8 - Szemerédi's Regularity Lemma

Jacques Verstraëte
jacques@ucsd.edu

1 Introduction

Szemerédi's Regularity Lemma [18] tells us that every graph can be partitioned into a constant number of sets of vertices in such a way that for most of the pairs of sets in the partition, the bipartite graph of edges between them has many of the properties that one would expect in a random bipartite graph with the same expected edge density. Szemerédi originally used his lemma to prove his celebrated theorem that sets of integers of positive density contain arbitrarily long arithmetic progressions [19]. Since the first version of the regularity lemma, the lemma has been extended a generalized and applied in many different areas of mathematics, in particular, in Green and Tao's proof [13] that the primes contain arbitrarily long arithmetic progressions. The lemma has also been generalized by a number of authors (see Gowers [11]) to hypergraphs.

1.1 Szemerédi's Regularity Lemma

Let A and B be parts of a bipartite graph. The *density* of the pair (A, B) is

$$d(A, B) = \frac{e(A, B)}{|A||B|}.$$

Let $\varepsilon \in [0, 1/2]$. The pair (A, B) is said to be ε*-regular* if for all $X \subset A$ and $Y \subset B$ with $|X| > \varepsilon|A|$ and $|Y| > \varepsilon|B|$,

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

An *ε-regular partition* of a graph G is an equipartition $\{V_1, V_2, \dots, V_k\}$ of $V(G)$ such that (V_i, V_j) is ε-regular for all but at most εk^2 pairs $\{i, j\} \subset [k]$. Szemerédi's Regularity Lemma [18] is as follows:

Theorem 1. (Szemerédi's Regularity Lemma) *Let $\varepsilon \in \mathbb{R}^+$ and $m \in \mathbb{Z}^+$. Then there exists an integer $M = M(\varepsilon, m)$ such that every graph has an ε-regular partition into k parts for some $k : m \leq k \leq M$.*

The proof of the regularity lemma gives a tower-type dependence of $M(\varepsilon, 2)$ on ε, namely $M(\varepsilon, 2)$ is at most a tower of twos of height about ε^{-5}. A breakthrough of Gowers [12] shows that this tower-type dependency is necessary in the regularity lemma.

1.2 Elementary properties of regular pairs

An ε-regular pair (A, B) with density d behaves similarly to a random bipartite graph with expected density d. Here is a summary of elementary property of regular pairs.

Proposition 2. *Let (A, B), (A', B') be ε-regular pairs of density d and $X \subseteq A$.*

1. *If $Y \subset A$ and $|Y| > \varepsilon|A|$, then $|\{x \in B : |N(x) \cap Y| < (d - \varepsilon)|Y|\}| < \varepsilon|B|$.*
2. *At least $(1 - \varepsilon)|A|$ vertices of A have at least $(d - \varepsilon)|B|$ neighbors in B.*
3. *$(A, B \cup B')$ is ε-regular of density at least d.*
4. *If $|X| > \delta|A|$, then (X, B) is $\varepsilon(1 + \frac{1}{\delta})$-regular of density at least $d - \varepsilon$.*

Proof. For the first statement, let $D = \{x \in B : |\Gamma(x) \cap Y| < (d - \varepsilon)|Y|\}$. Then $d(B, D) < (d - \varepsilon)$. By ε-regularity, therefore, $|D| < \varepsilon|B|$, as required. The second statement follows from the first with $Y = A$. The last two statements are exercises from the definition of ε-regular pairs. □

1.3 Cluster graphs

Let G be a graph and $\mathcal{P} = \{V_1, V_2, \dots, V_k\}$ a partition of $V(G)$. Then the *cluster graph with threshold d* is the graph $R = R_d(G, \mathcal{P})$ defined by $V(R) = [k]$ and

$$E(R) = \{\{i, j\} : (V_i, V_j) \text{ is } \varepsilon\text{-regular with } d(V_i, V_j) \geq d\}.$$

Theorem 1 can be rephrased as saying that for $\varepsilon > 0$, every graph has a cluster graph with $k = O_\varepsilon(1)$ vertices and at least $(1-\varepsilon)k^2$ edges. In applications, we choose the threshold $d > \varepsilon$ appropriately. We give a basic example of the use of the cluster graph.

Proposition 3. *For all $\gamma > 0$, there is $\delta > 0$ such that any n-vertex graph G with*

$$e(G) > (\tfrac{1}{4} + \gamma)n^2$$

contains an edge in at least δn triangles.

Proof. Let $\varepsilon = \frac{1}{3}\gamma$ and let $\mathcal{P} = \{V_1, V_2, \dots, V_k\}$ be an ε-regular partition of G, where $k : \varepsilon^{-1/2} \le k \le M$, guaranteed by Theorem 1, Let $R = R_{2\varepsilon}(G, \mathcal{P})$ and $\delta = \frac{\varepsilon^2}{M}$.

If R contains a triangle, say with vertex set $\{1,2,3\}$, then (V_i, V_j) is ε-regular with density at least 2ε, for $\{i,j\} \subset [3]$. By Lemma 2.2, for $i \neq j$, at least $(1-\varepsilon)|V_i|$ vertices of V_i have more than $\varepsilon|V_j|$ neighbors in V_j. Since $\varepsilon < \frac{1}{2}$, there exists a vertex $x \in V_1$ with more than $\varepsilon|V_i|$ neighbors in V_i for $i \in \{2,3\}$. Let $X = N(x) \cap V_2$ and $Y = N(x) \cap V_2$. Since (V_2, V_3) is ε-regular of density at least 2ε, $|d(X,Y) - d(V_2,V_3)| < \varepsilon$, which implies $d(X,Y) > \varepsilon$. In particular, there are at least $\varepsilon|X||Y|$ edges between X and Y, and therefore some $y \in Y$ has at least $\varepsilon|X| > \varepsilon^2|V_2|$ neighbors in X. The edge $\{x,y\}$ is therefore in at least $\varepsilon^2|V_2| \ge \varepsilon^2\frac{n}{M} = \delta n$ triangles.

If R contains no triangle, then by Mantel's Theorem, $e(R) \le \frac{1}{4}k^2$. For simplicity, assume $|V_i| = m = \frac{n}{k}$ for all i. We have

$$e(G) = \sum_{i=1}^{k} e(V_i) + \sum_{\{i,j\} \in R} e(V_i, V_j) + \sum_{\{i,j\} \notin R} e(V_i, V_j).$$

The last sum is at most $\varepsilon k^2 m^2 + 2\varepsilon\binom{k}{2}m^2 < 2\varepsilon n^2$, since there are at most εk^2 irregular pairs (V_i, V_j), each contributing at most $|V_i||V_j|$ edges, and at most $\binom{k}{2}$ pairs (V_i, V_j) of density at most 2ε, each contributing at most $2\varepsilon|V_i||V_j|$ edges. Also, $e(V_i) < |V_i|^2 < \varepsilon n^2$ since $k \ge \varepsilon^{-1/2}$. Therefore

$$e(G) < 3\varepsilon n^2 + \sum_{\{i,j\} \in R} e(V_i, V_j).$$

The last sum is at most

$$e(R) \cdot \left(\frac{n}{k}\right)^2 < \tfrac{1}{4}n^2.$$

We conclude since $\varepsilon = \frac{1}{3}\gamma$,

$$e(G) < \tfrac{1}{4}n^2 + 3\varepsilon n^2 = (\tfrac{1}{4} + \gamma)n^2.$$

This contradiction completes the proof. □

2 Applications

The regularity lemma is useful for counting embeddings of graphs in dense graphs. An embedding of a graph G into a graph H is a graph isomorphism from H to a subgraph of G isomorphic to H. In general, we write $H \to G$ for an embedding of H into G. We are interested in determining $\|H \to G\|$, the number of isomorphic copies of H in G. We need the following notation. Let $G(t)$ denote the inflation of t: i.e. replace vertices with independent sets of t vertices, called clusters and a complete bipartite graph between t-sets whenever the corresponding vertices were joined in G.

Lemma 4. (Embedding Lemma) *Let R be a graph on $\{1, 2, \ldots, k\}$ and let H be a subgraph of $R(t)$ of maximum degree Δ. Let $d, \varepsilon \in \mathbb{R}^+$, and $\ell \in \mathbb{Z}^+$ satisfy*

$$\begin{aligned} \varepsilon < \eta &= \frac{(d-\varepsilon)^\Delta}{\Delta+2} \\ (t-1) &\leq \eta\ell. \end{aligned}$$

Let G be the graph G obtained by replacing $V(R)$ with disjoint sets $V_1, \ldots, V_k$ of size ℓ, and an ε-regular pair (V_i, V_j) of density d for all edges $\{i, j\} \in R$. Then

$$\|H \to G\| \geq (\eta\ell)^{|V(H)|}.$$

Proof. For convenience, let $V(H) = \{x_1, x_2, \ldots, x_h\}$, let the clusters of $R(t)$ be $U_1, U_2, \ldots, U_k$, and $\delta = (d - \varepsilon)$. Let $\psi : \{1, 2, \ldots, h\} \to \{1, 2, \ldots, k\}$ be defined by $\psi(j) = i$ iff $x_j \in U_i$. Initially, let $C_{1,j} = V_{\psi(j)}$ for $j = 1, 2, \ldots, k$. Suppose we have defined $C_{i,j}$ for $j \geq i \geq 1$. Define $C_{i+1,i+1}$ as follows: choose $v_i \in C_{i,i}$ with

$$d(v_i, C_{i,j}) > \delta|C_{i,j}|$$

for every $j > i$ such that $x_i x_j \in E(H)$. Let $C_{i+1,j} = C_{i,j} \cap N(v_i)$ if $\{x_i, x_j\} \in E(H)$ and $C_{i+1,j} = C_{i,j}$ otherwise. Then $|C_{i,j}|$ is reduced by a factor δ at most Δ times, since H has degree at most Δ. If $|C_{i,i}| \geq \varepsilon\ell$ for all i, then Lemma 2.1 shows that for all i there is a set of at least $(1-\varepsilon)^\Delta \ell \geq \ell - \Delta\varepsilon\ell$ choices of $v_i \in V_{\psi(i)}$. The number of these choices which are in $C_{i,i}$ is at least $|C_{i,i}| - \Delta\varepsilon\ell - (t-1)$. Since $|C_{i,i}| \geq \delta^\Delta$, it is enough that

$$\delta^\Delta - \Delta\varepsilon\ell - (t-1) \geq \eta\ell$$

to obtain $\|H \to G\| \geq (\eta\ell)^h$. This is satisfied under the conditions $t - 1 < \eta\ell$ and $\varepsilon < \eta$. □

The graph G in Lemma 4 is of density roughly d. In the application of the lemma, we usually start with the graph G, find an ε-regular partition using the regularity lemma, form the cluster graph R and the graph $R(t)$, and then lift back the structure of $R(t)$ to G using the embedding lemma. The Erdős-Stone Theorem [5] is a consequence of this procedure.

Theorem 5. *For $\gamma \in \mathbb{R}^+$ and $r, t \in \mathbb{Z}^+$, there exists $n_0(r,t,\gamma)$ such that if G is any n-vertex graph with at least $(1 - \frac{1}{r-1} + \gamma)\binom{n}{2}$ edges and $n > n_0(r,t,\gamma)$ vertices, then G contains $K_r(t)$.*

This is proved in a very similar way to Proposition 3: if an appropriate cluster graph of G contains K_r, then apply Lemma 4 to find $K_r(t)$ in G, otherwise, count edges as in Proposition 3. This proof of Erdős-Stone gives a very poor $n_0(r,t,\gamma)$, due to the tower-type dependency in the function $M(\varepsilon, m)$ on ε in Theorem 1. By contrast, using direct methods we obtained the optimal $t = \Theta(\log n)$.

2.1 Half jumbled graphs

Let $\delta > 0$. An n-vertex graph G of density p is *δ-half-jumbled* if whenever $X, Y \subset V(G)$ are disjoint,

$$e(X,Y) \leq (1+\delta)p|X||Y|.$$

If G is a random n-vertex graph with edge-probability p, then for any $X \subseteq V(G)$, $\mathbb{E}(e(X)) = p\binom{|X|}{2}$, so the above expression measures in some sense the deviation over the expected number of edges. The following result shows that we can find a K_r in half-jumbled graphs of arbitrarily low density, as opposed To Turán's Theorem, which requires the density to be at least $1 - \frac{1}{r-1}$.

Theorem 6. *For each $r \geq 3$ and $p, q > 0$, if G is an n-vertex δ-half-jumbled graph where $\delta = \frac{1-q}{r-2}$, then G contains a K_r.*

Proof. Fix $\varepsilon = \frac{pq}{8(r-2)}$. Let R be a cluster graph of G with threshold 2ε from an ε-regular partition $\{V_1, V_2, \ldots, V_k\}$, and suppose $V_1, V_2, \ldots, V_k$ all have size $\frac{n}{k}$. If $K_r \subset R$, then we apply the embedding lemma. Otherwise,

$$e(R) \leq \frac{1}{2}\Big(1 - \frac{1}{r-1}\Big)k^2$$

by Turán's Theorem. There are at most $4\varepsilon n^2$ edges of G that are not in regular pairs of density at least 2ε – see Proposition 3 for similar details. So there are at least $(p - 8\varepsilon)\binom{n}{2}$ edges in pairs (V_i, V_j) with $\{i, j\} \in R$. However, since and G is δ-half-jumbled, $e(V_i, V_j) \leq (1 + \delta)p|V_i||V_j|$, and therefore

$$\sum_{\{i,j\} \in R} p(1+\delta)|V_i||V_j| \geq (p - 8\varepsilon)\binom{n}{2}.$$

The upper bound on $e(R)$ gives

$$1 - \frac{1}{r-1} > \frac{p - 8\varepsilon}{p(1+\delta)}.$$

Since $\varepsilon = \frac{pq}{8(r-2)}$ and $\delta = \frac{1-q}{r-2}$, this is a contradiction. □

In fact, the above proof combined with the embedding lemma shows that the graph G in the theorem contains $\Omega_p(n^r)$ copies of K_r. Also, this result says that for any $\delta < 1$, if G is an n-vertex δ-half-jumbled triangle-free graph then $e(G) = o(n^2)$. Jumbled graphs come up in the context of pseudorandom graphs, a topic to which we return shortly.

2.2 Bounded degree graph Ramsey Numbers

For any graph G, the Ramsey number $r(G)$ is the maximum number n such that there exists an edge-colouring of the complete graph K_n in which no subgraph $G \subset K_n$ is monochromatic. In this section, we present a result of Chvátal, Rödl, Szemerédi and Trotter [4]:

Theorem 7. *Let $d > 0$ and let G be a graph of maximum degree at most d. Then there exists a constant $c = c(d)$ such that $r(G) \leq c(d)|G|$.*

Proof. Colour K_n red and blue, and let H denote the red graph. Take an ε-regular partition $(V_1, V_2 \dots, V_k)$, and let R be the graph on $\{v_1, v_2, \dots, v_k\}$ formed by joining v_i to v_j if (V_i, V_j) is ε-regular. Since the partition was ε-regular, R has at least $(1-\varepsilon)\binom{k}{2}$ edges and, if $\varepsilon < 1/r - 1$, contains K_r, by Turán's Theorem. We will choose $r = R(K_{d+1}, K_{d+1})$, the Ramsey number for K_{d+1}. We now recolour the edges of $K_r \subset R$ green or yellow, according as the density of the corresponding pair (V_i, V_j) is at least or more than $1/2$. By the choice of $r = R(d+1, d+1)$, there is a monochromatic K_{d+1} or monochromatic K_{d+1} in this coloured $K_r \subset R$. By the embedding lemma, applied with $d = 1/2$ and $t = 1$, this means that either the blue graph or the red graph contains a K_r. □

The above proof gives weak bounds on $c(d)$, namely $c(d)$ is at most a tower of twos of height polynomial in d. A stronger approach to the problem was presented before using Dependent Random Choice, which gave $c(d) \leq 4^d$; however the above proof is a good illustration of the use of regularity.

2.3 The Removal Lemma

One of the powerful consequences of the regularity lemma is the *Removal Lemma*:

Lemma 8. *Let G be an n-vertex graph containing $o(n^k)$ complete graphs of order k. Then $o(n^2)$ edges may be deleted from G to obtain a graph containing no complete subgraph of order k.*

Proof. Consider the cluster graph arising from an ε-regular partition with M parts, consisting of pairs of density at least δ. If the cluster graph contains a complete graph of order k, then as in the embedding lemma, we find $f(\varepsilon, \delta)n^k$ complete graphs of order k in G provided δ is large enough relative to ε, which is a contradiction. Therefore the cluster graph does not contain a complete graph of order k amongst ε-regular pairs of density at least δ. Delete all edges of G lying in parts of the regular partition, or in irregular pairs or in pairs of density less than δ. In total, at most $(\delta + \varepsilon)n^2 + n^2/M$ edges have been deleted from G. We are left with a graph which has no complete graph of order k since the cluster graph of corresponding pairs has no complete graph of order k. This is valid for any $\varepsilon > 0$, and therefore $o(n^2)$ edges may be deleted from G to get a graph with no complete subgraph of order k. □

The explicit dependency of the number of edges deleted relative to the density of the host graph G has been examined by Fox [7], who recently obtained the best bounds.

2.3.1 Arithmetic Progressions

The *upper density* of a set A of integers is

$$\limsup_{n\to\infty} \frac{|A \cap [N]|}{N}.$$

Erdős and Turán [6] conjectured that any set of integers of positive upper density contains a k-term arithmetic progression. A famous theorem of Roth [14] states that this is true when $k = 3$, Szemerédi [19] and Fürstenburg [9] gave independent proofs of the full conjecture of Erdős and Turán. The current best bounds on the size of a set $A \subset [N]$ with no three-term progression are

$$|A| \leq \frac{N}{(\log N)^{1-o(1)}}$$

due to Sanders [16], and come from a culmination of ideas in Discrete Fourier Analysis, which go back to the original theorem of Roth. Green [12] was the first to prove a version of Roth's Theorem in the primes, namely that every subset of the primes of positive relative density contains a 3-term progression. Since the primes up to N have counting function $\pi(N) \sim N/\log N$, by the prime number theorem, the bound on $|A|$ due to Sanders falls just short of the primes. The best lower bound comes from the Behrend construction [2] (see Notes Part 6), namely $|A| = N \exp(-O(\sqrt{\log N}))$. Szemerédi [19] used the regularity lemma to prove his theorem. We show how to get the case $k = 3$ from the regularity lemma, as done by Ruzsa and Szemerédi [15].

Theorem 9. *A set $A \subset [N]$ without three-term progressions satisfies $|A| = o(N)$.*

Proof. Let G_A be the 3-partite graph with parts $X = [N]$, $Y = [2N]$ and $Z = [3N]$ in which the edge set is

$$\{\{x, x+a\}, \{x+a, x+2a\}, \{x, x+2a\} : x \in [n], a \in A\}.$$

This graph has exactly $3|A|N$ edges, and as we have verified before, G_A comprises a union of edge-disjoint triangle with vertices $\{x, x+a, x+2a\}$ and these are the only triangles in G_A. By the Removal Lemma, we can delete $o(N^2)$ edges of G_A

to destroy all the triangles. Since the triangles are all edge-disjoint, there must be $o(N^2)$ of them, which shows $N|A| = o(N^2)$, and hence $|A| = o(N)$. □

The version of the removal lemma due to Fox [7] gives better bounds on $|A|$ than those from the regularity lemma directly. However, the bound still looks like $O(n/\log^* n)$ at best. The heart of the proof is in creating the 3-partite graph G_A, which can be viewed as a hypergraph H_A of triples $\{x, x+a, x+2a\}$. The following is sometimes referred to as the $(6,3)$-Theorem. An (s,t)-*configuration* in a 3-graph H is a set of t edges whose union has at most s vertices.

Theorem 10. *Let H be an n-vertex 3-graph without a $(6,3)$-configuration. Then $e(H) = o(n^2)$.*

Let $\log^{(k)}(n)$ be the k-fold iterate of $\log n$, so $\log^{(2)}(n) = \log\log n$ for instance. It is an open problem to determine whether there exists k such that the maximum number of edges in an n-vertex 3-graph with no $(6,3)$-configuration is $O(n/\log^{(k)}(n))$; as we mentioned before the regularity lemma only gives $O(n/\log^* n)$.

The situation for k-term progressions with $k > 3$ is substantially more difficult. The following celebrated theorem was proved by Szemerédi [19]:

Theorem 11. (Szemerédi's Theorem) *Any set $A \subset [N]$ not containing a k-term progression satisfies $|A| = o(N)$.*

Green's version of Roth's Theorem [12] in the primes was then generalized by Green and Tao [13] to k-term progressions in the primes, and many quantitative bounds on the number of k-term progressions and other arithmetic structures have been found subsequent to their result.

Theorem 12. (Green-Tao Theorem) *If $A \subset [N]$ is a set of prime numbers containing no k-term progression, then $|A| = o(\pi(N))$.*

Without doing this theorem justice, we mention that one idea is to show that the primes are sufficiently randomly distributed to ensure that Szemerédi's Lemma can be transferred to sets of positive relative density in the primes. It turns out the k-term progression problem can also be reduced as above to an extremal problem on graphs or hypergraphs, but unfortunately there is no known way to obtain an easy reduction to which the regularity lemma can be applied effectively to prove the theorem. In the case of 4-term progressions, it is straightforward to verify that if the graph G_A in the proof above contains four distinct triangles covering at most seven vertices, then A contains a 4-term arithmetic progression. It is an open question of Brown, Erdős and Sós [3] to show $e(G_A) = o(N^2)$ in this case. This is known as the $(7,4)$-problem, and can be stated in hypergraph language as follows:

Conjecture 13. *If H is a 3-graph on n vertices without a $(7,4)$-configuration, then $e(H) = o(n^2)$.*

In fact, it is conjectured [3] that for $k \geq 3$, any n-vertex 3-graph with no $(k, k-3)$-configuration has $o(n^2)$ edges.

2.3.2 Corners

Fill an $n \times n$ matrix with zeros and ones. A *corner* consists of one positions $(i,j), (i \pm k, j), (i, j \pm k)$ in the matrix. How many ones can we put in an $n \times n$ matrix to avoid corners? This seemingly simple problem is actually very difficult. However, using a similar approach to the progressions we see that a matrix with no corner has $o(n^2)$ ones:

Theorem 14. *If A is an $n \times n$ zero-one matrix with no corners, then A contains $o(n^2)$ ones.*

Proof. Let V be the set of all n horizontal and n vertical lines of the matrix $R \cup C$, together with all $2n-1$ forty-five degree diagonals L of the matrix. Form a graph on V in which $r \in R$ is joined to $c \in C$ if position (r,c) is one, join $c \in C$ to $l \in L$ if the intersection of L and C is a one position, and finally do the same for r and l. The only triangles in this graph have vertex set $\{r, c, l\}$ such that r, c and l intersect in a common position, otherwise a corner is produced. Therefore there are exactly n^2 triangles, and they are all edge-disjoint. Applying the Removal Lemma, we get that the graph has $o(n^2)$ edges, and hence there are $o(n^2)$ non-zero one positions. □

Once again, it is an open problem to determine whether there exists k such that the maximum number of ones in a corner-free $n \times n$ matrix is $O(n/\log^{(k)}(n))$; the regularity lemma only gives $O(n/\log^* n)$.

3 The Proof of the regularity lemma

Start with the graph G, and a given $\varepsilon > 0$. Let $\mathcal{P}_0 = \{V_1, V_2, \dots, V_l\}$ be some equipartition of $V(G)$. At any stage, we have a equipartition $\mathcal{P}_i$ and a measure $q(\mathcal{P}_i)$ of the quality of $\mathcal{P}_i$ relative to ε-regularity, where $0 \le q(\mathcal{P}) \le 1$ for any partition $\mathcal{P}$. If $\mathcal{P}_i$ is ε-regular with at most N parts, we are done, otherwise we seek to refine $\mathcal{P}_i$ to an equipartition $\mathcal{P}_{i+1}$ with $q(\mathcal{P}_{i+1}) \ge q(\mathcal{P}_i) + \varepsilon^5$ such that if $p_i = |\mathcal{P}_i|$, then $p_{i+1} \le p_i 4^{p_i}$. Then after $t \le \varepsilon^{-5}$ steps, we must arrive at an ε-regular partition $\mathcal{P}_t$ such that $|\mathcal{P}_t| \le f \circ f \circ \dots \circ f(l)$ (composition t times) where $f(x) = x4^x$.

3.1 Energy functions

The key is in defining the quantity $q(\mathcal{P})$. For $A, B \in \mathcal{P}$ define the *index*

$$q(A, B) = \frac{1}{n^2} d(A, B)^2 |A||B|.$$

If $\mathcal{A}$ and $\mathcal{B}$ are partitions of A and B let

$$q(\mathcal{A}, \mathcal{B}) = \sum_{X \in \mathcal{A}} \sum_{Y \in \mathcal{B}} q(X, Y).$$

Then *energy function* of $\mathcal{P}$ is

$$q(\mathcal{P}) = \sum_{\{A,B\} \subset \mathcal{P}} q(A, B).$$

It is easy to see that $q(\mathcal{P}) \le 1$ for any partition $\mathcal{P}$. A key property of the energy function is the following simple lemma:

Lemma 15. *If $\mathcal{Q}$ refines $\mathcal{P}$ then $q(\mathcal{Q}) \le q(\mathcal{P})$.*

Proof. The Cauchy-Schwarz Inequality states that for real x_i, y_i, $\sum x_i^2 \sum y_i^2 \ge \sum x_i^2 y_i^2$. If a_i, b_i are positive reals, then setting $x_i = a_i/\sqrt{b_i}$ and $y_i = \sqrt{b_i}$ we obtain $\sum \frac{a_i^2}{b_i} \ge \left(\sum a_i\right)^2 / \sum b_i$. Now for each $A \in \mathcal{P}$, there is a partition $\mathcal{Q}_A$ of A comprising the parts of $\mathcal{Q}$ in A. Fixing $A, B \in \mathcal{P}$, enumerate the pairs in $\mathcal{Q}_A \times \mathcal{Q}_B$ and let a_i be the number of edges in the ith pair and let b_i be the product of the sizes of the parts of the ith pair. Then the index of the ith pair is exactly $a_i^2/n^2 b_i$. By the Cauchy-Schwarz inequality in the form above,

$$q(\mathcal{Q}_A, \mathcal{Q}_B) = \frac{1}{n^2} \sum_{i=1}^{m} \frac{a_i^2}{b_i} \ge \frac{1}{n^2} \frac{\left(\sum_{i=1}^m a_i\right)^2}{\sum_{i=1}^m b_i} = \frac{e(A, B)^2}{|A||B|n^2} = q(A, B).$$

Summing over all A, B in $\mathcal{P}$, we obtain $q(\mathcal{Q}) \geq q(\mathcal{P})$. This proves the lemma. $\square$

3.2 Refining a partition

We would like the energy function to strictly increase by ε^5 by refining $\mathcal{P}$ to a partition $\mathcal{Q}$ if $\mathcal{P}$ is not ε-regular. This is done by focusing on a single irregular pair (A, B) in $\mathcal{P}$, and then repeatedly refining $\mathcal{P}$ each time we encounter an irregular pair.

Lemma 16. *If (A, B) is not ε-regular, then there exist partitions $\mathcal{A}$ and $\mathcal{B}$ of A and B into two parts each such that*

$$n^2 q(\mathcal{A}, \mathcal{B}) > n^2 q(A, B) + \varepsilon^4 |A||B|.$$

Proof. There exist $A_1 \subset A, B_1 \subset B$ with $|A_1| > \varepsilon|A|$ and $|B_1| > \varepsilon|B|$ and $\delta = |d(A_1, B_1) - d(A, B)| > \varepsilon$. Let $\mathcal{A} = (A_1, A_2)$ and $\mathcal{B} = (B_1, B_2)$ where $A_2 = A \backslash A_1$ and $B_2 = B \backslash B_1$. Let $a_{i+2j-2} = e(A_i, B_j)$ and $b_{i+2j-2} = |A_i||B_j|$. Let $a = e(A, B)$ and $b = |A||B|$. By the Cauchy-Schwarz inequality $\sum_{i \geq 2} a_i^2 / b_i \geq (\sum_{i \geq 2} a_i)^2 / \sum_{i \geq 2} b_i$,

$$\begin{aligned} n^2 q(\mathcal{A}, \mathcal{B}) &= \sum_{i=1}^{4} \frac{a_i^2}{b_i} \\ &\geq \frac{a_1^2}{b_1} + \frac{(a_2 + a_3 + a_4)^2}{b_1 + b_2 + b_3} \\ &= \frac{a_1^2}{b_1} + \frac{(a - a_1)^2}{b - b_1}. \end{aligned}$$

A calculation shows this is

$$\frac{a^2}{b} + \delta^2 b_1 + \frac{(a_1 b - a b_1)^2}{b^2 (b - b_1)} \geq \frac{a^2}{b} + \delta^2 b_1.$$

Since $\delta > \varepsilon$ and $b_1 > \varepsilon^2 b$, we obtain

$$n^2 q(\mathcal{A}, \mathcal{B}) > n^2 q(A, B) + \varepsilon^4 |A||B|.$$

This proves the lemma. $\square$

3.3 Proof of Theorem 1

Let $\mathcal{P}_0$ be an arbitrary equipartition of $V(G)$ into two parts and set $\ell = \max\{m, 2\varepsilon^{-5}\}$. Let $f(x) = x4^x$ and let $f^i(x)$ denote the i-fold iterate of $f(x)$ starting with $f^0(2) = \ell$.

Having produced an equipartition $\mathcal{P}_i$ of $V(G)$ with at most $f^i(2)$ parts, we show that either $\mathcal{P}_i$ is ε-regular, or there is an equipartition $\mathcal{P}_{i+1}$ of $\mathcal{P}_i$ such that $\mathcal{P}_{i+1}$ has at most $f^{i+1}(2)$ parts and $q(\mathcal{P}_{i+1}) > q(\mathcal{P}_i) + \frac{1}{\ell}$. This is enough to finish the proof, since $q(\mathcal{P}) \leq 1$ for every partition $\mathcal{P}$, and therefore the process terminates in at most ℓ steps.

Suppose $\mathcal{P} = \mathcal{P}_i$ is not ε-regular, and has $k \leq f^i(2)$ parts. For simplicity, suppose all parts in $\mathcal{P}$ have size exactly n/k. For each irregular pair (A,B) in $\mathcal{P}$, there exists a partition $\mathcal{P}_A$ of A and $\mathcal{P}_B$ of B such that $n^2 q(\mathcal{P}_A, \mathcal{P}_B) > n^2 q(A,B) + \varepsilon^4|A||B|$. Since $\mathcal{P}$ is not ε-regular, there are at least εk^2 irregular pairs. Let $\mathcal{Q}$ be the coarsest common refinement of all the partitions $\mathcal{P}_A$ for A in $\mathcal{P}$. Then by Lemma 16,

$$\begin{aligned} q(\mathcal{Q}) &= \sum_{\{A,B\}\subset\mathcal{P}} q(\mathcal{P}_A,\mathcal{P}_B) \\ &\geq \sum_{\{A,B\}\subset\mathcal{P}} q(A,B) + \varepsilon k^2 \cdot \frac{1}{n^2} \sum_{\{A,B\}\subset\mathcal{P}} \varepsilon^4|A||B| \quad > \quad q(\mathcal{P}) + \varepsilon^5. \end{aligned}$$

Let $\mathcal{R}$ be a refinement of $\mathcal{Q}$ whose parts are each either singletons or a set of size $\lceil n/k4^k \rceil$, such that as many parts $\mathcal{R}$ have size exactly $\lceil n/k4^k \rceil$. Now $\mathcal{Q}$ has at most 2^k parts contained in each $A \in \mathcal{P}$. Then $q(\mathcal{R}) \geq q(\mathcal{Q})$ by Lemma 15, and the number of singletons in $\mathcal{R}$ is at most $n/2^{k-1}$. Let $\mathcal{S}$ be the equipartition of $V(G)$ obtained by distributing the singletons over the parts of $\mathcal{R}$ as evenly as possible. Since there are at most $n/2^{k-1}$ singletons in $\mathcal{R}$, $q(\mathcal{S}) \geq q(\mathcal{R}) - 2^{1-k} \geq q(\mathcal{P}) + \frac{1}{\ell}$ since we easily have $2^{1-k} \geq \ell$. Then $\mathcal{S}$ is the desired partition $\mathcal{P}_{i+1}$. □

4 Scott's Regularity Lemma

The regularity lemma as stated in Theorem 1 is not effective for n-vertex graphs with $o(n^2)$ edges –sparse graphs – since for any $\varepsilon > 0$, every $n \times n$ bipartite graph with $o(n^2)$ edges is ε-regular. An effective regularity lemma for sparse graphs was recently discovered by Scott [17].

If G is an n-vertex graph of density p, then pair (A,B) of disjoint subsets of $V(G)$ is said to be *relatively ε-regular* if for all $X \subset A$ and $Y \subset B$ with $|X| > \varepsilon|A|$ and $|Y| > \varepsilon|B|$, $|d(A,B) - d(X,Y)| \leq p\varepsilon$. Note that if p is small, possibly depending on n, then this is much stronger than ε-regularity. A *relatively ε-regular partition* is an equipartition $\{V_1, V_2, \ldots, V_k\}$ of $V(G)$ such that (V_i, V_j) is relatively ε-regular for all but εk^2 pairs $\{i,j\} \subset [k]$. The following was proved by Scott:

Theorem 17. *Let $\varepsilon \in \mathbb{R}^+$ and $\ell \in \mathbb{Z}^+$, and let G be a graph of density p. Then there exists $L = L(\varepsilon, \ell)$ such that G has a relatively ε-regular partition with k parts where $\ell \leq k \leq L$.*

The proof follows the same template as the proof of the regularity lemma, except that a different energy function is employed. If we merely defined the energy function using the index $q(A,B) = \frac{1}{p^2n^2}d(A,B)^2|A||B|$, we would run into trouble with very dense pairs (A,B), and the energy function $q(\mathcal{P}) = \sum_{A,B\in\mathcal{P}} q(A,B)$ would not be bounded. One therefore uses instead the index

$$q(A,B) = \min\Big\{\frac{1}{p^2n^2}d(A,B)^2|A||B|, \frac{8}{n^2\varepsilon^2}(\frac{d(A,B)}{p^2} - \frac{8}{\varepsilon^2})|A||B|\Big\}.$$

This has the effect of discounting very dense pairs, and now $q(\mathcal{P}) = \sum_{A,B\in\mathcal{P}} q(A,B)$ can be shown to be bounded above by a function of ε alone. The proof then follows that of Theorem 1 exactly, except that Lemmas 15 and 16 are slightly more technical.

4.1 An application

One of the key problems in applying Theorem 17 is that there is no analog of the embedding lemma. It is once more useful to construct the cluster graph R of a given relatively ε-regular partition of a graph, but it is no longer true that if R contains a triangle, then the original graph contains a triangle. In fact the original graph may be so sparse as to contain no triangles, despite R being a complete graph. For instance, the original graph may be a random triangle-free graph. However, an application was found by Allen, Keevash, Sudakov and the author [1]:

Theorem 18. *For large enough n, every extremal $\{C_4, C_5\}$-free n-vertex graph is bipartite.*

A possible extension of Turán's Theorem to sparse graphs is suggested as follows. For any $r \geq 3$ there exists t such that whenever F_1 is a graph with $\chi(F_1) = r+1$, and F_2 is a graph with $\mathrm{ex}(n, F_2) = \Omega(n^{2-1/t})$, then from every extremal $\{F_1, F_2\}$-free graph G we may delete $o(e(G))$ edges to obtain an r-partite graph. If $\chi(F_2) > 2$, this is true by the Erdős-Stone Theorem, so the problem is really when F_2 is bipartite, say $F_2 = K_{t,t}$. The above theorem shows that this holds when $r = 2$ and $F_1 = C_5$, $F_2 = C_4$ and $t = 2$, since $\mathrm{ex}(n, C_4) \sim n^{3/2}$.

References

[1] P. Allen, P. Keevash, B. Sudakov and J. Verstraete, Turan numbers of bipartite graphs plus an odd cycle, J. Combinatorial Theory Ser. B 106 (2014), 134–162.

[2] F. Behrend, On sets of integers which contain no three terms in arithmetical progression, Proc. Nat. Acad. Sci. U. S. A. 32, (1946) 331–332.

[3] W. Brown, P. Erdős and V. Sós, On the existence of triangulated spheres in 3-graphs and related problems, Periodica Mathematica Hungaria, 3 (1973), 221–228.

[4] V. Chvátal, V. Rödl, E. Szemerédi, W. Trotter, The Ramsey number of a graph with bounded maximum degree, J. Combin. Theory Ser. B 34 (1983), 239–243.

[5] P. Erdős and A. H. Stone, On the structure of linear graphs, Bulletin of the American Mathematical Society 52 (1946), 1087–1091.

[6] P. Erdos, P. Turán, On some sequences of integers, Journal of the London Mathematical Society 11 (1936) (4) 261–264.

[7] J. Fox, A new proof of the graph removal lemma, Ann. of Math. 174 (2011) 561–579.

[8] P. Frankl, V. Rödl, Extremal problems on set systems, Random Structures and Algorithms 20 (2002), no. 2, 131–164.

[9] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, J. D'Analyse Math. 31 (1977) 204–256.

[10] W. T. Gowers, Lower bounds of tower type for Szemerédi's uniformity lemma, Geometric and Functional Analysis 7 (2) (1997) 322–337.

[11] W. T. Gowers, Hypergraph regularity and the multidimensional Szemerédi theorem, Annals of Mathematics, 166 (2007), 897–946.

[12] B. Green, Roth's theorem in the primes, Ann. of Math. (2) 161 (2005), no. 3, 1609–1636.

[13] B. Green, T. Tao, The primes contain arbitrarily long arithmetic progressions, Annals of Mathematics 167 (2) (2008), 481–547.

[14] K. F. Roth, On certain sets of integers, I, Journal of the London Mathematical Society 28 (1953) 104–109.

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

[16] T. Sanders, On Roth's theorem on progressions, Ann. of Math. (2) 174 (2011), no. 1, 619–636.

[17] A. Scott, Szemerédi's Regularity Lemma for matrices and sparse graphs, Combin. Probab. Comput. 20 (2011) 455–466.

[18] E. Szemerédi, Regular partitions of graphs, in: Problèmes Combinatoires et Théorie des Graphes, Orsay, in: Colloq. Int. Cent. Natl. Rech. Sci., vol. 260, 1976, pp. 399–401.

[19] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, Acta Arithmetica 27 (1975) 199–245.