Applications of the Szemerédi Regularity Lemma

Shagnik Das

Opening remarks

It is difficult, if not impossible, to overstate the importance and utility of Szemerédi's celebrated Regularity Lemma in Extremal Graph Theory, and yet I have never been one to shy away from a challenge. To phrase it in terms of wider appeal, if we were doctors[1], the Regularity Lemma would be our antibiotics, the first and last line of defence against a bacterial catastrophe.[2] If we were engineers, the Regularity Lemma would be a bottomless can of WD-40 combined with an endless roll of duct tape, the solution to all our problems.[3] If we were movie producers, the Regularity Lemma would be Meryl Streep, guaranteeing our success.[4]

I could continue with these preliminary metaphors,[5] but during my formative years I was often told[6] in school that one should "show, not tell." Thus, rather than simply go on and on about how great the Regularity Lemma is, I intend to give you a glimpse of the Lemma in action through a couple of carefully-chosen applications. This selection is far from exhaustive[7] — think of it more as an appetiser in the feast that is the Regularity Lemma. For a more thorough exploration of the subject, I encourage you to turn to the excellent survey of Komlós and Simonovits[8].

The Regularity Lemma (and its lemmas)

It would be logical, I suppose, to start by introducing the Regularity Lemma itself, and since this is meant to be a mathematical document, one ought to take the logical path. This section thus contains the necessary definitions, the statement of the Regularity Lemma, and a few useful lemmas that help in applications.

Regularity and partitions

Informally, the Regularity Lemma tells us that the vertices of *any* large graph can be partitioned into a bounded number of parts, with the subgraph between most pairs of parts looking random. In order to state the Regularity Lemma formally, we must define what it means to 'look random.'

Definition (Density)**.** *Given a graph $G = (V, E)$, and two disjoint non-empty sets of*

[1] The useful, medical kind, not the degreed philosophers.

[2] `http://www.bbc.com/news/health-21702647`

[3] `https://www.flickr.com/photos/dullhunk/7214525854`

[4] `http://www.imdb.com/name/nm0000658/awards`

[5] Actually, I couldn't, which is why I started this new paragraph.

[6] Ironically!

[7] About as far as San Francisco, USA, is from Cairns, Australia: 209 days of rowing.

[8] `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.31.2310`

vertices $X, Y \subset V$, we define the density *of the pair (X, Y) as*

$$d(X,Y) = \frac{e(X,Y)}{|X|\,|Y|} = \frac{|\{(x,y) : x \in X, y \in Y, \{x,y\} \in E\}|}{|X|\,|Y|}.$$

Definition (ε-regularity)**.** *Let $\varepsilon > 0$ be fixed. Given a graph $G = (V, E)$, and two disjoint non-empty sets of vertices $A, B \subset V$, we say the pair (A, B) is ε-regular if, for every $X \subseteq A$ with $|X| \geq \varepsilon |A|$ and every $Y \subseteq B$ with $|Y| \geq \varepsilon |B|$, we have*

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

In other words, the edges in an ε-regular pair are distributed very uniformly, with the density between any pair of reasonably large subsets of vertices being very close to the overall density of the pair. This uniform distribution of edges is typical in a random bipartite graph, and captures what we mean when we say a (bipartite) graph 'looks random.' It remains to define what kinds of partitions of the vertices we will be concerned with.

Definition (Equipartition)**.** *A partition $V = V_0 \sqcup V_1 \sqcup \ldots \sqcup V_k$ of a vertex set into disjoint parts, including an* exceptional set *V_0, is said to be an* equipartition *if $|V_1| = |V_2| = \ldots = |V_k|$.*

Note that we do not require the exceptional set V_0 to be the same size as the other parts, as that would require the number of vertices $|V|$ to be divisible by $k+1$.[9] Finally, we have everything in place to define the kind of partition that the Regularity Lemma will provide us.

Definition (ε-regular partition)**.** *Let $\varepsilon > 0$ be fixed, and let $G = (V, E)$ be a graph. An ε-regular partition is an equipartition $V = V_0 \sqcup V_1 \sqcup \ldots \sqcup V_k$ such that $|V_0| \leq \varepsilon |V|$, and all but at most εk^2 pairs (V_i, V_j), $1 \leq i < j \leq k$, are ε-regular.*

Note that the parameter ε plays three roles[10] here: bounding the size of the exceptional set V_0, bounding the number of irregular pairs, and controlling the regularity of the regular pairs.

The statement of the Lemma

Observe that in an ε-regular partition, we have control over the distribution of edges between the ε-regular pairs, but not over the edges within any of the parts, involving the exceptional set, or in irregular pairs. Thus, in light of the three roles described above, the smaller ε is, the greater our control over the distribution of edges in an ε-regular partition.

---

[9] In fact, the exceptional set is only included to avoid these divisibility issues. One could state the Regularity Lemma without an exceptional set, and instead require $||V_i| - |V_j|| \leq 1$ for all $1 \leq i < j \leq k$, but the form we will use is a little more convenient to work with.

[10] Some versions of the Regularity Lemma use three different parameters for these, but we will not require any such separation.

On the other hand, observe that there are a couple of partitions that are trivially ε-regular. The first is when $V_1 = V$; that is, when all the vertices are contained in a single part where their edges are unrestricted. Thus, in order to have a useful partition, we would like there to be many parts, so that most of the edges lie in an ε-regular pair between different parts, where we have some control over their distribution. However, another trivial ε-regular partition is when we consider every single vertex as its own part. In this case, every pair has density either 0 or 1, and is trivially ε-regular for every $\varepsilon > 0$, without giving any structural information about the graph at all. Thus we want each part to have enough vertices for us to be able to make use of the regularity. This is equivalent to not having *too* many parts.

The Regularity Lemma tells us that we can have everything we want[11]: every (large enough) graph has an ε-regular partition with small ε and a large but bounded number of parts.

Theorem 1 (The Regularity Lemma; Szemerédi, 1978)**.** *For every $\varepsilon > 0$ and every $t \in \mathbb{N}$, there exists an integer $T = T(t, \varepsilon)$ such that every graph $G = (V, E)$ on at least T vertices has an ε-regular partition $V = V_0 \sqcup V_1 \sqcup \ldots \sqcup V_k$, where $t \leq k \leq T$.*

Useful little lemmas

As indicated above, the ε-regular partition of a graph G guaranteed by the Regularity Lemma provides us with control over the distribution of the edges of G. The following lemmas, used in most applications of the Regularity Lemma, translate that control into more explicitly combinatorial terms.

Lemma 1 (Degree Lemma)**.** *Let (A, B) be an ε-regular pair with $d(A, B) = d$. For any $Y \subseteq B$, $|Y| \geq \varepsilon |B|$, we have*

$$\left|\{a \in A : \deg(a, Y) < (d - \varepsilon)|Y|\}\right| < \varepsilon |A|,$$

where $\deg(a, Y)$ *counts the number of neighbours of a in Y.*

This first lemma states that in an ε-regular pair of density d, most vertices in one part have close to the expected number of neighbours in any large enough subset of the other part.[12]

Proof. Let $X = \{a \in A : \deg(a, Y) < (d - \varepsilon)|Y|\}$. Observe that

$$e(X, Y) = \sum_{a \in X} \deg(a, Y) < (d - \varepsilon)|X||Y|,$$

and so $d(X, Y) < d - \varepsilon$. Thus $|d(X, Y) - d(A, B)| > \varepsilon$, violating the ε-regularity of (A, B) if $|X| \geq \varepsilon |A|$. Hence we must have $|X| < \varepsilon |A|$, as desired. □

[11] How rarely we hear these words!

[12] Here we only impose a lower bound on $\deg(a, Y)$, which is the bound we will need, but one can prove a corresponding upper bound very similarly.

The next lemma extends Lemma 1 by looking not just at the neighbours of one vertex, but the common neighbours of an s-tuple of vertices. In a random bipartite graph of density d on the vertices $A \sqcup B$,[13] we would expect a set of s vertices in A to have $d^s |Y|$ common neighbours in any set $Y \subset B$. This lemma again shows that this random behaviour carries over to ε-regular pairs as well. Note that in order to get this more detailed information (common neighbourhoods, and not just neighbourhoods), we require the subset Y to be larger than before.

Lemma 2 (Common Neighbourhood Lemma)**.** *Let (A, B) be an ε-regular pair with $d(A, B) = d$, and let $s \in \mathbb{N}$ be a positive integer. For any subset $Y \subseteq B$ with $(d - \varepsilon)^{s-1} |Y| \geq \varepsilon |B|$, we have*

$$\left| \{ \vec{a} = (a_1, a_2, \ldots, a_s) \in A^s : |Y \cap N(\vec{a})| < (d - \varepsilon)^s |Y| \} \right| < s\varepsilon |A|^s ,$$

where $N(\vec{a}) = \cap_{i=1}^s N(a_i)$ is the common neighbourhood of the vertices in $\vec{a}$.

Proof. This can be proven by induction on s, using Lemma 1 for the base case and Lemma 3 in the induction step. The details are left to the reader. □

The previous two lemmas show that the degrees of vertices and subsets of vertices in an ε-regular pair behave as they would in a random bipartite graph. Another key feature of random graphs is that their induced subgraphs are also random graphs, inheriting the same distribution. Our final lemma shows that a similar inheritance takes place in ε-regular pairs.

Lemma 3 (Slicing Lemma)**.** *Let (A, B) be an ε-regular pair with $d(A, B) = d$, and let $\alpha > \varepsilon$. If $X \subseteq A$, $|X| \geq \alpha |A|$, and $Y \subseteq B$, $|Y| \geq \alpha |B|$, then (X, Y) is an ε'-regular pair, where $\varepsilon' = \max \left\{ \frac{\varepsilon}{\alpha}, 2\varepsilon \right\}$, and $d(X, Y) = d'$ for some d' with $|d' - d| \leq \varepsilon$.*

Proof. First observe that since $|X| \geq \varepsilon |A|$, and $|Y| \geq \varepsilon |B|$, the ε-regularity of (A, B) gives $|d(X, Y) - d(A, B)| = |d' - d| \leq \varepsilon$.

Now for any $S \subseteq X$, $|S| \geq \varepsilon' |X| \geq \frac{\varepsilon}{\alpha} |X| \geq \varepsilon |A|$, and $T \subseteq Y$, $|T| \geq \varepsilon' |Y| \geq \varepsilon |B|$, the ε-regularity of (A, B) gives $|d(S, T) - d(A, B)| \leq \varepsilon$. Hence, by the triangle inequality,

$$|d(S, T) - d(X, Y)| \leq |d(S, T) - d(A, B)| + |d(A, B) - d(X, Y)| \leq 2\varepsilon \leq \varepsilon'.$$

Hence it follows that (X, Y) is ε'-regular, as desired. □

Note that, as one might expect, repeated applications of Lemma 3, which take us to smaller and smaller subsets of the original ε-regular pair (A, B), provide worse and worse control over the parameters of regularity. However, having some regularity allows one to apply Lemmas 1 and 2 to these pairs of small subsets.

[13] In the binomial random bipartite graph model with density d, every possible edge appears independently with probability d.

Arithmetic progressions and Roth's Theorem

One may wonder why Theorem 1, if it is as significant as I have earlier claimed, is merely called a 'Lemma.' The answer lies in the history of the Lemma/Theorem. In 1975, Szemerédi proved what is now known as Szemerédi's Theorem, resolving an old conjecture of Erdős and Turán regarding arithmetic progressions in subsets of integers. In proving this theorem, Szemerédi used a weaker form of Theorem 1, which appeared as a lemma. It was some time before the true value of the Regularity Lemma was realised, but in 1979 Szemerédi published the full version of Theorem 1, and the rest, as they say, is history.

One ought to pay tribute to the origins of the Regularity Lemma, and so the first application we shall study will be Szemerédi's Theorem (Theorem 2 below). While the full proof of Szemerédi's Theorem is famously complex[14], we will restrict our attention to the first non-trivial case, a result that was earlier proven by Roth.

Arithmetic progressions

The Erdős–Turán conjecture concerns one of the fundamental objects in Combinatorial Number Theory, the arithmetic progression.

Definition (Arithmetic progression)**.** *An arithmetic progression of length k, or k-AP, with common difference d, is a sequence of k integers $(a_0, a_1, \ldots, a_{k-1})$ such that, for every $1 \le i \le k-1$, $a_i - a_{i-1} = d$.*

For any natural numbers $a, d \in \mathbb{N}$, there is a uniquely-defined k-AP with common difference d starting at a, namely $(a, a+d, \ldots, a+(k-1)d)$. Even if we restrict ourselves to the finite domain $[n]$, the first n positive integers, we still find a large number of k-APs (when n is large).

Being extremal combinators, Erdős and Turán asked the natural question: how large can a subset $A \subset [n]$ be if it does not contain any k-APs? They conjectured in 1936 that a k-AP-free set must have density approaching 0 (as n tends to infinity), and this is what Szemerédi proved.

Theorem 2 (Szemerédi, 1975)**.** *Fix $k \in \mathbb{N}$ and $\alpha > 0$. There exists some $n_0 = n_0(k, \alpha)$ such that if $n \ge n_0$ and $A \subseteq [n]$, $|A| \ge \alpha n$, then A contains a k-AP.*

Observe that any two distinct numbers form a 2-AP, and hence the first non-trivial (and therefore interesting) case is when $k = 3$. This case was settled in 1953 by Roth, using techniques from analytic number theory.

Theorem 3 (Roth, 1953)**.** *For every $\alpha > 0$, there is some $n_0 = n_0(3, \alpha)$ such that if $n \ge n_0$ and $A \subseteq [n]$, $|A| \ge \alpha n$, then A contains a 3-AP.*

Although the Regularity Lemma came many years later, it provides a very clean and simple proof of Theorem 3, as we shall see below.

[14] Szemerédi proved the case $k = 4$ of Theorem 2 in 1969, a good six years before settling the general case, which perhaps gives some indication of its difficulty.

Proof of Roth's Theorem

At first sight, the Regularity Lemma does not seem like the right tool to use for Roth's Theorem, as the former deals with graphs while the latter is a statement about subsets of integers. The connection between the two comes through the construction of an auxiliary graph, where 3-APs in our subset correspond to triangles in our graph. We will then need the following consequence of Theorem 1, which we shall prove later.

Theorem 4 (Triangle Removal Lemma)**.** *For every $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) > 0$ such that if G is an n-vertex graph such that at least εn^2 have to be removed from G before it becomes triangle-free, then G has at least δn^3 triangles.*

Using this theorem, we can prove Roth's Theorem.

Proof of Theorem 3. Given $\alpha > 0$, let $\delta = \delta\left(\frac{\alpha}{36}\right)$ be as given by Theorem 4, and set $n_0 = n_0(3, \alpha) = \frac{\alpha}{216\delta} + 1$. Suppose we have some $n \geq n_0$, and a set $A \subseteq [n]$ with $|A| = \alpha n$. We will show that A must contain a 3-AP.

Define an auxiliary three-partite graph G on the vertices $V = V_1 \sqcup V_2 \sqcup V_3$, where $V_i = \{(j, i) : 1 \leq j \leq in\}$ for each $1 \leq i \leq 3$. Hence $V_1 \cong [n]$, $V_2 \cong [2n]$, and $V_3 \cong [3n]$, and we have $6n$ vertices in total. We now add edges as follows: for every $x \in [n]$ and $a \in A$, add a triangle of edges between the vertices $(x, 1)$, $(x + a, 2)$ and $(x + 2a, 3)$ (in V_1, V_2 and V_3 respectively). Observe that these triangles are all edge-disjoint, and hence we have $3n\,|A| = 3\alpha n^2$ edges. Furthermore, since these triangles are edge-disjoint, we have to remove different edges to destroy each of these triangles, and hence have to remove at least $\alpha n^2 = \frac{\alpha}{36}(6n)^2$ edges to make G triangle-free. By Theorem 4, G has at least $\delta(6n)^3 = 216\delta n^3$ triangles in total.

We now claim that these triangles in G correspond to 3-APs in A. Suppose we have a triangle on the vertices $(x, 1)$, $(y, 2)$ and $(z, 3)$ (each set V_i is an independent set, any triangle must consist of one vertex from each part). Since $(x, 1) \sim (y, 2)$, we must have $y = x + a_0$ for some $a_0 \in A$. Since $(x, 1) \sim (z, 3)$, we must also have $z = x + 2a_1$ for some $a_1 \in A$. Finally, since $(y, 2) \sim (z, 3)$, we have $z = y + a_2$ for some $a_2 \in A$. Eliminating x, y and z, we find $a_0 + a_2 = 2a_1$, or $a_2 - a_1 = a_1 - a_0$. Hence, if $a_1 > a_0$, (a_0, a_1, a_2) forms a 3-AP with common difference $d = a_1 - a_0$. If $a_1 < a_0$, then (a_2, a_1, a_0) forms a 3-AP with common difference $d = a_0 - a_1$.

The only case in which we do not find a 3-AP in A is when $a_0 = a_1$. However, these triangles are precisely the edge-disjoint triangles we started with, of which there are αn^2. Since G contains at least $216\delta n^3 > \alpha n^2$ triangles, we can find one for which $a_0 \neq a_1$, giving the desired 3-AP in A. □

Note that the proof actually shows G has $\Omega(n^3)$ triangles for which $a_0 \neq a_1$, which corresponds to A containing $\Omega(n^2)$ 3-APs (since each 3-AP contributes n different triangles, one for each choice of $x \in [n]$). Hence any set $A \subset [n]$ of positive density contains very many 3-APs. It remains one of the major open problems in Combinatorial Number Theory to determine how large the largest 3-AP-free subset $A \subset [n]$ can be. The current best-known bounds are of the form

$$n \exp\left(-c\sqrt{\log n}\right) \leq |A| \leq \frac{C\left(\log\log n\right)^4}{\log n} n, \text{ for some constants } c, C > 0.$$

Proof of the Triangle Removal Lemma

It remains to prove the Triangle Removal Lemma, which we shall do now, so that it[15] will remain no longer.

Proof of Theorem 4. Let ε and G be as in the statement of the theorem. First note that since we must remove εn^2 edges from G to make it triangle-free, G must have at least one triangle. Since $\delta n^3 \leq 1$ when $n \leq \delta^{-1/3}$, we are done unless $n > \delta^{-1/3}$. By choosing δ sufficiently small, we can ensure that n is large enough for the following arguments to apply.

In particular, we may assume $n \geq T = T(t, \varepsilon_0)$, where $\varepsilon_0 = \frac{\varepsilon}{10}$ and $t = \frac{1}{\varepsilon_0}$. Applying Theorem 1, we obtain an ε_0-regular parition $V_0 \sqcup V_1 \sqcup \ldots \sqcup V_k$ of $V(G)$, where the number of parts satisfies $t \leq k \leq T$. Our next step is typical of proofs using the Regularity Lemma: we find a 'clean' subgraph $G_0 \subset G$, which only contains edges in dense regular pairs of the ε_0-regular partition of G. Let G_0 be the subgraph of G that remains after we delete the edges:

(i) incident to the exceptional set V_0,

(ii) contained inside one of the parts V_i, $1 \leq i \leq k$,

(iii) contained in an irregular pair (V_i, V_j), $1 \leq i < j \leq k$, and

(iv) contained in an ε_0-regular pair (V_i, V_j) of density at most ε.

Observe that all edges in G_0 are contained in ε_0-regular pairs (V_i, V_j) of density greater than ε. We now bound the number of edges of G that were deleted. In Step (i), each exceptional vertex has degree at most n, and there are at most $\varepsilon_0 n$ exceptional vertices, and hence we lost at most $\varepsilon_0 n^2$ edges. In Step (ii), observe that $|V_i| \leq \frac{n}{k}$, and hence there are at most $\sum_{i=1}^{k} \binom{|V_i|}{2} \leq k \binom{\frac{n}{k}}{2} \leq \frac{n^2}{2k} \leq \frac{n^2}{2t} < \varepsilon_0 n^2$ edges deleted. Since there are at most $\varepsilon_0 k^2$ irregular pairs, we lose at most $\varepsilon_0 k^2 \left(\frac{n}{k}\right)^2 = \varepsilon_0 n^2$ edges in Step (iii). Finally, there can be at most $\binom{k}{2}$ regular pairs of low density, each of which has at most $\varepsilon \left(\frac{n}{k}\right)^2$ edges, and hence we lost at most $\varepsilon \binom{k}{2} \left(\frac{n}{k}\right)^2 \leq \frac{1}{2}\varepsilon n^2$ edges in Step (iv).

In total, therefore, G_0 has at most $\left(\frac{1}{2}\varepsilon + 3\varepsilon_0\right) n^2 < \varepsilon n^2$ fewer edges than G. As we have removed fewer than εn^2 edges, G_0 must still contain a triangle. Each vertex of this triangle must come from a different non-exceptional part of the partition; without loss of generality we may assume the parts are V_1, V_2 and V_3. We further know that each pair (V_i, V_j), $1 \leq i < j \leq 3$, is ε_0-regular with density greater than ε. We will now show that there must in fact be at least δn^3 triangles between these parts, for some appropriate value of δ.

Indeed, by Lemma 1, we know that all but at most $\varepsilon_0 |V_1|$ vertices in V_1 have at least $(\varepsilon - \varepsilon_0) |V_2|$ neighbours in V_2. Similarly, all but at most $\varepsilon_0 |V_1|$ vertices in V_1 have at least the same number[16] of neighbours in V_3. Putting these two statements together,

[15] The overdue proof, not the Lemma itself, which, like all mathematical truths, will endure for eternity, if not longer.

[16] Recall that $|V_1| = |V_2| = |V_3|$.

at least $(1-2\varepsilon_0)\,|V_1|$ vertices in V_1 have at least $(\varepsilon-\varepsilon_0)\,|V_2|$ neighbours in V_2 and at least the same number in V_3.

Let v_1 be any such vertex in V_1, and let N_2 and N_3 be its neighbourhoods in V_2 and V_3 respectively. Since $|N_i| \geq (\varepsilon-\varepsilon_0)\,|V_i| > \varepsilon_0\,|V_i|$ for $i \in \{2,3\}$, the ε_0-regularity of (V_2, V_3) implies $d(N_2, N_3) \geq d(V_2, V_3) - \varepsilon_0 \geq \varepsilon - \varepsilon_0$. Hence there are at least $(\varepsilon-\varepsilon_0)\,|N_2|\,|N_3|$ edges between N_2 and N_3, each of which extends to a triangle when adding v. Moreover, for each different choice of v, these triangles are distinct.

Altogether, therefore, we find at least

$$(1-2\varepsilon_0)\,|V_1|\,(\varepsilon-\varepsilon_0)\,((\varepsilon-\varepsilon_0)\,|V_2|)^2 = (1-2\varepsilon_0)(\varepsilon-\varepsilon_0)^3\,|V_1|^3$$

triangles between V_1, V_2 and V_3. Finally, note that $|V_1| = \frac{n-|V_0|}{k} \geq \frac{1-\varepsilon_0}{T}n$, and hence we have at least δn^3 triangles, provided we choose $\delta = (1-2\varepsilon_0)(1-\varepsilon_0)^3(\varepsilon-\varepsilon_0)^3 T^{-3}$. (Note that for this choice of δ, when $n > \delta^{-1/3}$ we have $n > T$, as required to apply Theorem 1.) □

From the combinatorial point of view, the Triangle Removal Lemma would certainly be a very interesting statement even without any number theoretic applications, and hence at the very least merits a few additional remarks. The first is to note that there is nothing special about the triangle, and one can prove a Removal Lemma for general graphs along the same lines — a worthwhile exercise!

Second, when one is proving something using the Regularity Lemma, one should always bear in mind that the Regularity Lemma is a very powerful tool.[17] The price one pays for this is that the constants involved are typically terrible. Thus, once you have proven the statement to be true, the next goal is to find a proof without the Regularity Lemma, which often will result in much better dependence of the parameters.[18] While we have not discussed the sizes of the parameters in the Regularity Lemma in this note, it is worth pointing out that there are now more direct proofs of the Removal Lemma with much better quantitative bounds.

Turán numbers and the Erdős–Stone–Simonovits Theorem

For our second[19] application, we return to our roots and visit one of the fundamental areas of Extremal Graph Theory: Turán Theory. As with our first example, the theorems here predate Theorem 1, but looking back through our Regularity-tinted glasses, we find this is precisely the kind of problem the Regularity Lemma helps us solve.

Turán numbers

The quintessential extremal question in Graph Theory asks, for a fixed graph F, how many edges an n-vertex graph can have without containing a copy of F.

[17] It is like playing a video game with the easiest level of difficulty and all its cheat codes activated.

[18] From a purely mathematical perspective, it is also nice to use the 'right' tools for the job, and, as one of my professors used to say, not use a sledgehammer to open a walnut.

[19] And last, because I want to finish this note before the course ends.

Definition (Turán number)**.** *The* Turán number *(or* extremal number*) of a graph F, denoted* $\mathrm{ex}(n, F)$*, is the maximum number of edges in an n-vertex F-free graph.*

Turán greatly extended Mantel's Theorem by determining the Turán numbers of all complete graphs, thus kickstarting the field of Extremal Graph Theory. In Turán's Theorem, the following graphs play a pivotal role.

Definition (Turán graph)**.** *The r-partite* Turán graph *on n vertices, denoted $T_{n,r}$, is the complete r-partite n-vertex graph, with vertex parts as equal as possible (and hence of size $\lfloor \frac{n}{r} \rfloor$ or $\lceil \frac{n}{r} \rceil$).*

Since $T_{n,r}$ is r-colourable, and K_{r+1} is not, it follows that $K_{r+1} \not\subseteq T_{n,r}$. Turán showed that $T_{n,r}$ is the largest graph not containing K_{r+1}, extending Mantel's result (which showed $T_{n,2} = K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}$ is the largest K_3-free graph).

Theorem 5 (Turán, 1941)**.** *For all $r, n \in \mathbb{N}$, $\mathrm{ex}(n, K_{r+1}) = e(T_{n,r})$. Furthermore, if G is a K_{r+1}-free graph with n vertices and $\mathrm{ex}(n, K_{r+1})$ edges, then $G \cong T_{n,r}$.*

The Erdős–Stone–Simonovits Theorem

Once we have Turán's Theorem, the natural[20] question to ask is what happens for other graphs; while it is certainly worth knowing the Turán numbers of all complete graphs, the fact remains that most graphs are incomplete. The Erdős–Stone–Simonovits Theorem provides an asymptotic answer, showing that the Turán number $\mathrm{ex}(n, F)$ is, in general, controlled by the chromatic number $\chi(F)$.

Theorem 6 (The Erdős–Stone–Simonovits Theorem; Erdős–Simonovits, 1966)**.** *Let F be a graph with $\chi(F) \geq 2$. Then $\mathrm{ex}(n, F) = \left(1 - \frac{1}{\chi(F)-1} + o(1)\right) \binom{n}{2}$.*

Note that if $\chi(F) = 2$ (that is, if F is bipartite), then Theorem 6 only tells us that $\mathrm{ex}(n, F) = o(n^2)$. Even determining the order of magnitude, let alone the asymptotics, of the Turán numbers of bipartite graphs remains one of the outstanding problems of Extremal Graph Theory, but that is a story for another day[21].

When $\chi(F) \geq 3$, Theorem 6 is much more satisfying, as it provides the asymptotic value of $\mathrm{ex}(n, F)$. The lower bound comes from the Turán graphs: since $\chi(T_{n,\chi(F)-1}) = \chi(F) - 1 < \chi(F)$, $F \not\subseteq T_{n,\chi(F)-1}$. The lower bound then follows from a straightforward calculation that shows $e(T_{n,\chi(F)-1}) = \left(1 - \frac{1}{\chi(F)-1} + o(1)\right) \binom{n}{2}$.

The work, then, comes in proving the upper bound: that dense enough graphs must contain a copy of F. However, this is where the strange naming of the theorem is explained. In 1966, Erdős and Simonovits observed that a theorem of Erdős and Stone, twenty years old at the time, easily gives a matching upper bound.

Theorem 7 (Erdős–Stone, 1946)**.** *For any $r, s \in \mathbb{N}$ and $\varepsilon > 0$, there exists an $n_0 = n_0(r, s, \varepsilon)$ such that if $n \geq n_0$ and G is a graph with n vertices and at least $\left(1 - \frac{1}{r-1} + \varepsilon\right) \binom{n}{2}$ edges, G contains the Turán graph $T_{rs,r}$.*

[20]Where here "natural" is defined as "what follows in this note."

[21]Indeed, for another course.

With this theorem in hand, we can now prove the Erdős–Stone–Simonovits Theorem.

Proof of Theorem 6. We already have the lower bound

$$\mathrm{ex}(n, F) \geq e(T_{n,\chi(F)-1}) = \left(1 - \frac{1}{\chi(F)-1} + o(1)\right)\binom{n}{2},$$

and now seek to show $\mathrm{ex}(n, F) \leq \left(1 - \frac{1}{\chi(F)-1} + o(1)\right)\binom{n}{2}$. This is equivalent to showing that for every $\varepsilon > 0$ and n sufficiently large, $\mathrm{ex}(n, F) \leq \left(1 - \frac{1}{\chi(F)-1} + \varepsilon\right)\binom{n}{2}$.

Fix $\varepsilon > 0$, let $r = \chi(F)$, and set $s = |V(F)|$. Observe that $F \subset T_{rs,r}$, since we can embed each colour class of an r-colouring of F into a different part of $T_{rs,r}$. By Theorem 7, if $n \geq n_0(r, s, \varepsilon)$, then any n-vertex graph G with at least $\left(1 - \frac{1}{r-1} + \varepsilon\right)\binom{n}{2}$ edges contains $T_{rs,r}$, and hence F, as a subgraph. Thus $\mathrm{ex}(n, F) < \left(1 - \frac{1}{r-1} + \varepsilon\right)\binom{n}{2}$, as desired. □

Proof of the Erdős–Stone Theorem

All the hard work, then, must come in the proof of Theorem 7. However, with the Regularity Lemma, it becomes a slice of cake! We will use the following proposition, whose proof we defer to the end of this section.

Proposition 1. *Let $r, s \in \mathbb{N}$ and $\varepsilon \in \left(0, \frac{1}{2}\right)$ be fixed. For each $1 \leq \ell \leq r$, set $\varepsilon_\ell = \left(\frac{\varepsilon}{4}\right)^{(\ell-1)s}$, $n_\ell = s\varepsilon_\ell^{-1}$ and $d_\ell = \frac{\varepsilon}{4} + \sum_{i=2}^{\ell} \varepsilon_i$. If $V_1, V_2, \ldots, V_r$ are disjoint sets of vertices in a graph G such that each set has size $|V_i| \geq n_r$ and each pair (V_i, V_j), $1 \leq i < j \leq r$, is ε_r-regular with density at least d_r, then $T_{rs,r} \subset G$.*

Proof of Theorem 7. If $r = 1$, then any s vertices provide a copy of $T_{s,1}$, and so we may take $n_0 = s$. If $s = 1$, then $T_{r,r} = K_{r-1}$, and so we are done by Theorem 5. Hence we may assume $r, s \geq 2$. Without loss of generality, we may further assume $\varepsilon < \frac{1}{2}$, as we may certainly reduce ε if needed. Set $n_0(r, s, \varepsilon) = (1 - \varepsilon_r)^{-1} n_r T(\varepsilon_r^{-1}, \varepsilon_r)$, where ε_r and n_r are as in Proposition 1, and $T(t, \varepsilon)$ is the function from Theorem 1. Let $n \geq n_0$, and let G be an n-vertex graph with at least $\left(1 - \frac{1}{r-1} + \varepsilon\right)\binom{n}{2}$ edges. Our goal is to show $T_{rs,r} \subset G$.

Applying Theorem 1 to G with parameters ε_r and $t = \varepsilon_r^{-1}$, we obtain an ε_r-regular partition $V_0 \sqcup V_1 \sqcup \ldots \sqcup V_k$ of $V(G)$, where $\varepsilon_r^{-1} \leq k \leq T(\varepsilon_r^{-1}, \varepsilon_r)$. We find a 'clean' subgraph G' of G by removing the edges:

(i) incident to the exceptional set V_0,

(ii) contained inside one of the parts V_i, $1 \leq i \leq k$,

(iii) contained in an irregular pair (V_i, V_j), $1 \leq i < j \leq k$, and

(iv) contained in an ε_r-regular pair (V_i, V_j) of density at most $\frac{3\varepsilon}{4}$.

In this process, we remove at most $\left(\frac{3\varepsilon}{8} + 3\varepsilon_r\right) n^2 \leq \frac{31\varepsilon}{32}\binom{n}{2}$ edges from G,[22] and hence $e(G') \geq \left(1 - \frac{1}{r-1} + \frac{\varepsilon}{32}\right)\binom{n}{2} > e(T_{n,r-1})$. By Theorem 5, $K_r \subset G'$.

The only edges of G' are contained in ε_r-regular pairs of density greater then $\frac{3\varepsilon}{4}$, which itself is at least d_r (where d_r is as in Proposition 1). Hence each vertex of the clique K_r must come from a different part in the partition, which we may assume to be the first r non-exceptional parts. Thus for $1 \leq i < j \leq r$, the pair (V_i, V_j) is ε_r-regular with density at least d_r. Moreover, we have $|V_i| = \frac{n-|V_0|}{k} \geq \frac{1-\varepsilon_r}{T} n \geq n_r$.

Therefore the conditions of Proposition 1 are satisfied, implying $T_{rs,r} \subset G' \subseteq G$. $\square$

To complete the chain of proofs leading to Theorem 6, we now prove Proposition 1.

Proof of Proposition 1. We prove this proposition by induction on r. The base case $r = 1$ is trivial, as $|V_1| \geq n_1 = s$, and any s vertices in V_1 form a copy of $T_{s,1}$.

Now suppose $r \geq 2$. Our plan is to find s suitable vertices in V_r, and then use induction on their common neighbourhoods in $V_1, V_2, \ldots, V_{r-1}$. By Lemma 2, there are at most $s\varepsilon_r |V_r|^s$ s-tuples of vertices in V_r with fewer than $(d_r - \varepsilon_r)^s n_r$ common neighbours in V_i, for each $1 \leq i \leq r-1$. Hence there are at most $(r-1)s\varepsilon_r |V_r|^s$ s-tuples of vertices in V_r with fewer than $(d_r - \varepsilon_r)^s n_r$ common neighbours in some V_i, $1 \leq i \leq r-1$. The number of s-tuples of vertices in V_r with a repeated vertex is at most $\binom{s}{2}|V_r|^{s-1} < s\varepsilon_r |V_r|^s$. Hence the number of s-tuples of s distinct vertices in V_r with at least $(d_r - \varepsilon_r)^s n_r$ common neighbours in V_i for each $1 \leq i \leq r-1$ is at least $(1 - rs\varepsilon_r)|V_r|^s > 0$.

Fix any such set S_r of s vertices, and for each $1 \leq i \leq r-1$, let V_i' be the set of common neighbours of S_r in V_i. We wish to apply induction to the sets $V_1', V_2', \ldots, V_{r-1}'$, finding subsets $S_1, S_2, \ldots, S_{r-1}$ that form a copy of $T_{(r-1)s,r-1}$. Adding the common neighbours S_r would then give the desired copy of $T_{rs,r}$ in G. First, though, we must ensure that the conditions of Proposition 1 are satisfied.

Observe that $d_r - \varepsilon_r = d_{r-1} \geq \frac{\varepsilon}{4}$. Hence each set has size at least $(d_r - \varepsilon_r)^s n_r \geq \left(\frac{\varepsilon}{4}\right)^s n_r = n_{r-1}$. Furthermore, by Lemma 3, each pair (V_i', V_j'), $1 \leq i < j \leq r-1$, is ε'-regular of density d', where $\varepsilon' = \max\left\{\frac{\varepsilon_r}{(d_r-\varepsilon_r)^s}, 2\varepsilon\right\} = \frac{\varepsilon_r}{(d_r-\varepsilon_r)^s} \leq \varepsilon_r \left(\frac{\varepsilon}{4}\right)^{-s} = \varepsilon_{r-1}$ and $d' \geq d_r - \varepsilon_r = d_{r-1}$. Thus the conditions are indeed satisfied. $\square$

Concluding statement

It is my sincere hope that this little note has helped developed your appreciation of the Szemerédi Regularity Lemma, and also shown you how the Regularity Lemma can be used to prove other theorems. Here, despite seeing a couple of famous consequences, we have barely scratched the surface of applications of the Regularity Lemma.[23] In closing, I wholeheartedly encourage you to explore the subject further!

[22] See the proof of Theorem 4 for details.

[23] If the set of applications of the Regularity Lemma were the Black Knight[24], it would no doubt say, "'tis but a scratch," and invite us to delve further, but it is not, so it will not.

[24] If it were instead the Dark Knight, it would give rise to one of the most epic film trilogies of our time. I, for one, would pay good money to see a series of movies on the Regularity Lemma, especially if DiCaprio were to play the role of ε.