2

Graph Regularity Method

Chapter Highlights

- Szemerédi's graph regularity lemma: partitioning an arbitrary graph into a bounded number of parts with random-like edges between parts
- Graph regularity method: recipe and applications
- Graph removal lemma
- Roth's theorem: a graph theoretic proof using the triangle removal lemma
- Strong regularity and induced graph removal lemma
- Graph property testing
- Hypergraph removal lemma and Szemerédi's theorem

In this chapter, we discuss a powerful technique in extremal graph theory developed in the 1970's, known as Szemerédi's graph regularity lemma. The graph regularity method has wide ranging applications, and is now considered a central technique in the field. The regularity lemma produces a "rough structural" decomposition of an arbitrary graph (though it is mainly useful for graphs with quadratically many edges). It then allows us to model an arbitrary graph by a random graph.

The regularity method introduces us to a central theme of the book: **the dichotomy of structure and pseudorandomness**. This dichotomy is analogous to the more familiar concept of "signal and noise," namely that a complex system can be decomposed into a structural piece with plenty of information content (the signal) as well as a random-like residue (the noise). This idea will show up again later in Chapter 6 when we discuss Fourier analysis in additive combinatorics.

In general, we face two related challenges:

- How to decompose an object into a structured piece and a random-like piece?
- How to analyze the resulting components and their interactions?

We begin the chapter with the statement and the proof of the graph regularity lemma. We then prove Roth's theorem using the regularity method. This proof, due to Ruzsa and Szemerédi (1978), is not the original proof by Roth (1953), whose original Fourier analytic proof we will see in Chapter 6. Nevertheless, it is important for being historically one of the first major applications of the graph regularity method. Similar to the proof of Schur's theorem in Chapter 0, this graph theoretic proof of Roth's theorem demonstrates a fruitful connection between graph theory and additive combinatorics.

By **the regularity method**, we mean both the graph regularity lemma as well as methods for applying it. Rather than some specific set of theorems, graph regularity should be viewed as a general technique malleable to adaptations. Do not get bogged down by specific choices of parameters in the statements and proofs below, and rather, focus on the main ideas and techniques.

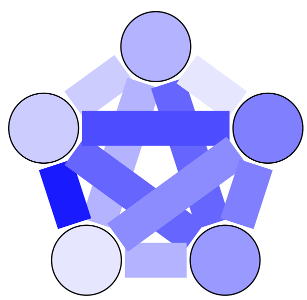
Many students experience a steep learning curve when studying the regularity method. The technical details can obscure the underlying intuition. Also, the style of arguments may be quite different from the type of combinatorial proofs they encountered earlier in their studies (e.g., the type of proofs from earlier in this book). Section 2.7 contains important exercises on applying the graph regularity method, which are essential for understanding the material.

2.1 Szemerédi's Graph Regularity Lemma

In this section, we state and prove the graph regularity lemma. Let us first give an informal statement.

Graph Regularity Lemma (Informal). The vertex set of every graph can be partitioned into a bounded number of parts so that the graph looks random-like between most pairs of parts.

Following is an illustration of what the outcome of the partition looks like. Here the vertex set of a graph is partitioned into five parts. Between a pair of parts (including between a part and itself) is a random-like graph with a certain edge-density (e.g., 0.4 between the first and second parts, 0.7 between the first and third parts, . . .).



Definition 2.1.1 (Edge density)

Let X and Y be sets of vertices in a graph G. Let $e_G(X,Y)$ be the number of edges between X and Y; that is,

$$\boldsymbol{e_G(X,Y)} := |\{(x,y) \in X \times Y : xy \in E(G)\}| .$$

Define the ***edge density*** between X and Y in G by

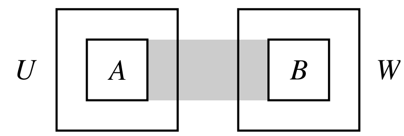
$$\boldsymbol{d_G(X,Y)} := \frac{e_G(X,Y)}{|X|\,|Y|}.$$

We drop the subscript G if the context is clear.

We allow X and Y to overlap in the preceding definition. For intuition, it is mostly fine to picture the bipartite setting, where X and Y are automatically disjoint.

What should it mean for a graph to be "random-like"? We will explore the concept of

pseudorandom graphs in depth in Chapter 3. Given vertex sets X and Y, we would like the edge density between them to not change much even if we restrict X and Y to smaller subsets. Intuitively, this says that the edges are somewhat evenly distributed.



Definition 2.1.2 (ε-regular pair)
Let G be a graph and $U, W \subseteq V(G)$. We call (U, W) an ***ε-regular pair*** in G if for all $A \subseteq U$ and $B \subseteq W$ with $|A| \geq \varepsilon |U|$ and $|B| \geq \varepsilon |W|$, one has

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

If (U, W) is not ε-regular, then we say that their irregularity is ***witnessed*** by some $A \subseteq U$ and $B \subseteq W$ satisfying $|A| \geq \varepsilon |U|$, $|B| \geq \varepsilon |W|$, and $|d(A, B) - d(U, W)| > \varepsilon$.

We need the hypotheses $|A| \geq \varepsilon |U|$ and $|B| \geq \varepsilon |W|$ since the definition would be too restrictive otherwise. For example, by taking $A = \{x\}$ and $B = \{y\}$, $d(A, B)$ could end up being both 0 (if $xy \notin E$) and 1 (if $xy \in E$).

***Remark* 2.1.3 (Different roles of ε).** The ε in $|A| \geq \varepsilon |U|$ and $|B| \geq \varepsilon |W|$ plays a different role from the ε in $|d(A, B) - d(U, W)| \leq \varepsilon$. However, it is usually not important to distinguish these εs. So we use only one ε for convenience of notation.

The "random-like" intuition is justified as random graphs indeed satisfy the above property. (This can be proved by the Chernoff bound; more on this in the next chapter.)

The following exercises can help you check your understanding of ε-regularity.

Exercise 2.1.4 (Basic inheritance of regularity). Let G be a graph and $X, Y \subseteq V(G)$. If (X, Y) is an $\varepsilon\eta$-regular pair, then (X', Y') is 2ε-regular for all $X' \subseteq X$ with $|X'| \geq \eta |X|$ and $Y' \subseteq Y$ with $|Y'| \geq \eta |Y|$.

Exercise 2.1.5 (An alternate definition of regular pairs). Let G be a graph and $X, Y \subseteq V(G)$. Say that (X, Y) is ***ε-homogeneous*** if for all $A \subseteq X$ and $B \subseteq Y$, one has

$$|e(A, B) - |A| |B| d(X, Y)| \leq \varepsilon |X| |Y| .$$

Show that if (X, Y) is ε-regular, then it is ε-homogeneous. Also, show that if (X, Y) is ε^3-homogeneous, then it is ε-regular.

Exercise 2.1.6 (Robustness of regularity). Prove that for every $\varepsilon' > \varepsilon > 0$, there exists $\delta > 0$ so that given an ε-regular pair (X, Y) in some graph, if we modify the graph by adding/deleting $\leq \delta |X|$ vertices to/from X, adding/deleting $\leq \delta |Y|$ vertices to/from Y, and adding/deleting $\leq \delta |X| |Y|$ edges, then the resulting new (X, Y) is still ε'-regular.

Next, let us define what it means for a vertex partition to be ε-regular.

Definition 2.1.7 (ε-regular partition)
Given a graph G, a partition $\mathcal{P} = \{V_1, \dots, V_k\}$ of its vertex set is an ***ε-regular partition*** if

$$\sum_{\substack{(i,j)\in[k]^2 \\ (V_i,V_j) \text{ not } \varepsilon\text{-regular}}} |V_i||V_j| \le \varepsilon |V(G)|^2.$$

In other words, all but at most an ε-fraction of pairs of vertices of G lie between ε-regular parts.

Remark **2.1.8.** When $|V_1| = \cdots = |V_k|$, the inequality says that at most εk^2 of pairs (V_i, V_j) are not ε-regular.

Also, note that the summation includes $i = j$. If none of the V_is are too large, say $|V_i| \le \varepsilon n$ for each i, then the terms with $i = j$ contribute $\le \sum_i |V_i|^2 \le \varepsilon n \sum_i |V_i| = \varepsilon n^2$, which is neglible.

We are now ready to state Szemerédi's graph regularity lemma.

Theorem 2.1.9 (Szemerédi's graph regularity lemma)
For every $\varepsilon > 0$, there exists a constant M such that every graph has an ε-regular partition into at most M parts.

Proof of the Graph Regularity Lemma

Proof idea. We will generate the desired vertex partition according to the following algorithm:

(1) Start with the trivial partition of $V(G)$. (The trivial partition has a single part consisting of the whole set.)
(2) While the current partition $\mathcal{P}$ is not ε-regular:
   (a) For each (V_i, V_j) that is not ε-regular, find a witnessing pair in V_i and V_j
   (b) Refine $\mathcal{P}$ using all the witnessing pairs. (Here given two partitions $\mathcal{P}$ and $\mathcal{Q}$ of the same set, we say that $\mathcal{Q}$ ***refines*** $\mathcal{P}$ if each part of $\mathcal{Q}$ is contained in a part of $\mathcal{P}$. In other words, we divide each part of $\mathcal{P}$ further to obtain $\mathcal{Q}$.)

We repeat step (2) until the partition is ε-regular, at which point the algorithm terminates. The resulting partition is always ε-regular by design. It remains to show that the number of iterations is bounded as a function of ε. To see this, we keep track of a quantity that necessarily increases at each iteration of the procedure. This is called an **energy increment argument**. (The reason that we call it an "energy" is because it is the L^2 norm of a vector of edge-densities, and the kinetic energy in physics is also an L^2 norm.) ∎

Definition 2.1.10 (Energy)
Let G be an n-vertex graph (whose dependence we drop from the notation). Let $U, W \subseteq V(G)$. Define

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

For partitions $\mathcal{P}_U = \{U_1, \dots, U_k\}$ of U and $\mathcal{P}_W = \{W_1, \dots, W_l\}$ of W, define

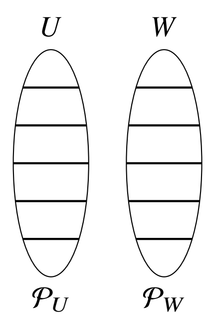
$$\boldsymbol{q(\mathcal{P}_U, \mathcal{P}_W)} := \sum_{i=1}^{k} \sum_{j=1}^{l} q(U_i, W_j).$$

Finally, for a partition $\mathcal{P} = \{V_1, \dots, V_k\}$ of $V(G)$, define its ***energy*** to be

$$\boldsymbol{q(\mathcal{P})} := q(\mathcal{P}, \mathcal{P}) = \sum_{i=1}^{k} \sum_{j=1}^{k} q(V_i, V_j) = \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{|V_i|\,|V_j|}{n^2} d(V_i, V_j)^2.$$

Since the edge density is always between 0 and 1, we have $0 \le q(\mathcal{P}) \le 1$ for all partitions $\mathcal{P}$. The following lemmas show that the energy cannot decrease upon refinement, and furthermore, it must increase substantially at each step of the preceding algorithm.

Lemma 2.1.11 (Energy never decreases under refinement)
Let G be a graph, $U, W \subseteq V(G)$, $\mathcal{P}_U$ a partition of U, and $\mathcal{P}_W$ a partition of W. Then $q(\mathcal{P}_U, \mathcal{P}_W) \ge q(U, W)$.



Proof. Let $n = v(G)$. Let $\mathcal{P}_U = \{U_1, \dots, U_k\}$ and $\mathcal{P}_W = \{W_1, \dots, W_l\}$. Choose $x \in U$ and $y \in W$ uniformly and independently at random. Let U_i be the part of $\mathcal{P}_U$ that contains x and W_j be the part of $\mathcal{P}_W$ that contains y. Define the random variable $Z := d(U_i, W_j)$. We have

$$\mathbb{E}[Z] = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{|U_i|}{|U|} \frac{|W_j|}{|W|} d(U_i, W_j) = d(U, W) = \sqrt{\frac{n^2}{|U||W|} q(U, W)}.$$

We have

$$\mathbb{E}[Z^2] = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{|U_i|}{|U|} \frac{|W_j|}{|W|} d(U_i, W_j)^2 = \frac{n^2}{|U||W|} q(\mathcal{P}_U, \mathcal{P}_W).$$

By convexity, $\mathbb{E}[Z^2] \ge \mathbb{E}[Z]^2$, which implies $q(\mathcal{P}_U, \mathcal{P}_W) \ge q(U, W)$. $\square$

Lemma 2.1.12 (Energy never decreases under refinement)
Given two vertex partitions $\mathcal{P}$ and $\mathcal{P}'$ of some graph, if $\mathcal{P}'$ refines $\mathcal{P}$, then $q(\mathcal{P}) \le q(\mathcal{P}')$.

Proof. The conclusion follows by applying Lemma 2.1.11 to each pair of parts of $\mathcal{P}$. In more detail, letting $\mathcal{P} = \{V_1, \dots, V_m\}$, and supposing $\mathcal{P}'$ refines each V_i into a partition $\mathcal{P}'_{V_i} = \{V'_{i1}, \dots, V'_{ik_i}\}$ of V_i, so that $\mathcal{P}' = \mathcal{P}'_{V_1} \cup \cdots \cup \mathcal{P}'_{V_m}$, we have

$$q(\mathcal{P}) = \sum_{i,j} q(V_i, V_j) \le \sum_{i,j} q(\mathcal{P}'_{V_i}, \mathcal{P}'_{V_j}) = q(\mathcal{P}'). \qquad \square$$

Lemma 2.1.13 (Energy boost for an irregular pair)
Let G be an n-vertex graph. If (U, W) is not ε-regular, as witnessed by $A \subseteq U$ and $B \subseteq W$, then

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

This is the "red bull lemma," giving an energy boost when feeling irregular.

Proof. Define Z as in the proof of Lemma 2.1.11 for $\mathcal{P}_U = \{A, U \setminus A\}$ and $\mathcal{P}_W = \{B, W \setminus B\}$. Then

$$\operatorname{Var}(Z) = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2 = \frac{n^2}{|U|\,|W|} \left(q(\mathcal{P}_U, \mathcal{P}_W) - q(U, W) \right).$$

We have $Z = d(A, B)$ with probability $\ge |A|\,|B| / (|U|\,|W|)$ (corresponding to the event $x \in A$ and $y \in B$). So

$$\begin{aligned} \operatorname{Var}(Z) &= \mathbb{E}[(Z - \mathbb{E}[Z])^2] \\ &\ge \frac{|A|}{|U|} \frac{|B|}{|W|} (d(A, B) - d(U, W))^2 \\ &> \varepsilon \cdot \varepsilon \cdot \varepsilon^2. \end{aligned}$$

Putting the two inequalities together gives the claim. $\square$

The next lemma, corresponding to step (2)(b) of the preceding algorithm, shows that we can put all the witnessing pairs together to obtain an energy increment.

Lemma 2.1.14 (Energy boost for an irregular partition)
If a partition $\mathcal{P} = \{V_1, \dots, V_k\}$ of $V(G)$ is not ε-regular, then there exists a refinement $\mathcal{Q}$ of $\mathcal{P}$ where every V_i is partitioned into at most 2^{k+1} parts, and such that

$$q(\mathcal{Q}) > q(\mathcal{P}) + \varepsilon^5.$$

Proof. Let

$$R = \{(i, j) \in [k]^2 : (V_i, V_j) \text{ is } \varepsilon\text{-regular}\} \qquad \text{and} \qquad \overline{R} = [k]^2 \setminus R.$$

For each pair (V_i, V_j) that is not ε-regular, find a pair $A^{i,j} \subseteq V_i$ and $B^{i,j} \subseteq V_j$ that witnesses the irregularity. Do this simultaneously for all $(i, j) \in \overline{R}$. Note that for $i \ne j$, we can take $A^{i,j} = B^{j,i}$ due to symmetry. When $i = j$, we should allow for the possibility of $A^{i,i}$ and $B^{i,i}$ to be distinct.

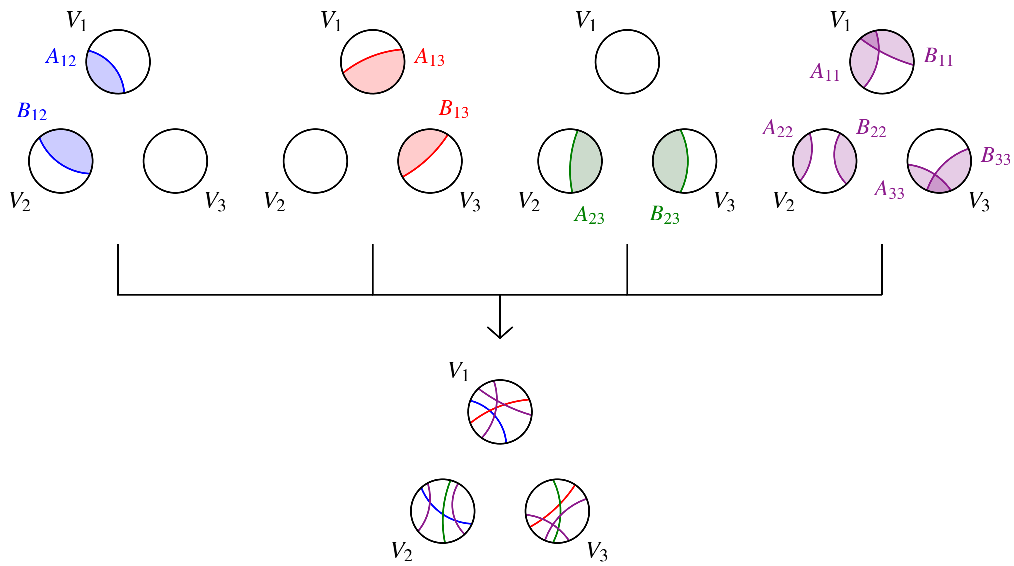


Figure 2.1 In the proof of Lemma 2.1.14, we refine the partition by taking a common refinement using witnesses of irregular pairs.

Let $\mathcal{Q}$ be a common refinement of $\mathcal{P}$ by all the $A^{i,j}$ and $B^{i,j}$ (i.e., the parts of $\mathcal{Q}$ are maximal subsets that are not "cut up" into small pieces by any element of $\mathcal{P}$ or by the $A^{i,j}$ and $B^{i,j}$; intuitively, imagine regions of a Venn diagram). See Figure 2.1 for an illustration. There are $\le k+1$ such distinct nonempty sets inside each V_i. So $\mathcal{Q}$ refines each V_i into at most 2^{k+1} parts. Let $\mathcal{Q}_i$ be the partition of V_i given by $\mathcal{Q}$. Then, using the monotonicity of energy under refinements (Lemma 2.1.11),

$$\begin{aligned} q(\mathcal{Q}) &= \sum_{(i,j)\in[k]^2} q(\mathcal{Q}_i,\mathcal{Q}_j) \\ &= \sum_{(i,j)\in R} q(\mathcal{Q}_i,\mathcal{Q}_j) + \sum_{(i,j)\in \overline{R}} q(\mathcal{Q}_i,\mathcal{Q}_j) \\ &\ge \sum_{(i,j)\in R} q(V_i,V_j) + \sum_{(i,j)\in \overline{R}} q(\{A^{i,j}, V_i\setminus A^{i,j}\},\{B^{i,j}, V_j\setminus B^{i,j}\}). \end{aligned}$$

By Lemma 2.1.13, the energy boost lemma, the preceding sum is

$$> \sum_{(i,j)\in[k]^2} q(V_i,V_j) + \sum_{(i,j)\in\overline{R}} \varepsilon^4 \frac{|V_i|\,|V_j|}{n^2}.$$

The first sum equals $q(\mathcal{P})$, and the second sum is $> \varepsilon^5$ by Lemma 2.1.13 since $\mathcal{P}$ is not ε-regular. This gives the desired inequality. □

***Remark* 2.1.15 (Refinements should be done simultaneously).** Here is a subtle point in the preceding proof. The refinement $\mathcal{Q}$ must be obtained in a single step by refining $\mathcal{P}$ using all the witnessing sets $A^{i,j}$ *simultaneously*. If instead we pick out a pair $A^{i,j} \subseteq V_i$ and

$A^{j,i} \subseteq V_j$, refine the partition using just this pair, and then iterate using another irregular pair $(V_{i'}, V_{j'})$, the energy boost step would not work. This is because ε-regularity (or lack thereof) is not well-preserved under taking refinements.

Proof of the graph regularity lemma (Theorem 2.1.9). Start with a trivial partition of the vertex set of the graph. Repeatedly apply Lemma 2.1.14 whenever the current partition is not ε-regular. By Lemma 2.1.14, the energy of the partition increases by more than ε^5 at each iteration. Since the energy of the partition is ≤ 1, we must stop after $< \varepsilon^{-5}$ iterations, terminating in an ε-regular partition.

If a partition has k parts, then Lemma 2.1.14 produces a refinement with $\leq k2^{k+1}$ parts. We start with a trivial partition with one part, and then refine $< \varepsilon^{-5}$ times. Observe the crude bound $k2^{k+1} \leq 2^{2^k}$. So the total number of parts at the end is $\leq \operatorname{tower}(\lceil 2\varepsilon^{-5} \rceil)$, where

$$\textbf{tower}(k) := 2^{2^{\cdot^{\cdot^{\cdot^{2}}}}} \Big\}\text{height } k. \qquad \square$$

***Remark* 2.1.16** (The proof does not guarantee that the partition becomes "more regular" after each step.). Let us stress what the proof is *not* saying. It is *not* saying that the partition gets more and more regular under each refinement. Also, it is *not* saying that partition gets more regular as the energy gets higher. Rather, the energy simply bounds the number of iterations.

The bound on the number of parts guaranteed by the proof is a constant for each fixed $\varepsilon > 0$, but it grows extremely quickly as ε gets smaller. Is the poor quantitative dependence somehow due to a suboptimal proof strategy? Surprisingly, the tower-type bound is necessary, as shown by Gowers (1997).

Theorem 2.1.17 (Lower bound on the number of parts in a regularity partition)
There exists a constant $c > 0$ such that for all sufficiently small $\varepsilon > 0$, there exists a graph with no ε-regular partition into fewer than $\operatorname{tower}(\lceil \varepsilon^{-c} \rceil)$ parts.

We do not include the proof here. See Moshkovitz and Shapira (2016) for a short proof. The general idea is to construct a graph that roughly reverse engineers the proof of the regularity lemma. So there is essentially a unique ε-regular partition, which must have many parts.

***Remark* 2.1.18** (Irregular pairs are necessary in the regularity lemma). Recall that in Definition 2.1.7 of an ε-regular partition, we are allowed to have some irregular pairs. Are irregular pairs necessary? It turns that we must permit them. Exercise 2.1.24 gives an example of a canonical example (a "half graph") where every regularity partition has irregular pairs.

The regularity lemma is quite flexible. For example, we can start with an arbitrary partition of $V(G)$ instead of the trivial partition in the proof, in order to obtain a partition that is a refinement of a given partition. The exact same proof with this modification yields the following.

Theorem 2.1.19 (Regularity starting with an arbitrary initial partition)
For every $\varepsilon > 0$ and k, there exists a constant M such that for every graph G and a partition $\mathcal{P}_0$ of $V(G)$ with at most k parts, there exists an ε-regular partition $\mathcal{P}$ of $V(G)$ that is a refinement of $\mathcal{P}_0$, and such that each part of $\mathcal{P}_0$ is refined into at most M parts.

Here is another strengthening of the regularity lemma. We impose the additional requirement that vertex parts should be as equal in size as possible. We say that a partition is ***equitable*** if all part sizes are within one of each other; that is, $\left||V_i| - |V_j|\right| \le 1$. In other words, a partition of a set of size n into k parts is equitable if every part has size $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$.

Theorem 2.1.20 (Equitable regularity lemma)
For all $\varepsilon > 0$ and m_0, there exists a constant M such that every graph has an ε-regular equitable partition of its vertex set into k parts with $m_0 \le k \le M$.

***Remark* 2.1.21.** The lower bound m_0 requirement on the number of parts is somewhat superficial. The reason for including it here is that it is often convenient to discard all the edges that lie within individual parts of the partition, and since there are most n^2/k such edges, they contribute negligibly if the number of parts k is not too small, which is true if we require $m_0 \ge 1/\varepsilon$ in the equitable regularity lemma statement.

There are several ways to guarantee equitability. One method is sketched in what follows. We equitize the partition at every step of the refinement iteration, so that at each step in the proof, we both obtain an energy increment and also end up with an equitable partition.

Proof sketch of the equitable regularity lemma (Theorem 2.1.20). Here is a modified algorithm:

(1) Start with an arbitrary equitable partition of the graph into m_0 parts.
(2) While the current equitable partition $\mathcal{P}$ is not ε-regular:
    (a) (Refinement/energy boost) Refine the partition using pairs that witness irregularity (as in the earlier proof). The new partition $\mathcal{P}'$ divides each part of $\mathcal{P}$ into $\le 2^{|\mathcal{P}|}$ parts.
    (b) (Equitization) Modify $\mathcal{P}'$ into an equitable partition by arbitrarily chopping each part of $\mathcal{P}'$ into parts of size $|V(G)|/m$ (for some appropriately chosen $m = m(|\mathcal{P}'|, \varepsilon)$) plus some leftover pieces, which are then combined together and then divided into parts of size $|V(G)|/m$.

The refinement step (2)(a) increases energy by $\ge \varepsilon^5$ as before. The energy might go down in the equitization step (2)(b), but it should not decrease by much, provided that the m chosen in that step is large enough (say, $m = \lfloor 100 |\mathcal{P}'| \varepsilon^{-5} \rfloor$). So overall, we still have an energy increment of $\ge \varepsilon^5/2$ at each step, and hence the process still terminates after $O(\varepsilon^{-5})$ steps. The total number of parts at the end is bounded. □

Exercise 2.1.22. Complete the details in the preceding proof sketch.

Exercise 2.1.23 (Making each part ε-regular to nearly all other parts)**.** Prove that for all $\varepsilon > 0$ and m_0, there exists a constant M so that every graph has an equitable vertex partition into k parts, with $m_0 \le k \le M$, such that each part is ε-regular with all but at most εk other parts.

The important example in the next exercise shows why we must allow irregular pairs in the graph regularity lemma.

Exercise 2.1.24 (Unavoidability of irregular pairs)**.** Let the ***half-graph*** H_n be the bipartite graph on $2n$ vertices $\{a_1, \dots, a_n, b_1, \dots, b_n\}$ with edges $\{a_i b_j : i \le j\}$.

(a) For every $\varepsilon > 0$, explicitly construct an ε-regular partition of H_n into $O(1/\varepsilon)$ parts.

(b) Show that there is some $\varepsilon > 0$ such that for every integer k and sufficiently large multiple n of k, every partition of the vertices of H_n into k equal-sized parts contains at least εk pairs of parts none of which are ε-regular.

The next exercise should remind you of the iteration technique from the proof of the graph regularity lemma.

Exercise 2.1.25 (Existence of a regular pair of subsets)**.** Show that there is some absolute constant $C > 0$ such that for every $0 < \varepsilon < 1/2$, every graph on n vertices contains an ε-regular pair of vertex subsets each with size at least δn, where $\delta = 2^{-\varepsilon^{-C}}$.

Hint: Density increment (don't use energy).

This exercise asks for two different proofs of the following theorem.

Given a graph G, we say that $X \subseteq V(G)$ is ***ε-regular*** if the pair (X, X) is ε-regular; that is, for all $A, B \subseteq X$ with $|A|, |B| \ge \varepsilon |X|$, one has $|d(A, B) - d(X, X)| \le \varepsilon$.

Theorem 2.1.26 (ε-regular subset)

For every $\varepsilon > 0$, there exists $\delta > 0$ such that every graph contains an ε-regular subset of vertices that is an $\ge \delta$ fraction of the vertex set.

Exercise 2.1.27 (ε-regular subset)**.**

(a) Prove Theorem 2.1.26 using Szemerédi's regularity lemma, showing that one can obtain the ε-regular subset by combining a suitable subcollection of parts from some regularity partition.

(b*) Give an alternative proof of the theorem with $\delta = \exp(-\exp(\varepsilon^{-C}))$ for some constant C.

Exercise 2.1.28* (Regularity partition into regular sets)**.** Show that for every $\varepsilon > 0$ there exists M so that every graph has an ε-regular partition into at most M parts, with every part being ε-regular with itself.

2.2 Triangle Counting Lemma

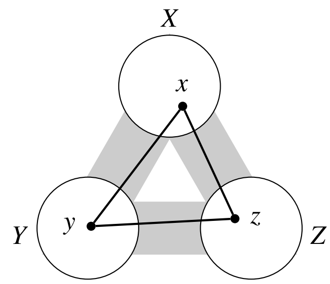
Szemerédi's regularity lemma gave us a vertex partition of a graph. How can we use this partition?

In this section, we begin by establishing the **triangle counting lemma**. Given three vertex sets X, Y, Z, pairwise ε-regular in G, we can approximate it by a random tripartite graph on

X, Y, Z with the same edge densities between parts. By comparing G to its random model approximation, we expect the number of triples $(x, y, z) \in X \times Y \times Z$ forming a triangle in G to be roughly

$$d(X,Y)d(X,Z)d(Y,Z)|X||Y||Z|.$$

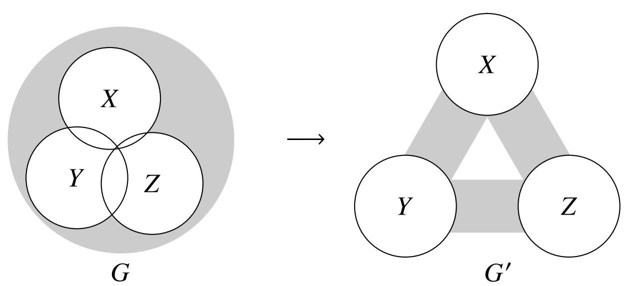
The triangle counting lemma makes this intuition precise.



Theorem 2.2.1 (Triangle counting lemma)
Let G be a graph and X, Y, Z be subsets of the vertices of G such that (X, Y), (Y, Z), (Z, X) are all ε-regular pairs for some $\varepsilon > 0$. If $d(X, Y), d(X, Z), d(Y, Z) \geq 2\varepsilon$, then

$$|\{(x, y, z) \in X \times Y \times Z : xyz \text{ is a triangle in } G\}| \geq (1 - 2\varepsilon)(d(X, Y) - \varepsilon)(d(X, Z) - \varepsilon)(d(Y, Z) - \varepsilon)\,|X|\,|Y|\,|Z|\,.$$

***Remark* 2.2.2.** The vertex sets X, Y, Z do not have to be disjoint, but one does not lose any generality by assuming that they are disjoint in this statement. Indeed, starting with $X, Y, Z \subseteq V(G)$, one can always create an auxiliary tripartite graph G' with vertex parts being disjoint replicas of X, Y, Z and the edge relations in $X \times Y$ being the same for G and G', and likewise for $X \times Z$ and $Y \times Z$. Under this auxiliary construction, a triple in $X \times Y \times Z$ forms a triangle in G if and only it forms a triangle in G'.

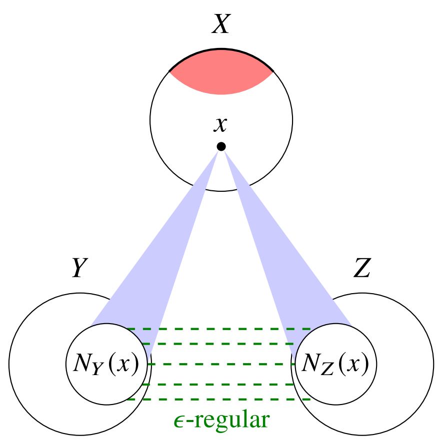


Now we show that in an ε-regular pair (X, Y), almost all vertices of X have roughly the same number of neighbors in Y. (The next lemma only states a lower bound on degree, but the same argument also gives an analogous upper bound.)

Lemma 2.2.3 (Most vertices have roughly the same degree)
Let (X, Y) be an ε-regular pair. Then fewer than $\varepsilon|X|$ vertices in X have fewer than $(d(X, Y) - \varepsilon)|Y|$ neighbors in Y. Likewise, fewer than $\varepsilon|Y|$ vertices in Y have fewer than $(d(X, Y) - \varepsilon)|X|$ neighbors in X.

Proof. Let A be the subset of vertices in X with $< (d(X, Y) - \varepsilon)|Y|$ neighbors in Y. Then $d(A, Y) < d(X, Y) - \varepsilon$, and thus $|A| < \varepsilon|X|$ by Definition 2.1.2 as (X, Y) is an ε-regular pair. The other claim is similar. □

Proof of Theorem 2.2.1. By Lemma 2.2.3, we can find $X' \subseteq X$ with $|X'| \geq (1 - 2\varepsilon)|X|$ such that every vertex $x \in X'$ has $\geq (d(X, Y) - \varepsilon)|Y|$ neighbors in Y and $\geq (d(X, Z) - \varepsilon)|Z|$ neighbors in Z. Write $N_Y(x) = N(x) \cap Y$ and $N_Z(x) = N(x) \cap Z$.



For each such $x \in X'$, we have $|N_Y(x)| \geq (d(X, Y) - \varepsilon)|Y| \geq \varepsilon|Y|$. Likewise, $|N_Z(x)| \geq \varepsilon|Z|$. Since (Y, Z) is ε-regular, the edge density between $N_Y(x)$ and $N_Z(x)$ is $\geq d(Y, Z) - \varepsilon$. So for each $x \in X'$, the number of edges between $N_Y(x)$ and $N_Z(x)$ is

$$\geq (d(Y, Z) - \varepsilon)|N_Y(x)||N_Z(x)| \geq (d(X, Y) - \varepsilon)(d(X, Z) - \varepsilon)(d(Y, Z) - \varepsilon)|Y||Z|.$$

Multiplying by $|X'| \geq (1 - 2\varepsilon)|X|$, we obtain the desired lower bound on the number of triangles. □

***Remark* 2.2.4.** We only need the lower bound on the triangle count for our applications in this chapter, but the same proof can also be modified to give an upper bound, which we leave as an exercise.

2.3 Triangle Removal Lemma

The triangle removal lemma (Ruzsa and Szemerédi 1978) is one of the first major applications of the regularity method. Informally, the triangle removal lemma says that a graph with few triangles can be made triangle-free by removing a few edges. Here, "few triangles" means a subcubic number of triangles (i.e., asymptotically less than the maximum possible number) and "few edges" means a subquadratic number of edges.

Theorem 2.3.1 (Triangle removal lemma)
For all $\varepsilon > 0$, there exists $\delta > 0$ such that any graph on n vertices with fewer than δn^3 triangles can be made triangle-free by removing fewer than εn^2 edges.

The triangle removal lemma can be equivalently stated as:

An n-vertex graph with $o(n^3)$ triangles can be made triangle-free by removing $o(n^2)$ edges.

Our proof of Theorem 2.3.1 demonstrates how to apply the graph regularity lemma. Here is a representative "recipe" for the regularity method.

***Remark* 2.3.2** (Regularity method recipe). Typical applications of the regularity method proceed in the following steps:

(1) ***Partition*** the vertex set of a graph using the regularity lemma.
(2) ***Clean*** the graph by removing edges that behave poorly in the regularity partition. Most commonly, we remove edges that lie between pairs of parts with
   (a) irregularity, or
   (b) low-density, or
   (c) one of the parts too small.
   This ends up removing a negligible number of edges.
(3) ***Count*** a certain pattern in the cleaned graph using a counting lemma.

To prove the triangle removal lemma, after cleaning the graph (which removes few edges), we claim that the resulting cleaned graph must be triangle-free, or else the triangle counting lemma would find many triangles, contradicting the hypothesis.

Proof of the triangle removal lemma (Theorem 2.3.1). Suppose we are given a graph on n vertices with $< \delta n^3$ triangles, for some parameter δ we will choose later. Apply the graph regularity lemma, Theorem 2.1.9, to obtain an $\varepsilon/4$-regular partition of the graph with parts $V_1, V_2, \cdots, V_m$. Next, for each $(i, j) \in [m]^2$, remove all edges between V_i and V_j if

(a) (V_i, V_j) is not $\varepsilon/4$-regular, or
(b) $d(V_i, V_j) < \varepsilon/2$, or
(c) $\min\{|V_i|, |V_j|\} < \varepsilon n/(4m)$.

Since the partition is $\varepsilon/4$-regular (recall Definition 2.1.7), the number of edges removed in (a) from irregular pairs is

$$\le \sum_{\substack{i,j \\ (V_i,V_j) \text{ not } (\varepsilon/4)\text{-regular}}} |V_i||V_j| \le \frac{\varepsilon}{4} n^2.$$

The number of edges removed in (b) from low-density pairs is

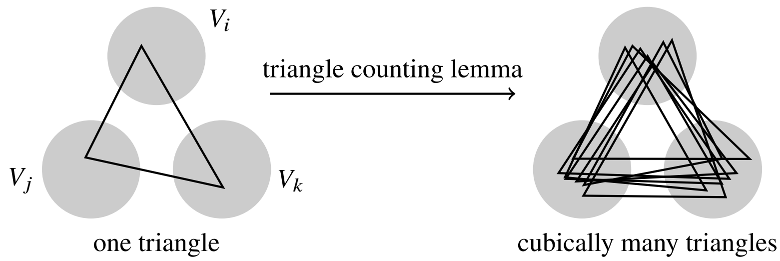
$$\le \sum_{\substack{i,j \\ d(V_i,V_j)<\varepsilon/2}} d(V_i, V_j)|V_i||V_j| \le \frac{\varepsilon}{2} \sum_{i,j} |V_i||V_j| = \frac{\varepsilon}{2} n^2.$$

The number of edges removed in (c) with an endpoint in a small part is

$$< m \cdot \frac{\varepsilon n}{4m} \cdot n = \frac{\varepsilon}{4} n^2.$$

In total, we removed $< \varepsilon n^2$ edges from the graph.

We claim that the remaining graph is triangle-free, provided that δ was chosen appropriately small. Indeed, suppose there remains a triangle whose three vertices lie in V_i, V_j, V_k (not necessarily distinct parts).



Because edges between the pairs described in (a) and (b) were removed, V_i, V_j, V_k satisfy the hypotheses of the triangle counting lemma (Theorem 2.2.1),

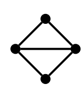
$$\begin{aligned}\#\{\text{triangles in } V_i \times V_j \times V_k\} &\geq \left(1 - \frac{\varepsilon}{2}\right)\left(\frac{\varepsilon}{4}\right)^3 |V_i|\left|V_j\right||V_k| \\ &\geq \left(1 - \frac{\varepsilon}{2}\right)\left(\frac{\varepsilon}{4}\right)^3 \left(\frac{\varepsilon n}{4m}\right)^3,\end{aligned}$$

where the final step uses (c) above. Then as long as

$$\delta < \frac{1}{6}\left(1 - \frac{\varepsilon}{2}\right)\left(\frac{\varepsilon}{4}\right)^3 \left(\frac{\varepsilon}{4m}\right)^3,$$

we would contradict the hypothesis that the original graph has $< \delta n^3$ triangles. (The extra factor of 6 above is there to account for the possibility that $V_i = V_j = V_k$.) Since m is bounded for each fixed ε, we see that δ can be chosen to depend only on ε. □

The next corollary of the triangle removal lemma will soon be used to prove Roth's theorem. Here "diamond" refers to the following graph, consisting of two triangles sharing an edge.



Corollary 2.3.3 (Diamond-free lemma)
Let G be an n-vertex graph where every edge lies in a unique triangle. Then G has $o(n^2)$ edges.

Proof. Let G have m edges. Because each edge lies in exactly one triangle, the number of triangles in G is $m/3 = O(n^2) = o(n^3)$. By the triangle removal lemma (see the statement after Theorem 2.3.1), we can remove $o(n^2)$ edges to make G triangle-free. However, deleting an edge removes at most one triangle from the graph by assumption, so $m/3$ edges need to be removed to make G triangle-free. Thus $m = o(n^2)$. □

***Remark* 2.3.4** (Quantitative dependencies in the triangle removal lemma). Since the above proof of the triangle removal lemma applies the graph regularity lemma, the resulting bounds from the proof are quite poor: it shows that one can pick $\delta = 1/\text{tower}(\varepsilon^{-O(1)})$. Using a different but related method, Fox (2011) proved the triangle removal lemma with a slightly better dependence $\delta = 1/\text{tower}(O(\log(1/\varepsilon)))$. In the other direction, we know that the triangle removal lemma does not hold with $\delta = \varepsilon^{c\log(1/\varepsilon)}$ for a sufficiently small constant $c > 0$. The construction comes from the Behrend construction of large 3-AP-free sets that we will soon see in Section 2.5. Our knowledge of the quantitative dependence in Corollary 2.3.3 comes from the same source; specifically, we know that the $o(n^2)$ can be sharpened to $n^2/e^{\Omega(\log^*(1/\varepsilon))}$ (where $\log^*$, the iterated logarithm function, is the number of iterations of log that one needs to take to bring a number to at most 1) but the statement is false if the $o(n^2)$ is replaced by $n^2 e^{-C\sqrt{\log n}}$ for some sufficiently large constant C. It is a major open problem to close the gap between the upper and lower bounds in these problems.

The triangle removal lemma was historically first considered in the following equivalent formulation.

Theorem 2.3.5 ((6, 3)-theorem)
Let H be an n-vertex 3-uniform hypergraph without a subgraph having six vertices and three edges. Then H has $o(n^2)$ edges.

Exercise 2.3.6. Deduce the (6, 3)-theorem from Corollary 2.3.3, and vice versa.

The following conjectural extension of the (6, 3)-theorem is a major open problem in extremal combinatorics. The conjecture is attributed to Brown, Erdős, and Sós (1973).

Conjecture 2.3.7 ((7, 4)-conjecture)
Let H be an n-vertex 3-uniform hypergraph without a subgraph having seven vertices and four edges. Then H has $o(n^2)$ edges.

2.4 Graph Theoretic Proof of Roth's Theorem

We will now prove Roth's theorem, which we saw in Chapter 0 and is restated below. The proof below, due to Ruzsa and Szemerédi (1978) connects graph theory and additive combinatorics, akin to the proof of Schur's theorem in Chapter 0.

We write ***3-AP*** for "3-term arithmetic progression." We say that A is ***3-AP-free*** if there are no $x, x+y, x+2y \in A$ with $y \neq 0$.

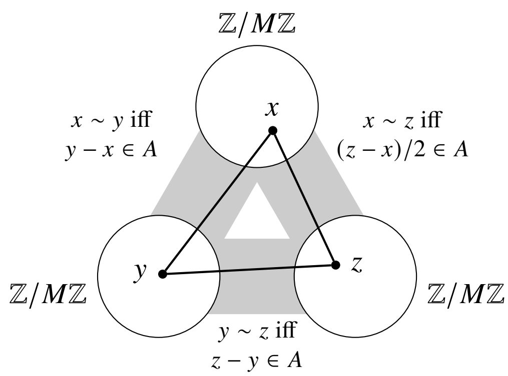
Theorem 2.4.1 (Roth's theorem)
Let $A \subseteq [N]$ be 3-AP-free. Then $|A| = o(N)$.

Proof. Embed $A \subseteq \mathbb{Z}/M\mathbb{Z}$ with $M = 2N+1$ (to avoid wraparounds). Since A is 3-AP-free in $\mathbb{Z}$, it is 3-AP-free in $\mathbb{Z}/M\mathbb{Z}$ as well.

Now, we construct a tripartite graph G whose parts X, Y, Z are all copies of $\mathbb{Z}/M\mathbb{Z}$. The edges of the graph are (since M is odd, we are allowed to divide by 2 in $\mathbb{Z}/M\mathbb{Z}$):

- $(x, y) \in X \times Y$ whenever $y - x \in A$;

- $(y, z) \in Y \times Z$ whenever $z - y \in A$;
- $(x, z) \in X \times Z$ whenever $(z - x)/2 \in A$.



In this graph, $(x, y, z) \in X \times Y \times Z$ is a triangle if and only if

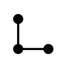
$$y - x, \frac{z - x}{2}, z - y \in A.$$

The graph was designed so that the above three numbers form an arithmetic progression in the listed order. Since A is 3-AP-free, these three numbers must all be equal. So, every edge of G lies in a unique triangle, formed by setting the three numbers above to equal.

The graph G has exactly $3M = 6N + 3$ vertices and $3M |A|$ edges. Corollary 2.3.3 implies that G has $o(N^2)$ edges. So $|A| = o(N)$. □

Now we prove a higher-dimensional generalization of Roth's theorem.

A ***corner*** in $\mathbb{Z}^2$ is a three-element set of the form $\{(x, y), (x + d, y), (x, y + d)\}$ with $d > 0$.



(Note that one could relax the assumption $d > 0$ to $d \neq 0$, allowing "negative" corners. As shown in the first step in the proof below, the assumption $d > 0$ is inconsequential.)

Theorem 2.4.2 (Corner-free)
Every corner-free subset of $[N]^2$ has size $o(N^2)$.

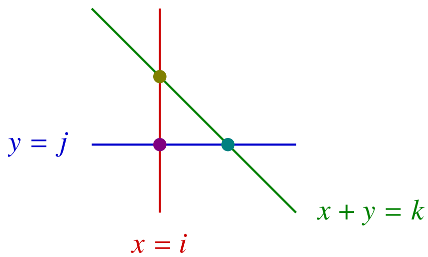
***Remark* 2.4.3** (History)**.** The theorem is due to Ajtai and Szemerédi (1974), who originally proved it by invoking the full power of Szemerédi's theorem. Here we present a much simpler proof using the triangle removal lemma due to Solymosi (2003).

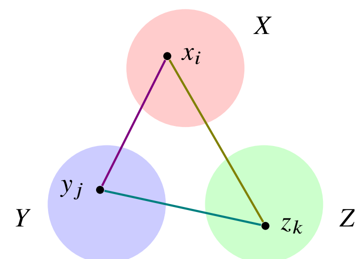
Proof. First we show how to relax the assumption in the definition of a corner from $d > 0$ to $d \neq 0$.

Let $A \subseteq [N]^2$ be a corner-free set. For each $z \in \mathbb{Z}^2$, let $A_z = A \cap (z - A)$. Then $|A_z|$ is the number of ways that one can write $z = a + b$ for some $(a, b) \in A \times A$. So $\sum_{z \in [2N]^2} |A_z| = |A|^2$, so there is some $z \in [2N]$ with $|A_z| \geq |A|^2 / (2N)^2$. To show that $|A| = o(N^2)$, it suffices to show that $|A_z| = o(N^2)$. Moreover, since $A_z = z - A_z$, it being corner-free implies that it does not contain three points $\{(x, y), (x + d, y), (x, y + d)\}$ with $d \neq 0$.

Write $A = A_z$ from now on. Build a tripartite graph G with parts $X = \{x_1, \ldots, x_N\}$,

$Y = \{y_1, \ldots, y_N\}$ and $Z = \{z_1, \ldots, z_{2N}\}$, where each vertex x_i corresponds to a vertical line $\{x = i\} \subseteq \mathbb{Z}^2$, each vertex y_j corresponds to a horizontal line $\{y = j\}$, and each vertex z_k corresponds to a slanted line $\{y = -x + k\}$ with slope -1. Join two distinct vertices of G with an edge if and only if the corresponding lines intersect at a point belonging to A. Then, each triangle in the graph G corresponds to a set of three lines of slopes $0, \infty, -1$ pairwise intersecting at a point of A.

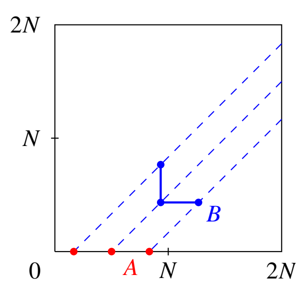




Since A is corner-free in the sense stated at the end of the previous paragraph, x_i, y_j, z_k form a triangle in G if and only if the three corresponding lines pass through the same point of A (i.e., forming a trivial corner with $d = 0$). Since there is exactly one line of each direction passing through every point of A, it follows that each edge of G belongs to exactly one triangle. Thus, by Corollary 2.3.3, $3|A| = e(G) = o(N^2)$. □

The upper bound on corner-free sets actually implies Roth's theorem, as shown below. So we now have a second proof of Roth's theorem. (Though, this second proof is secretly the same as the first proof.)

> **Proposition 2.4.4** (Corner-free sets vs. 3-AP-free sets)
> Let $r_3(N)$ be the size of the largest subset of $[N]$ which contains no 3-term arithmetic progression, and $r_\llcorner(N)$ be the size of the largest subset of $[N]^2$ which contains no corner. Then, $r_3(N)N \le r_\llcorner(2N)$.



Proof. Given a 3-AP-free set $A \subseteq [N]$ of size $r_3(N)$, define a set

$$B := \left\{(x, y) \in [2N]^2 : x - y \in A\right\}.$$

Each element $a \in A$ gives rise to $\ge N$ different elements (x, y) of B with $x - y = a$. So $|B| \ge N|A|$. Furthermore, B is corner-free, since each corner $(x + d, y), (x, y), (x, y + d)$ in B gives rise to a 3-AP $x - y - d, x - y, x - y + d$ with common difference d. So $|B| \le r_\llcorner(2N)$. Thus $r_3(N)N \le |A|N \le |B| \le r_\llcorner(2N)$. □

***Remark* 2.4.5** (Quantitative bounds)**.** Both proofs above rely on the graph regularity lemma, and hence give poor quantitative bounds. They tell us that a 3-AP-free $A \subseteq [N]$ has $|A| \leq N/(\log^* N)^c$, where $\log^* N$ is the iterated logarithm (the number of times the logarithm function must be applied to bring N to less than or equal to 1). Later in Chapter 6 we discuss Roth's original Fourier analytic proof, which uses different methods (though sharing the structure and randomness dichotomy theme) and gives much better quantitative bounds.

The current best upper bound on the size of a 3-AP-free subset of $[N]$ is $N/(\log N)^{1+c}$ for some constant $c > 0$ (Bloom and Sisask 2020). The current best upper bound on the size of corner-free subsets of $[N]^2$ is $N^2/(\log \log N)^c$ for some constant $c > 0$ (Shkredov 2006). Both use Fourier analysis.

For the next exercise, apply the triangle removal lemma to an appropriate graph.

Exercise 2.4.6* (Arithmetic triangle removal lemma)**.** Show that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $A \subseteq [n]$ has fewer than δn^2 many triples $(x, y, z) \in A^3$ with $x + y = z$, then there is some $B \subseteq A$ with $|A \setminus B| \leq \varepsilon n$ such that B is sum-free (i.e., no $x, y, z \in B$ with $x + y = z$).

2.5 Large 3-AP-Free Sets: Behrend's Construction

How can we construct a large 3-AP-free subset of $[N]$?

We can do it greedily. Starting with 0 (which produces a nicer pattern), we successively put in each positive integer if adding it does not create a 3-AP with the already chosen integers. This would produce the following sequence:

$$0 \quad 1 \quad 3 \quad 4 \quad 9 \quad 10 \quad 12 \quad 13 \quad 27 \quad 28 \quad 30 \quad 31 \quad \ldots.$$

The above sequence is known as a ***Stanley sequence***. It consists of all nonnegative integers whose ternary representations have only the digits 0 and 1. (Why?) Up to $N = 3^k$, the subset $A \subseteq [N]$ so constructed has size $|A| = 2^k = N^{\log_3 2}$.

For quite some time, people thought the above example was close to the optimal. Salem and Spencer (1942) then found a much larger 3-AP-free subset of $[N]$, with size $N^{1-o(1)}$. Their result was further improved by Behrend (1946), whose construction we present below. This construction has not yet been substantially improved (see Elkin (2011) and Green and Wolf (2010) for some lower order improvements).

Behrend's construction has surprising applications, such as in the design of fast matrix multiplication algorithms (Coppersmith and Winograd 1990).

Theorem 2.5.1 (Behrend's construction)

There exists a constant $C > 0$ such that for every positive integer N, there exists a 3-AP-free $A \subseteq [N]$ with $|A| \geq N e^{-C\sqrt{\log N}}$.

The rough idea is to first find a high-dimensional sphere with many lattice points via the pigeonhole principle. The sphere contains no 3-AP due to convexity. We then project these lattice points onto $\mathbb{Z}$ in a way that creates no additional 3-APs. This is done by treating the coordinates as the base-q expansion of an integer with some large q.

Proof. Let m and d be two positive integers depending on N to be specified later. Consider the lattice points of $X = [m]^d$ that lie on a sphere of radius $\sqrt{L}$:

$$X_L := \left\{(x_1, \ldots, x_d) \in X : x_1^2 + \cdots + x_d^2 = L\right\}.$$

Then, $X = \bigcup_{i=1}^{dm^2} X_i$. So by the pigeonhole principle, there exists an $L \in [dm^2]$ such that $|X_L| \geq m^d/(dm^2)$. Define the base $2m$ digital expansion

$$\phi(x_1, \ldots, x_d) := \sum_{i=1}^{d} x_i (2m)^{i-1}.$$

Then ϕ is injective on X. Furthermore, $x, y, z \in [m]^d$ satisfy $x + z = 2y$ if and only if $\phi(x) + \phi(z) = 2\phi(y)$ (there are no wraparounds in base $2m$ with all coordinates in $[m]$). Since X_L is a subset of a sphere, it is 3-AP-free. Thus $\phi(X) \subseteq [(2m)^d]$ is a 3-AP-free set of size $\geq m^d/(dm^2)$. We can optimize the parameters and take $m = \lfloor e^{\sqrt{\log N}}/2 \rfloor$ and $d = \lfloor \sqrt{\log N} \rfloor$, thereby producing a 3-AP-free subset of $[N]$ with of size $\geq Ne^{-C\sqrt{\log N}}$, where C is some absolute constant. □

The Behrend construction also implies lower bound constructions for the other problems we saw earlier. For example, since we used the diamond-free lemma (Corollary 2.3.3) to deduce an upper bound on the size of 3-AP-free set, turning this implication around, we see that having a large 3-AP-free set implies the following quantitative limitation on the diamond-free lemma.

Corollary 2.5.2 (Lower bound for the diamond-free lemma)

For every $n \geq 3$, there is some n-vertex graph with at least $n^2 e^{-C\sqrt{\log n}}$ edges where every edge lies on a unique triangle. Here C is some absolute constant.

Proof. In the proof of Theorem 2.4.1, starting from a 3-AP-free set $A \subseteq [N]$, we constructed a graph with $6N+3$ vertices and $(6N+3)\,|A|$ edges such that every edge lies in a unique triangle. Choosing $N = \lfloor (n-3)/6 \rfloor$ and letting A be the Behrend construction of Theorem 2.5.1 with $|A| \geq Ne^{-C\sqrt{\log N}}$, we obtain the desired graph. □

***Remark* 2.5.3** (More lower bounds from Behrend's construction)**.** The same graph construction also shows, after examining the proof of Corollary 2.3.3, that in the triangle removal lemma, Theorem 2.3.1, one cannot take $\delta = e^{-c(\log(1/\varepsilon))^2}$ if the constant $c > 0$ is too small.

In Proposition 2.4.4 we deduced an upper bound $r_3(N)N \leq r_\llcorner(2N)$ on corner-free sets using 3-AP-free sets. The Behrend construction then also gives a corner-free subset of $[N]^2$ of size $\geq N^2 e^{-C\sqrt{\log N}}$.

Exercise 2.5.4 (Modifying Behrend's construction)**.** Prove that there is some constant $C > 0$ so that for all N, there exists $A \subseteq [N]$ with $|A| \geq N \exp(-C\sqrt{\log N})$ so that there do not exist $w, y, x, z \in A$ not all equal and satisfying $x + y + z = 3w$.

Exercise 2.5.5* (Avoiding 5-term quadratic configurations)**.** Prove that there is some constant $C > 0$ so that for all N, there exists $A \subseteq [N]$ with $|A| \geq N \exp(-C\sqrt{\log N})$ so that there does not exist a nonconstant quadratic polynomial P so that $P(0), P(1), P(2), P(3), P(4) \in A$.

2.6 Graph Counting and Removal Lemmas

In this section, we generalize the triangle counting lemma from triangles to other graphs and discuss applications.

Graph Counting Lemma

Let us first illustrate the technique for K_4. Similar to the triangle counting lemma, we embed the vertices of K_4 one at a time. At each stage we ensure that many eligible vertices remain for the yet to be embedded vertices.

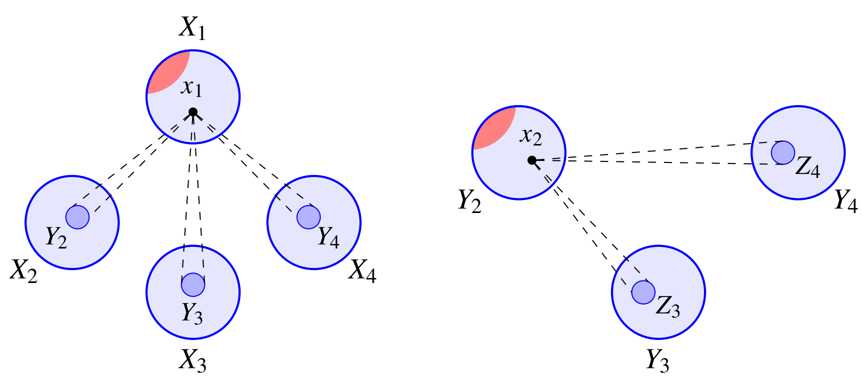
Proposition 2.6.1 (K_4 counting lemma)
Let $0 < \varepsilon < 1$. Let $X_1, \dots, X_4$ be vertex subsets of a graph G such that (X_i, X_j) is ε-regular with edge-density $d_{ij} := d(X_i, X_j) \geq 3\sqrt{\varepsilon}$ for each pair $i < j$. Then the number of quadruples $(x_1, x_2, x_3, x_4) \in X_1 \times X_2 \times X_3 \times X_4$ such that $x_1x_2x_3x_4$ is a clique in G is

$$\geq (1 - 3\varepsilon)(d_{12} - 3\varepsilon)(d_{13} - \varepsilon)(d_{14} - \varepsilon)(d_{23} - \varepsilon)(d_{24} - \varepsilon)(d_{34} - \varepsilon) |X_1| |X_2| |X_3| |X_4| .$$

Proof. We repeatedly apply the following statement, which is a simple consequence of the definition of ε-regularity (and a small extension of Lemma 2.2.3):

Given an ε-regular pair (X, Y), and $B \subseteq Y$ with $|B| \geq \varepsilon |Y|$, the number of vertices in X with $< (d(X, Y) - \varepsilon) |B|$ neighbors in B is $< \varepsilon |X|$.

The number of vertices X_1 with $\geq (d_{1i} - \varepsilon) |X_i|$ neighbors in X_i for each $i = 2, 3, 4$ is $\geq (1 - 3\varepsilon) |X_1|$. Fix a choice of such an $x_1 \in X_1$. For each $i = 2, 3, 4$, let Y_i be the neighbors of x_1 in X_i, so that $|Y_i| \geq (d_{1i} - \varepsilon) |X_i|$.



The number of vertices in Y_2 with $\geq (d_{2i} - \varepsilon) |Y_i|$ common neighbors in Y_i for each $i = 3, 4$ is $\geq |Y_2| - 2\varepsilon |X_2| \geq (d_{12} - 3\varepsilon) |X_2|$. Fix a choice of such an $x_2 \in Y_2$. For each $i = 3, 4$, let Z_i be the neighbors of x_2 in Y_i.

For each $i = 3, 4$, $|Z_i| \geq (d_{1i} - \varepsilon)(d_{2i} - \varepsilon) |X_i| \geq \varepsilon |X_i|$, and so

$$\begin{aligned} e(Z_3, Z_4) &\geq (d_{34} - \varepsilon) |Z_3| |Z_4| \\ &\geq (d_{34} - \varepsilon) \cdot (d_{13} - \varepsilon)(d_{23} - \varepsilon) |X_3| \cdot (d_{14} - \varepsilon)(d_{24} - \varepsilon) |X_4| . \end{aligned}$$

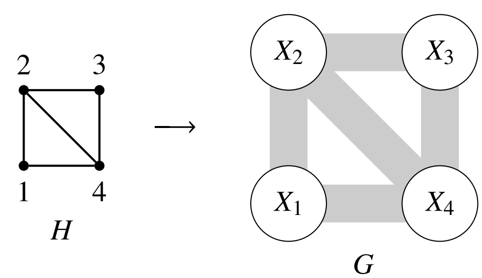
Any edge between Z_3 and Z_4 forms a K_4 together with x_1 and x_2. Multiplying the above quantity with the earlier lower bounds on the number of choices of x_1 and x_2 gives the result. □

The same strategy works more generally for counting any graph. To find copies of H, we embed vertices of H one at a time.

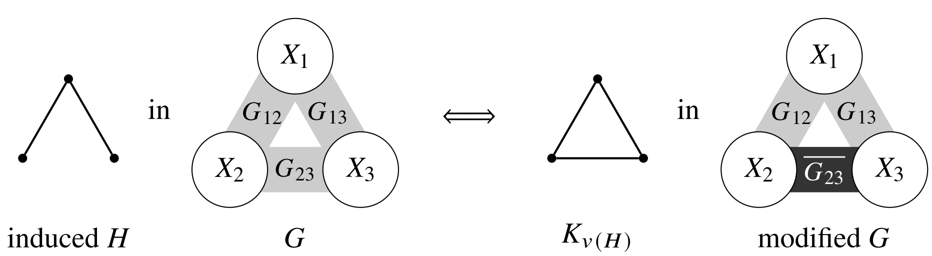
Theorem 2.6.2 (Graph counting lemma)
For every graph H and real $\delta > 0$, there exists an $\varepsilon > 0$ such that the following is true.
Let G be a graph, and $X_i \subseteq V(G)$ for each $i \in V(H)$ such that for each $ij \in E(H)$, (X_i, X_j) is an ε-regular pair with edge density $d_{ij} := d(X_i, X_j) \geq \delta$. Then the number of graph homomorphisms $H \to G$ where each $i \in V(H)$ is mapped to X_i is

$$\geq (1-\delta) \prod_{ij \in E(H)} (d_{ij} - \delta) \prod_{i \in V(H)} |X_i|.$$

***Remark* 2.6.3.** (a) For a fixed H, as $|X_i| \to \infty$ for each i, all but a negligible fraction of such homomorphisms from H are injective (i.e., yielding a copy of H as a subgraph).
(b) It is useful (and in fact equivalent) to think about the setting where G is a multipartite graph with parts X_i, as illustrated below.



In the multipartite setting, we see that the graph counting lemma can be adapted to variants such as counting induced copies of H. Indeed, an induced copy of H is the same as a $v(H)$-clique in an auxiliary graph G' obtained by replacing the bipartite graph in G between X_i and X_j by its complementary bipartite graph between X_i and X_j for each $ij \notin E(H)$.



(c) We will see a different proof in Section 4.5 using the language of graphons. There, instead of embedding H one vertex at a time, we compare the density of H and $H \setminus \{e\}$.

We establish the following stronger statement, which has the additional advantage that one can choose the regularity parameter ε to depend on the maximum degree of H rather than H itself. You may wish to skip reading the proof, as it is notationally rather heavy. The main ideas were already illustrated in the K_4 counting lemma.

Theorem 2.6.4 (Graph counting lemma)
Let H be a graph with maximum degree $\Delta \geq 1$ and $c(H)$ connected components. Let $\varepsilon > 0$. Let G be a graph. Let $X_i \subseteq V(G)$ for each $i \in V(H)$. Suppose that for each $ij \in E(H)$, (X_i, X_j) is an ε-regular pair with edge density $d_{ij} := d(X_i, X_j) \geq (\Delta+1)\varepsilon^{1/\Delta}$. Then the number of graph homomorphisms $H \to G$ where each $i \in V(H)$ is mapped to X_i is

$$\geq (1 - \Delta\varepsilon)^{c(H)} \prod_{ij \in E(H)} (d_{ij} - \Delta\varepsilon^{1/\Delta}) \cdot \prod_{i \in V(H)} |X_i| \,.$$

Furthermore, if $|X_i| \geq v(H)/\varepsilon$ for each i, then there exists such a homomorphism $H \to G$ that is injective (i.e., an embedding of H as a subgraph).

Proof. Let us order and label the vertices of H by $1, \dots, v(H)$ arbitrarily. We will select vertices $x_1 \in X_1, x_2 \in X_2, \dots$ in order. The idea is to always make sure that they have enough neighbors in G so that there are many ways to continue the embedding of H. We say that a partial embedding $x_1, \dots, x_{s-1}$ (here *partial embedding* means that $x_i x_j \in E(G)$ whenever $ij \in E(H)$ for all the x_is chosen so far) is *abundant* if for each $j \geq s$, the number of valid extensions $x_j \in X_j$ (meaning that $x_i x_j \in E(G)$ whenever $i < s$ and $ij \in E(H)$) is $\geq |X_j| \prod_{i<s: ij \in E(H)} (d_{ij} - \varepsilon)$.

For each $s = 1, 2, \dots, v(H)$ in order, suppose we have already fixed an abundant partial embedding $x_1, \dots, x_{s-1}$. For each $j \geq s$, let

$$Y_j = \{x_j \in X_j : x_i x_j \in E(G) \text{ whenever } i < s \text{ and } ij \in E(H)\}$$

be the set of valid extensions of the jth vertex in X_j given the partial embeddings of $x_1, \dots, x_{s-1}$, so that the abundance hypothesis gives

$$|Y_j| \geq |X_j| \prod_{\substack{i<s \\ ij \in E(H)}} (d_{ij} - \varepsilon) \geq (\varepsilon^{1/\Delta})^{|\{i<s: ij \in E(H)\}|} |X_j| \geq \varepsilon |X_j|.$$

Thus, as in the proof of Proposition 2.6.1 for K_4, the number of choices $x_s \in X_s$ that would extend $x_1, \dots, x_{s-1}$ to an abundant partial embedding is

$$\begin{aligned} &\geq |Y_s| - |\{i > s : si \in E(H)\}| \, \varepsilon \, |X_s| \\ &\geq |X_s| \prod_{\substack{i<s \\ is \in E(H)}} (d_{ij} - \varepsilon) - |\{i > s : si \in E(H)\}| \, \varepsilon \, |X_s| \,. \end{aligned} \tag{$\dagger$}$$

If none of $1, \dots, s-1$ is a neighbor of s in H, then the first term in ($\dagger$) is $|X_s|$, and so

$$(\dagger) \geq (1 - \Delta\varepsilon) \, |X_s| \,.$$

Otherwise we can absorb the second term into the product and obtain

$$(\dagger) \geq |X_s| \prod_{\substack{i<s \\ is \in E(H)}} (d_{ij} - \varepsilon) - (\Delta - 1)\varepsilon \, |X_s| \geq |X_s| \prod_{\substack{i<s \\ is \in E(H)}} (d_{ij} - \Delta\varepsilon^{1/\Delta}).$$

Fix such a choice of x_s. And now we move onto embedding the next vertex x_{s+1}.

Multiplying together these lower bounds for the number of choices of each x_s over all $s = 1, \dots, v(H)$, we obtain the lower bound on the number of homomorphisms $H \to G$.

Finally, note that in both cases $(\dagger) \ge \varepsilon |X_s|$, and so if $|X_s| \ge v(H)/\varepsilon$, then $(\dagger) \ge v(H)$ and so we can choose each x_s to be distinct from the previously embedded vertices $x_1, \dots, x_{s-1}$, thereby yielding an injective homomorphism. □

Graph Removal Lemma

As an application, we have the following graph removal lemma, generalizing the triangle removal lemma, Theorem 2.3.1. The proof is basically the same as Theorem 2.3.1 except with the above graph counting lemma taking the role of the triangle counting lemma, so we will not repeat the proof here.

Theorem 2.6.5 (Graph removal lemma)
For every graph H and constant $\varepsilon > 0$, there exists a constant $\delta = \delta(H, \varepsilon) > 0$ such that every n-vertex graph G with fewer than $\delta n^{v(H)}$ copies of H can be made H-free by removing fewer than εn^2 edges.

The next exercise asks you to show that, if H is bipartite, then one can prove the H-removal lemma without using regularity, and thereby getting a much better bound.

Exercise 2.6.6 (Removal lemma for bipartite graphs with polynomial bounds)**.** Prove that for every bipartite graph H, there is a constant C such that for every $\varepsilon > 0$, every n-vertex graph with fewer than $\varepsilon^C n^{v(H)}$ copies of H can be made H-free by removing at most εn^2 edges.

Erdős–Stone–Simonovits Theorem

As another application, let us give a different proof of the Erdős–Stone–Simonovits theorem from Section 1.5, restated below, which gives the asymptotics (up to a $+o(n^2)$ error term) for $\mathrm{ex}(n, H)$, the maximum number of edges in an n-vertex H-free graph. We saw a proof in Section 1.5 using supersaturation and the hypergraph KST theorem. The proof below follows the partition-clean-count strategy in Remark 2.3.2 combined with an application of Turán's theorem. A common feature of many regularity applications is that they "boost" an exact extremal graph theoretic result (e.g., Turán's theorem) to an asymptotic result involving more complex derived structures (e.g., from the existence of a copy of K_r to embedding a complete r-partite graph).

Theorem 2.6.7 (Erdős–Stone–Simonovits theorem)
Fix graph H with at least one edge. Then

$$\mathrm{ex}(n, H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right) \frac{n^2}{2}.$$

Proof. Fix $\varepsilon > 0$. Let G be any n-vertex graph with at least $\left(1 - \frac{1}{\chi(H)-1} + \varepsilon\right) \frac{n^2}{2}$ edges. The theorem is equivalent to the claim that for $n = n(\varepsilon, H)$ sufficiently large, G contains H as a subgraph.

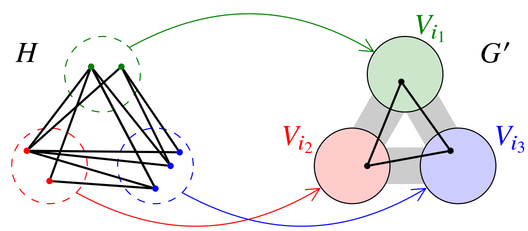
Apply the graph regularity lemma to obtain an η-regular partition $V(G) = V_1 \cup \cdots \cup V_m$

for some sufficiently small $\eta > 0$ only depending on ε and H, to be decided later. Then the number m of parts is also bounded for fixed H and ε.

Remove an edge $(x, y) \in V_i \times V_j$ if

(a) (V_i, V_j) is not η-regular, or
(b) $d(V_i, V_j) < \varepsilon/8$, or
(c) $\min\{|V_i|, |V_j|\} < \varepsilon n/(8m)$.

Then, as in Theorem 2.3.1, the number of edges in (a) is $\le \eta n^2 \le \varepsilon n^2/8$, the number of edges in (b) is $< \varepsilon n^2/8$, and the number of edges in (c) is $< m\varepsilon n^2/(8m) \le \varepsilon n^2/8$. Thus, the total number of edges removed is $\le (3/8)\varepsilon n^2$. After removing all these edges, the resulting graph G' has still has $> \left(1 - \frac{1}{\chi(H)-1} + \frac{\varepsilon}{4}\right)\frac{n^2}{2}$ edges.



By Turán's theorem (Corollary 1.2.6), G' contains a copy of $K_{\chi(H)}$. Suppose that the $\chi(H)$ vertices of this $K_{\chi(H)}$ land in $V_{i_1}, \dots, V_{i_{\chi(H)}}$ (allowing repeated indices). Since each pair of these sets is η-regular, has edge density $\ge \varepsilon/8$, and each has size $\ge \varepsilon n/(8m)$, by applying the graph counting lemma, Theorem 2.6.2, we see that as long as η is sufficiently small in terms of ε and H, and n is sufficiently large, there exists an injective embedding of H into G' where the vertices of H in the rth color class are mapped into V_{i_r}. So G contains H as a subgraph. □

2.7 Exercises on Applying Graph Regularity

The regularity method can be difficult at first to grasp conceptually. The following exercises are useful for gaining familiarity in applying the regularity lemma. For these exercises, you are welcome to use the equitable form of the graph regularity lemma (Theorem 2.1.20), which is more convenient to apply.

Exercise 2.7.1 (Ramsey–Turán)**.**

(a) Show that for every $\varepsilon > 0$, there exists $\delta > 0$ such that every n-vertex K_4-free graph with at least $(\frac{1}{8} + \varepsilon)n^2$ edges contains an independent set of size at least δn.
(b) Show that for every $\varepsilon > 0$, there exists $\delta > 0$ such that every n-vertex K_4-free graph with at least $(\frac{1}{8} - \delta)n^2$ edges and independence number at most δn can be made bipartite by removing at most εn^2 edges.

Exercise 2.7.2 (Nearly homogeneous subset)**.** Show that for every H and $\varepsilon > 0$ there exists $\delta > 0$ such that every graph on n vertices without an induced copy of H contains an induced subgraph on at least δn vertices whose edge density is at most ε or at least $1 - \varepsilon$.

Exercise 2.7.3 (Ramsey numbers of bounded degree graphs). Show that for every Δ there exists a constant C_Δ so that if H is a graph with maximum degree at most Δ, then every 2-edge-coloring of a complete graph on at least $C_\Delta v(H)$ vertices contains a monochromatic copy of H.

Exercise 2.7.4 (Counting H-free graphs).

(a) Show that the number of n-vertex triangle-free graphs is $2^{(1/4+o(1))n^2}$.

(b) More generally, show that for any fixed graph H, the number of n-vertex H-free graphs is $2^{\mathrm{ex}(n,H)+o(n^2)}$.

Exercise 2.7.5* (Induced Ramsey). Show that for every graph H there is some graph G such that if the edges of G are colored with two colors, then some induced subgraph of G is a monochromatic copy of H.

Exercise 2.7.6* (Finding a degree-regular subgraph). Show that for every $\alpha > 0$, there exists $\beta > 0$ such that every graph on n vertices with at least αn^2 edges contains a d-regular subgraph for some $d \geq \beta n$. (Here *d-regular* refers to every vertex having degree d.)

2.8 Induced Graph Removal and Strong Regularity

Recall that H is an ***induced subgraph*** of G if one can obtain H from G by deleting vertices from G (but you are not allowed to simply remove edges from G). We say that G is ***induced H-free*** if G contains no induced subgraph isomorphic to H. (See Notation and Conventions.)

The following removal lemma for induced subgraphs is due to Alon, Fischer, Krivelevich, and Szegedy (2000).

Theorem 2.8.1 (Induced graph removal lemma)

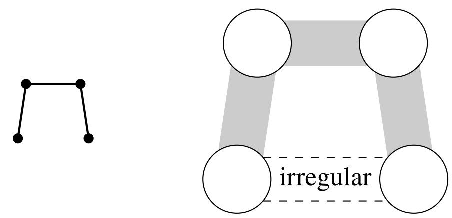
For any graph H and $\varepsilon > 0$, there exists $\delta > 0$ such that if an n-vertex graph has fewer than $\delta n^{v(H)}$ induced copies of H, then it can be made induced H-free by adding and/or deleting fewer than εn^2 edges.

Remark* 2.8.2.** Given two graphs on the same vertex set, the minimum number of edges that one needs to add/delete to obtain the second graph from the first graph is called the ***edit distance between the two graphs. The induced graph removal lemma can be rephrased as saying that every graph with few induced copies of H is close in edit distance to an induced H-free graph.

Unlike the previous graph removal lemma, for the induced version, it is important that we allow both adding and deleting edges. The statement would be false if we only allow edge deletion but not addition. For example, suppose $G = K_n \setminus K_3$ (i.e., a complete graph on n vertices with three edges of a single triangle removed). If H is an empty graph on three vertices, then G has exactly one induced copy of H, but G cannot be made induced H-free by only deleting edges.

To see why the earlier proof of the graph removal lemma (Theorem 2.6.5) does not apply in a straightforward way to prove the induced graph removal lemma, let us attempt to follow the earlier strategy and see where things go wrong.

First we apply the graph regularity lemma. Then we need to *clean* up the graph. In the induced graph removal lemma, edges and nonedges play symmetric roles. We can handle low-density pairs (edge density less than ε) by removing edges between such pairs. Naturally, for the induced graph removal lemma, we also need to handle high-density pairs (density more than $1 - \varepsilon$), and we can add all the edges between such pairs. However, it is not clear what to do with irregular pairs. Earlier, we just removed all edges between irregular pairs. The problem is that this may create many induced copies of H that were not present previously (see illustration below). Likewise, we cannot simply add all edges between irregular pairs.



Perhaps we can always find a regularity partition without irregular pairs? Unfortunately, this is false, as shown in Exercise 2.1.24. One must allow for the possibility of irregular pairs.

Strong Regularity Lemma

We will iterate the regularity partitioning lemma to obtain a stronger form of the regularity lemma. Recall the energy $q(\mathcal{P})$ of a partition (Definition 2.1.10) as the mean-squared edge density between parts.

Theorem 2.8.3 (Strong regularity lemma)
For any sequence of constants $\varepsilon_0 \geq \varepsilon_1 \geq \varepsilon_2 \geq \ldots > 0$, there exists an integer M so that every graph has two vertex partitions $\mathcal{P}$ and $\mathcal{Q}$ so that

(a) $\mathcal{Q}$ refines $\mathcal{P}$,
(b) $\mathcal{P}$ is ε_0-regular and $\mathcal{Q}$ is $\varepsilon_{|\mathcal{P}|}$-regular,
(c) $q(\mathcal{Q}) \leq q(\mathcal{P}) + \varepsilon_0$, and
(d) $|\mathcal{Q}| \leq M$.

***Remark* 2.8.4.** One should think of the sequence $\varepsilon_1, \varepsilon_2, \ldots$ as rapidly decreasing. This strong regularity lemma outputs a refining pair of partitions $\mathcal{P}$ and $\mathcal{Q}$ such that $\mathcal{P}$ is regular, $\mathcal{Q}$ is *extremely* regular, and $\mathcal{P}$ and $\mathcal{Q}$ are close to each other (as captured by $q(\mathcal{P}) \leq q(\mathcal{Q}) \leq q(\mathcal{P}) + \varepsilon_0$; see Lemma 2.8.7 below). A key point here is that we demand $\mathcal{Q}$ to be extremely regular relative to the number of parts of $\mathcal{P}$. The more parts $\mathcal{P}$ has, the more regular $\mathcal{Q}$ should be.

Proof. We repeatedly apply the following version of Szemerédi's regularity lemma:

Theorem 2.1.19 (restated): For all $\varepsilon > 0$ and k, there exists an integer $M_0 = M_0(k, \varepsilon)$ so that for all partitions $\mathcal{P}$ of $V(G)$ with at most k parts, there exists a refinement $\mathcal{P}'$ of $\mathcal{P}$ with each part in $\mathcal{P}$ refined into $\leq M_0$ parts so that $\mathcal{P}'$ is ε-regular.

By iteratively applying the above regularity partition, we obtain a sequence of partitions

$\mathcal{P}_0, \mathcal{P}_1, \ldots$ of $V(G)$ starting with $\mathcal{P}_0 = \{V(G)\}$ being the trivial partition. Each $\mathcal{P}_{i+1}$ is $\varepsilon_{|\mathcal{P}_i|}$-regular and refines $\mathcal{P}_i$. The regularity lemma guarantees that we can have $|\mathcal{P}_{i+1}| \le M_0(|\mathcal{P}_i|, \varepsilon_{|\mathcal{P}_i|})$.

Since $0 \le q(\cdot) \le 1$, there exists $i \le \varepsilon_0^{-1}$ so that $q(\mathcal{P}_{i+1}) \le q(\mathcal{P}_i) + \varepsilon_0$. Then setting $\mathcal{P} = \mathcal{P}_i$ and $\mathcal{Q} = \mathcal{P}_{i+1}$ satisfies the desired requirements. Indeed, the number of parts of $\mathcal{Q}$ is bounded by a function of the sequence $(\varepsilon_0, \varepsilon_1, \ldots)$ since there are a bounded number of iterations and each iteration produced a refining partition with a bounded number of parts. □

Remark* 2.8.5** (Bounds in the strong regularity lemma)**.** The bound on M produced by the proof depends on the sequence $(\varepsilon_0, \varepsilon_1, \ldots)$. In the application below, we use $\varepsilon_i = \varepsilon_0/\text{poly}(i)$. Then the size of M is comparable to applying M_0 to ε_0 in succession $1/\varepsilon_0$ times. Note that M_0 is a tower function, and this makes M a tower function iterated i times. This iterated tower function is called the ***wowzer function:

$$\mathbf{wowzer}(k) := \text{tower}(\text{tower}(\cdots(\text{tower}(2))\cdots))$$

(with k applications of tower). The wowzer function is one step up from the tower function in the Ackermann hierarchy. It grows extremely quickly.

***Remark* 2.8.6** (Equitability)**.** We can further ensure that the parts have nearly equal size. This can be done by adapting the ideas sketched in the proof sketch of Theorem 2.1.20.

The following lemma explains the significance of the inequality $q(\mathcal{Q}) \le q(\mathcal{P}) + \varepsilon$ from earlier.

Lemma 2.8.7 (Energy and approximation)
Let $\mathcal{P}$ and $\mathcal{Q}$ both be vertex partitions of a graph G, with $\mathcal{Q}$ refining $\mathcal{P}$. For each $x \in V(G)$, write V_x for the part of $\mathcal{P}$ that x lies in and W_x for the part of $\mathcal{Q}$ that x lies in. If

$$q(\mathcal{Q}) \le q(\mathcal{P}) + \varepsilon^3,$$

then

$$\left|d(V_x, V_y) - d(W_x, W_y)\right| \le \varepsilon$$

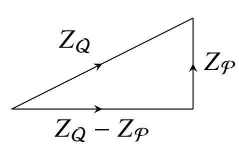
for all but εn^2 pairs $(x, y) \in V(G)^2$.

Proof. Let $x, y \in V(G)$ be chosen uniformly at random. As in the proof of Lemma 2.1.11, we have $q(\mathcal{P}) = \mathbb{E}[Z_{\mathcal{P}}^2]$, where $Z_{\mathcal{P}} = d(V_x, V_y)$. Likewise, $q(\mathcal{Q}) = \mathbb{E}[Z_{\mathcal{Q}}^2]$, where $Z_{\mathcal{Q}} = d(W_x, W_y)$.

We have

$$q(\mathcal{Q}) - q(\mathcal{P}) = \mathbb{E}[Z_{\mathcal{Q}}^2] - \mathbb{E}[Z_{\mathcal{P}}^2] = \mathbb{E}[(Z_{\mathcal{Q}} - Z_{\mathcal{P}})^2],$$

where the final step above is a "Pythagorean identity."



Indeed, the identity $\mathbb{E}[Z_Q^2] - \mathbb{E}[Z_\mathcal{P}^2] = \mathbb{E}[(Z_Q - Z_\mathcal{P})^2]$ is equivalent to $\mathbb{E}[Z_\mathcal{P}(Z_Q - Z_\mathcal{P})] = 0$, which is true since as x and y each vary over their own parts of $\mathcal{P}$, the expression $Z_Q - Z_\mathcal{P}$ averages to zero.

So $q(Q) \le q(\mathcal{P}) + \varepsilon^3$ is equivalent to $\mathbb{E}[(Z_Q - Z_\mathcal{P})^2] \le \varepsilon^3$, which in turn implies, by Markov's inequality, that $\mathbb{P}(|Z_Q - Z_\mathcal{P}| > \varepsilon) \le \varepsilon$, which is the same as the desired conclusion. □

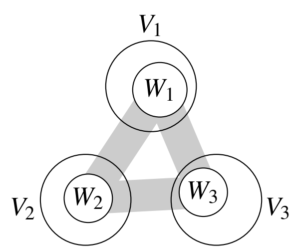
Exercise 2.8.8. Let $0 < \varepsilon < 1$. Using the notation of Lemma 2.8.7, show that if $|d(V_x, V_y) - d(W_x, W_y)| \le \varepsilon$ for all but εn^2 pairs $(x, y) \in V(G)^2$, then $q(Q) \le q(\mathcal{P}) + 2\varepsilon$.

We now deduce the following form of the strong regularity lemma, which considers only select subsets of vertex parts but does not require irregular pairs.

Theorem 2.8.9 (Strong regularity lemma)

For any sequences of constants $\varepsilon_0 \ge \varepsilon_1 \ge \varepsilon_2 \ge \cdots > 0$, there exists a constant $\delta > 0$ so that every n-vertex graph has an equitable vertex partition $V_1 \cup \cdots \cup V_k$ and a subset $W_i \subseteq V_i$ for each i satisfying

(a) $|W_i| \ge \delta n$,
(b) (W_i, W_j) is ε_k-regular for all $1 \le i \le j \le k$, and
(c) $\left|d(V_i, V_j) - d(W_i, W_j)\right| \le \varepsilon_0$ for all but $< \varepsilon_0 k^2$ pairs $(i, j) \in [k]^2$.



***Remark* 2.8.10.** It is significant that *all* (rather than nearly all) pairs (W_i, W_j) are regular. We will need this fact in our applications below.

Proof sketch.. Here we show how to prove a slightly weaker result where $i \le j$ in (b) is replaced by $i < j$. In other words, this proof does not promise that each W_i is ε_k-regular. To obtain the stronger conclusion as stated (requiring each W_i to be regular with itself), we can adapt the ideas in Exercise 2.1.27. We omit the details.

By decreasing the ε_is if needed (we can do this since a smaller sequence of ε_is yields a stronger conclusion), we may assume that $\varepsilon_i \le 1/(10i^2)$ and $\varepsilon_i \le \varepsilon_0/4$ for every $i \ge 1$.

Let us apply the strong regularity lemma, Theorem 2.8.3, with equitable partitions (see above Remark 2.8.6). That is, we have (we make the simplifying assumption that all partitions are exactly equitable, to avoid unimportant technicalities):

- an equitable ε_0-regular partition $\mathcal{P} = \{V_1, \ldots, V_k\}$ of $V(G)$ and
- an equitable ε_k-regular partition Q refining $\mathcal{P}$

satisfying

- $q(Q) \le q(\mathcal{P}) + \varepsilon_0^3/8$, and
- $|Q| \le M = M(\varepsilon_0, \varepsilon_1, \ldots)$.

Inside each part V_i, let us choose a part W_i of Q uniformly at random. Since $|Q| \le M$, the equitability assumption implies that each part of Q has size $\ge \delta n$ for some constant $\delta = \delta(\varepsilon_0, \varepsilon_1, \dots)$. So (a) is satisfied.

Since Q is ε_k-regular, all but at most an ε_k-fraction of pairs of parts of Q are ε_k-regular. Summing over all $i < j$, using linearity of expectations, the expected the number of pairs (W_i, W_j) that are not ε_k-regular is $\le \varepsilon_k k^2 \le 1/10$. It follows that with probability $\ge 9/10$, (W_i, W_j) is ε_k-regular for all $i < j$, so (b) is satisfied (this argument ignores $i = j$ as mentioned at the beginning of the proof).

Let X denote the number of pairs $(i, j) \in [k]^2$ with $\left|d(V_i, V_j) - d(W_i, W_j)\right| > \varepsilon_0$. Since $q(Q) \le q(\mathcal{P}) + (\varepsilon_0/2)^3$, by Lemma 2.8.7 and linearity of expectations, $\mathbb{E}X \le (\varepsilon_0/2)k^2$. So by Markov's inequality, $X \le \varepsilon_0 k^2$ with probability $\ge 1/2$, so that (c) is satisfied.

It follows that (a) and (b) are both satisfied with probability $\ge 1 - 1/10 - 1/2$. Therefore, there exist valid choices of W_is. □

Induced Graph Removal Lemma

As with earlier regularity applications, we follow the partition-clean-count recipe from Remark 2.3.2.

Proof of the induced graph removal lemma (Theorem 2.8.1). Apply Theorem 2.8.9 to obtain a **partition** $V_1 \cup \cdots \cup V_k$ of the vertex set of the graph, along with $W_k \subseteq V_k$, so that:

(a) (W_i, W_j) is ε'-regular for every $i \le j$, with some sufficiently small constant $\varepsilon' > 0$ depending on ε and H,

(b) $\left|d(V_i, V_j) - d(W_i, W_j)\right| \le \varepsilon/8$ for all but $< \varepsilon k^2/8$ pairs $(i, j) \in [k]^2$, and

(c) $|W_i| \ge \delta_0 n$, for some constant δ_0 depending only on ε and H.

Now we **clean** the graph. For each pair $i \le j$ (including $i = j$),

- if $d(W_i, W_j) \le \varepsilon/8$, then remove all edges between (V_i, V_j), and
- if $d(W_i, W_j) \ge 1 - \varepsilon/8$, then add all edges between (V_i, V_j).

Note that we are not simply add/removing edges within each pair (W_i, W_j), but rather all of (V_i, V_j). To bound the number of edges add/deleted, recall (b) from the previous paragraph. If $d(W_i, W_j) \le \varepsilon/8$ and $\left|d(V_i, V_j) - d(W_i, W_j)\right| \le \varepsilon/8$, then $d(V_i, V_j) \le \varepsilon/4$, and the number of edges in all such (V_i, V_j) is at most $\varepsilon n^2/4$. Likewise for $d(W_i, W_j) \ge 1 - \varepsilon/8$. For the remaining $< \varepsilon k^2/8$ pairs (i, j) not satisfying $\left|d(V_i, V_j) - d(W_i, W_j)\right| \le \varepsilon/8$, the total number of edges among all such pairs is at most $\varepsilon n^2/8$. All together, we added/deleted $< \varepsilon n^2$ edges from G. Call the resulting graph G'. There are no irregular pairs (W_i, W_j) for us to worry about.

It remains to show that G' is induced H-free. Suppose otherwise. Let us **count** induced copies of H in G as in the proof of the graph removal lemma, Theorem 2.6.5. We have some induced copy of H in G', with each vertex $v \in V(H)$ embedded in $V_{\phi(v)}$ for some $\phi\colon V(H) \to [k]$.

Consider a pair of distinct vertices u, v of H. If $uv \in E(H)$, there must be an edge in G' between $V_{\phi(u)}$ and $V_{\phi(v)}$ (here $\phi(u)$ and $\phi(v)$ are not necessarily different). So we must not have deleted all the edges in G between $V_{\phi(u)}$ and $V_{\phi(v)}$ in the cleaning step. By the cleaning algorithm above, this means that $d_G(W_i, W_j) > \varepsilon/8$. Likewise, if $uv \notin E(H)$ for any pair of distinct $u, v \in V(H)$, we have $d_G(W_i, W_j) < 1 - \varepsilon/8$.

Since (W_i, W_j) is ε'-regular in G for every $i \leq j$, provided that ε' is small enough (in terms of ε and H), the graph counting lemma, (Theorem 2.6.2 with the induced variation as in Remark 2.6.3(b)) applied to G gives

$$\#\text{ induced copies of } H \text{ in } G \geq (1-\varepsilon)\left(\frac{\varepsilon}{10}\right)^{\binom{v(H)}{2}} (\delta_0 n)^{v(H)} =: \delta n^{v(H)}$$

(recall $|W_i| \geq \delta_0 n$). Setting δ as above, this contradicts the hypothesis that G has $< \delta n^{v(H)}$ copies of H. Thus G' must be induced H-free. □

Infinite Graph Removal Lemma

Finally, let us prove a graph removal lemma with an infinite number of forbidden induced subgraphs (Alon and Shapira 2008). Given a (possibly infinite) set $\mathcal{H}$ of graphs, we say that G is ***induced $\mathcal{H}$-free*** if G is induced H-free for every $H \in \mathcal{H}$.

Theorem 2.8.11 (Infinite graph removal lemma)
For each (possibly infinite) set of graphs $\mathcal{H}$ and $\varepsilon > 0$, there exist h_0 and $\delta > 0$ so that if G is an n-vertex graph with fewer than $\delta n^{v(H)}$ induced copies of H for every $H \in \mathcal{H}$ with at most h_0 vertices, then G can be made induced $\mathcal{H}$-free by adding/removing fewer than εn^2 edges.

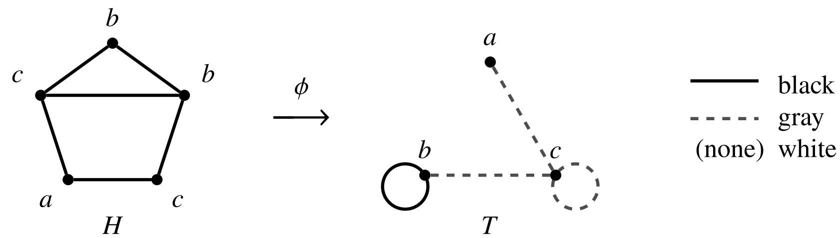
***Remark* 2.8.12.** The presence of h_0 may seem a bit strange at first. In the next section, we will see a reformulation of this theorem in the language of property testing, where h_0 comes up naturally.

Proof. The proof is mostly the same as the proof of the induced graph removal lemma that we just saw. The main tricky issue here is how to choose the regularity parameter ε' for every pair (W_i, W_j) in condition (a) of the earlier proof. Previously, we did not use the full strength of Theorem 2.8.9, which allowed ε' to depend on k, but now we are going to use it. Recall that we had to make sure that this ε' was chosen to be small enough for the H-counting lemma to work. Now that there are possibly infinitely many graphs in $\mathcal{H}$, we cannot naively choose ε' to be sufficiently small. The main point of the proof is to reduce the problem to a finite subset of $\mathcal{H}$ for each k.

Define a ***template*** T to be an edge-coloring of the looped k-clique (i.e., a complete graph on k vertices along with a loop at a every vertex) where each edge is colored by one of {white, black, gray}. We say that a graph H is ***compatible*** with a template T if there exists a map $\phi\colon V(H) \to V(T)$ such that for every distinct pair u, v of vertices of H:

- if $uv \in E(H)$, then $\phi(u)\phi(v)$ is colored black or gray in T; and
- if $uv \notin E(H)$, then $\phi(u)\phi(v)$ is colored white or gray in T.

That is, a black edge in a template means an edge of H, a white edge means a nonedge of H, and a gray edge is a wildcard. An example is shown below.



As another example, every graph is compatible with every completely gray template.

For every template T, pick some ***representative*** $H_T \in \mathcal{H}$ compatible with T, as long as such a representative exists (and ignore T otherwise). A graph in $\mathcal{H}$ is allowed to be the representative of more than one template. Let $\mathcal{H}_k$ be a set of all $H \in \mathcal{H}$ that arise as the representative of some k-vertex template. Note that $\mathcal{H}_k$ is finite since there are finitely many k-vertex templates. We can pick each $\varepsilon_k > 0$ to be small enough so that the conclusion of the counting step later can be guaranteed for all elements of $\mathcal{H}_k$.

Now we proceed nearly identically as in the proof of the induced removal lemma, Theorem 2.8.1, that we just saw. In applying Theorem 2.8.9 to obtain the partition $V_1 \cup \cdots \cup V_k$ and finding $W_i \subseteq V_i$, we ensure the following condition instead of the earlier (a):

(a) (W_i, W_j) is ε_k-regular for every $i \le j$.

We set h_0 to be the maximum number of vertices of a graph in $\mathcal{H}_k$.

Now we do the cleaning step. Along the way, we create a k-vertex template T with vertex set $[k]$ corresponding to the parts $\{V_1, \ldots, V_k\}$ of the partition. For each $1 \le i \le j \le n$,

- if $d(W_i, W_j) \le \varepsilon/4$, then remove all edges between (V_i, V_j) from G, and color the edge ij in template T white;
- if $d(W_i, W_j) \ge 1 - \varepsilon/4$, then add all edges between (V_i, V_j), and color the edge ij in template T black;
- otherwise, color the edge in ij in template T gray.

Finally, suppose some induced $H \in \mathcal{H}$ remains in G'. Due to our cleaning procedure, H must be compatible with the template T. Then the representative $H_T \in \mathcal{H}_k$ of T is a graph on at most h_0 vertices, and furthermore, the counting lemma guarantees that, provided $\varepsilon_k > 0$ is small enough (subject to a finite number of pre-chosen constraints, one for each element of $\mathcal{H}_k$), the number of copies of H_T in G is $\ge \delta n^{v(H_T)}$ for some constant $\delta > 0$ that only depends on ε and $\mathcal{H}$. This contradicts the hypothesis, and thus G' is induced $\mathcal{H}$-free. □

All the techniques above work nearly verbatim for a generalization to colored graphs.

Theorem 2.8.13 (Infinite edge-colored graph removal lemma)

For every $\varepsilon > 0$, positive integer r, and a (possibly infinite) set $\mathcal{H}$ of r-edge-colored graphs, there exists some h_0 and $\delta > 0$ such that if G is an r-edge-coloring of the complete graph on n vertices with $< \delta n^{v(H)}$ copies of H for every $\mathcal{H}$ with at most h_0 vertices, then G can be made $\mathcal{H}$-free by recoloring $< \varepsilon n^2$ edges (using the same palette of r colors throughout).

The induced graph removal lemma corresponds to the special case $r = 2$, with the two colors representing edges and nonedges respectively.

2.9 Graph Property Testing

We are given random query access to a very large graph. The graph may be too large for us to see every vertex or edge. What can we learn about the graph by sampling a constant number of vertices and the edges between them?

For example, we cannot distinguish two graphs if they only differ on a small number of vertices or edges. We also need some error tolerance.

A ***graph property*** $\mathcal{P}$ is simply a set of isomorphism classes of graphs. The graph properties that we usually encounter have some nice name and/or compact description, such as *triangle-free*, *planar*, and 3-*colorable*.

We say that an n-vertex graph G is ***ε-far*** from property $\mathcal{P}$ if one cannot change G into a graph in $\mathcal{P}$ by adding/deleting εn^2 edges.

The following theorem gives a straightforward algorithm, with a probabilistic guarantee, on testing triangle-freeness. It allows us to distinguish two types of graphs from each other:

triangle-free vs. **far from triangle-free**.

Theorem 2.9.1 (Triangle-freeness is testable)
For every $\varepsilon > 0$, there exists $K = K(\varepsilon)$ so that the following algorithm satisfies the probabilistic guarantees below.

Input: A graph G.

Algorithm: Sample K vertices from G uniformly at random without replacement (if G has fewer than K vertices, then return the entire graph). If G has no triangles among these K vertices, then output that G is triangle-free; else output that G is ε-far from triangle-free.

Probabilistic guarantees:

(a) If the input graph G is triangle-free, then the algorithm always correctly outputs that G is triangle-free;

(b) If the input graph G is ε-far from triangle-free, then with probability ≥ 0.99 the algorithm outputs that G is ε-far from triangle-free;

(c) We do not make any guarantees when the input graph is neither triangle-free nor ε-far from triangle-free.

Remark 2.9.2. This is an example of a ***one-sided tester***, meaning that it always outputs a correct answer when G satisfies property $\mathcal{P}$ and only has a probabilistic guarantee when G does not satisfy property G. (In contrast, a two-sided tester would have probabilistic guarantees for both situations.)

For a one-sided tester, there is nothing special about the number 0.99 above in (b). It can be any positive constant $\delta > 0$. If we run the algorithm m times, then the probability of success improves from $\geq \delta$ to $\geq 1 - (1 - \delta)^m$, which can be made arbitrarily close to 1 if we choose m large enough.

The probabilistic guarantee turns out to be essentially a rephrasing of the triangle removal lemma.

Proof. If the graph G is triangle-free, the algorithm clearly always outputs correctly. On the other hand, if G is ε-far from triangle-free, then by the triangle removal lemma (Theorem 2.3.1), G has $\geq \delta\binom{n}{3}$ triangles with some constant $\delta = \delta(\varepsilon) > 0$. If we sample three vertices from G uniformly at random, then then they form a triangle with probability $\geq \delta$. And if run $K/3$ independent trials, then the probability that we see a triangle is $\geq 1 - (1 - \delta)^{K/3}$, which is ≥ 0.99 as long as K is a sufficiently large constant (depending on δ, which in turn depends on ε).

In the algorithm as stated in the theorem, K vertices are sampled without replacement. Above we had K independent trials of picking a triple of vertices at random. But this difference hardly matters. We can couple the two processes by adding additional random vertices to the latter process until we see K distinct vertices. □

Just as the guarantee of the above algorithm is essentially a rephrasing of the triangle removal lemma, other graph removal lemmas can be rephrased as graph property testing theorems. For the infinite induced graph removal lemma, Theorem 2.8.11, we can rephrase the result in terms of graph property testing for hereditary properties.

A graph property $\mathcal{P}$ is ***hereditary*** if it is closed under vertex-deletion: if $G \in \mathcal{P}$, then every induced subgraph of G is in $\mathcal{P}$. Here are some examples of hereditary graph properties: *H-free*, *induced H-free*, *planar*, *3-colorable*, *perfect*. Every hereditary property $\mathcal{P}$ can be characterized as the set of induced $\mathcal{H}$-free graphs for some (possibly infinite) family of graphs $\mathcal{H}$; we can take $\mathcal{H} = \{H : H \notin \mathcal{P}\}$.

Theorem 2.9.3 (Every hereditary graph property is testable)
For every hereditary graph property $\mathcal{P}$, and constant $\varepsilon > 0$, there exists a constant $K = K(\mathcal{P}, \varepsilon)$ so that the following algorithm satisfies the probabilistic guarantees listed below.

Input: A graph G.

Algorithm: Sample K vertices from G uniformly at random without replacement and let H be the induced subgraph on these K vertices. If $H \in \mathcal{P}$, then output that G satisfies $\mathcal{P}$; else output that G is ε-far from $\mathcal{P}$.

Probabilistic guarantees:

(a) If the input graph G satisfies $\mathcal{P}$, then the algorithm always correctly outputs that G satisfies $\mathcal{P}$;

(b) If the input graph G is ε-far from $\mathcal{P}$, then with probability ≥ 0.99 the algorithm outputs that G is ε-far from $\mathcal{P}$;

(c) We do not make any guarantees when the input graph is neither in $\mathcal{P}$ nor ε-far from $\mathcal{P}$.

Proof. If $G \in \mathcal{P}$, then since $\mathcal{P}$ is hereditary, $H \in \mathcal{P}$, and so the algorithm always correctly outputs that $G \in \mathcal{P}$. So suppose G is ε-far from $\mathcal{P}$. Let $\mathcal{H}$ be such that $\mathcal{P}$ is the set of induced $\mathcal{H}$-free graphs. By the infinite induced graph removal lemma, there is some h_0 and $\delta > 0$ so that G has $\geq \delta\binom{n}{v(H)}$ copies of some $H \in \mathcal{H}$ with at most h_0 vertices. So with probability $\geq \delta$, a sample of h_0 vertices sees an induced subgraph not satisfying $\mathcal{P}$. Running K/h_0 independent trials, we see some induced subgraph not satisfying $\mathcal{P}$ with probability $\geq 1 - (1 - \delta)^{K/h_0}$, which can be made arbitrarily close to 1 by choosing K to be sufficiently large. As with earlier, this implies the result about choosing K random points without replacement. □

2.10 Hypergraph Removal and Szemerédi's Theorem

We showed earlier how to deduce Roth's theorem from the triangle removal lemma. However, the graph removal lemma, or the graph regularity method more generally, is insufficient for understanding longer arithmetic progressions.

Szemerédi's theorem follows as a corollary of a hypergraph generalization of the triangle removal lemma. (Note that historically, Szemerédi's theorem was initially shown using other methods; see the discussion in Section 0.2). The hypergraph removal lemma turns out to be substantially more difficult. The following theorem was proved by Rödl et al. (2005) and Gowers (2007). The special case of the tetrahedron removal lemma in 3-graphs was proved earlier by Frankl and Rödl (2002).

Theorem 2.10.1 (Hypergraph removal lemma)

For every r-graph H and $\varepsilon > 0$, there exists $\delta > 0$ so that every n-vertex r-graph with $< \delta n^{v(H)}$ copies of H can be made H-free by removing $< \varepsilon n^r$ edges.

Recall that Szemerédi's theorem says that for every fixed $k \geq 3$, every k-AP-free subset of $[N]$ has size $o(N)$. We will prove it as a corollary of the hypergraph removal lemma for $H = K_k^{(k-1)}$, the complete $(k-1)$-graph on k vertices (also known as a ***simplex***; when $k = 3$ it is called a ***tetrahedron***). For concreteness, we will show how the deduction works in the case $k = 4$ (it is straightforward to generalize).

Here is a corollary of the tetrahedron removal lemma. It is analogous to Corollary 2.3.3.

Corollary 2.10.2

If G is a 3-graph such that every edge is contained in a unique tetrahedron (i.e., a clique on four vertices), then G has $o(n^3)$ edges.

Proof of Szemerédi's theorem for 4-APs. Let $A \subseteq [N]$ be 4-AP-free. Let $M = 6N + 1$. Then A is also a 4-AP-free subset of $\mathbb{Z}/M\mathbb{Z}$ (there are no wraparounds). Build a 4-partite 3-graph G with parts W, X, Y, Z, all of which are M-vertex sets indexed by the elements of $\mathbb{Z}/M\mathbb{Z}$. We define edges as follows, where w, x, y, z range over elements of W, X, Y, Z, respectively:

$$\begin{aligned} wxy \in E(G) &\iff 3w + 2x + y \in A, \\ wxz \in E(G) &\iff 2w + x - z \in A, \\ wyz \in E(G) &\iff w - y - 2z \in A, \\ xyz \in E(G) &\iff -x - 2y - 3z \in A. \end{aligned}$$

What is important here is that the ith expression does not contain the ith variable.

The vertices $xyzw$ form a tetrahedron if and only if

$$3w + 2x + y, 2w + x - z, w - y - 2z, -x - 2y - 3z \in A.$$

However, these values form a 4-AP with common difference $-x - y - z - w$. Since A is

4-AP-free, the only tetrahedra in A are trivial 4-APs (those with common difference zero). For each triple $(w, x, y) \in W \times X \times Y$, there is exactly one $z \in \mathbb{Z}/M\mathbb{Z}$ such that $x+y+z+w = 0$. Thus, every edge of the hypergraph lies in exactly one tetrahedron.

By Corollary 2.10.2, the number of edges in the hypergraph is $o(M^3)$. On the other hand, the number of edges is exactly $4M^2 |A|$ (for example, for every $a \in A$, there are exactly M^2 triples $(w, x, y) \in (\mathbb{Z}/M\mathbb{Z})^3$ with $3w + 2x + y = a$). Therefore $|A| = o(M) = o(N)$. □

The hypergraph removal lemma is proved using a substantial and difficult generalization of the graph regularity method to hypergraphs. We will not be able to prove it in this book. In the next section, we sketch some key ideas in hypergraph regularity.

It is instructive to work out the proof in the special cases below. For the next two exercises, you should assume Corollary 2.10.2.

Exercise 2.10.3 (3-dimensional corners)**.** Suppose $A \subseteq [N]^3$ contains no four points of the form

$$(x, y, z),\ (x + d, y, z),\ (x, y + d, z),\ (x, y, z + d), \quad \text{with } d > 0.$$

Show that $|A| = o(N^3)$.

Exercise 2.10.4 (Multidimensional Szemerédi for axis-aligned squares)**.** Suppose $A \subseteq [N]^2$ contains no four points of the form

$$(x, y),\ (x + d, y),\ (x, y + d),\ (x + d, y + d), \quad \text{with } d \neq 0.$$

Show that $|A| = o(N^2)$.

Exercise 2.10.5 (Multidimensional Szemerédi theorem from the hypergraph removal lemma)**.** Generalizing the previous exercise, prove the multidimensional Szemerédi theorem (Theorem 0.2.6) using the hypergraph removal lemma.

2.11 Hypergraph Regularity

Hypergraph regularity is substantially more difficult to prove than graph regularity. We only sketch some key ideas here. For concreteness, we focus our discussion on 3-graphs. Throughout this section, G will be a 3-graph with vertex set V.

What should correspond to an "ε-regular pair" from the graph regularity lemma? Here is an initial attempt.

Definition 2.11.1 (Initial attempt at 3-graph regularity)

Given vertex subsets $V_1, V_2, V_3 \subseteq V$, we say that (V_1, V_2, V_3) is ***ε-regular*** if, for all $A_i \subseteq V_i$ such that $|A_i| \geq \varepsilon |V_i|$, we have

$$|d(V_1, V_2, V_3) - d(A_1, A_2, A_3)| \leq \varepsilon.$$

Here, the edge density $d(X, Y, Z)$ is the fraction of elements of $X \times Y \times Z$ that are edges of G.

By following the proof of the graph regularity lemma nearly verbatim, we can show the following.

Proposition 2.11.2 (Initial attempt at 3-graph regularity partition)
For all $\varepsilon > 0$, there exists $M = M(\varepsilon)$ such that every 3-graph has a partition into at most M parts so that all but at most an ε-fraction of triples of vertices lie in ε-regular triples of vertex parts.

Can this result be used to prove the hypergraph removal lemma? Unfortunately, no.

Recall that our graph regularity recipe (Remark 2.3.2) involves three steps: partition, clean, and count. It turns out that no counting lemma is possible for the above notion of 3-graph regularity.

The notion of ε-regularity is supposed to model pseudorandomness. So why don't we try truly random hypergraphs and see what happens? Let us consider two different random 3-graph constructions:

(a) First pick constants $p, q \in [0, 1]$. Build a random graph $G^{(2)} = \mathbf{G}(n, p)$, an ordinary Erdős–Rényi graph. Then construct $G^{(3)}$ by including each triangle of $G^{(2)}$ as an edge of $G^{(3)}$ with probability q. Call this 3-graph X.

(b) For each possible edge (i.e. triple of vertices), include the edge with probability $p^3 q$, independent of all other edges. Call this 3-graph Y.

The edge density in both X and Y are close to $p^3 q$, even when restricted to linearly sized triples of vertex subsets. So both graphs satisfy our above notion of ε-regularity with high probability. However, we can compute the tetrahedron densities in both of these graphs and see that they do not match.

The tetrahedron density in X is around q^4 times the K_4 density in the underlying random graph $G^{(2)}$. The K_4 density in $G^{(2)}$ is around p^6. So the tetrahedron density in X is around $p^6 q^4$.

On the other hand, the tetrahedron density in Y is around $(p^3 q)^4$, different from $p^6 q^4$ earlier. So we should not expect a counting lemma with this notion of ε-regularity. (Unless the 3-graph we are counting is linear, as in the exercise below.)

Exercise 2.11.3. Under the notion of 3-graph regularity in Definition 2.11.1, formulate and prove an H-counting lemma for every linear 3-graph H. Here a hypergraph is said to be ***linear*** if every pair of its edges intersects in at most one vertex.

As hinted by the first random hypergraph above, a more useful notion of hypergraph regularity should involve both vertex subsets as well as subsets of vertex-pairs (i.e., an underlying 2-graph).

Given a 3-graph G, a regularity decomposition will consist of

(1) a partition of $\binom{V}{2}$ into 2-graphs $G_1^{(2)} \cup \cdots \cup G_l^{(2)}$ so that G sits in a random-like way on top of most triples of these 2-graphs (we won't try to make it precise), and

(2) a partition of V that gives an extremely regular partition for all 2-graphs $G_1^{(2)}, \ldots, G_l^{(2)}$ (this should be somewhat reminiscent of the strong graph regularity lemma from Section 2.8).

For such a decomposition to be applicable, it should come with a corresponding *counting lemma*.

There are several ways to make the above notions precise. Certain formulations make the regularity partition easier to prove while the counting lemma harder, and some vice versa. The interested readers should consult Rödl et al. (2005), Gowers (2007) (see Gowers (2006) for an exposition of the case of 3-uniform hypergraphs), and Tao (2006) for three different approaches to the hypergraph regularity lemma.

***Remark* 2.11.4** (Quantitative bounds). Whereas the proof of the graph regularity lemma gives tower-type bounds tower($\varepsilon^{-O(1)}$), the proof of the 3-graph regularity lemma has wowzer-type bounds. The 4-graph regularity lemma moves us one more step up in the Ackermann hierarchy (i.e., iterating wowzer), and so on. Just as with the tower-type lower bound (Theorem 2.1.17) for the graph regularity lemma, Ackermann-type bounds are necessary for hypergraph regularity as well (Moshkovitz and Shapira 2019).

Further Reading

For surveys on the graph regularity method and applications, see Komlós and Simonovits (1996) and Komlós, Shokoufandeh, Simonovits, and Szemerédi (2002).

The survey *Graph Removal Lemmas* by Conlon and Fox (2013) discusses many variants, extensions, and proof techniques of graph removal lemmas.

For a well-motivated introduction to the hypergraph regularity lemma, see the article *Quasirandomness, Counting and Regularity for 3-Uniform Hypergraphs* by Gowers (2006).

Chapter Summary

- **Szemerédi's graph regularity lemma.** For every $\varepsilon > 0$, there exists a constant M such that every graph has an ε-regular partition into at most M parts.
  - Proof method: **energy increment**.
- Regularity method recipe: **partition**, **clean**, **count**.
- **Graph counting lemma.** The number of copies of H among ε-regular parts is similar to random.
- **Graph removal lemma.** Fix H. Every n-vertex graph with $o(n^{v(H)})$ copies of H can be made H-free by removing $o(n^2)$ edges.
- **Roth's theorem** can be proved by applying the triangle removal lemma to a graph whose triangles correspond to 3-APs.
- **Szemerédi's theorem** follows from the **hypergraph removal lemma**, whose proof uses the **hypergraph regularity method** (not covered in this book).
- **Induced removal lemma.** Fix H. Every n-vertex graph with $o(n^{v(H)})$ induced copies of H can be made induced H-free by adding/removing $o(n^2)$ edges
  - Proof uses a **strong regularity lemma**, which involves iterating the earlier graph regularity lemma.
- Every hereditary graph property is **testable**.
  - One can distinguish graphs that have property $\mathcal{P}$ from those that are ε-far from property $\mathcal{P}$ (far in the sense of edit distance $\geq \varepsilon n^2$) by sampling a subgraph induced by a constant number of random vertices.
  - The probabilistic guarantee is essentially equivalent to removal lemmas.

18.225 Graph Theory and Additive Combinatorics
Fall 2023