (henceforth, 3-AP). The other says that if $\varepsilon > 0$ is fixed and N_0 is large enough, then any $S \subseteq [N]$ with $N \geq N_0$ and $|S| \geq \varepsilon N$ contains a three-term arithmetic progression.
   (a) Prove that these two statements are equivalent.
   (b) State two such versions of Szemerédi's theorem, and prove that they are equivalent.
   $\star$ (c) State two such versions of van der Waerden's theorem, and prove that they are equivalent.

4. The following result is called the *diamond-free lemma.* For every $\varepsilon > 0$ there exists some $n_0 \in \mathbb{N}$ such that the following holds for every $n \geq n_0$ and every n-vertex graph G. If every edge of G lies in exactly one triangle, then G has at most $\varepsilon\binom{n}{2}$ edges.

$\star$ means that this problem is harder than the others. Also, stars are additive: two extra stars in a part of a starred multi-part problem correspond to three normal stars.

$\oplus$ means that this problem is not directly related to the content of the class, and is for general breadth and edification.

(a) Prove the diamond-free lemma, using the triangle removal lemma.

(b) Prove Roth's theorem using the diamond-free lemma. In other words, we don't need the full force of the triangle removal lemma to prove Roth's theorem.

Tomorrow, you'll see several other equivalent formulations of the diamond-free lemma.

5. In this problem you will find a construction of a large 3-AP-free subset of $[N]$, originally due to Behrend. I sketched this construction in my colloquium in Week 1.

   (a) Let m, d be positive integers to be chosen later, and let $\Gamma_m^d := \{0, \ldots, m-1\}^d$ be the $m \times m \times \cdots \times m$ grid in $\mathbb{R}^d$. Say that a subset $X \subseteq \Gamma_m^d$ is 3-AP-free if there do not exist $x, y, z \in X$ with $x + z = 2y$. Consider the map $\varphi : \Gamma_m^d \to \{0, \ldots, (2m)^d - 1\}$ given by
   $$\varphi((x_1, \ldots, x_d)) = \sum_{i=1}^{d} x_i (2m)^{i-1}.$$
   In other words, φ converts a vector to an integer by treating the vector as a base-$(2m)$ representation of an integer.
   Prove that if $X \subseteq \Gamma_m^d$ is 3-AP-free, then so is $\varphi(X) \subseteq \{0, \ldots, (2m)^d - 1\}$.

   (b) Prove that there is a $(d-1)$-dimensional sphere $S \subseteq \mathbb{R}^d$ centered at the origin with
   $$|S \cap \Gamma_m^d| \geq \frac{m^d}{dm^2}.$$
   Hint: Pigeonhole principle.

   (c) Let $X = S \cap \Gamma_m^d$. Prove that X is 3-AP-free, and conclude that so is $1 + \varphi(X) \subseteq [(2m)^d]$.

   $\star$ (d) Let $N = (2m)^d$. Pick m and d to make $|\varphi(X)|$ as large as you can.
   Hint: You should be able to get $|\varphi(X)| \geq N/2^{C\sqrt{\log N}}$ for some constant $C > 0$.

$\star$ 6. In this problem you will construct a "bad graph" for the triangle removal lemma, namely you will show that δ cannot be taken too large as a function of ε in the triangle removal lemma.

   (a) In class, we proved Roth's theorem from the triangle removal lemma. Using the same construction, as well as the previous problem, show that for every N, there is an N-vertex graph G_0 with at least $N^2/2^{C\sqrt{\log N}}$ triangles, and at least $N^2/2^{C\sqrt{\log N}}$ edges must be deleted to make G_0 triangle-free, where $C > 0$ is an absolute constant.

   (b) Given a graph G and an integer s, let $G[s]$ denote the *s-blowup* of G, which obtained from G by replacing every vertex by an independent set of size s, and replacing each edge of G by a copy of $K_{s,s}$. Prove that if G has t triangles, then $G[s]$ has ts^3 triangles.

   $\star$ (c) Prove that if at least m edges must be removed to make G triangle-free, then at least ms^2 edges must be removed to make $G[s]$ triangle-free.

   $\star$ (d) Using the previous three parts, show that the following holds for every $\varepsilon > 0$ and every sufficiently large N. There exists an N-vertex G with at most $\delta\binom{N}{3}$ triangles, but at least $\varepsilon\binom{N}{2}$ edges must be removed to make G triangle-free, where $\delta \leq \varepsilon^{C\log(1/\varepsilon)}$.

1. Let $\varepsilon > 0$ and $0 < \alpha \le \frac{1}{2}$. Suppose that a pair of vertex sets (A, B) is ε-regular, and let $X \subseteq A, Y \subseteq B$ satisfy $|X| \ge \alpha|A|, |Y| \ge \alpha|B|$. Prove that (X, Y) is (ε/α)-regular.

   Why is the assumption $\alpha \le \frac{1}{2}$ necessary?

2. On yesterday's homework, you proved the *diamond-free lemma*: For every $\varepsilon > 0$ there exists some $n_0 \in \mathbb{N}$ such that the following holds for every $n \ge n_0$ and every n-vertex graph G. If every edge of G lies in exactly one triangle, then G has at most $\varepsilon\binom{n}{2}$ edges.

   $\star$ (a) Prove that the diamond-free lemma is equivalent to the following statement, called the *induced matching theorem*.

   For every $\varepsilon > 0$ there exists some $n_0 \in \mathbb{N}$ such that the following holds for every $n \ge n_0$ and every n-vertex graph G. If the edges of G can be decomposed into n induced matchings, then G has at most $\varepsilon\binom{n}{2}$ edges.

   (A *matching* is a collection of edges which have no vertices in common. It is *induced* if there are no other edges of G going between its vertices.)

   (b) Prove that the diamond-free lemma is equivalent to the following statement, called the *(6,3) theorem*. This was the original problem Ruzsa and Szemerédi set out to solve when they formulated the triangle removal lemma. Recall that a *3-uniform hypergraph* consists of a set of vertices, and some triples of vertices called *hyperedges*.

   For every $\varepsilon > 0$ there exists some $n_0 \in \mathbb{N}$ such that the following holds for every $n \ge n_0$ and every n-vertex 3-uniform hypergraph $\mathcal{G}$. If that every 6-tuple of vertices of $\mathcal{G}$ contains at most 2 edges, then $\mathcal{G}$ has at most εn^2 edges.

   ? (c) Prove the *(7,4) conjecture*: If every 7-tuple of vertices of $\mathcal{G}$ contains at most 3 edges, then $\mathcal{G}$ has at most εn^2 edges.

   In fact, the $(k + 3, k)$ conjecture is open for all $k \ge 4$: If every $(k + 3)$-tuple of vertices of $\mathcal{G}$ contains fewer than k edges, then $\mathcal{G}$ has at most εn^2 edges.

3. Let $S \subseteq [N]^2$. A *corner* in S is three points of the form $(x, y), (x + d, y), (x, y + d)$ for some $d \ne 0$.

   $\star$ (a) Prove the following result, called the *corners theorem*.

   For every $\varepsilon > 0$ there exists some $N_0 \in \mathbb{N}$ such that for every $N \ge N_0$ and every $S \subseteq [N]^2$ with $|S| \ge \varepsilon N^2$ contains a corner.

   (b) Deduce Roth's theorem from the corners theorem.

4. A pair (A, B) of vertex sets is called *ε-homogeneous* if for all $A' \subseteq A, B' \subseteq B$, we have that

$$|e(A', B') - d(A, B)|A'||B'|| < \varepsilon|A||B|.$$

   Prove that if (A, B) is ε-regular, then it is ε-homogeneous. Conversely, prove that if (A, B) is ε^3-homogeneous, then it is ε-regular.

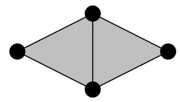
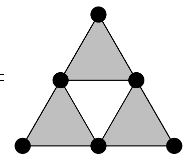
$\star$ means that this problem is harder than the others. Also, stars are additive: two extra stars in a part of a starred multi-part problem correspond to three normal stars.

✢ means that this problem is not directly related to the content of the class, and is for general breadth and edification.

? means that this is an open problem.

✥$\star$ 5. This problem is stolen verbatim from my Extremal graph theory class; you should probably skip it if you weren't in that class.

$\star$ (a) Consider the following two 3-partite 3-graphs:

$K^{(3)}_{1,1,2} =$  $T =$ 

Prove that $\mathrm{ex}(n, K^{(3)}_{1,1,2}) = \Theta(n^2)$ and $\mathrm{ex}(n, T) = \Theta(n^2)$.

(b) Prove that $\mathrm{ex}(n, \{K^{(3)}_{1,1,2}, T\}) = o(n^2)$. This shows that the *compactness conjecture* fails for hypergraphs.
Hint: Use the (6,3) theorem from Problem 2(b).

✥$\star$ 6. In this problem you will construct an ε-regular pair in a graph without using randomness at all. This problem requires some knowledge of how quadratic residues work mod p.

(a) Fix an odd prime p. Prove that for any $T \subseteq \mathbb{Z}/p\mathbb{Z}$, we have that

$$\sum_{z \in \mathbb{Z}/p\mathbb{Z}} \left| \sum_{t \in T} e^{2\pi i t z/p} \right|^2 = p|T|.$$

(b) Let $\chi : \mathbb{Z}/p\mathbb{Z} \to \{-1, 0, 1\}$ be the *quadratic character* mod p, namely the function

$$\chi(x) = \begin{cases} 1 & \text{if } x \text{ is a quadratic residue mod } p, \\ -1 & \text{if } x \text{ is a quadratic non-residue mod } p, \\ 0 & \text{if } x = 0. \end{cases}$$

Prove the *Gauss sum formula*,

$$\left| \sum_{z \in \mathbb{Z}/p\mathbb{Z}} \chi(z) e^{2\pi i z/p} \right| = \sqrt{p}.$$

$\star$ (c) Prove that for all $X, Y \subseteq \mathbb{Z}/p\mathbb{Z}$, we have that

$$\left| \sum_{x \in X} \sum_{y \in Y} \chi(x - y) \right| \leq \sqrt{p|X||Y|}.$$

Hint: Use parts (a) and (b), as well as Cauchy–Schwarz.

(d) Define the following graph, called the *(bipartite) Paley graph*. Its vertex set is $A \sqcup B$, where $A = B = \mathbb{Z}/p\mathbb{Z}$. For vertices $a \in A, b \in B$, we join them by an edge if and only if $b - a$ is a quadratic residue mod p.

Fix some $\varepsilon > 0$. Prove that if p is sufficiently large with respect to ε, then the pair (A, B) is ε-regular.

1. Read and understand the statements of the general counting lemma and removal lemma from Section 4 of today's notes.

   (a) Prove the K_4 counting lemma.

   Hint: The hereditary property of regularity, from problem 1 in yesterday's homework, may be useful.

   ⋆ (b) Prove that general counting lemma, or convince yourself that the same idea will work in general.

   (c) Prove the general removal lemma.

2. Prove that for every $0 < \varepsilon_1 < \varepsilon_2$, there exists some $\delta > 0$ so that the following holds. Suppose (A, B) is ε_1-regular, and suppose we "perturb" the graph by adding or subtracting at most $\delta|A|$ vertices from A, adding or subtracting at most $\delta|B|$ vertices from B, and adding or subtracting at most $\delta|A||B|$ edges, then the resulting pair is ε_2-regular.

3. In the statement of Szemerédi's regularity lemma (and thus in the definition of an ε-regular partition), we had to allow a small fraction of the pairs of parts to be not ε-regular. For many years after Szemerédi proved the regularity lemma, people were unsure if such a condition was necessary, or if we could ensure that *all* pairs of parts are ε-regular. It turns out this condition is necessary.

   (a) The *half graph* with parameter k, denoted H_k, is the bipartite graphs with parts $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_k\}$, where a_i is adjacent to b_j if and only if $i < j$. Suppose that $m \mid k$, and we partition the vertices of H_k into $A_1 \sqcup \cdots \sqcup A_m \sqcup B_1 \sqcup \cdots \sqcup B_m$, where

   $$A_1 = \{a_1, \ldots, a_{k/m}\}, A_2 = \{a_{k/m+1}, \ldots, a_{2k/m}\}, \ldots, A_m = \{a_{(m-1)k/m+1}, \ldots, a_k\}$$

   and similarly for $B_1, \ldots, B_m$. This is a partition of $V(H_k)$ into $2m$ equally-sized parts; prove that at least m of the $(2m)^2$ pairs of parts are not ε-regular, assuming k is sufficiently large with respect to ε.

   ⋆ (b) Prove that for every fixed $\varepsilon > 0$ and every fixed $m \in \mathbb{N}$, if k is large enough, then for every partition $V(H_k) = V_1 \sqcup \cdots \sqcup V_m$,

   $$\sum_{\substack{(i,j)\in[m]^2 \\ (V_i,V_j) \text{ not } \varepsilon\text{-regular}}} |V_i||V_j| \geq ck,$$

   where $c > 0$ is an absolute constant. In other words, it is impossible to get rid of the assumption that some pairs of parts are irregular.

⋆ 4. In this problem, you'll prove the Erdős–Stone theorem, which was the main result in my Extremal graph theory class (but this problem will be interesting whether or not you were in that class).

   (a) Turán's theorem implies the following. If $\varepsilon > 0$ and $r \in \mathbb{N}$ are fixed and n is sufficiently large, then any n-vertex graph G with

   $$e(G) \geq \left(1 - \frac{1}{r-1} + \varepsilon\right)\binom{n}{2}$$

edges has a copy of K_r.

If you haven't seen Turán's theorem before, prove this! Or accept that it is true.

(b) Let $K_r[s]$ denote the s-blowup of K_r, also known as the complete r-partite graph with parts of size s. Prove that if $\varepsilon > 0$ and $r, s \in \mathbb{N}$ are fixed, n is large enough, and G is an n-vertex graph with

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

then G contains a copy of $K_r[s]$.

Hint: Apply the regularity lemma, delete edges as in the proof of the removal lemma, apply Turán's theorem, then apply the $K_r[s]$ counting lemma.

5. Let G be an n-vertex graph, and let Γ be an m-vertex graph. For $\varepsilon > 0$, an *ε-approximate homomorphism* $G \to \Gamma$ is a function $\varphi : V(G) \to V(\Gamma)$ that sends all but at most $\varepsilon\binom{n}{2}$ edges of G to edges of Γ. More formally, it has the property that

$$|\{(u, v) \in E(G) : (\varphi(u), \varphi(v)) \notin E(\Gamma)\}| \leq \varepsilon\binom{n}{2}.$$

(a) Prove the *strong triangle removal lemma*: For every $\varepsilon > 0$, there exists some $\delta > 0$ and some $m \in \mathbb{N}$ so that the following holds for all $n \in \mathbb{N}$. If G is an n-vertex graph with at most $\delta\binom{n}{3}$ triangles, then there exists a triangle-free graph Γ on m vertices such that G has an ε-approximate homomorphism to Γ.

(b) Conclude from part (a) the *triangle-free lemma*: For every $\varepsilon > 0$, there exists some $m \in \mathbb{N}$ so that the following holds for all $n \in \mathbb{N}$. If G is an n-vertex triangle-free graph, then there exists some triangle-free graph Γ on m vertices such that G has an ε-approximate homomorphism to Γ.

$\star\star$ (c) Fix $m \in \mathbb{N}$ and $\alpha, \varepsilon \in (0, 1)$. Prove that for sufficiently large n, there exists some n-vertex G with $\alpha\binom{n}{2}$ edges so that G has no ε-approximate homomorphism to *any* m-vertex graph Γ.

[In general, ε-approximate homomorphisms are pretty rare, so the triangle-free lemma is pretty amazing: the class of triangle-free graphs is rich in ε-approximate homomorphisms.]

$\star$ 6. In this problem, you'll count how many triangle-free graphs there are. Recall Mantel's theorem, which says that every n-vertex triangle-free graph has at most $n^2/4$ edges.

(a) Fix $\varepsilon > 0$. Prove that if n is sufficiently large, then there are at least $2^{(\frac{1}{4}-\varepsilon)n^2}$ labeled n-vertex triangle-free graphs.

$\star\star$ (b) Prove a matching upper bound: if n is sufficiently large, then there are at most $2^{(\frac{1}{4}+\varepsilon)n^2}$ labeled n-vertex triangle-free graphs.